



# A composition algorithm based on crossmodal taste-music correspondences

Bruno Mesz<sup>1</sup>, Mariano Sigman<sup>2</sup> and Marcos A. Trevisan<sup>3\*</sup>

<sup>1</sup> Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Bernal, Buenos Aires, Argentina

<sup>2</sup> Laboratory of Integrative Neuroscience, Physics Department, FCEN, University of Buenos Aires. Ciudad Universitaria, Buenos Aires, Argentina

<sup>3</sup> Laboratory of Dynamical Systems, IFIBA-Physics Department, FCEN, University of Buenos Aires. Ciudad Universitaria, Buenos Aires, Argentina

## Edited by:

Robert J. Zatorre, McGill University, Canada

## Reviewed by:

Mikko E. Sams, Helsinki University of Technology, Finland

Jorge Moll, D'Or Institute for Research and Education (IDOR), Brazil

## \*Correspondence:

Marcos A. Trevisan, Laboratorio de Sistemas Dinámicos, Departamento de Física, IFIBA, Pabellón I, Ciudad Universitaria, 1428EGA, Buenos Aires, Argentina.  
e-mail: marcos@df.uba.ar

While there is broad consensus about the structural similarities between language and music, comparably less attention has been devoted to semantic correspondences between these two ubiquitous manifestations of human culture. We have investigated the relations between music and a narrow and bounded domain of semantics: the words and concepts referring to taste sensations. In a recent work, we found that taste words were consistently mapped to musical parameters. Bitter is associated with low-pitched and continuous music (*legato*), salty is characterized by silences between notes (*staccato*), sour is high pitched, dissonant and fast and sweet is consonant, slow and soft (Mesz et al., 2011). Here we extended these ideas, in a synergistic dialog between music and science, investigating whether music can be algorithmically generated from taste-words. We developed and implemented an algorithm that exploits a large corpus of classic and popular songs. New musical pieces were produced by choosing fragments from the corpus and modifying them to minimize their distance to the region in musical space that characterizes each taste. In order to test the capability of the produced music to elicit significant associations with the different tastes, musical pieces were produced and judged by a group of non-musicians. Results showed that participants could decode well above chance the taste-word of the composition. We also discuss how our findings can be expressed in a performance bridging music and cognitive science.

**Keywords:** music, taste, cross-modal associations, musical algorithm, semantics

## INTRODUCTION

Music and language involve the production and interpretation of organized complex sound sequences (Patel, 2003; Jentschke et al., 2005, 2008; Shukla et al., 2007; Zatorre et al., 2007). While there is broad consensus about the structural similarities between language and music, comparably less attention has been devoted to semantic correspondences between these two ubiquitous manifestations of human culture.

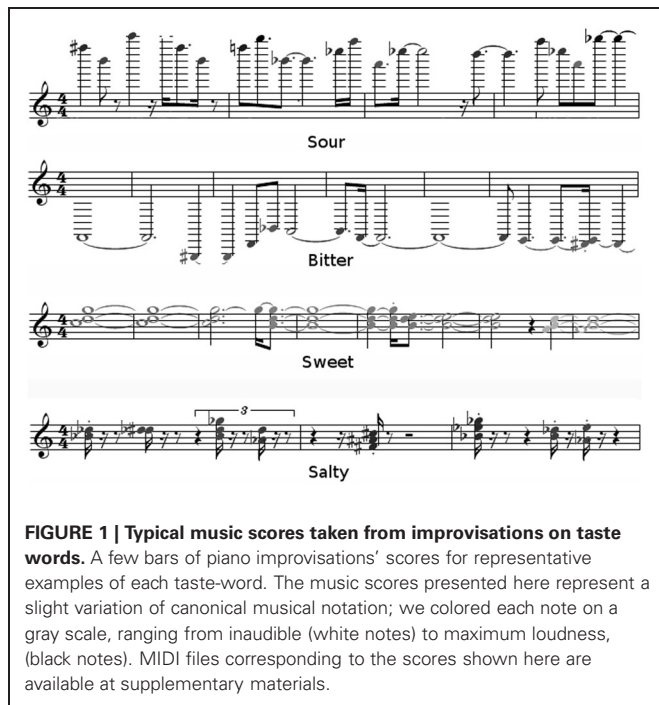
From a semantical perspective, a common view is that language is connected primarily with propositional content, while music is thought to refer to the inner world of affects and perceptions (Jackendoff and Lerdahl, 2006). Koelsch and collaborators showed that, instead, musical expressions have the capacity to activate brain mechanisms related to the processing of semantic meaning (Koelsch et al., 2004). Sound symbolism theory of language (Ramachandran and Hubbard, 2003; Simner et al., 2010) suggests that, along the history of language, the assignation of sense to vocal sounds relies on shared cross-modal associations, which consistently map sounds to experiences from other sensory modalities.

Cross-modal correspondences have been reported between audition and vision, audition and touch, audition and smell (Driver and Spence, 2000), interconnecting audition with virtually every other sense. These correspondences involve basic sound

features common to speech and music, such as pitch and visual size (Parise and Spence, 2008; Evans and Treisman, 2010) or brightness and pitch (Spence, 2011). It seems plausible that this mechanism extends beyond low-level sensory features, intervening in the assignation of “meaning” to music, at least in the broad semiotic sense of meaning: a musical event pointing to something different than itself. Language may hence share with music a privileged window into the mind in its proposed double role of mediator and shaper of concepts (Chomsky, 1988).

Specifically, associations between music, language and taste have been reported in several recent studies (Crisinel and Spence, 2009, 2010a,b; Simner et al., 2010; Mesz et al., 2011). Beyond the strictly acoustic level, some conventional metaphors like “sour note” or “sweet voice” involve taste-sound mappings. A well-established taste-sound synaesthetic metaphor in music vocabulary is the use of the Italian term *dolce* (sweet) to indicate soft and delicate playing. Although very infrequent, other taste words appear as expressive indications in music, for example “*âpre*” (bitter) in *La puerta del vino* of Debussy (Debussy, 1913). This example exhibits a low pitch register and a moderate dissonance.

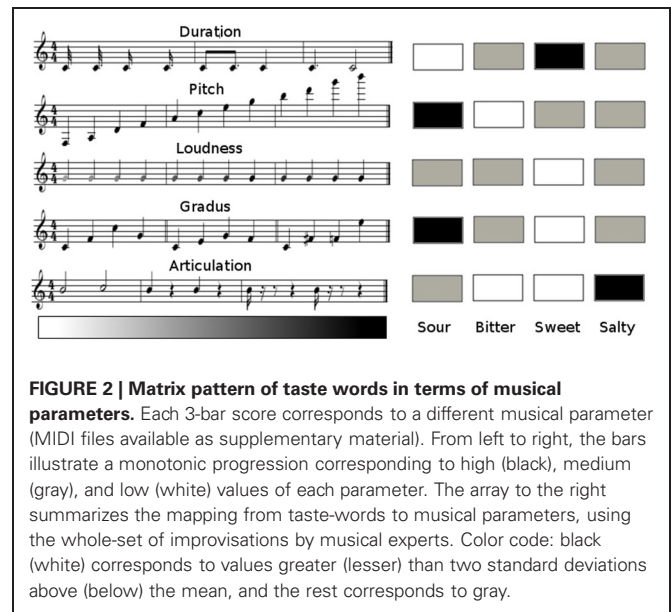
In Mesz et al. (2011) we investigated how taste words elicited specific representations in musical space. Highly trained musicians improvised on a MIDI keyboard upon the presentation of the four canonical taste words: sweet, sour, bitter, and



salty (Figure 1). Right after the improvisations, they also were asked to write down the words that came to their minds during the performance. Our results showed that sweetness is related to melodiousness and softness, sourness with dissonance or inharmonicity and the musical representation of “bitter” is low-pitched and slightly dissonant. This was quantified by measuring five canonical musical parameters: *loudness*, *pitch*, *duration*, *articulation* (ranging from a continuum of notes, corresponding to *legato*, low articulation, to sharply detached notes, corresponding to *staccato*, high values of articulation) and Euler’s Gradus Suavitatis, henceforth called *gradus*, a measure associated with psychoacoustic pleasantness (detailed explanation of the meaning of each musical parameter is supplied as supplementary material, see *glossary in the supplementary material*). Note that *gradus* was measured only for the upper voice, which usually carries the main melodic component.

Taste-words were mapped to consistent directions in this five-dimensional musical space. Bitter is low-pitched and legato, salty is staccato, sour is high pitched, dissonant and fast and sweet is consonant, slow and soft (Figure 2). When we played back this improvisations to naive listeners they could decode well above chance the taste-word that had triggered the improvisation revealing a symmetrical mapping between musical forms and semantics, in the specific domain of taste-words.

These studies reveal the pertinence of music as a vehicle to investigate semantic organization. A synergistic dialog between art and science calls also for the creation of music from taste words. Here we embark on this venture, developing, and implementing an algorithm that creates flavored musical pieces. The algorithm applies a sequence of musical operations to an original improvisation, transforming it according to motives extracted from a large corpus of classic and popular songs. The successive



steps are sketched in Figure 3 and described in section “Materials and Methods.” With this procedure we intend to produce music within specific regions of musical space, in the form of pieces that stimulate a referential listening, bringing the listener memories of other pieces, a mosaic of reminiscences remixing well-known music. Our aim is to use the improvisations as seeds pointing to relevant musical directions and increase their expressive power incorporating gestures from the musical tradition.

To test the performance of this algorithmically composed music, we designed and ran an experiment in which we asked a group of participants to judge the taste-words evoked by four musical pieces produced by the algorithm, one for each taste-word.

## MATERIALS AND METHODS

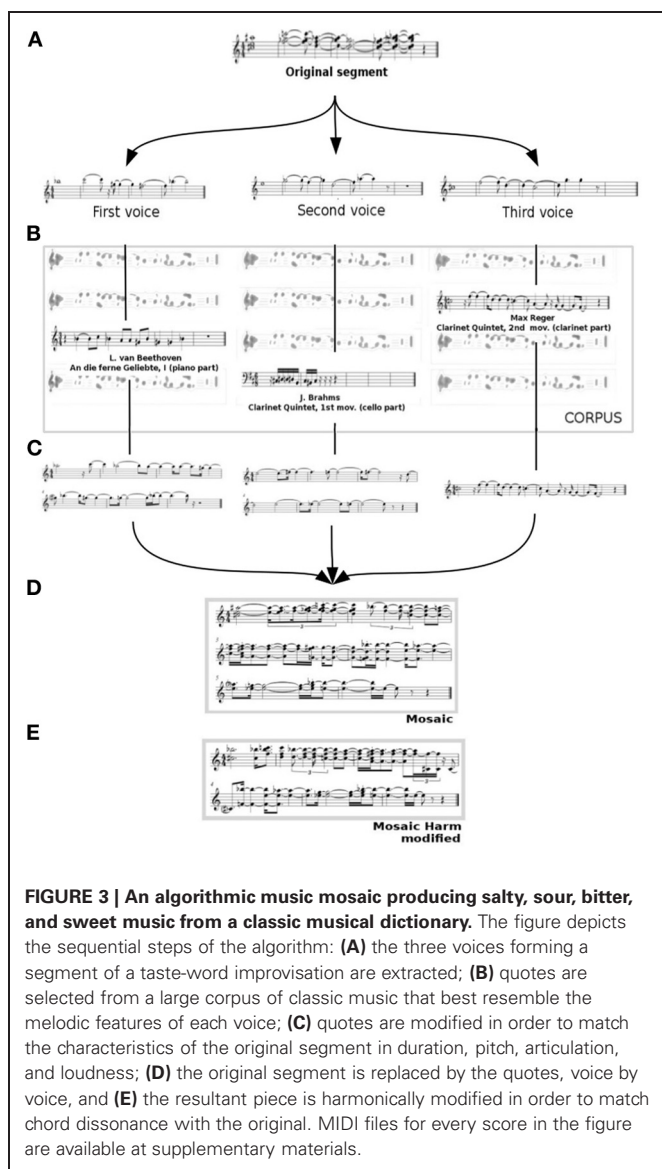
### ALGORITHM DESCRIPTION

#### Segmentation

The original score is separated in successive segments that vary from 3 to 10 s, about the size of the estimated auditory sensory memory (Fraisse, 1982). For this sake we use a boundary detection function implemented in the MIDIToolBox (Toolbox URL), which assigns a local boundary strength to each note in the score. Boundary strength of a note depends on various musical factors such as a leap in pitch to the following note, a prolonged duration and/or being followed by a long rest. Boundaries can be detected as statistical outliers in this distribution, as conventionally done in stream segmentation (the original score and segmentation points can be found at supplementary materials, indicated by a double bar symbol above the staff). A segment is shown in Figure 3A.

#### Voice extraction

Each polyphonic segment is split in monophonic *voices*. At every multiple note onset, all simultaneous notes ( $N_1, N_2, \dots, N_m$ ) are ordered decreasingly in pitch (as shown in Figure 3A for the



first three simultaneous notes  $N_1$ ,  $N_2$ , and  $N_3$  of the original segment). Voices are easily constructed by peeling progressively the highest melody.  $N_1$  is assigned to voice  $V_1$ ,  $N_2$  to voice  $V_2$  and so on.

#### Quotation from the database of classical and popular MIDI files

Each voice in the segment is replaced by the most similar melodic fragment from a large database of classical and popular music MIDI files, segmented as in section “Segmentation” (Figure 3B). We use two databases, a local one of about 250 pieces and the Petrucci Music Library (Petrucci URL), hosting more than 150,000 pieces. We then search throughout the database for the sequence with most similar melodic contour to the original voice. This is done as follows: we first determine the sequence of melodic intervals between successive notes of the voice. We then search over the corpus for an identical interval sequence. If several matches are found, one is selected at random and the segment

that contains it is the chosen quote (we quote a whole segment, which is usually a musical phrase or sub-phrase, in order to keep the context and make the musical reference more recognizable). If no sequence is found, we consider the sub-sequence obtained deleting the last interval, and proceed iteratively until a match is reached. In the case of using the Petrucci Music Library we employ an on-line search engine (Engine URL).

Note that although the melodic contour of the quote is similar to that of the voice, the pitch, duration and the rest of the parameters are not. Since we want to preserve these parameters, we perform the following transformations on the quotes to match the characteristics of the original segment (Figure 3C):

- a transposition to match the pitch;
- a change of time scale to match the duration;
- a change in the articulation, to match the interval between successive notes and
- a change in velocity to match the mean loudness.

We use only a fraction of the voices (normally 3 or 4), which reduces the polyphonic density and increases counterpoint clarity.

#### Mosaic and harmonic control

Mosaic and harmonic control is the final step of the algorithm. First, the quotes are superposed, synchronizing their beginnings (Figure 3D). We aim also to control the degree of vertical dissonance, which we measure by weighing the components of the interval vector of each chord according to a measure of the dissonance of the corresponding harmonic interval (see harmonic dissonance in the glossary in the supplementary material). Then, whenever some chord in the mosaic exceeds the dissonance bounds of the chords of the original segment, we replace it by a chord within these bounds, keeping the upper voice of the mosaic to maintain its *gradus* while eventually moving the other voices, always less than one octave (Figure 3E).

In a second version of the algorithm we just take as target parameter range the mean and standard deviation of parameters over all improvisations of a given taste and apply the operations of segmentation and transformations to a single piece, permuting finally the transformed segments. The idea here is to preserve to a certain extent the identity of the piece, giving a “flavored variation.” Some examples of variations of “Isn’t she lovely” by Stevie Wonder figure in the supplementary materials as “isl\_taste.mid.”

#### PARTICIPANTS

A total of 27 participants (12 females and 15 males,  $25 \pm 8$  years old) with no hearing impairments and no formal musical education participated in our experiment. Participants signed a written consent form.

#### EXPERIMENT

Participants judged the taste-words evoked by four musical pieces produced by the algorithm. Participants were asked to listen to the first minute of each of the composition and then assign the taste-word that they thought was in correspondence with the music. After listening to all musical pieces participants could change their

assignments since they were informed that they had to respond to each of the four taste words only once.

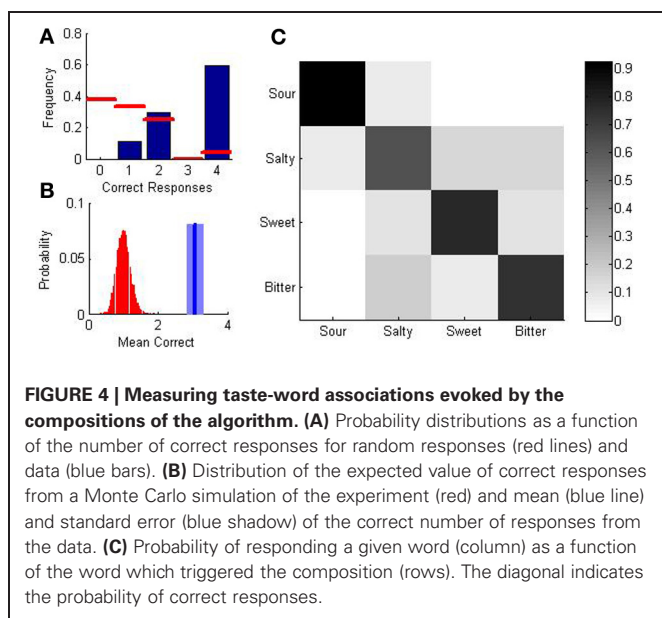
## RESULTS

Participants decoded well above chance the taste-word of the composition. All 27 subjects gave at least one correct response and 16 out of 27 participants (59%) showed a perfect performance.

The red bars in **Figure 4A** illustrate the probabilities of obtaining 0, 1, 2, 3, of 4 correct responses under the hypothesis that responses are random. The most likely response is 0 correct (with  $9/24$ ,  $p = 0.375$ ). The probability decreases monotonically with the number of responses except for three correct responses which, due to the characteristics of the response (assigning one taste-word to each composition, without repetition) is impossible. If three responses are correct the fourth also has to be correct.

The number of correct responses averaged over all subjects was  $3.07 \pm 0.22$ , over three times larger than the expected value of random assignments, which is 1. Since the distribution of responses is non-Gaussian, in order to verify the statistical significance of these observations we performed a Monte Carlo simulation of  $10^8$  experiments with 27 random measures. The distribution of responses compared to the empiric distribution can be seen in **Figure 4B**. Not a single simulated random experiment yielded an average value of performance greater than 3, which implies that the probability that our observed set of measures results from random assignments is  $p < 10^{-8}$ .

The stimulus-response matrix (**Figure 4C**) shows the specific mappings of compositions to taste-words. For this experiment, whose aim was to investigate the semantic robustness of quotations, we only explored one composition per taste word. Hence, specific patterns should be taken with caution and may not be indicative of an ensemble of compositions. However, the trend observed here of very accurate performance (with a minimum for salty representations) replicates precisely the decoding of an ensemble of improvisations (Mesz et al., 2011).



## DISCUSSION

The results of our experiment allow us to close the loop started in Mesz et al. (2011), where taste-words elicited musical improvisations within specific clusters in musical space. In this work we follow the opposite path, showing that algorithmically composed music within those regions elicit a robust association with the different tastes by an audience of non-experts.

Although literature on the specific subject of this work is scarce, it is interesting to contrast our results with those of Simner et al. (2010), that explores the relation between vocal sounds and tastants (substances with different tastes) at different concentrations. The acoustical parameters used in that work were the first two vocal tract resonant frequencies (formants)  $F_1$  and  $F_2$ , the sound discontinuity and spectral balance. Notably, they found that sour is mapped to parameters of higher frequency than the other tastes (for  $F_1$ ,  $F_2$  and spectral balance), a property we also found in music for the pitch. Moreover, sound discontinuity, a parameter analogous to high articulation (staccato) in music, is higher for salty and sour than for the other two tastes at medium concentrations of the tastants, a result which we find again in the articulation dimension.

Simner et al. also consider the incidence of the vocal phonemes of taste names on their results and the question of the presence of these taste-sound mappings in the genesis of the taste vocabulary in English. In Spanish, we have the words *ácido* (sour), *salado* (salty), *amargo* (bitter), and *dulce* (sweet). Considering just the stressed vowels in each taste name, we have the phoneme /a/ for all tastes except *dulce* (sweet) which has a stressed /u/. /a/ has in Spanish a mean  $F_1$  of 700 Hz and  $F_2$  of 1460 Hz, while /u/ has mean  $F_1$  of 260 Hz and  $F_2$  of 613 Hz, being the lowest vowel of all in  $F_1$ ,  $F_2$  and spectral balance. Since soft vocalizations have lower  $F_1$  than loud ones (Traunmüller and Eriksson, 2010), and also soft loudness lowers the spectrum of piano timbre, (Askenfelt URL), this is in agreement both with the hypothesis that the /u/ in *dulce* reflects the sweet-soft loudness mapping and that our subjects may have used associations involving the taste names for their production and labeling tasks. Note also that the word *ácido* (sour) has the /i/ that has in Spanish the highest  $F_2$  and spectral mean of all vowels, paralleling the higher frequencies and loudness of the sour piano improvisations. If we look at the consonants, we observe that the taste names in Spanish have exclusively consonants articulated with the front part of the mouth and lips, except for *amargo* (bitter) that has a /g/ which involves the back of the palate. One of the performers told us that she used consciously the mapping from the back of the mouth and throat (which she associated with bitter taste), to low pitch, so at least in this case a connection between mouth position and pitch was explicitly present. However, in a recent study (Crisinel et al., 2011) the authors failed to observe a significant relationship between flavor localization in the mouth and sweet/bitter ratings of *soundscape*s presented simultaneously with food taking, although they observe that increased bitter ratings were associated with a trend to localize the flavor in the back of their mouth.

In a study on vocal expression of affects, Scherer (Scherer, 1986) considers the acoustic correlates in the voice to orofacial movements produced by stereotypical reactions to basic

tastes: sweet—retracted mouth corners, expression of a smile, sucks and licking; sour—lip pursing, nose wrinkling, blinking; bitter—depressed mouth corners and elevated upper lip, flat tongue visible, contraction of platysma muscle, retching, or spitting (Steiner, 1979). He concludes that for unpleasant tastes the consequences would be stronger resonances in the high frequency region, so a brighter timbre, as in the relatively loud piano music for sour, bitter, and salty in our experiment. For pleasant tastes, the vocal outcome would be some damping of higher frequency resonances (as in the soft piano playing of “sweet improvisations”), and also a clear harmonic structure, which relates to sensory consonance (Cariani URL), measured here by gradus and harmonic dissonance (see glossary in the supplementary material) and which is maximum (i.e., minimum gradus) for sweet music.

In the light of this discussion, it is interesting to note the case of the taste-music synaesthete E. S. (Beeli et al., 2005; Hänggi et al., 2008), since the experiences of synaesthetes often reflect the implicit associations made by all people (Simner, 2009 review). E.S. experiences tastes in her tongue listening to musical intervals, in the following pairs (listed by decreasing gradus of the interval): minor second, sour; major second, bitter; minor third, salty; major third, sweet. This is in concordance with the results we obtained for our measure of consonance (gradus, see **Figure 2**).

We also explored in detail the correspondences of the words annotated by the improvisers during the production experiment (Mesz et al., 2011) with respect to affective content. Affect in words is represented in the table ANEW (Bradley and Lang, 1999) in a three-dimensional space of valence (pleasant/unpleasant), arousal (calm/ex-cited) and dominance (controlled/in-control), with ratings from 1 to 9. Restricting ourselves to terms appearing more than once for different subjects, we have the following list of vectors in the space of (valence, arousal, dominance):

**Sweet:** tenderness (7.29, 4.68, 5.33); peace (8.20, 2.05, 5.20); soft (7.38, 4.32, 5.86). Mean: (7.62, 3.68, 5.46).  
**Salty:** joy (8.62, 6.73, 6.46); cold (4.14, 5, 4.73). Mean: (6.38, 5.86, 5.59).  
**Sour:** sour (4.18, 5.23, 4.50); cruel (1.97, 5.68, 4.24). Mean: (3.07, 5.45, 4.37).  
**Bitter:** dark (4.71, 4.28, 4.84); pain (2.13, 6.50, 3.71). Mean: (3.42, 5.39, 4.27).

In the musical domain, Luck et al. (Luck et al., 2008) relate similar affective dimensions to musical features in a listening experiment where participants had to rate keyboard improvisations along the dimensions of *pleasantness* (pleasant/unpleasant), *activity* (active/inactive), and *strength* (strong/weak). They conclude that higher activity is associated with higher note density, greater pulse clarity, higher mean velocity, and higher levels of sensory dissonance. On the other hand, higher pleasantness is associated with lower note density, higher tonal clarity and lower pulse clarity and higher strength ratings correspond to higher mean velocity, higher note density, and higher dissonance.

Although their terminology is not exactly the same as that in the ANEW, assuming that valence-pleasantness and arousal-activity are related pairs we find a semantic-musical affective correspondence: the semantically least arousing and most pleasing

taste (sweet) gives improvisations with low velocity, low dissonance and long duration (which in the case of melodies corresponds to low note density, that is, fewer notes by unit time).

Some more complex cross-modal mappings are also present, involving other senses. For instance, sour was associated verbally with luminous and bitter with dark, and their characterizations by high- and low-pitched music, respectively, agree with the direct mapping from brightness to pitch (Spence, 2011), in such a way that the correspondence may have been mediated by a magnitude visual-auditory mapping matching high magnitudes across dimensions (Walsh, 2003). Moreover, from a psychoacoustical standpoint, sweet improvisations are related to high values of psychoacoustical *pleasantness*, which correspond to soft sound intensity and low roughness (Fastl and Zwicker, 2007), that is, to low dissonance (Plomp and Levelt, 1965). The word *sour*, on the other hand, elicited loud, dissonant, and high-pitched improvisations which correspond to high values of sensory sharpness (Fastl and Zwicker, 2007), a psychoacoustical magnitude that is inversely related to pleasantness.

Throughout this discussion we connected the results of our experiments using piano music with those of an eclectic repertoire of experiments using food, tastants, written words, instrumental sounds, voice sounds, and visual cues related to voice production. This sort of consistency helps building confidence on the existence of strong, multiple associations between auditory-taste sensory channels that are worth exploring via non-standard tools. A key part of our project is to express these results as an installation. A series of short piano pieces composed with the algorithm will be played on a classic piano for an audience located in a customized room. Musical pieces and food will be presented to the audience in sequential order. Some pieces will be composed and executed in real time, based on time-intensity curves given by expert food evaluators, and read by the pianist from interactive scores projected on screens using a system designed by Luciano Azzigotti (Azzigotti URL). Feedback from the participants will serve to investigate temporal and across-people fluctuations in the perception of music and taste sensation. Will we find new dimensions in musical space emerging from a very large number of participants? Will the information of these participants be able to feed back to the algorithm to identify the minimal number of features which evoke coherently a taste-like association? Might the successive iterations of words presented and words proposed by the participants converge or cycle? Will there be moments of coherent synchronization between participants? These questions will be asked in the form of a performance, both a musical expression and a quantitative experiment.

## ACKNOWLEDGMENTS

We thank Ivonne Bordelois and Guillermo Hough for enlightening discussions, and María Ceña for her ideas on the design of the installation. This work was partially funded by the University of Buenos Aires and CONICET.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at [http://www.frontiersin.org/Human\\_Neuroscience/10.3389/fnhum.2012.00071/abstract](http://www.frontiersin.org/Human_Neuroscience/10.3389/fnhum.2012.00071/abstract)

## REFERENCES

- Askenfelt URL: [http://www.speech.kth.se/music/5\\_lectures/askenfelt/askenfelt.html](http://www.speech.kth.se/music/5_lectures/askenfelt/askenfelt.html)
- Azzigotti URL: <http://zztt.org/en/onkimi>
- Beeli, G., Esslen, M., and Jänke, L. (2005). Synaesthesia: when coloured sounds taste sweet. *Nature* 434, 38.
- Bradley, M. M., and Lang, P. J. (1999). *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings*, Technical Report C-1. The Center for Research in Psychophysiology, University of Florida.
- Cariani URL: [http://homepage.mac.com/cariani/CarianiNewWebsite/Publications\\_files/Cariani-ICMPC-8-2004-Proceedings.pdf](http://homepage.mac.com/cariani/CarianiNewWebsite/Publications_files/Cariani-ICMPC-8-2004-Proceedings.pdf)
- Chomsky, N. (1988). *Language and Problems of Language*. Cambridge, MA: MIT Press.
- Crisinel, A.-S., Cossier, S., King, S., Jones, R., Petrie, J., and Spence, C. (2011). A bittersweet symphony: systematically modulating the taste of food by changing the sonic properties of the soundtrack playing in the background. *Food Qual. Prefer.* 24, 201–204.
- Crisinel, A. S., and Spence, C. (2009). Implicit association between basic tastes and pitch. *Neurosci. Lett.* 464, 39–42.
- Crisinel, A. S., and Spence, C. (2010a). As bitter as a trombone: synesthetic correspondences in nonsynesthetes between tastes/flavors and musical notes. *Atten. Percept. Psychophys.* 72, 1994–2002.
- Crisinel, A. S., and Spence, C. (2010b). A sweet sound? Food names reveal implicit associations between taste and pitch. *Perception* 39, 417–425.
- Debussy, C. (1913). *Préludes*, Book 2: III, [http://imslp.org/wiki/Préludes\\_\(Book\\_2\)\\_\(Debussy,\\_Claude\)](http://imslp.org/wiki/Préludes_(Book_2)_(Debussy,_Claude)).
- Driver, J., and Spence, C. (2000). Multisensory perception: beyond modularity and convergence. *Curr. Biol.* 10, R731–R735.
- Engine URL: <http://www.peachnote.com>
- Evans, K. K., and Treisman, A. (2010). Natural cross-modal mappings between visual and auditory features. *J. Vis.* 10, 6.1–12.
- Fastl, H., and Zwicker, E. (2007). *Psychoacoustics. Facts and Models*. Berlin: Springer.
- Fraisse, P. (1982) “Rhythm and tempo,” in *Psychology of Music*, ed D. Deutsch (New York, NY: Academic), 149–180.
- Hänggi, J., Beeli, G., Oechslin, M. S., and Jaencke, L. (2008). The multiple synaesthete E.S.: neuroanatomical basis of interval-taste and tone-colour synaesthesia. *Neuroimage* 43, 192–203.
- Jackendoff, R., and Lerdahl, F. (2006). The capacity for music: what is it, and what’s special about it? *Cognition* 100, 33–72.
- Jentschke, S., Koelsch, S., and Friederici, A. D. (2005). Investigating the relationship of music and language in children: influences of musical training and language impairment. *Ann. N.Y. Acad. Sci.* 1060, 231–242.
- Jentschke, S., Koelsch, S. S., Sallat, S., and Friederici, A. D. (2008). Children with specific language impairment also show impairment of music-syntactic processing. *J. Cogn. Neurosci.* 20, 1940–1951.
- Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., and Friederici, A. D. (2004). Music, language and meaning: brain signatures of semantic processing. *Nat. Neurosci.* 7, 302–307.
- Luck, G., Toiviainen, P., Erkkilä, J., Lartillot, O., Riikkilä, K., Mäkelä, A., Pyhälä, K., Raine, H., Varkila, L., and Värri, J. (2008). Modelling the relationships between emotional responses to, and musical content of, music therapy improvisations. *Psychol. Music* 36, 25–45.
- Mesz, B., Trevisan, M. A., and Sigman, M. (2011). The taste of music. *Perception* 40, 209–219.
- Parise, C., and Spence, C. (2008). Synaesthetic congruency modulates the temporal ventriloquism effect. *Neurosci. Lett.* 442, 257–261.
- Patel, A. D. (2003). Language, music, syntax and the brain. *Nat. Neurosci.* 6, 674–681.
- Petrucci URL: <http://imslp.org/wiki>
- Plomp, R., and Levelt, W. J. M. (1965). Tonal consonance and critical bandwidths. *J. Acoust. Soc. Am.* 38, 548–560.
- Ramachandran, V. S., and Hubbard, E. M. (2003). Hearing colors, tasting shapes. *Sci. Am.* 288, 52–59.
- Scherer, K. R. (1986). Vocal affect expression: a review and a model for future research. *Psychol. Bull.* 99, 143–165.
- Shukla, M., Nespor, M., and Mehler, J. (2007). An interaction between prosody and statistics in the segmentation of fluent speech. *Cogn. Psychol.* 54, 1–32.
- Simner, J. (2009). Synaesthetic visuo-spatial forms: viewing sequences in space. *Cortex* 45, 1138–1147.
- Simner, J., Cuskey, C., and Kirby, S. (2010). What sound does that taste? Cross-modal mappings across gustation and audition. *Perception* 39, 553–569.
- Spence, C. (2011). Crossmodal correspondences: a tutorial review. *Atten. Percept. Psychophys.* 73, 971–995.
- Steiner, J. E. (1979). “Human facial expressions in response to taste and smell stimulation,” in *Advances in Child Development and Behavior*, Vol. 13, eds H. W. Reese and L. P. Lipsitt (New York, NY: Academic Press), 257–295.
- Toolbox URL: <https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/miditoolbox>
- Traunmüller, H., and Eriksson, A. (2010). Acoustic effects of variation in vocal effort by men, women, and children. *J. Acoust. Soc. Am.* 107, 3438–3451.
- Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* 7, 483–488.
- Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8, 547–558.

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 14 September 2011; accepted: 13 March 2012; published online: 27 April 2012.

Citation: Mesz B, Sigman M and Trevisan MA (2012) A composition algorithm based on crossmodal taste-music correspondences. *Front. Hum. Neurosci.* 6:71. doi: 10.3389/fnhum.2012.00071  
Copyright © 2012 Mesz, Sigman and Trevisan. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.