RUNNING HEAD: IMPACT OF MUSIC ON WINE PERCEPTION

Analysing the impact of music on the perception of red wine via the Temporal Dominance of Sensations

Qian Janice Wang[1,2], Bruno Mesz[3], Pablo Riera[4], Marcos Trevisan[5], Mariano Sigman[6,7], Apratim Guha[8], & Charles Spence[1]

[1] Crossmodal Research Laboratory, Department of Experimental Psychology, Oxford University, Oxford, UK
[2] Department of Food Science, Aarhus University, Aarslev, Denmark
[3] MUNTREF Tecnópolis, Universidad Nacional de Tres de Febrero, Buenos Aires, Argentina
[4] Laboratorio de Inteligencia Artificial Aplicada, Instituto de Ciencias de la Computación, Universidad de Buenos Aires, CONICET, Argentina
[5] Department of Physics, University of Buenos Aires and Institute of Physics Buenos Aires (IFIBA) CONICET
[6] Laboratorio de Neurociencia, CONICET, Universidad Torcuato Di Tella, C1428BIJ Buenos Aires, Argentina
[7] Facultad de Lenguas y Educación, Universidad Nebrija, Madrid, Spain
[8] Production, Operations and Decision Sciences Area, XLRI, Xavier School of Management, Jamshedpur, India

DATE: MARCH 2019

CORRESPONDENCE TO: Qian Janice Wang, Department of Food Science, Aarhus University, Kirstinebjergvej 10, Aarslev, DK-5792, Denmark. TEL: +4587154886 E-MAIL: qianjanice.wang@food.au.dk

ABSTRACT

Several studies have examined how music may affect the evaluation of food and drink, but the vast majority have not observed how this interaction unfolds in time. This seems to be quite relevant since both music and the consumer experience of food/drink are time-varying in nature. In the present study we sought to fix this gap, using the Temporal Dominance of Sensations (TDS), a method developed to record the dominant sensory attribute at any given moment in time, to examine the impact of music on the wine taster's perception. More specifically, we assessed how the same red wine might be experienced differently when tasters were exposed to various sonic environments (two pieces of music plus a silent control condition). The results revealed diverse patterns of dominant flavours for each sound condition, with significant differences in flavour dominance in each music condition as compared to the silent control condition. Moreover, musical correspondence analysis revealed that differences in perceived dominance of acidity and bitterness in the wine were correlated in the temporality of the experience, with changes in basic auditory attributes. Potential implications for the role of attention in auditory flavour modification and opportunities for future studies are discussed.

KEYWORDS: CROSSMODAL CORRESPONDENCES; TEMPORAL DOMINANCE OF SENSATIONS; ATTENTION; WINE EVALUATION; MUSIC

I. Introduction

Over the last decade, a rapidly-growing body of empirical research has demonstrated the existence of what is often called “sonic seasoning”, whereby soundtracks with congruent taste/flavour attributes have been shown to influence the perception of what we eat and drink (e.g., Crisinel et al., 2012; Reinoso Carvalho et al., 2017; Wang & Spence, 2016). Crisinel et al. (2012) first demonstrated sonic seasoning in a study in which the participants were given samples of bittersweet cinder toffee to evaluate the taste while listening to one of two soundscapes. The soundscapes had been specially composed to match either sweetness or bitterness. Listening to the higher pitched sweet soundscape resulted in the toffee being rated as tasting significantly sweeter and less bitter than while listening to the lower-pitched bitter soundscape instead.

One plausible theory behind sonic seasoning relies on the role of attention; more specifically, the claim is that sound-flavour correspondences may help direct (either automatically, or voluntarily) our attention to certain aspects of the flavour mixture in a food or drink (see Spence et al., in press, for a review). Attention is intrinsic to how we perceive sensory inputs (Chen et al., 2013), and likely plays a crucial role in determining what we perceive in food (Spence et al., 2000). For instance, Stevenson (2012) illustrated the role of attention in flavour perception in his review by arguing that the reason why we perceive flavour as coming from the mouth, even though smells are captured by the olfactory receptors in the nasal cavity, is due to attentional capture by somatosensory cues (i.e., over olfaction; though see Spence, 2016, for a critical evaluation of the claim). Moreover, when it comes to flavour mixtures, attended flavour elements become relatively more salient than relatively less attended elements (Ashkenazi & Marks, 2004; Marks & Wheeler, 1998; Rabin & Cain, 1989). Crossmodal correspondences between pitch and spatial location have been shown to modulate attentional orienting (e.g., Chiou & Rich, 2012; Klapetek et al., 2012; Mossbridge et al., 2011; Parrot et al., 2015), so it is conceivable that auditory stimuli (specifically taste/flavour-congruent soundtracks) may also be able to shift our attention towards specific taste/flavours. Moreover, since music and food/drink are both time-varying in nature, it seems only appropriate to take temporality into account when studying the impact of music on the eating/drinking experience.

It should be noted that multisensory congruency has bee documented to influence attentional selection in the case of perceptually ambiguous stimuli (van Ee et al., 2009). In their study, Ee and colleagues demonstrated that congruent auditory or tactile information aided attentional control over competing visual stimuli. Therefore, we might potentially view the phenomenon of “sonic seasoning” as a specific case of attentional selection, whereby sound-flavour congruence aids attentional selection for congruent flavours present in a mixture (e.g., wine).

In the present study, we used the method of Temporal Dominance of Sensations (TDS). TDS is a relatively recent technique used to record several sensory attributes simultaneously over time. TDS was first introduced at the 5th Pangborn symposium (Pineau et al., 2003) and quickly caught on within the food science community. This technique requires the participant to keep assessing the most dominant flavour attribute, among multiple possible flavour attributes, over the period of assessment. The participant was instructed to assess which flavour attribute is perceived as dominant at any given point in time. TDS has been used to characterise beverages such as blackcurrant squash (Ng et al., 2012) and wine (Meillon et al., 2009; Sokolowsky & Fischer, 2012). Recently, Kantono and his colleagues used TDS to study how liked/disliked music might influence the perception of gelato in terms of basic tastes (Kantono et al., 2016, 2018). We decided to use TDS in order to explore the effect of music on how the taste of wine, a beverage with a complex array of flavours, is perceived. In fact, various multisensory interaction effects on wine flavour perception have been demonstrated with coloured lights (Oberfeld et al., 2009; Spence, Velasco, & Knoeferle, 2014), tactile stimuli (Wang & Spence, 2018), music (Wang & Spence, 2015, 2017), and combinations of light, soundscapes, and other ambient sensory effects (Velasco et al., 2013). If music does indeed direct the taster's attention to specific flavours, those attended flavours would be more salient/dominant (Ashkenazi & Marks, 2004; Marks & Wheeler, 1998), and different patterns of dominant flavours would be expected under different auditory conditions. Furthermore, we conducted content analysis on the soundtracks used in the study, in order to understand which auditory properties might lead to differences in the temporal patterns of the dominant flavour attributes.

II. Methods

Participants

A total of 39[1] participants took part in the study. Of these, 21 (11 women, 10 men), aged 21-69 years (M=37.6, SD=12.8), participated in the main study and 18 (4 women, 14 men), aged 21-41 years (M=27.1, SD=5.5), took part in a control experiment to assess the hypothesis that the "musical flavours" dominate over the perceived wine flavours. The participants were recruited at the University Tres de Febrero (UNTREF, Buenos Aires, Argentina). All the participants gave their informed consent to take part in the study. None of the participants reported a cold or any other known impairment of their sense of smell, taste, or hearing at the time of

[1] The number of participants was determined by a convenience sampling. As there was no precedence for music-food TDS analysis at the time of study, it was difficult to run a power analysis. However, the sample sizes used in the studies are in line with typical TDS panel sizes (Meillon et al., 2009; Pineau et al., 2009).

the study. The study was approved by the Central University Research Ethics Committee of Oxford University (MSD-IDREC-C1-2014-205).

Auditory stimuli

Two pieces of music were chosen for the study that varied in tempo, mode, and instrumentation. The first piece was Brian Eno's "Discreet Music", and the second piece was Mussorgsky's "A Night on Bald Mountain". A 45 second excerpt was taken from the beginning of each piece. Note that both of these pieces have been used in various wine-music demos and talks related to sound-taste interactions previously (e.g., Burzynska, 2018). A pre-test (N=19) revealed that both pieces of music led to a significant change in the perceived fruitiness and tannin levels of a light red wine (Georges du Boeuf's Beaujolais-Village, 2014), with the Eno soundtrack enhancing the perceived fruitiness and decreasing the perceived tannin levels, as ompared to the Mussorgsky soundtrack. The Eno soundtrack is fairly static throughout, with moderate tempo (70 beats per minute) and consonant harmony, whereas the Mussorgsky soundtrack has dynamic changes in orchestration, register, and loudness; fast tempo (121 beats per minute), and both consonant and dissonant harmonies.

Wine

The wine used in the study was a Pinot Noir produced in Argentina – Manos Negras Red Soil Select Pinot Noir, 2014. The wine has 13.9% alcohol by volume, 3.92 pH, and 5.70 g/L of Total Acidity (TA), and has been aged for 12 months in 20% new French oak casks[2]. A Pinot Noir was used because relatively light-bodied red wines have been commonly used previously in sound-taste demonstrations. The wine was served in 15 mL samples, inside clear plastic 50 mL cups, at room temperature

[2] http://www.manosnegras.com.ar/images/fichas/Pinot-Noir-Red-Soil-Select-EN.pdf.

(between 16 and 22 degrees Celsius).

Design and procedure

Each participant was seated in front of a computer monitor with headphones and a cup of water to cleanse their palate. The experiment was programmed using the Sensomaker tool for the sensorial characterisation of food products (Pinheiro et al., 2013).

At the onset of each trial, the participants were given a sample of wine by the experimenter. They were instructed to start the trial as soon as the wine entered their mouth, and to hold it there for the 45-second duration of the trial. The choice of 45 seconds was informed by other TDS studies (e.g., Kantono et al., 2016, 2019). During this time, the TDS computerised system displayed the entire list of 8 adjectives in two columns (red fruit, tannins, alcohol, woody, sweet, acidic, spicy, and bitter) to the participants. The attributes were selected on the basis of a similar TDS study on Pinot Noir wines (Visalli, 2016). The participants were instructed to click on the start button as soon as the wine sample entered their mouth, and then to consider which attribute was perceived as the most dominant. Each time they felt that their perception had changed, they were to click on the new attribute that they perceived to be most dominant. The participants were free to select an attribute several times during the course of the trial, or not at all. The participants first practiced with a weak yerba mate tea solution 1-3 times to ensure that they could operate the TDS software with ease.

The order in which the adjectives were presented was randomised for each participant in order to avoid any order effect of the list of attributes. However, for a specific participant, the order was always the same and so learning the terms and scoring was facilitated. We used a within-participant full factorial design, with each participant tasting three wine samples in the three auditory conditions, without knowing that the wines were indeed the same. The participants always tasted the wine in the silent condition first, but the order of the two music soundtracks was randomised. In those trials involving a soundtrack, the experimenter started the music at the same time as the participant clicked on the start button. Participants were informed at the start of the soundtrack trials that background music would be presented, but they were not informed about the purpose of the music nor how the music were selected. The entire experiment lasted for around 10 minutes and the participants were debriefed afterwards.

Data Analysis

TDS curves were produced by the SensoMaker software. They were averaged over all participants and smoothing was applied. Each graph had two additional lines. One, the "chance level" is the dominance rate that an attribute would be chosen by chance, in this case equal to 1/8, since there are 8 attributes. The second line shows the "significance level", which is the minimum value for the dominance level to be considered significantly greater than chance, calculated using the confidence interval of a binomial proportion based on a normal approximation (Pineau et al., 2009). In order to understand specifically how music influenced TDS responses, pairwise correlations between the musical features (frequency content, intensity, and musical segmentation) and the reported dominant tastes were calculated.

III. Results

Figure 1 shows the TDS curves for each of the three auditory conditions. Concentrating on dominance ratings above the significance levels, three major differences can immediately be seen. First, onset time for acidity occurred at around 9 seconds in the silent baseline condition, whereas acidity peaked at 23 seconds for the Eno Soundtrack and at 27 seconds for the Mussorgsky Soundtrack. Second, bitterness was prominent during 25-38 seconds for the Eno Soundtrack, whereas it was prominent during 8-14 seconds for the Mussorgsky Soundtrack (and was barely registered at 29 seconds in the silent condition). Finally, in the silent condition, both alcohol and astringency were at significant dominance levels (alcohol between 5 and 10 seconds, astringency at 10-20 seconds, then 35-45 seconds), but was not significant when either of the two soundtracks were played. For the Eno soundtrack, acidity was registered before bitterness, whereas for the Mussorgsky soundtrack, bitterness was registered before acidity.

< INSERT FIGURE 1 ABOUT HERE>

For each sound condition, the total number of citations (i.e., number of times chosen) as well as the duration of dominance for each of the 8 descriptors were calculated. Since none of the measures were normally distributed according to the Shapiro-Wilk test, we used non-parametric rank-sum tests. A Friedman test revealed there to be no significant differences in the total number of adjectives participants' used for each of the three auditory conditions ($\chi^2(2) = 4.00$, p =.14). Neither were there any significant differences in dominance durations of acidity ($\chi^2(2) = 2.00$, p = .37), alcohol ($\chi^2(2)$ = 2.92, p = .23), woodiness ($\chi^2(2) = 3.86$, p = .14), sweetness ($\chi^2(2) = 1.81$, p = .41), spiciness ($\chi^2(2) = 2.68$, p = .26), or red fruit ($\chi^2(2) = 1.27$, p = .53).

There was, however, a significant difference in dominance durations for bitterness

($\chi^2(2) = 7.75$, $p = .021$) and astringency ($\chi^2(2) = 7.55$, $p = .023$). Post hoc analysis with Wilcoxon signed-rank tests revealed that compared to the silent condition, there were significant increases in bitterness dominance durations for both the Eno soundtrack ($Z = -2.28$, $p = .023$) and the Mussorgsky soundtrack ($Z = -2.30$, $p = .021$). There were, however, no difference in bitterness between the two soundtrack conditions ($Z = -0.024$, $p = .98$).

There were also significant reductions in the durations of astringency dominance for both the Eno soundtrack ($Z = -2.04$, $p = .042$) and the Mussorgsky soundtrack ($Z = -2.19$, $p = .029$), as compared to the silent condition. There were, once again, no differences between the two soundtrack conditions ($Z = -0.75$, $p = .45$).

In order to determine the effect of each soundtrack on the perceived taste of the wine more clearly, TDS difference curves were plotted (see Figure 2) to reveal the net influence of background music on wine perception, while controlling for how the wine tastes in the silent control condition. To calculate the difference curves, the differences between each soundtrack and the silent condition were plotted at points where they were significantly different from zero by comparing two binomial proportions (Pineau et al., 2009). The effect of the Eno soundtrack, compared to the silent baseline condition, highlights an enhancement of bitterness and a reduction of alcohol in the 0-15 second timeframe, and then a reduction in alcohol around the 30 second mark. The effect of the Mussorgsky soundtrack, compared to the silent condition, is a longer and more prominent enhancement of bitterness during the 0-15 second timeframe along with a reduction in acidity and astringency. There follows an enhancement in acidity around the 25-30 seconds timeframe.

< INSERT FIGURE 2 ABOUT HERE>

Analysis of individual musical features in relation to taste:

Pairwise Pearson correlations were computed between the TDS curves for the tastes that reach significance in the presence of music (bitterness and acidity, see Figure 1), and different types of time-varying musical features: 1) acoustic features: frequency content (measured by spectral centroid), and sound intensity (measured by root mean square energy, RMSE); 2) psychoacoustic features: roughness, brightness, inharmonicity; 3) emotion-related features: valence, activity, and tension. These features include many of the major factors in crossmodal taste-sound correspondences that have been documented to date (see Knöferle & Spence, 2012, for a review): spectral centroid (also called spectral balance) is related to timbre brightness/sharpness; energy is related to loudness; roughness, brightness and inharmonicity are timbre features while valence is connected with pleasantness. Another relevant feature, tempo, was almost constant for each musical excerpt. Pitch was not analysed since the music is polyphonic (that is, several pitches are present at the same time). We used MIRToolbox (Lartillot & Toiviainen, 2007) for the computation of the musical curves from the audio files, with an overlapping running window (window length = 0.05 s, overlap = 0.025 s). If there were to be an influence of a musical feature on taste, we would expect a positive time delay between the music onset time and its effect on the TDS response (due to the time required for auditory processing, choosing an attribute, and finally clicking on it). This delay was estimated as the averaged time $\underline{T_1}$ of first response, corresponding to the time it took the participant to choose a first attribute after the start of the music. We calculated similar average delays for the two music pieces (Mussorgsky: $\underline{T_{1Muss}}$ = 8.4 seconds with standard deviation σ_{Muss} = 4.2 seconds, Eno: $\underline{T_{1Eno}}$ = 8.1 seconds with standard deviation $\sigma_{Eno} = 2.6$ seconds) , and for the silent condition ($\underline{T_{1\ silence}} = 6.9$ seconds with standard deviation $\sigma_{silence} = 1.9$ seconds).

In Figure 3, we plotted the pairwise correlations among the musical parameters and taste attributes for different positive delay windows of the taste curves. We take the relevant lags as those in the intervals ($\underline{T_{1music}} - \sigma_{music}$, $\underline{T_{1music}} + \sigma_{music}$) and ($\underline{T_{1silence}} - \sigma_{silence}$, $\underline{T_{1silence}} + \sigma_{silence}$) marked by the coloured vertical bars in the figure. The specific pairs of music parameters and taste curves plotted in Figure 3 were the only ones for which significant correlations were found in the music condition (in blue), while no significant correlation was observed in the silent control condition (in orange) for the relevant delay windows. Moreover, they have important and perceptible variations during the music: the spectral centroid ranged from 946 Hz to 3336 Hz in the Mussorgsky soundtrack, and from 790 Hz to 1091 Hz in the Eno Soundtrack; RMSE varied from 0.0006 to 0.09 in the Mussorgsky Soundtrack (-45 dB to -3 dB), and from 0.003 to 0.05 in the Eno Soundtrack (-34 dB to -11 dB). Another important factor, sensorial dissonance, measured by psychoacoustic roughness (Johnson-Laird et al., 2012; Bigand et al, 1996), gave very similar correlations to those of the RMSE, which are not shown in Figure 3. Note that even if there were a significant correlation between a music parameter and a taste curve in the silent condition, this would be merely coincidental (since in fact people were not listening to any music!). We also considered the significant correlations that appear outside the relevant delay window to be meaningless in this context: since the delay between the curves would be either too short or too long in comparison to the estimated response delay, we do not consider these correlations as representing an influence of music on taste.

In agreement with previous results in the literature (Knöferle & Spence, 2012; Mesz et al., 2011), when the Eno soundtrack was playing in the background, the dominance of acidity was positively correlated with high spectral centroid, and the dominance of bitterness was correlated with low spectral centroid. Furthermore, the dominance of bitterness was positively correlated with sound intensity (RMSE) and sensorial dissonance for the Eno soundtrack. On the other hand, for the Mussorgsky soundtrack, a correlation was observed between the dominance of bitterness and high spectral centroid, contrary to what we observed for the Eno soundtrack.

< INSERT FIGURE 3 ABOUT HERE>

Influence of musical structure relation on taste

We also explored a possible correspondence between the structural segments in the musical excerpts and regions where significant taste evaluations were prominent. A method for novelty-based segmentation was used to help in locating those points in time from a music signal that would correspond to the changes on instrumentation and dynamics (Müller, 2015). Novelty was computed by inspecting the recurrence matrix of the Mel Frequency Cepstral Coefficients (MFCC) audio descriptor and measuring the edges in the block-like structures typically found on recurrence matrix. This was achieved by convolving an edge detection kernel over the recurrence matrix following the principal diagonal direction. The output of this process was a signal that indicates the novelty as a continuous signal, and the peaks of this novelty signal gave the boundaries of the musical segments (Fuller, 2000).

We plotted acidity and bitterness taste curves and overlaid novelty boundaries on top (see Figure 4). Only for the Mussorgsky soundtrack did we find significant differences between bitterness and acidity, which occur after the boundaries with a delay within one standard deviation of $\underline{T_1}$. Between the first and second boundaries there was significantly more bitterness than acidity, and between the second and third boundaries, there was significantly more dominance in acidity than bitterness. The three boundaries correspond to important points in the music (see video in the Supplementary Materials): the first boundary, at 4.08 seconds, marks the entrance of excited glissandos in the high register of the woodwind (simultaneously there are chromatic triplets in the violins and a rumbling bass line in the low strings). It is also during this period where bitterness is perceived as significantly more dominant than acidity. The second boundary, at 14.39 seconds, coincides with the first theme in bassoons, trombones, tuba, violas and low strings. Next, the horns and trumpets enter playing a D, creating a dissonance with the C at the melody instruments. The dissonance continues until it is halted by two loud chords, occurring during the period when perceived acidity was more dominant than bitterness. At the third boundary, the orchestra decays together with the acidity curve, then two more accentuated chords lead to a trill and a general pause (see videos of the music versus the taste curves in the Supplementary Materials).

< INSERT FIGURE 4 ABOUT HERE>

< INSERT FIGURE 5 ABOUT HERE>

IV. Discussion

The results of the present study demonstrate that tasting wine while listening to different soundtracks leads to different perceptions of dominant flavours. Overall, the onset of acidity was earlier in the silent condition than in either of the soundtrack conditions, and astringency was less noticeable when there was music playing. Bitterness was more prominent in the beginning when the wine was tasted when listening to the Mussorgsky soundtrack, whereas for the Brian Eno soundtrack, bitterness in the wine came after the initial registration of acidity. Analysing dominance durations supported results from TDS difference curves,, where bitterness was dominant for longer – but astringency shorter – during the two soundtrack conditions when compared to the silent condition. Furthermore, while alcohol was dominant in the 10-20 second interval in the silent condition, it was not at significant dominance level during either of the soundtracks. This implies that music could potentially distract participants from perceiving alcohol accurately, especially when music was presented in combination with the cognitively demanding task (Stafford et al., 2012, 2013).

There were no significant differences in the number of adjectives participants selected – on average the number was around four, or half of the eight available adjectives. For this group of participants, basic tastes such as acidity and bitterness were used more often (Dur_{acid} = 9.1 seconds, Dur_{bitter} = 7.4 seconds) than more descriptive terms such as red fruit and woody (Dur_{red_fruit} = 1.6 seconds, Dur_{woody} = 3.7 seconds). This might either be attributable to the fact that the participants simply did not taste the more descriptive attributes in the wine, or that basic tastes were simply more dominant (or more easily came to mind) in the wine, especially under experimental conditions.

What are the mechanisms underlying the changes in taste perception in the presence of music? A plausible hypothesis concerning the influence of the music on the flavour in this experiment is that, at least at some moments, the flavour labels are associated with the music, independently of the taste of the wine, and then attributed to the wine. To examine this hypothesis, we performed a control test in which a group of participants (N=18) had to evaluate the music alone, without drinking, using the TDS protocol with the same labels used for the wine. The dominance regions of different musical “flavours” for this test are shown in Figure 5. While, in the case of the Eno soundtrack, there is no immediate correspondence between the music only TDS test and the music + wine TDS test, for Mussorgsky’s music, we found an overlap of the acidity-dominant regions in both tests (from 23 seconds to 27 seconds), with a delay between them within one standard deviation of $\underline{T_1}$ (see Figure 5).

A number of possible explanations can be advanced for the overlap of semantic regions for music and taste. On the most superficial level, the participants could have been applying a recognition heuristic (Goldstein & Gigerenzer, 2002). Recognition heuristics consist, in general, in choosing a known alternative over an unknown one; in our case, this implies that having picked a descriptor significant in the music, a participant would prefer to apply this label (e.g., acidity) to an ambiguous, unrecognized taste. However, if the participants merely applied labels from the music to the wine, we would have expected to have seen sweet being chosen for the Eno soundtrack, both when there was music alone or with the combination music and wine. This implies that the recognition heuristic was not the *only* mechanism involved in the participants’ TDS ratings.

Besides semantics, more general crossmodal associations between musical features and taste can be relevant for altering wine’s flavours. As we hypothesised, specific crossmodal associations can shift people’s attention to certain flavours in the wine, which could then make it appear dominant. For instance, acidity/sourness has been reliably associated with high spectral centroid (Knoeferle et al., 2015; Simner et al., 2010) and correspondingly, a higher spectral centroid in the music was correlated with the dominance of perceived acidity in the Eno soundtrack. Is should be said that some conflicting evidence also appeared for the Mussorgsky soundtrack, where spectral centroid was positively correlated with bitterness, but according to literature, bitterness is associated with low spectral centroid (Knoeferle et al., 2015). However, it is worth noting that this initial section was also highly dissonant, which is correlated in the literature with both bitterness and acidity. Therefore, the Mussorgsky soundtrack, while technically having a high spectral centroid, contained auditory attributes associated with bitterness as well as sourness.

An important theoretical question here regarding the attention hypothesis concerns the temporal resolution of auditory and flavour attention. In the analysis of musical structural correlations to perceived tastes, we took into account the delay time between the participant hearing a feature in a given piece of music, to the participant choosing a specific flavour as dominant. This lag time was approximated by the average time it took the participant to choose an attribute following the onset of the music. For both soundtracks, the lag time was around 8 seconds (8.1 seconds for the Eno soundtrack, and 8.4 seconds for the Mussorgsky soundtrack). For the silent condition, the average lag time was not significantly shorter at 6.9 seconds. This agrees with the average time before first citation in other TDS studies involving wine (Galmarini et al., 2018; Meillon et al., 2009). That is to say, it would appear to take participants around 7 to 8 seconds to perceive and select a dominant taste in wine, no matter whether music is present or not (which is understandable as it takes approximately 200 milliseconds for people to register auditory loudness, von Békésy, 1963). Interestingly, this delay time also agrees with the time required before participant can make emotional judgments about a piece of music, which is also 8 seconds (Bachorik et al., 2009). Therefore, it is plausible that differences in TDS ratings could also be due to sensation transference (Biggs et al., 2016; Kantono et al., 2019; Wang & Spence, 2018), where participants could have transferred emotions experienced from listening to the music to the TDS task (for instance, negative feelings from the music could have resulted in ratings of acidity or bitterness). However, the difference in TDS responses between the music only condition in the control study and the music + wine condition in the main study does not entirely support the sensation transfer hypothesis.

It is worth stressing that the present study has several limitations. First, we had a fairly small sample size (N=21 for the main study, N=18 for the control study), so that any variability in individual taste perception could have altered the pattern of results obtained. Furthermore, we did not control for the levels of wine expertise, although all participants were self-identified wine novices. Wine novices might make less consistent ratings when it comes to wine compared to wine experts (Tempere et al., 2016). Another limitation with the design of the present study was that the silent condition was always presented first, so when participants experienced the soundtrack conditions they had already tasted the wine on one previous occasion (even if they might not have known that it was the same wine). Research has shown that when it comes to tasting several wines in a flight, the order in which the wines are tasted can play a large role in the judgment of the wines (Mantonakis et al., 2009). Therefore, the TDS difference curves between wines tasted while listening to music compared to wines tasted in silence could be due, in part, to the fact that the silent condition always came first. However, it should be stressed that the order of appearance of the two soundtracks was fully randomised, so between-music comparisons are not compromised by any possible order effects.

Moreover, we did not account for the participants' level of musical expertise. The same segment of music might possibly be associated with different tastes depending on the participants' musical expertise. For instance, Wang et al. (2015) found that a high-pitched piano piece with dissonant tonality was associated with sweetness by musical novices, but with bitterness by those with musical training. This could be due to the fact that people attend to different elements in the music depending on their level of expertise; for example, novices may tend to focus on timbre, whereas experts may tend to focus on harmony (Wolpert, 1990). Furthermore, Reinoso Carvalho and colleagues (2015) have demonstrated that using participants' individual music-chocolate associations produced more robust crossmodal effects (i.e. modulations in chocolate ratings) compared to the music-chocolate matches designed by the experimenters. Given the evidence for individual differences, it is possible that, in the present study, different participants could have attended to different tastes while listening to the same segment of music.

Looking to the future, the fruitful use of TDS in the present study opens many potential avenues of research, with a focus on the temporal aspects of the music listening experience as well as the tasting experience. As discussed in Spence and Wang (2015), most off-the-shelf music have not been ideal for research purposes since stylistic changes often occur and, unless one is careful, there is a real danger in a piece of music corresponding to different tastes/flavours. For instance, both Queen's Bohemian Rhapsody as well as the second movement of Mozart's Piano Sonata No. 12 in F Major (K332) vary between major and minor modes, which would correspond to sweet and sour/bitter tastes, respectively (Knöferle & Spence, 2012). Learning more about the temporal characteristics of such "sonic seasoning" effects could free researchers from such constraints, as well as enable experience designers to create more fluid and sophisticated experiences which take advantage of the evolving nature of both the listening and the eating/drinking experience.

REFERENCES

Ashkenazi, A., & Marks, L. E. (2004). Effect of endogenous attention on detection of weak gustatory and olfactory flavors. *Perception & Psychophysics*, **66**, 596-608.

Bachorik, J. P., Bangert, M., Loui, P., Lark, K., Berger, J. Rowe, R., & Schlaug, G. (2009). Emotion in motion: Investigating the time-course of emotional judgments of musical stimuli. *Music Perception*, **26**, 355-364.

Bigand, E., Parncutt, R., & Lerdahl, F. (1996). Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. *Perception & Psychophysics*, **58**, 125-141.

Burzynska, J. (2018). Assessing oenosthesia: Blending wine and sound. *International Journal of Food Design*, **3**, 83-101.

Chen, K., Zhou, B., Chen, S., He, S., & Zhou, W. (2013). Olfaction spontaneously highlights visual saliency map. *Proceedings of the Royal Society*, **280**:20131729.

Chiou, R., & Rich, A. N. (2012). Cross-modality correspondence between pitch and spatial location modulates attentional orienting. *Perception*, **41**, 339-353.

Crisinel, A.-S., Cosser, S., King, S., Jones, R., Petrie, J., & Spence, C. (2012). A bittersweet symphony: Systematically modulating the taste of food by changing the sonic properties of the soundtrack playing in the background. *Food Quality and Preference*, **24**, 201-204.

Foote, J. (2000). Automatic audio segmentation using a measure of audio novelty. *Proceedings of the 2000 IEEE International Conference on Multimedia and Expo*, 452-455.

Galmarini, M. V., Dufau, L., Loiseau, A.-L., Visalli, M., & Schlich, P. (2018). Wine and cheese: Two products or one association? A new method for assessing wine-cheese pairing. *Beverages*, **4**:13.

Goldstein, D. G., & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, **109**, 75-90.

Goodstein, E. S., Bohlscheid, J. C., Evans, M., & Ross, C. F. (2014). Perception of flavor finish in model white wine: A time-intensity study. *Food Quality and Preference*, **36**, 50-60.

Gray, W. B. (2007). Music to drink wine by: Vintner insists music can change wine's flavors. *San Francisco Chronicle*, **November 11th**. http://www.sfgate.com/wine/article/Music-to-drink-wine-by-Vintner-insists-music-can-3235602.php.

Johnson-Laird, P. N., Kang, O. E., & Leong, Y. C. (2012). On musical dissonance. *Music Perception: An Interdisciplinary Journal*, **30**, 19-35.

Kantono, K., Hamid, N., Shepherd, D., Lin, Y. H. T., Brard, C., Grazioli, G., & Carr, B. T. (2018). The effect of music on gelato perception in different eating contexts. *Food Research International*, **113**, 43-56.

Kantono, K., Hamid, N., Shepherd, D., Lin, Y. H. T., Skiredj, S., & Carr, B. T. (2019). Emotional and electrophysiological measures correlate to flavour perception in the presence of music. *Physiology & Behavior*, **199**, 154-164.

Kantono, K., Hamid, N., Shepherd, D., Yoo, M. J. Y., Grazioli, G., & Carr, B. T.

(2016). Listening to music can influence the hedonic and sensory perceptions of gelati. *Appetite*, **100**, 244-255.

Klapetek, A., Ngo, M. K., & Spence, C. (2012). Does crossmodal correspondence modulate the facilitatory effect of auditory cues on visual search? *Attention, Perception, & Psychophysics*, **74**, 1154-1167.

Knoeferle, K. M., Woods, A., Käppler, F., & Spence, C. (2015). That sounds sweet: Using crossmodal correspondences to communicate gustatory attributes. *Psychology & Marketing*, **32**, 107-120.

Knöferle, K., & Spence, C. (2012). Crossmodal correspondences between sounds and tastes. *Psychonomic Bulletin & Review*, **19**, 992-1006.

Kontukoski, M., Luomala, H., Mesz, B., Sigman, M., Trevisan, M., Rotola-Pukkila, M., & Hopia A. (2015). Sweet and sour: Music and taste associations. *Nutrition and Food Science*, **45,** 357-376.

Lartillot, O., & Toiviainen, P. (2007, September), a Matlab toolbox for musical feature extraction from audio. In *International conference on digital audio effects* (pp. 237-244).

Mantonakis, A., Rodero, P., Lesschaeve, I., & Hastie, R. (2009). Order in choice: Effects of serial position on preferences. *Psychological Science*, **20**, 1309-1312.

Marks, L. E., & Wheeler, M. E. (1998). Attention and the detectability of weak taste stimuli. *Chemical Senses*, **23**, 19-29.

Meillon, S., Urbano, C., & Schlich, P. (2009). Contribution of the Temporal Dominance of Sensations (TDS) method to the sensory description of subtle differences in partially dealcoholized red wines. *Food Quality and Preference*, **20**, 490-499.

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

Mossbridge, J. A., Grabowecky, M., & Suzuki, S. (2011). Changes in auditory frequency guide visual-spatial attention. *Cognition*, **121**, 133-139.

Müller, M. (2015). *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications*. Cham, Switzerland: Springer International Publishing.

Ng, M., Lawlor, J. B., Chandra, S., Chaya, C., Hewson, L., & Hort, J. (2012). Using Quantitative Descriptive Analysis and Temporal Dominance of Sensations analysis as complementary methods for profiling commercial blackcurrant squashes. *Food Quality and Preference*, **25**, 121-134.

Nunes, C. A., & Pinheiro, A. C. M. (2015). *Sensomaker user guide*. Downloaded from http://ufla.br/sensomaker/.

Oberfeld, D., Hecht, H., Allendorf, U., & Wickelmaier, F. (2009). Ambient lighting modifies the flavor of wine. *Journal of Sensory Studies*, **24**, 797.

Parrott, S., Guzman-Martinez, E., Ortega, L., Grabowecky, M., Huntington, M. D., & Suzuki, S. (2015). Direction of auditory pitch-change influences visual search for slope from graphs. *Perception*, **44**, 764-778.

Pineau, N., Schlich, P., Cordelle, S., Mathonnière, C., Issanchou, S., Imbert, A., Rogeaux, M., Etiévant, P., & Köster, E. (2009). Temporal dominance of sensations: Construction of the TDS curves and comparison with time-intensity. *Food Quality* *and Preference*, **20**, 450-455.

Pinheiro, A. C. M., Nunes, C. A., & Vietoris, V. (2013). Sensomaker: A tool for sensorial characterization of food products. *Ciência e Agrotecnologia*, **37**, 199-201.

Rabin, M. D., & Cain, W. S. (1989). Attention and learning in the perception of odor mixtures. In D. G. Laing, W. S., Cain, R. L. McBride, & B. W. Ache (Eds.). *Perception of complex smells and tastes* (pp. 173-188). Sydney, Australia: Academic Press.

Reinoso Carvalho, F., van Ee, R., Rychtarikova M., Touhafi, A., Steenhaut, K., Persoone, D., Spence, C., & Leman, M. (2015). Does music influence the multisensory tasting experience? *Journal of Sensory Studies*, **30**, 404-412.

Reinoso Carvalho, F., 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.

Simner, J., Cuskley, C., & Kirby, S. (2010). What sound does that taste? Cross-modal mappings across gustation and audition. *Perception*, **39**, 553-569.

Sokolowsky, M., & Fischer, U. (2012). Evaluation of bitterness in white wine applying descriptive analysis, time-intensity analysis, and Temporal Dominance of Sensations analysis. *Analytica Chimica Acta*, **732**, 46-52.

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

Spence, C. (2016). Oral referral: On the mislocalization of odours to the mouth. *Food Quality & Preference*, **50**, 117-128.

Spence, C., Kettenmann, B., Kobal, G., & McGlone, F. P. (2000). Selective attention to the chemosensory modality. *Perception & Psychophysics*, **62**, 1265-1271.

Spence, C., Reinoso Carvalho, F., Velasco, C., & Wang, Q. J. (in press). Extrinsic auditory contributions to food perception & consumer behaviour: An interdisciplinary review. *Multisensory Research.*

Spence, C., Richards, L., Kjellin, E., Huhnt, A.-M., Daskal, V., Scheybeler, A., Velasco, C., & Deroy, O. (2013). Looking for crossmodal correspondences between classical music & fine wine. *Flavour*, **2**:29.

Spence, C., Velasco, C., & Knoeferle, K. (2014). A large sample study on the influence of the multisensory environment on the wine drinking experience. *Flavour,* **3**:8.

Stafford, L.D., Agobiani, E., Fernandes, M. (2013). Perception of alcohol strength impaired by low and high volume distraction. *Food Quality & Preference*, **28**, 470-474.

Stafford, L. D., Fernandes, M., & Agobiani, E. (2012). Effects of noise and distraction on alcohol perception. *Food Quality & Preference*, **24**, 218-224.

Stevenson, R. J. (2012). The role of attention in flavour perception. *Flavour*, **1**:2.

Tempere, S., Hamtat, M.-L., de Revel, G., & Sicard, G. (2016). Comparison of the ability of wine experts and novices to identify odorant signals: A new insight in wine expertise. *Australian Journal of Grape and Wine Research*, **22**, 190-196.

van Ee, R., van Boxtel, J. J. A., Parker, A. L., & Alais, D. (2009). Multisensory congruency as a mechanism for attentional control over perceptual selection. *Journal* *of Neuroscience*, **29**, 11641-11649.

Velasco, C., Jones, R., King, S., & Spence, C. (2013b). Assessing the influence of the multisensory environment on the whisky drinking experience. *Flavour*, **2**:23.

Visalli, M. (2016). *Results of the analysis of the wine-cheese data collected during the social program*. Presentation given at EuroSense 2016, Dijon, France.

Wang, Q. (J.), & Spence, C. (2015). Assessing the effect of musical congruency on wine tasting in a live performance setting. *i-Perception,* **6**:3.

Wang, Q. (J.), & Spence, C. (2016). 'Striking a sour note': Assessing the influence of consonant and dissonant music on taste perception. *Multisensory Research,* **29**, 195-208.

Wang, Q. (J.), & Spence, C. (2017). Assessing the influence of music on wine perception amongst wine professionals. *Food Science and Nutrition*, **52**, 211-217.

Wang, Q. (J.), & Spence, C. (2018). 'A sweet smile': More than merely a metaphor? *Emotion and Cognition*, **32**, 1052-1061.

Wang, Q. (J.), Wang, S., & Spence, C. (2016). "Turn up the taste": Assessing the role of taste intensity and emotion in mediating crossmodal correspondences between basic tastes and pitch. *Chemical Senses*, **29**, 345-356.

Wang, Q. (J.), Woods, A., & 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.

Wang, Q. J., & Spence, C. (2018). A smooth wine? Haptic influences on wine evaluation. *International Journal of Gastronomy and Food Science,* **14**, 9-13.

Wolpert, R. S. (1990). Recognition of melody, harmonic accompaniment, and instrumentation: Musicians vs. nonmusicians. *Music Perception*, **8**, 95-105.

FIGURE LEGENDS

Figure 1. Experiment 1's TDS graphical representation for Manos Negras Pinot Noir wine, in the three auditory conditions. Top: Eno soundtrack, Middle: Mussorgsky soundtrack, Bottom: silent control condition. The curves were averaged over all participant data and smoothed.

Figure 2. TDS difference curves between A) The Eno soundtrack and silence condition, and B) The Mussorgsky soundtrack and silence condition., Only significant differences in dominance ratings between the silent and soundtrack conditions are plotted. In other words, these TDS curves showcase the net influence of background music on wine perception.

Figure 3. Pearson correlations between musical features and lagged taste curves in the music condition (blue) and silent condition (orange). Stars denote significant correlations. Colored vertical bars delimit the span of relevant lag times (see text for details).

Figure 4. Musical segments and taste curves averaged across participants. Black vertical bars mark novelty peaks in the music. Asterisks mark points of significant difference between bitterness and acidity curves.

Figure 5. TDS bandplots curves of results with music and wine versus just music, for the Mussorgsky soundtrack and Eno soundtrack. Only dominant flavours significantly greater than zero are plotted.

Figure 1

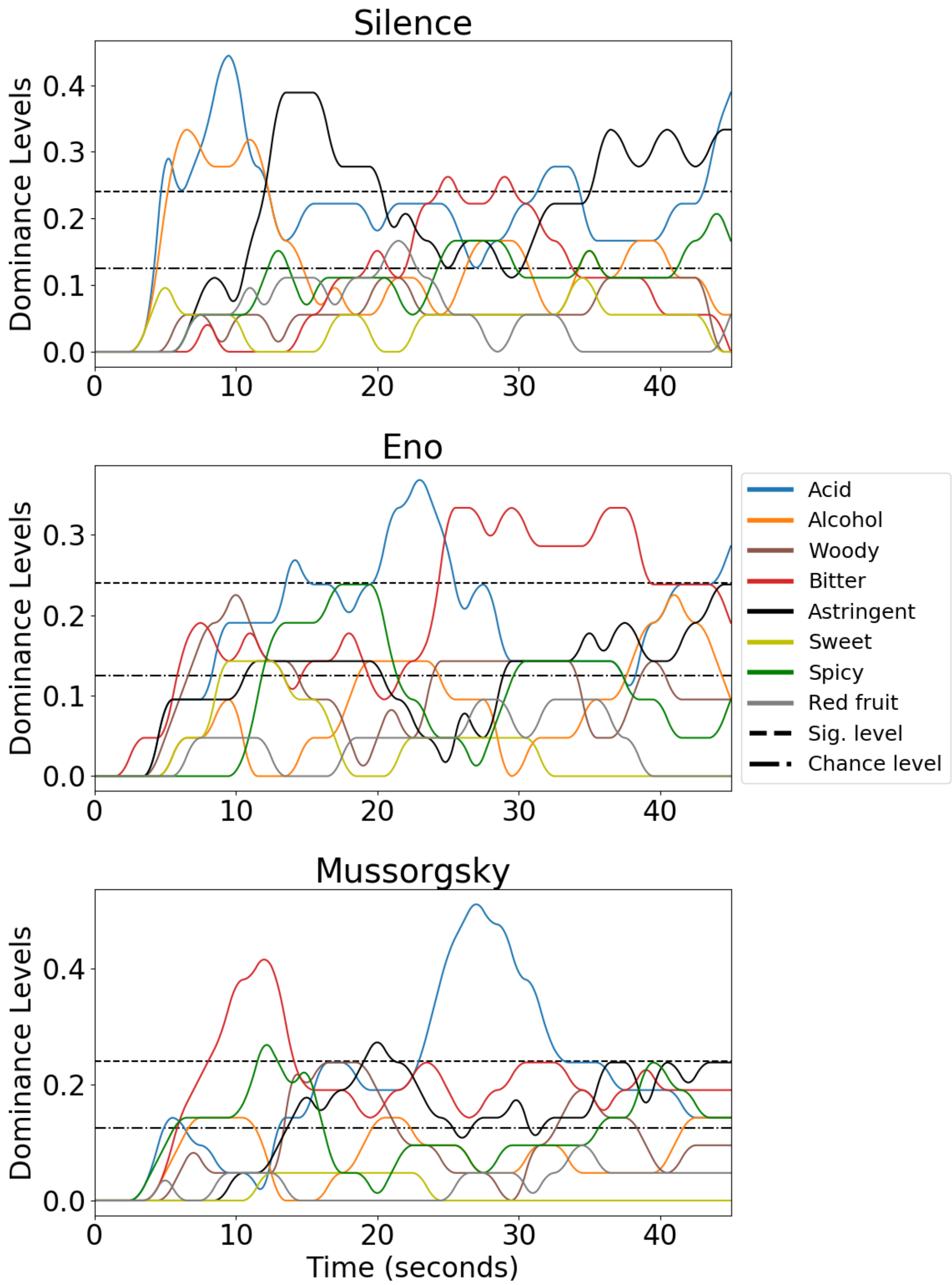

Silence
Eno
Mussorgsky
Dominance Levels
Time (seconds)
0.0
0.1
0.2
0.3
0.4
0
10
20
30
40
Acid
Alcohol
Woody
Bitter
Astringent
Sweet
Spicy
Red fruit
Sig. level
Chance level

Figure 2

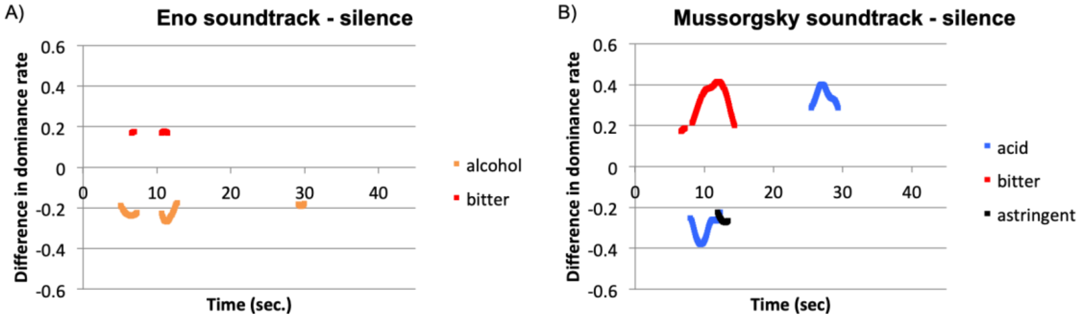

A)
Eno soundtrack - silence
Difference in dominance rate
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
10
20
30
40
Time (sec.)
alcohol
bitter
B)
Mussorgsky soundtrack - silence
Difference in dominance rate
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
10
20
30
40
Time (sec)
acid
bitter
astringent

Figure 3

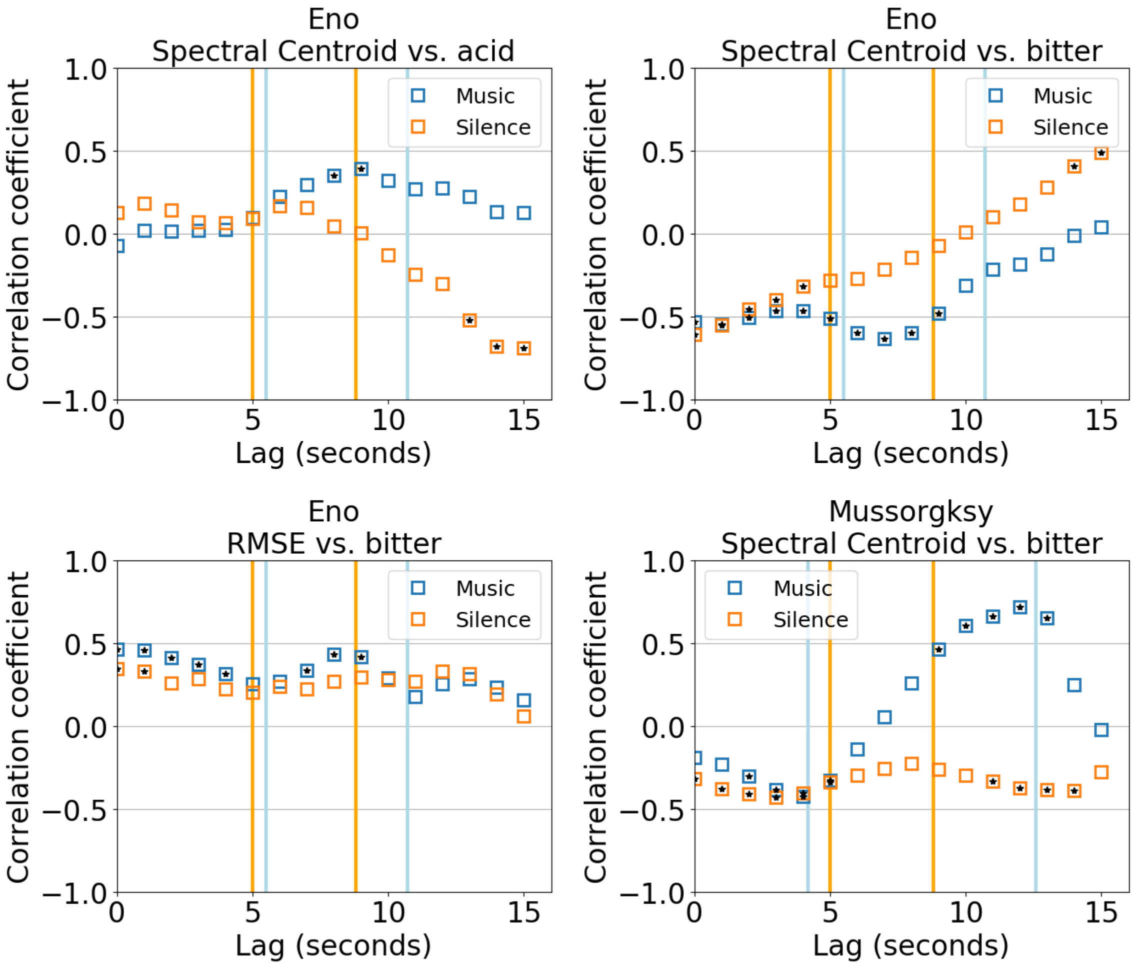

Eno
Spectral Centroid vs. acid
Eno
Spectral Centroid vs. bitter
Eno
RMSE vs. bitter
Mussorgksy
Spectral Centroid vs. bitter
Music
Silence
Correlation coefficient
Lag (seconds)
1.0
0.5
0.0
−0.5
−1.0
0
5
10
15

Figure 4

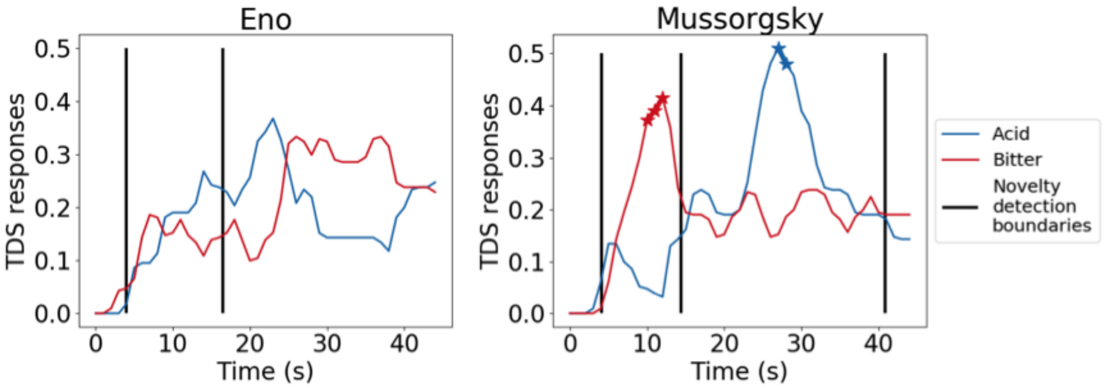
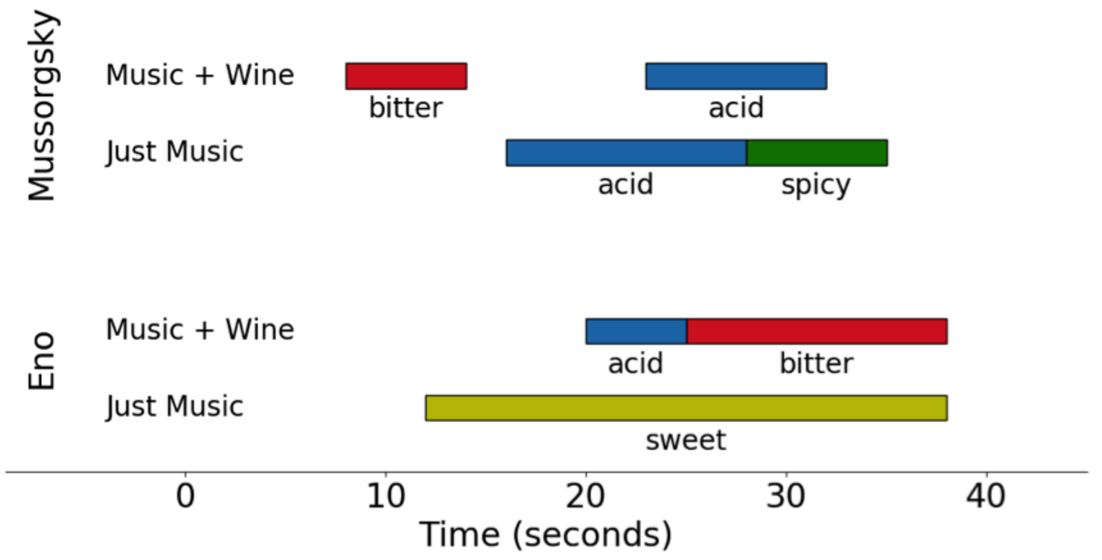

Eno
Mussorgsky
TDS responses
Time (s)
0.0
0.1
0.2
0.3
0.4
0.5
0
10
20
30
40
Acid
Bitter
Novelty detection boundaries

Figure 5


Mussorgsky
Music + Wine
bitter
acid
Just Music
acid
spicy
Eno
Music + Wine
acid
bitter
Just Music
sweet
0
10
20
30
40
Time (seconds)