2015

An exploration of semiotics of new auditory displays: A comparative analysis with visual displays

Myounghoon Jeon
Michigan Technological University

AN EXPLORATION OF SEMIOTICS OF NEW AUDITORY DISPLAYS: A COMPARATIVE ANALYSIS WITH VISUAL DISPLAYS

Myounghoon Jeon

Michigan Technological University
Mind Music Machine Laboratory
Houghton, MI 49931 USA
mjeon@mtu.edu

ABSTRACT

Communicability is an important factor of user interfaces. To address communicability, extensive research has been done on visual displays, whereas relatively little research has been done on auditory displays. The present paper attempts to analyze semiotics of novel auditory displays (spearcon, spindex, and lyricon) using Peirce's classification of signs: icon, symbol, and index. After the aesthetic developmental patterns of the visual counterparts are presented, semiotics of auditory cues is discussed with future design directions.

1. INTRODUCTION

Auditory displays can be defined as all intentional, non-speech audio that is designed to transmit information between a system and a user [1]. Even though it refers to non-speech audio, just as with speech, efficient and effective communicability is one of the most important factors to assess auditory displays. For the last two decades, there have been continuous efforts to guide researchers and practitioners to design auditory displays in a more systematic way in the ICAD (international community on auditory display) community [e.g., 2, 3]. However, more theoretical design background is still required, which will lead them to going beyond their empirical knowledge or personal preferences. In this aspect, a semiotic approach is expected to provide HCI and auditory displays with a better framework of communicability [4], but in the ICAD community only a couple of researchers have investigated its application to auditory displays [e.g., 5, 6].

The present paper attempts to exploratively analyze novel auditory cues (spearcon, spindex, and lyricon) used in the auditory user interfaces, compared to visual arts and displays – based on Peirce's semiotic framework [7], which has been widely applied to the analysis of fine arts and photographs, and even HCI [4]. This comparative analysis is expected to offer an opportunity to understand the status quo of auditory displays more systematically and shed light on a future research direction.

2. SIGNS: ICONS, INDICES, & SYMBOLS

Peirce classified signs as icons, indices, and symbols [7]. Icons refer to signs that work based on "similarity" between the sign and the referent (e.g., ordinary paintings of the object). Indices refer to signs that work based on "causality" (e.g., wet ground in the morning indicates rain at night), "proximity", or "trace" (e.g., an arrow for the next direction). Symbols, however, do not have such a natural connection between the sign and the referent. Symbols are the signs that are used based on convention and agreement. However, the relationship of these three categories is not just linear, but could be changed [7].

3. THE HISTORY OF PAINTINGS & PHOTOGRAPHS

A developmental pattern of visual arts does not follow the semantic similarity order (i.e., icons-indices-symbols), but follows a different order (i.e., icons-symbols-indices) [8][1]. Classical paintings are understood as "icons" – imitation of nature – until the 19th century. However, in the early 20th century painters attempted to manifest what is not seen beyond representing what can be seen. For example, through abstract (nonrepresentational) arts, paintings became "symbols" of higher level-world reality, abandoning resemblance to visible reality [8]. After the Second World War, paintings have returned to "trace" and "scent" of reality/nature by Abstract Expressionist or Informalism. This is a transition of paintings to "indices". According to Dubois [9], the trend of photographs also followed the same order. Similarly, the early theory of photographs considered photographs as a "picture" of the world or a "mirror" of reality (icon). In the 20th century, photographs were considered as a photographer's ideology and text, or a "transformation" of reality (symbol). Since 1980s, photographs have become a "trace" of reality (index). In sum, the icons-symbols-indices order seems to account for the developmental pattern of visual aesthetics.

4. VISUAL DISPLAYS

This Peirce's classification is so widely used that it has been applied to an analysis of visual displays of the contemporary user interfaces even though the order of the occurrence of each type is not clear and might not be the same as in visual arts. In the recent article, Nielsen [10] categorized visual icons into resemblance icons, arbitrary icons, and reference icons, which are equivalent to Peirce's classification.

[1] Here, some might postulate a dialectical development that the history goes to each extreme pole first and settles down somewhere in between. However, it might not necessarily work in such a way as we will see in the following sections. This is just a rough approximation of some researchers. Of course, various styles and hybrids of visual arts have co-existed at the same time.

4.1. Icons: Resemblance icons

Resemblance icons depict a physical object, which the icon is intended to represent. Using a picture of an envelope to represent a file of electronic mail would be a resemblance icon. The main challenge for a resemblance icon is to design an image that is visually similar to the intended object. It is not always easy given the small size of icons.

4.2. Symbols: Arbitrary icons

Arbitrary icons have arbitrary shapes that only have meaning by convention. Traffic signs are often arbitrary icons and may form a good source of computer icons because of their fairly standardized international use. For example, a warning triangle might be used as the icon for a warning message. Obviously, arbitrary icons are the hardest for users to learn, unless they are so widely used that the convention becomes second nature just as a language.

4.3. Indices: Reference icons

Reference icons describe an object, which by reference or analogy, represents the concept that the icon is intended to represent. To illustrate, using a picture of a clamp to represent a file-compression utility would be a reference icon (because it squeezes). It would be hard to come up with a good resemblance icon for file compression except through the use of a before–after combination of a large and a small document, but icons showing state changes are hard to understand.

The question is whether users will catch the mapping between the reference domain and the system domain. Are the two concepts sufficiently closely related in users' mental models that people will think of the system feature when they see the picture of the reference item? This question is precisely applied to auditory display. Will people think of the system feature when they hear the sound of the reference item?

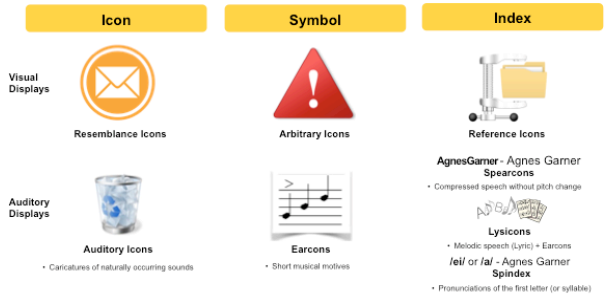

Figure 1: Comparisons of visual displays and auditory displays according to Peirce's semiotics framework.

5. AUDITORY DISPLAYS

Auditory displays have also been analyzed using Peirce's classification and seem to follow the similar developmental pattern of visual arts (i.e., icons-symbols-indices).

5.1. Icons: Auditory icons

Auditory icons are sounds that represent an object by capturing the object's essential features, such as functions and events [11]. Auditory icons can denote many objects in devices more clearly than some other auditory cues because the relation between the sound and the referent is direct (i.e., iconic relationship). For example, a typing sound can represent a typewriter or even a printer. Similar to resemblance icons, auditory icons typically require little training and are easily learned. Gaver [12] created an auditory icon-enhanced desktop. Other researchers have attempted to convert GUIs to non-visual interfaces using auditory icons [13]. However, it is sometimes difficult to match all functions of a device with proper auditory icons (e.g., "save" or "unit change").

5.2. Symbols: Earcons

Earcons are short, rhythmic musical motives, used to provide information to a user about some objects, operations or interactions [14]. Since earcons use an arbitrary mapping between the sound and the referent, they can be analogous to a language or a symbolic sign. This arbitrary mapping between earcons and represented items means that earcons can represent nearly any concept. However, this flexibility can also be a weakness because the arbitrary mapping of earcons to concepts requires user training. Brewster has conducted considerable research on applications of earcons and shown improved usability and user experience [e.g., 15].

5.3. Indices: Spearcons, lyricons, and spindexes

Since the birth of auditory icons and earcons, they have been analyzed in terms of semiotics and specifically, Peirce's framework [11, 14, 23, 25]. However, since then, we have had a number of novel auditory displays, which have not been analyzed in terms of semiotics perspective. Among new auditory cues that have recently appeared (e.g., auditory scroll bars, musicons, auditory emoticons, etc.), the present paper focuses on spearcons, spindexes, and lyricons, which are a type of hybrid auditory displays and placed between speech and non-speech cues. Such characteristics provide these auditory displays with a unique position crossing borders of the different semiotics categories. Therefore, it seems worth attempting to further analyze with the same framework used for the traditional auditory cues.

Spearcons are brief sounds that are produced by compressing spoken phrases, even to the point where the resulting sound might no longer be comprehensible as a particular spoken word [16] and thus, they are a non-speech sound cue. These sounds are analogous to fingerprints because of the unique acoustic relation between the spearcons and the original speech phrases. However, spearcons have no direct similarity to the referent (object) or the sounds that the referent generates. Spearcons are easy to learn because they derive from the original speech. Spearcons are easily created by algorithmically compressing Text-to-Speech. This allows the system to cope with dynamically changing items in the auditory system. For example, the spearcon for "save" can be easily extended into the spearcon for "save as." Spearcons have shown to enhance performance and preference for auditory menus [17].

Lyricons are devised by combining the concurrent two layers of musical speech (lyrics) and non-speech sounds (earcons) [18]. For example, the lyricon for FUNCTION ON of the electronic device can be composed of the speech, "func/tion/on" (lyric part) + the musical melody, "C(Do)/D(Re)/E(Mi)" (earcon part). This combination is expected to improve both semantics (with speech part) and aesthetics (with earcon part) of auditory user interfaces. It

can also improve learnability for the first-time users to operate the system more intuitively. Jeon and Sun [19] evaluated the effectiveness of lyricons compared to traditional earcons. Results showed that the average of accuracy rate of the lyricon group was almost double than that of the earcon group. The practical application of the lyricons implies *adaptability*; Once users get familiar with lyricons, they could use just the earcon part without the lyric part just as spearcons can be used without TTS [20].

A spindex (i.e., speech index) is a brief non-speech auditory cue based on the pronunciation of the first letter of each item [21]. To illustrate, the spindex cue for "Superstar" would be the sound /es/ or even /s/ based on the spoken sound of "S", the first letter of the item. The set of spindex cues in an alphabetical structure (e.g. address book on the mobile phone) is analogous to the visual index tabs that are often used to facilitate flipping to the right section of a thick reference book, such as a dictionary. Because spindex cues are part of the original word, they do not require much training. The subsequent studies showed that visually impaired users as well as sighted users can benefit from adding spindex cues to plain Text-to-Speech (TTS) menus, and they preferred the use of a spindex over plain TTS menus. [22].

5.4. Semiotic analysis of new auditory displays

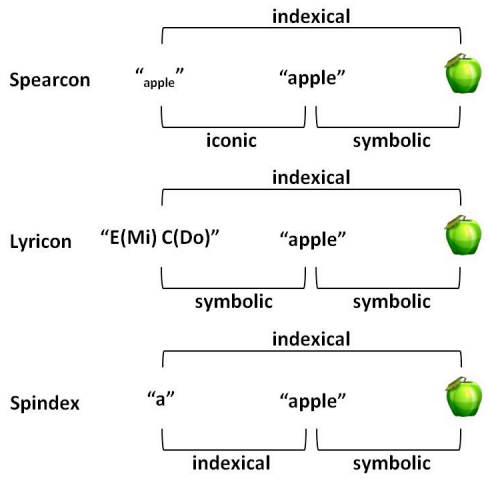

Figure 2: Relationships between auditory displays and an object according to Peirce's semiotics framework.

We can more deeply analyze new auditory displays using Peirce's framework. Based on its acoustic similarity (i.e., spearcons have almost same acoustical characteristics as the original speech. See [17] more details), a spearcon, "apple" has an iconic relationship with the original speech, "apple", which has a symbolic relationship with an actual apple (because speech is inherently symbolic). However, this symbolic relationship between the speech and an object is automatized based on life-long learning. With short training, the spearcon, "apple" can have an indexical relationship with an actual apple due to its trace of the actual apple even though users do not recognize the original speech (see Figure 1).

In the lyricon composed of "E(Mi)/C(Do)" + "apple", the earcon part, "E/C" has a symbolic (arbitrary) relationship with the lyric part, "apple". The lyric (i.e., melodic speech) part also has a symbolic relationship with an actual apple. However, with short training of the earcon part, "E/C" can have an indexical relationship [c.f., 23] with an actual apple because the earcon part would get to show trace of the lyric part based on the association obtained by a repetitive rehearsal while users use the system. In other words, even though the system removes the lyric part after the user is familiar with the system, the earcon part can remind the user of the lyric automatically.

The spindex, "a" has an indexical relationship with the speech, "apple" just like a visual arrow. Again, this speech has a symbolic relationship with an actual apple. Even though the spindex cue, "a" can have an indexical relationship with an apple, it might not be sufficiently indexical because it can also refer to other objects, which start with "a", such as an avocado. In other words, spearcons and lyricons can have a specific mapping between an auditory imagery and an actual object (apple), but spindex cues just match with a linguistic category (e.g., starting with "a"). Thus, the spindex does not provide sufficient indexicability about a specific object.

In sum, these three new hybrid auditory displays seem to belong to indexical cues when we try to apply Peirce's semiotics framework in that those cues have trace or information scent for the actual referent, rather than directly pointing to the actual referent.

Based on this analysis, our next step can be looking at how to strengthen the indexical relationship to give out more information trace to signs. For spearcons, optimizing a compression algorithm might be a good starting point (e.g., type, rate, etc.). Depending on users, tasks, and contexts, different compressions might give more indexicability. Given that much research on spearcons includes the address book application on the smartphone, an automatic reflection of a contact's gender, race, and age can make spearcons more indexical to a specific object (i.e., contact). For lyricons, more specified musical variables (e.g., pitch, timbre, rhythm, etc.) can be used to enhance lyricons' indexicability to trace a specific object. Moreover, we can design an earcon part by reflecting the innate accent and inflection of the lyric part. For example, if we design a lyricon of "apple," with two musical notes (e.g., E/C), we might use a higher pitch (E) and a lower pitch (C) in a sequence, rather than the reverse (C/E) to reflect the location of the accent of the original speech (see e.g., Deutsch's speech-to-song illusion, [26]). To enhance indexicability of the spindex, musical notes can be mapped onto each spindex cue, similar to auditory scrollbars. Spindex design can also be extended to the inclusion of the second syllable of the speech to distinguish a spindex from others (e.g., "ap" for apple vs. "av" for avocado).

6. CONCLUSION

Auditory displays inherently include a process of "translating" a data dimension into an auditory dimension. This representation and mapping process is a critical point for users to interpret the system. The more users correctly interpret the auditory signs of the system, the more communicable and usable the system would be. As Burks implies [24], elaborating "indexical" signs might be a right way to obtain the sophisticated mapping. An attempt to analyze new auditory displays through Peirce's semiotic prism helps us understand better the meaning of what has been done and glean some hints about what has to be done further and why. However, there might be some arguments whether this theory is the best option to explain the semiotics of new auditory displays – spearcons, lyricons, and spindexes. This question would be a good starting point for the next step of this research. I hope this attempt can

stimulate various discussions on semiotic aspects of auditory displays and contribute to improving both theory and practice in auditory display design.

7. REFERENCES

[1] B. N. Walker and G. Kramer, *Auditory Displays, Alarms, and Auditory Interfaces.* In W. Karwowski (Ed.), International Encyclopedia of Ergonomics and Human Factors (2nd ed.) (pp. 1021-1025). New York, USA: CRC Press, 2006.

[2] T. Hermann, A. Hunt, and J. G. Neuhoff, *The sonificatioin handbook*. Verlag, Germany: Logos, 2011.

[3] M. Jeon, "Tow or three things you need to know about AUI design or designers," in *Proc. of the Int. Conf. (ICAD2010,* Washington, USA, 2010.

[4] C. S. de Souza, The semiotic engineering of human-computer interaction. Cambridge, UK: MIT Press, 2005.

[5] P. Vickers, "Ways of listening and modes of being: Electroacoustic auditory display," *Journal of Sonic Studies,* vol. 2, No. 1, 2012.

[6] A. Pirhonen, E. Murphy, G. McAllister, W. Yu, "Non-speech sounds as elements of a use scenario: A semiotic perspective," in *Proc. of the Int. Conf. (ICAD2006)* (2006).

[7] C. S. Peirce, Logic as semiotic: The theory of signs. *Philosophical Writings,* 1902.

[8] J. K. Chin, *Image Humanities*, Seoul, Korea: Imagination of the Millennium, 2014.

[9] P. Dubois, L'acte photographique (K. R. Lee, trans.). Seoul, Korea: Photograph Art Books, 2005.

[10] J. Nielsen, Icon classification: Resemblance, reference, and arbitrary icons. *Nielsen Norman Group Article*.

[11] W. W. Gaver, "Auditory icons: Using sound in computer interfaces," *Human-Computer Interaction*, vol. 2, pp. 167-177, 1986.

[12] W.W. Gaver, "The SonicFinder, a prototype interface that uses auditory icons," *Human-Computer Interaction,* vol. 4, pp. 67-94, 1989.

[13] E. Mynatt, "Transforming graphical interfaces into auditory interfaces for blind users," *Human-Computer Interaction,* vol. 12, pp. 7-45, 1997.

[14] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, "Earcons and icons: Their structure and common design principles," *Human-Computer Interaction*, vol. 4, pp. 11-44, 1989.

[15] S. A. Brewster, P. C. Wright, and A. D. N. Edwards, "A detailed investigation into the effectiveness of earcons," in *Proc. of the Int. Conf. (ICAD1992),* 1992.

[16] B. N. Walker, A. Nance, and J. Lindsay, "Spearcons: Speech-based earcons improve navigation performance in auditory menus," in *Proc. of the Int. Conf. (ICAD2006),* 2006.

[17] B. N. Walker, J. Lindsay, A. Nance, Y. Nakano, D. K. Palladino, T. Dingler, and M. Jeon, "Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus," *Human Factors,* vol. 55, no. 1, pp. 157-182, 2013.

[18] M. Jeon, Lyricons (Lyrics + Earcons): Designing a new auditory cue combining speech and sounds. In C. Stephanidis (Ed.), *Posters, Part I, HCII 2013, CCIS,* vol. 373, Heidelberg, Germany: Springer, pp. 342-346, 2013.

[19] M. Jeon and Y. Sun, "Design and evaluation of lyricons (Lyrics + Earcons) for semantic and aesthetic improvements of auditory cues," in *Proc. of Int. Conf. (ICAD2014),* New York, USA, 2014.

[20] H. Suh, M. Jeon, and B. N. Walker, "Spearcons improve navigation performance and perceived speediness in Korean auditory menus," in *Proc. of Int. Conf. (HFES2012),* 2012.

[21] M. Jeon and B. N. Walker, "Spindex": Accelerated initial speech sounds improve navigation performance in auditory menus. *Proc. of Int. Conf. (HFES2009),* pp. 1081-1085, 2009.

[22] M. Jeon and B. N. Walker, ""Spindex" (Speech Index) improves acceptance and performance in auditory menu navigation for visually impaired and sighted users," *ACM Transactions on Accessible Computing,* vol. 3, no. 3, pp. 10:11-26, 2011.

[23] Y. Nam and J. Kim, "A semiotic analysis of sounds in personal computers: Toward a semiotic model of human-computer interaction," *Semiotica*, vol. 182, no. 1/4, pp. 269-284, 2010.

[24] A. W. Burks, "Icon, index, and symbol," *Philosophy and Phenomenological Research*, vol. 9, no. 4, pp. 673-689, 1949.

[25] D. Oswald, "Non-speech audio-semiotics: A review and revision of auditory icon and earcon theory," in *Proc. of Int. Conf. (ICAD2012)*, Atlanta, GA, USA, 2012.

[26] D. Deutsch, T. Henthorn, and R. Lapidis, “Illusory transformation from speech to song,” *J. Acoust. Soc. Am.*, vol. 129, no. 4, pp. 2245-2252, April 2011.