Let's Drink This Song Together: Interactive Taste-Sound Systems

Bruno Mesz
Department of Electronic Arts
Universidad Nacional de Tres de Febrero
Argentina
bruno.mesz@gmail.com

Kevin Herzog
Department of Electronic Arts
Universidad Nacional de Tres de Febrero
Argentina
kevh.bis@gmail.com

Juan Cruz Amusategui
Department of Electronic Arts
Universidad Nacional de Tres de Febrero
Argentina
juan.x.a@gmail.com

Lucas Samaruga
Department of Social Sciences
Universidad Nacional de Quilmes
Argentina
samaruga.lucas@gmail.com

Sebastián Tedesco
MUNTREF Arte y Ciencia
Univ. Nac. De Tres de Febrero
Argentina
sebastian.tedesco@gmail.com

ABSTRACT

We present three digital systems: the Augmented Glass, the Bone-Conduction Hookah, and the sound installation T2M, designed for displaying sound and taste stimuli, with applications in research on crossmodal taste-sound interactions, multisensory experiences and performances, entertainment and health.

CCS CONCEPTS

• **Human-centered computing → Human computer interaction (HCI) → Interaction devices → Sound-based input / output**

KEYWORDS

Crossmodal; music; taste; technology; multisensory; gastronomy.

ACM Reference format:

Bruno Mesz, Kevin Herzog, Juan Cruz Amusategui, Lucas Samaruga, Sebastián Tedesco. 2017. Let's Drink This Song Together: Interactive Taste-Sound Systems. *In Proceedings of 2nd ACM SIGCHI International Workshop on Multisensory Approaches to Human-Food Interaction (MHFI'17). ACM, New York, NY, USA, 5 pages.*
https://doi.org/10.1145/3141788.3141791

MHFI, Glasgow, United Kingdom

https://doi.org/10.1145/3141788.3141791

1 INTRODUCTION

A number of recent studies have analyzed the influence of music and sound on food or drink evaluation ([1], [2]). Crossmodal taste-sound congruence has been shown to have a perceptual impact on taste intensity and other qualities such as bitterness or sweetness ([3], [4]). As multisensorial integration is more likely to occur in conditions of spatial and temporal proximity between the stimuli in different modalities ([5], [6]), and music and taste are time-varying in nature, it seems interesting to develop technologies that favor these situations of synchronization and co-location, while allowing to keep track of multisensory interactive effects along time [7]. Some existing virtual reality technologies go in this direction, such as the straw-like interface of [8] or the Mouth Jockey [9].

The devices we present in Sections 2.1 and 2.2 (the Augmented Glass and the Bone-Conduction Hookah) are designed to deliver sound and/or subsonic vibrations exactly while the user is eating or drinking. The Hookah, moreover, locates vibrations on the lips, tongue and teeth, near where taste sensations appear subjectively to come (perhaps, vibration slightly modifies taste localization, see [10]).

Among several possible mechanisms underlying taste-sound correspondences, a relevant one is semantic matching [11]. This consists in the tendency to associate taste and sound stimuli that are described by the same or similar verbal labels (such as the word "sweet" used metaphorically to refer to the expressive character of a melody). Inspired by this fact, the multimedia system T2M seeks to exploit free verbal associations arising from taste experiences, using music-language semantic mappings to generate a sonic atmosphere for a food or drink tasting (see Section 2.3).

Along with the description of our digital systems, we describe in Section 2 a pilot study on the influence of music on the taste of coffee using a simple version of the Hookah; also, we mention possible uses of our systems for promoting healthy drinking or eating, and present some feedback from the users.

Some new technologies for human-food interaction change our food experiences by enhancing, or exploiting in some way, crossmodal taste-sound correspondences [7], [12]. Other applications of these technologies are connected to art, in the context of multisensory performances and installations involving food [13], [14]. In a similar vein, the devices and the multimedia system T2M presented here are mainly meant to be employed in experimental dining or degustations. We will briefly describe two of our projects in this field, the performance Beyond Intimacy and the installation Let´s Drink This Song Together.

2 THREE SYSTEMS FOR SOUND-TASTE INTERACTION

2.1: Augmented Glass

This device consists in a wine glass with sensors that are able to detect two main gestures of the user: when the cup is taken by the hand and when the liquid contacts the mouth while drinking. The sensing is performed through electrodes attached to the surface of the cup, next to the edge, and connected to capacitive sensors. These sensors generate an electric field and work by detecting other present fields interacting with it, for example that generated by the body. We extend the detection area to reach the hand holding the cup, using a special non-toxic metallic paint that is employed routinely in glassware. To avoid wiring, the sensor data are sent wirelessly to a computer. In this way the intervention of the object is minimal and its use is similar to that of a normal glass. When the hand and cup are both detected, a sound or vibration is activated (Fig. 1).

Figure 1: Augmented glass.

We implemented a pilot exploration of the Augmented Glass by making it activate motors and transducers placed in an upright piano (a video can be seen at https://www.youtube.com/watch?v=HWEBJT2vxCI), and asking users to drink red and white wine with different sounds (low and high pitched). From this initial experience we collected some opinions of the users, who in all cases enjoyed the interaction with the glasses: "this sound makes the wine feel like a Rutini [an expensive Argentine red wine], with strong body and expanding taste", "the low pitched sounds seem to give weight to the wine", "the pizzicatos [probably referring to the motors bouncing over the strings] bring forward sour fruit notes". Also high pitched sounds were judged to be more congruent to white wine, and low pitched sounds to red wine.

The initial motivation for designing the glass was a recent study [15] on the effect of music on wine taste, that employed time-based methodologies such as time-intensity curves and temporal dominance of sensations. Participants had to evaluate wine in short periods of time (30-45 sec.) while listening to soundtracks of classical and contemporary music. Significant effects of events such as climaxes and relaxations in the music on taste qualities were observed, so that the musical form organized, to some extent, the evolution of the taste curves. An illustrative video can be seen at https://www.youtube.com/watch?v=DIxBXXKZEFY.

The Augmented Glass could be useful, for this kind of studies, in achieving more precise synchronization between the music and the act of drinking, allowing to differentiate automatically the different stages of tasting and aftertaste, either by instructing participants to remove the glass from the lips each time they swallow, or detecting the swallowing by accelerometers placed on the neck. Other stimuli may be synchronized as well with each sip, such as lights, videos or smells.

Applications to healthy drinking behavior are also easy to implement: we imagine varying the music emotion or pleasantness with the number of sips, or giving an alert sound or vibration when some limit on this number has been exceeded.

2.2: Bone-conduction Hookah

This contraption, whose basic mechanism is similar to the straw-like user interface (SUI) [7], is meant to deliver vibrations on the lips, tongue and teeth while drinking or eating, and also transmit sound by teeth-bone conduction.

The Bone Conduction (BC) device in which the Hookah is based has three parts: a sound player, an amplifier and an actuator. The sound player can be a computer, a portable music player or a smartphone. We use a mono audio channel, and connect the device, via the mini plug output, to a class D amplifier giving a signal of 15 watts. This signal feeds the actuator, a continuous current motor attached to a straw, fork, spoon, etc.

2.2.1: Pilot study: coffee and sound. We are currently using the BC system in a study exploring the incidence of sound perceived through teeth conduction on coffee taste perception (see Fig. 3). We use two soundtracks previously found to be congruent with sweetness [16] and creaminess [17]. Participants taste five small cups of instantaneous coffee (without knowing it is the same coffee each time) in the following conditions: 1) perceiving the creamy soundtrack through the BC; 2) perceiving the sweet soundtrack through the BC; 3) listening to the creamy soundtrack with headphones; 4) listening to the sweet soundtrack with headphones; 5) in silence. They have to rate each time the creaminess, sweetness, bitterness and pleasantness of the coffee in 7-point scales, with ‘1’ being ‘not at all’, ‘4’ ‘neutral’ and ‘7’ ‘very much’. So far, 15 participants (6 females), aged from 21 to 37 years have completed the experiment. Analysis of results using RM-ANOVA, with conditions as factor, shows no statistically significant differences between the conditions. However, in all cases and for all dimensions substantial individual differences between the maximum and minimum value attributed to a given dimension was observed between the different conditions. For instance, for creaminess rating, the mean of these differences across participants is 3.33 (about half of the available punctuation range), with standard deviation 1.44. This is also confirmed by the reactions of the participants: ‘lots of change in sweetness and creaminess’, ‘arpeggiated sounds enhance the creaminess’, ‘I am sure that it is the same coffee but I perceive it differently each time’. So, although there is no reliable crossmodal effect in this pilot study, at the subjective level there is an impact of the sound on the taste, which we think is a relevant observation for aesthetic purposes. Since participants reported no differences between BC and headphones in the sound source localization or quality, we will add subsonic vibrations to the sound transmitted by BC in future experiments to look at their crossmodal and emotional effects on taste perception.

Also for this system, applications to healthy eating or drinking as mentioned in Section 2.1 could be easily implemented. The device could be employed for amusement, matching funny sounds to food and drink, or adapted to a toothbrush for fostering dental hygiene of children.

2.3: T2M

T2M is a generative text-based music system for interactive performances. T2M takes words from tweets produced by the audience or Twitter™ trending topics and composes music using Freesound™ sound-files (labeled with those words) as raw material. A demo using trending topics including the word “food” may be seen at https://www.youtube.com/watch?v=cRTWl5u-m8c

T2M explores the intra-sonic and extra-sonic semantics of sounds that are given by the tags of audio files in the Freesound database, with the purpose of generating soundscapes from textual input.

The motivation for this installation was research on music-taste crossmodal correspondences ([18]), which studied musical representations of the extra-musical semantic domain of taste words. Previous systems mapping text to music are for example [19] and [20]. Semantics of music in connection with language is a vast topic, see for instance [21]. Language is considered an important factor in crossmodal associations [22].

The base implementation for this installation (https://github.com/smrg-lm/ttm) is made with SuperCollider [23] as an external library (quark in SC jargon). It consist in a set of classes that organizes information retrieval, both from Freesound (http://www.freesound.org) and Twitter (https://twitter.com) using external command-line utilities on Linux, processing and storage of that information, sequencing, sound synthesis and visual feedback of original messages and retrieved sounds.

An autonomous T2M execution loop is schematically as follows: tweets query, text parsing and storage, sound query based on selected words, sound-files pull and storage, sound processing and reproduction and visual presentation of information.

Freesound querying and sound-file download depends on the Freesound.sc quark (https://github.com/g-roma/Freesound.sc) which is a higher level interface to communicate with the Freesound API from within sclang (SC programming language). Twitter CLI program (https://github.com/sferik/t) is used for querying tweets using Twitter’s API. A basic text parsing consisting in noun and verb selections is performed by a command-line utility provided by Apertium (http://wiki.apertium.org/wiki/Main_Page). See the flowchart in Fig. 2.

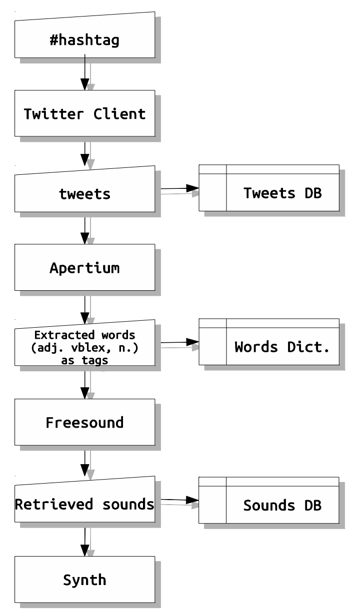

Figure 2: Flowchart of T2M.

We have presented T2M at the Museum of Contemporary Art in Mar del Plata, Argentina. Four specially designed food perfumes were successively spread in the auditorium, and the audience was required to tweet any associations or memories evoked by the aromas. The tweets were displayed on a screen while T2M generated a sonic atmosphere.

We envisage several aspects of the use of T2M in a gastronomic context. By asking people to exteriorize, via Twitter, free verbal associations produced by tasting a dish or a drink, we may encourage focusing attention on the sensorial evaluation of food. Moreover, by extracting terms related specifically to taste and flavor from the tweets, we can transform the musical flow in order to have music congruent (or contrasting) with the perceived taste, potentially modulating it ([3], [4]). This could be done for each guest, using individual loudspeakers. Another possibility is analyzing the affective content of the tweets using sentiment analysis and, by automatic evaluation of music emotion in the Freesound sound files, producing music sensitive to the emotional states of the diners. Again, using individual loudspeakers would permit to distinguish between different emotions expressed in the room, localizing the music near each commensal in a spatialized polyphony. Of course, some ordering of the sound events in time would be needed to avoid sonic chaos.

Figure 3: Pilot study with the bone conduction device.

4 FUTURE PERFORMANCES

We plan to use the Augmented Glasses for a multisensory performance (Beyond Intimacy). Our general concept will be exteriorizing the intimate sense of taste that inhabits the personal sphere of the inner body cavities, to generate a shared enveloping atmosphere of sound, dance, and smell. Beyond Intimacy will proceed in layers, beginning with the audience (limited to few people each time) drinking in silence, then activating piano sounds and smells, and finishing with live dance performance in the created multisensory landscape.

The Hookah design aims at a shared, convivial multisensory experience in an installation (Let´s Drink This Song Together). We decided to work based on a known object, the hookah, and transplant its shape and modality of use to our crossmodal object. This reference gives us two advantages; one is to implant a new functionality in a format that gives an orientation on its usage, the other is the possibility of transforming a subjective experience in a shared and relational one (Fig. 4)

Figure 4: Sketch of the Bone Conduction Hookah.

REFERENCES

[1] Spence, C. (2016). Sound–the forgotten flavour sense. Multisensory Flavor Perception: From Fundamental Neuroscience Through to the Marketplace, 81.

[2] Spence, C. (2012). Auditory contributions to flavour perception and feeding behaviour. *Physiology and behavior*, 107(4), 505-515.

[3] Reinoso Carvalho, F., Van Ee, R., Rychtarikova, M., Touhafi, A., Steenhaut, K., Persoone, D., ... and Leman, M. (2015). Does music influence the multisensory tasting experience?. *Journal of sensory studies*, 30(5), 404-412.

[4] Carvalho, F. R., Wang, Q. J., Van Ee, R., and Spence, C. (2016). The influence of soundscapes on the perception and evaluation of beers. *Food Quality and Preference*, 52, 32-41.

[5] Slutsky, D. A., and Recanzone, G. H. (2001). Temporal and spatial dependency of the ventriloquism effect. *Neuroreport*, 12(1), 7-10.

[6] Chen, L., and Vroomen, J. (2013). Intersensory binding across space and time: a tutorial review. *Attention, Perception, and Psychophysics*, 75(5), 790-811.

[7] Velasco, C., Carvalho, F. R., Petit, O., and Nijholt, A. (2016, November). A multisensory approach for the design of food and drink enhancing sonic systems. In *Proceedings of the 1st Workshop on Multi-sensorial Approaches to Human-Food Interaction* (p. 7). ACM.

[8] Yamamoto, A., ... and Inami, M. (2006, June). Straw-like user interface: virtual experience of the sensation of drinking using a straw. In *Proceedings of the 2006 ACM SIGCHI international conference on Advances in computer entertainment technology* (p. 50). ACM.

[9] Koizumi, N., Tanaka, H., Uema, Y., & Inami, M. (2011, November). Chewing jockey: augmented food texture by using sound based on the cross-modal effect. In Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology (p. 21). ACM.

[10] Todrank, J., and Bartoshuk, L. M. (1991). A taste illusion: taste sensation localized by touch. *Physiology and behavior*, 50(5), 1027-1031.

[11] Knöferle, K., and Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychonomic bulletin and review*, 1-15.

[12] Spence, C., Ranasinghe, N., Velasco, C., & Obrist, M. (2017). Digitizing the chemical senses: Possibilities & pitfalls. *International Journal of Human-Computer Studies*, 107, 62-74.

[13] Bruijnes, M., Huisman, G., & Heylen, D. K. J. (2016). Tasty Tech: human-food interaction and multimodal interfaces. DOI: http://dx.doi.org/10.1145/3007577.3007581

[14] Spence, C., & Piqueras-Fiszman, B. (2013). Technology at the dining table. Flavour, 2(1), 16.

[15] Wang, Q.J., Mesz, B., and Spence C. Analysing the impact of music on wine perception via TDS and TI (Poster presentation ot 12th Pangborn Sensory Science Symposium, August 2017).

[16] Wang, Q., Woods, A. T., & Spence, C. (2015). "What's your taste in music?" A comparison of the effectiveness of various soundscapes in evoking specific tastes. i-Perception, 6(6), 2041669515622001.

[17] Carvalho, F. R., Wang, Q. J., Van Ee, R., Persoone, D., & Spence, C. (2017). "Smooth operator": Music modulates the perceived creaminess, sweetness, and bitterness of chocolate. Appetite, 108, 383-390.

[18] Mesz, B., Trevisan, M. A., & Sigman, M. (2011). The taste of music. Perception, 40(2), 209-219.

[19] http://www.erikbunger.com/html/let_them_sing.html

[20] http://melobytes.com/app/melobytes

[21] Zbikowski, L. M. (2002). Conceptualizing music: Cognitive structure, theory, and analysis. Oxford University Press.

[22] Spence, C. (2011). Crossmodal correspondences: A tutorial review. Attention, Perception, & Psychophysics, 73(4), 971-995.

[23] Mccartney, J.(2002). Rethinking the Computer Music Language: SuperCollider. Computer Music Journal, 26(4), 61-68.