

available at www.sciencedirect.comwww.elsevier.com/locate/brainres

**BRAIN
RESEARCH**

Research Report
Interactive processing of timbre dimensions: A Garner interference study
Anne Caclin^{a,b,c,d,*}, Marie-Hélène Giard^{b,c,d}, Bennett K. Smith^{a,e}, Stephen McAdams^{a,e}
^aInstitut de Recherche et Coordination Acoustique/Musique, STMS-IRCAM-CNRS, Paris, F-75004, France^bBrain Dynamics and Cognition, INSERM, U821, Lyon, F-69500, France^cInstitut Fédératif des Neurosciences, Lyon, F-69000, France^dUniv. Lyon 1, Lyon, F-69000, France^eCentre for Interdisciplinary Research in Music Media and Technology (CIRMMT), Schulich School of Music, McGill University, Montréal, Québec, Canada

ARTICLE INFO
Article history:

Accepted 21 December 2006

Available online 29 December 2006

Keywords:

Audition

Timbre

Dimensional interaction

Garner interference

Reaction time

ABSTRACT

Timbre characterizes the identity of a sound source. Psychoacoustic studies have revealed that timbre is a multidimensional perceptual attribute with multiple underlying acoustic dimensions of both temporal and spectral types. Here we investigated the relations among the processing of three major timbre dimensions characterized acoustically by attack time, spectral centroid, and spectrum fine structure. All three pairs of these dimensions exhibited Garner interference: speeded categorization along one timbre dimension was affected by task-irrelevant variations along another timbre dimension. We also observed congruency effects: certain pairings of values along two different dimensions were categorized more rapidly than others. The exact profile of interactions varied across the three pairs of dimensions tested. The results are interpreted within the frame of a model postulating separate channels of processing for auditory attributes (pitch, loudness, timbre dimensions, etc.) with crosstalk between channels.

© 2006 Elsevier B.V. All rights reserved.

1. Introduction

We can readily recognize numerous sound sources (voices, musical instruments, etc.), with a rather fine-grained resolution, allowing for example to infer a person's emotional or physical state through variations in the quality of his/her voice. Timbre processing in the auditory system appears to be a key determinant of such abilities (McAdams, 1993; Handel, 1995). Timbre is the perceptual attribute that allows one to distinguish among tones equated for pitch, loudness and perceived duration (ANSI, 1973), and for example, each musical instrument has a characteristic timbre at a given

pitch. Timbre processing remains less precisely understood than that of other auditory attributes, at both behavioral and neurophysiological levels (Hajda et al., 1997; Menon et al., 2002), partly as a consequence of the difficulty of specifying the acoustic parameters underlying timbre perception.

1.1. Timbre, a multidimensional auditory attribute

Psychophysical work on timbre perception has revealed that timbre is a multidimensional attribute, which means that timbre relationships cannot be modeled using a single timbre scale, but that instead several dimensions are needed to

* Corresponding author. INSERM U821, 69675 Bron Cedex, France. Fax: +33 4 72 13 89 01.
E-mail address: caclin@lyon.inserm.fr (A. Caclin).

account for them. Evidence for the multidimensional nature of timbre comes from studies of instrument identification (Berger, 1964; Saldanha and Corso, 1964; Hajda, 1999), ratings on verbal scales (e.g. Kendall and Carterette, 1993a,b) and multidimensional scaling (MDS) of (dis)similarity ratings (e.g. Plomp, 1970; Wedin and Goude, 1972; Grey, 1977; McAdams et al., 1995). The latter has been widely used to analyze timbre relationships on the basis of dissimilarity ratings performed by listeners for all pairs from a set of sounds. MDS allows us to uncover from these ratings a low-dimensional (usually Euclidean) space, known as a *timbre space*, in which the distances between the sounds match as closely as possible the dissimilarity ratings. Distances in a timbre space are thus perceptual distances, and the axes of this space are perceptual dimensions. Acoustic correlates can then be proposed for each perceptual dimension.

The number of timbre dimensions, as well as their potential acoustic correlates, remain a subject of debate and seem to depend partly both on the set of sounds being used and on the group of listeners. Two- to four-dimensional spaces are typically reported for studies with natural instruments or synthetic tones imitating instruments of the orchestra. Consistently across studies, attack time and spectral centroid (amplitude-weighted mean frequency of the spectrum energy) are reported as major determinants of timbre (e.g. Grey, 1977; Krumhansl, 1989; McAdams et al., 1995; Samson et al., 1997; Lakatos, 2000; Marozeau et al., 2003). Other proposed acoustic correlates include spectral dimensions, such as spectrum fine structure (Krimphoff et al., 1994) or spectral spread (Marozeau et al., 2003), as well as spectro-temporal characteristics including spectral flux (Grey, 1977; McAdams et al., 1995). Recent psychophysical work using carefully controlled synthetic tones confirmed the perceptual salience of attack time, spectral centroid, and spectrum fine structure, whereas the salience of spectral flux appeared more restricted (Caclin et al., 2005).

This multidimensional account of timbre raises the question of whether timbre 'dimensions' are processed separately or holistically in the auditory system. There is indeed an apparent paradox between a global account of timbre as the perceptual identity of a specific sound source and this psychophysical account of timbre as a collection of timbre dimensions. We have already addressed the question of holistic versus separate processing of timbre dimensions at the level of auditory sensory memory, using the Mismatch Negativity (MMN) component of the auditory event-related potential (Caclin et al., 2006). The MMN is recorded after tones deviating along at least one acoustic dimension from repetitive 'standard' tones. The MMN is assumed to index a change-discrimination process comparing the incoming (deviant) tone with the representation of the past tones (standards) in sensory memory (see Näätänen and Winkler, 1999, for a review). The MMN can thus be used to probe the representations in auditory sensory memory (see Ritter et al., 1995). We have shown that three dimensions of timbre corresponding acoustically to attack time, spectral centroid, and spectrum fine structure (as described in Caclin et al., 2005) are mainly processed separately in auditory sensory memory, as MMNs elicited by changes along these different dimensions were additive and exhibited at least partly

anatomically distinct generators (Caclin et al., 2006). Yet this study using the MMN only probes stages of processing implicating sensory memory operations. The purpose of the present study was thus to explore whether this property of separability exists along the whole sensory processing pathway.

1.2. Garner interference as a probe of dimensional interactions

Garner (1974) and his followers have developed an experimental paradigm assessing the processing separability of different stimulus dimensions. The core idea in this approach is that if the processing of two dimensions is separate, then performance in a classification task based on a particular dimension of the stimulation (referred to as the *critical dimension*) should be unaffected by variations along the other dimension (the *separability property*). Should the reverse be observed, then there exist interactions during the processing of these two dimensions (referred to as the *integrality of the dimensions*). Garner's classical paradigm involves creating four stimuli by combining two values along the two dimensions of interest (see Fig. 1a). Participants are requested to categorize the stimuli according to the critical dimension in five different stimulus contexts: the other dimension is either kept at a constant value (two *baseline conditions*), or varies systematically with the dimension of interest (two *correlation conditions*), or varies independently of the dimension of interest (*filtering condition*). The respective roles of the two dimensions are then exchanged, leading to a total of 10 different experimental conditions. Most of the time, this paradigm is used in a speeded version, and the performance measured is reaction time (RT). *Separability* is revealed by equal reaction times in baseline, correlation, and filtering conditions for the two dimensions, whereas a decrease in performance in the filtering condition and/or a gain in the correlation conditions (both relative to baseline) for the two dimensions characterize *integral dimensions* (see Fig. 1b). For example in the visual modality, size and brightness are described as *separable dimensions*, whereas brightness and saturation exhibit *integrality* (Garner and Felfoldy, 1970).

When Garner interference is observed between two dimensions, one usually tries to determine the direction of the crosstalk between the channels of processing of these two dimensions and to interpret integrality in terms of levels of processing where interactions could occur (see for example Melara and Marks, 1990a). It must be emphasized that a major prerequisite for such interpretations is that performance in the baseline conditions for the two dimensions should be equal (see Melara and Mounts, 1993, 1994). Regarding the direction of crosstalk issue, if there is an effect of the experimental condition when judging any of the two dimensions, then the crosstalk between channels is considered *bidirectional*. Conversely, if judgments for only one dimension are affected by variations along the other dimension, then the crosstalk is likely to be *unidirectional* (Melara and Marks, 1990a).

Possible levels of processing where crosstalk can occur include perceptual processing, semantic evaluation of the

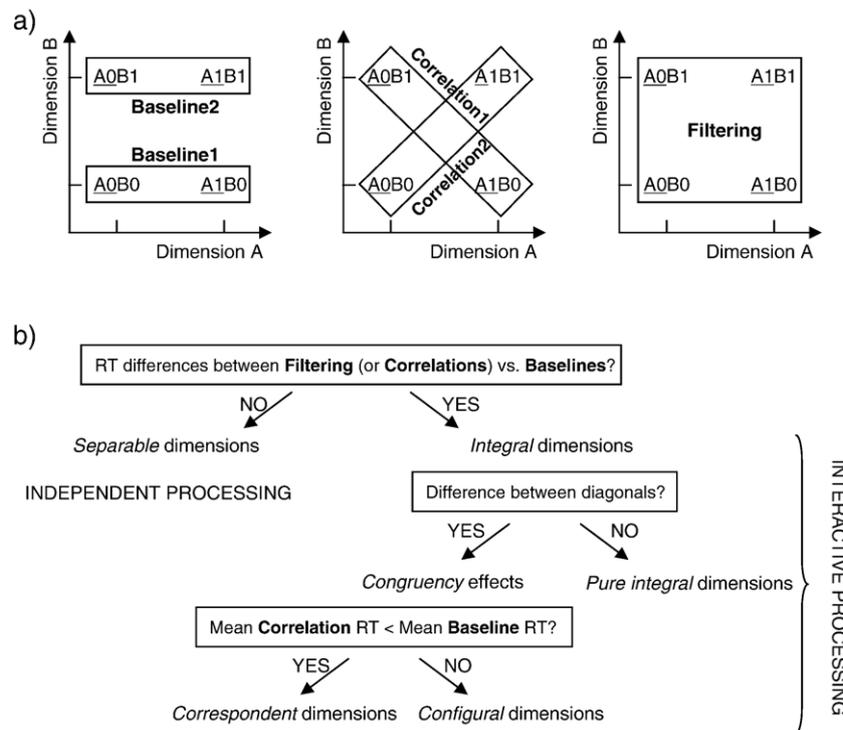


Fig. 1 – (a) Principle of Garner's interference paradigm. In this example, the participant has to classify the stimuli according to dimension A (as underlined), in five types of stimulus sequences represented by rectangles. (b) Identification of Garner's interference syndrome. The main criteria used to distinguish between interference profiles are represented.

stimuli, and response selection. There is no simple rule for choosing the most likely candidate in the general case, but analysis of *congruency* effects allows one to form a first hypothesis. In the context of Garner's interference studies, congruency effects measure the difference in performance for the two diagonals of the square formed by the four stimuli (see Fig. 1a). With the notations of Fig. 1a, this means that congruency analysis explores whether RTs are systematically smaller (or larger) for A0B1 and A1B0 than for A0B0 and A1B1, irrespective of the experimental condition (Baseline, Correlation, or Filtering). Such congruency effects have been observed, for example, for crossmodal associations involving visual vertical position and auditory pitch (Melara and O'Brien, 1987), where the high position/high pitch and low position/low pitch stimuli are classified faster than are the other two combinations. When congruency effects are observed, the dimensions are denoted as *correspondent* or *configural* (see Fig. 1b), depending on the performance in an additional divided attention task implicating both dimensions (the participant is requested to classify the stimuli according to the diagonal of the stimulus square to which they belong, i.e. A0B0 and A1B1 vs. A0B1 and A1B0, with the notations of Fig. 1a) and/or whether the mean performance in the correlation conditions is better or not than baseline performance (Macmillan and Kingston, 1995). For *correspondent* (vs. *configural*) dimensions, performance is better (vs. worst) in the filtering condition than in the divided attention condition and the mean performance in correlation conditions is better (vs. worst) than baseline performance (see Fig. 1b). Based on the syndrome of interference observed—*pure integrality* (without congruency effects), *correspondence*, or

configurality—different levels of interaction are generally proposed.

Correspondence is often observed when semantic associations between the values along the two dimensions exist and is therefore usually taken as a sign of interactions occurring mostly at a semantic level (Melara and O'Brien, 1987; Melara and Marks, 1990a,d), whereas *pure integrality* would reflect interactions at a perceptual stage of processing (Ben-Artzi and Marks, 1995). *Configurality* arises when the combination of values along both dimensions of interest elicits changes along another perceptible dimension of the stimulus (Pomerantz and Garner, 1973; Lockhead and King, 1977; Melara et al., 1992), and would reflect the cost of ignoring this additional irrelevant dimension. For example, variations in perceived surface area and shape of rectangles can emerge from controlled variations of the length of the vertical and horizontal sides of rectangles (see Garner, 1974).

Interactions among various auditory attributes have been investigated using Garner's paradigm. Pitch, loudness, and timbre have been described as *integral* or *correspondent* dimensions (Grau and Kemler-Nelson, 1988; Melara and Marks, 1990b,c; Krumhansl and Iverson, 1992), and this is taken as the consequence of interactions occurring mostly at early perceptual stages of processing. Congruency effects are observed for pitch and loudness, with high-pitched/loud and low-pitched/soft tones being classified faster than low-pitched/loud and high-pitched/soft tones (e.g. Melara and Marks, 1990c). Congruency effects involving timbre (taken as a whole) are more complex and depend on the exact acoustic attributes that have been varied to create timbre

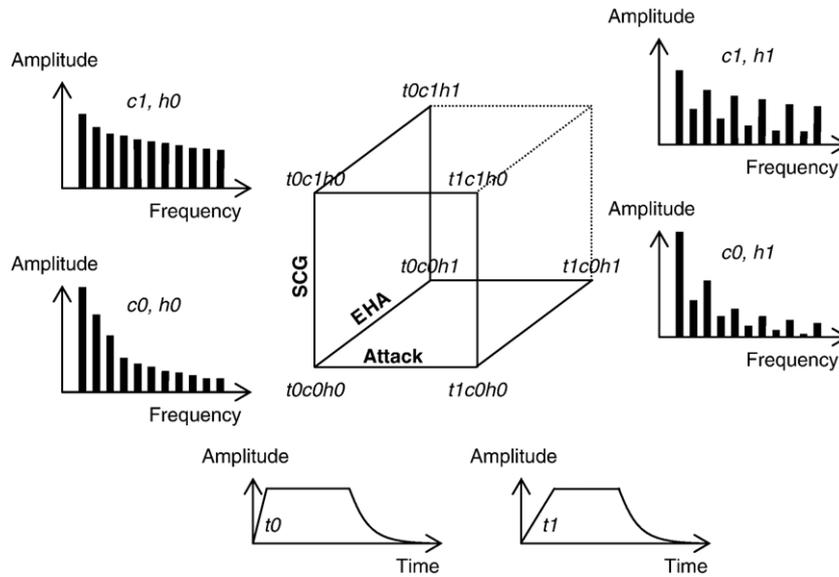


Fig. 2 – Stimulus construction. All stimuli have the same pitch, loudness and perceived duration but differ according to three timbre dimensions (Attack, SCG=Spectral Center of Gravity, and EHA=Even Harmonics Attenuation). Sounds are labeled as in **Table 1**: for Attack time, $t_0=15$ ms, $t_1=100$ ms; for SCG, $c_0=3$ (harmonic rank units), $c_1=4.5$; for EHA, $h_0=0$ dB, $h_1=-10$ dB (attenuation of even relative to odd harmonics). The three stimulus squares used in this study are represented with solid lines.

differences. Taken together these studies suggest that interactive (as opposed to separate) processing among auditory attributes is the rule.

1.3. Rationale of the study

The present study sought to clarify possible interactions in the processing of timbre dimensions. Stimuli were fully controlled synthetic tones adapted from previous work (Caclin et al., 2005, 2006), and the three acoustic cues known to give rise to the timbre dimensions investigated were attack time, spectral center of gravity (SCG), and spectrum fine structure. In a previous study, we found evidence for separate representations of these timbre dimensions in auditory sensory memory (Caclin et al.,

2006). The present study uses the Garner interference paradigm to test whether these timbre dimensions, two related to the frequency spectrum shape and one to the temporal envelope, exhibit interactive processing at other stages of analysis than sensory memory. So far, Garner interference has been observed for all pairs of auditory dimensions investigated (see the examples above for pitch, loudness, and timbre), and we therefore hypothesized that timbre dimensions would exhibit Garner interference. In the study reported here, however, we also intended to test whether interaction profiles would be similar or not across different pairs of timbre dimensions, to appreciate whether the mechanisms of such interactions are ubiquitous or not, particularly as concerns temporal-spectral and spectral-spectral interactions.

Table 1 – Stimuli used in each experimental condition (see Fig. 1a) for the three pairs of dimensions

| Pair of dimensions | Dimension to judge | Baseline 1 | Baseline 2 | Correlation 1 | Correlation 2 | Filtering |
|--------------------|--------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| Attack/SCG | Attack | <u>t0c0h0</u> <u>t1c0h0</u> | <u>t0c1h0</u> <u>t1c1h0</u> | <u>t0c1h0</u> <u>t1c0h0</u> | <u>t0c0h0</u> <u>t1c1h0</u> | <u>t*c*h0</u> <u>t*c*h0</u> |
| | SCG | <u>t0c0h0</u> <u>t0c1h0</u> | <u>t1c0h0</u> <u>t1c1h0</u> | <u>t0c1h0</u> <u>t1c0h0</u> | <u>t0c0h0</u> <u>t1c1h0</u> | <u>t*c*h0</u> <u>t*c*h0</u> |
| Attack/EHA | Attack | <u>t0c0h0</u> <u>t1c0h0</u> | <u>t0c0h1</u> <u>t1c0h1</u> | <u>t0c0h1</u> <u>t1c0h0</u> | <u>t0c0h0</u> <u>t1c0h1</u> | <u>t*c0h*</u> <u>t*c0h*</u> |
| | EHA | <u>t0c0h0</u> <u>t0c0h1</u> | <u>t1c0h0</u> <u>t1c0h1</u> | <u>t0c0h1</u> <u>t1c0h0</u> | <u>t0c0h0</u> <u>t1c0h1</u> | <u>t*c0h*</u> <u>t*c0h*</u> |
| SCG/EHA | SCG | <u>t0c0h0</u> <u>t0c1h0</u> | <u>t0c0h1</u> <u>t0c1h1</u> | <u>t0c0h1</u> <u>t0c1h0</u> | <u>t0c0h0</u> <u>t0c1h1</u> | <u>t0c*h*</u> <u>t0c*h*</u> |
| | EHA | <u>t0c0h0</u> <u>t0c0h1</u> | <u>t0c1h0</u> <u>t0c1h1</u> | <u>t0c0h1</u> <u>t0c1h0</u> | <u>t0c0h0</u> <u>t0c1h1</u> | <u>t0c*h*</u> <u>t0c*h*</u> |

Sounds are labeled as in Fig. 2: for Attack time, $t_0=15$ ms, $t_1=100$ ms, $t^*=t_0$ and t_1 ; for SCG (Spectral Center of Gravity), $c_0=3$ (harmonic rank units), $c_1=4.5$, $c^*=c_0$ and c_1 ; for EHA (Even-Harmonic Attenuation), $h_0=0$ dB, $h_1=-10$ dB (attenuation of even relative to odd harmonics), $h^*=h_0$ and h_1 . The relevant parameter for each task is underlined.

2. Results

Three groups of 21 participants were tested with one of three pairs of timbre dimensions: Attack/Spectral Center of Gravity (SCG), Attack/Even Harmonics Attenuation (EHA), and SCG/EHA (see Fig. 2). Each participant was tested successively for the two criterial dimensions in the five conditions of the Garner interference paradigm (see Fig. 1a and Table 1 for details about the experimental conditions). The sounds were designed to give rise to equal performance in all Baseline conditions. Data from participants who made more than 20 errors (out of 48 trials) during any of the 10 blocks performed were excluded from the analysis. This concerned only one participant in each of the three groups. Mean RTs (across the 20 remaining participants per group) for the different experimental conditions are presented in Fig. 3a.

2.1. RT analysis

For the two criterial dimensions of each of the three pairs of dimensions investigated, RTs were analyzed using a one-way ANOVA with Condition as a five-level (Baseline 1, Baseline 2, Correlation 1, Correlation 2, and Filtering, see Table 1 for details) within-subject factor. In each case, there was a significant effect of the Condition factor on RTs (see Table 2), which was explored using a Tukey–Kramer HSD post-hoc test. The results of these tests are summarized in Fig. 3b. For Attack judgments in the Attack/SCG pair, this effect of Condition on RTs arose from slower RTs in the Filtering condition relative to the others ($p < 0.02$), and faster RTs in the Correlation 1 condition relative to both Baselines ($p < 0.05$). Other comparisons were not significant ($p > 0.30$). For SCG judgments in the Attack/SCG pair, the effect of the condition factor was due to slower RTs in the Filtering condition relative to the others ($p < 0.007$). Other comparisons were not significant ($p > 0.17$).

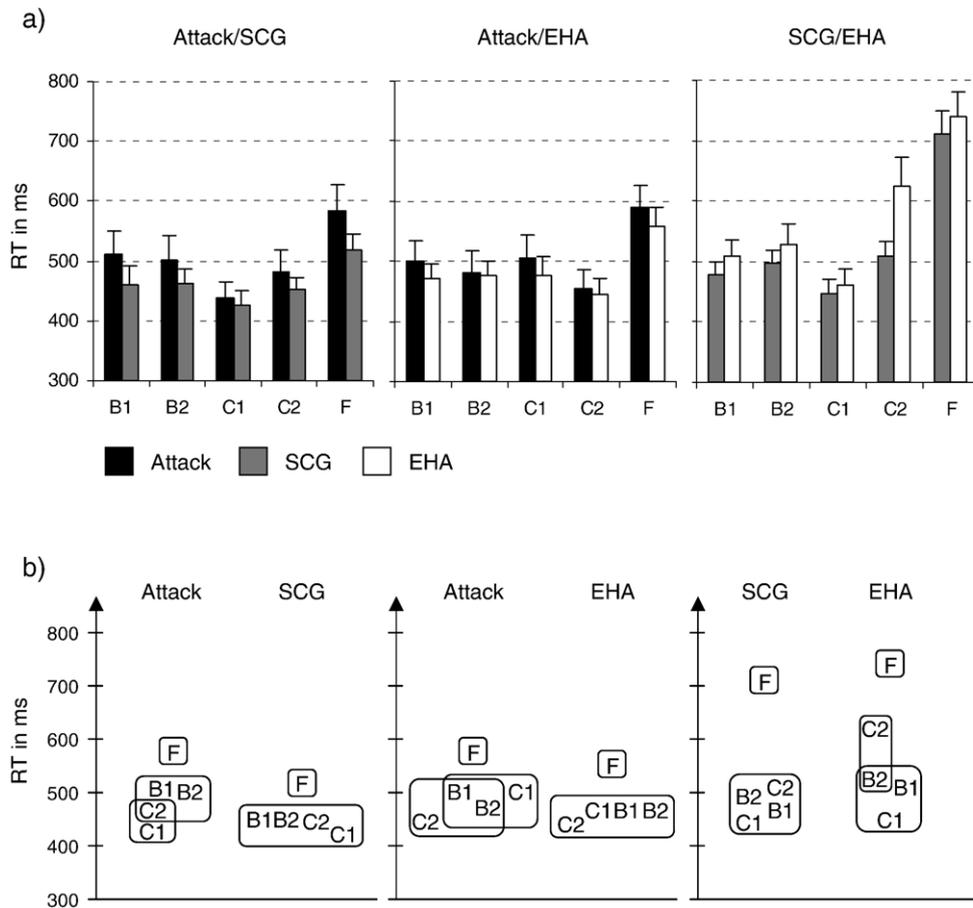


Fig. 3 – Reaction times. (a) RTs are plotted separately for each dimension in each pair of dimensions, with error bars representing the standard error of the mean. B = Baseline, C = Correlation, F = Filtering (see Table 1 for details about experimental conditions). Correlation condition names were randomly assigned at the beginning of the experiment, and it turned out that the Congruent Correlation condition was C1 for Attack/SCG and SCG/EHA, but C2 for Attack/EHA. Note that for each pair of dimensions the baseline conditions (B1 and B2) include different sounds according to the dimension to be judged, whereas the same sounds are presented for the correlation and filtering conditions, irrespective of the actual dimension to be judged (see Table 1). (b) Summary of the comparison between the conditions (Tukey–Kramer HSD post-hoc tests, see Text) for each dimension in each pair. Conditions are ordered vertically along the RT axis, and rectangles denote groups with indistinguishable RTs.

Table 2 – Results of the ANOVAs with Condition as a within-subject factor on RTs and on error rates for the three pairs of dimensions

| Pair of dimensions | Dimension to judge | RT analysis | | | Error analysis | | |
|--------------------|--------------------|-------------|--------------------|---------------------------------|----------------|--------------------|---------------------------------|
| | | F(4,76) | ε (GG) | <i>p</i> | F(4,76) | ε (GG) | <i>p</i> |
| Attack/SCG | Attack | 10.688 | 0.750 | ≤ 0.0001 | 7.646 | 0.710 | 0.0003 |
| | SGG | 8.753 | 0.658 | 0.0002 | 3.927 | 0.568 | 0.023 |
| Attack/EHA | Attack | 17.067 | 0.621 | ≤ 0.0001 | 4.072 | 0.475 | 0.027 |
| | EHA | 8.160 | 0.615 | 0.0004 | 2.854 | 0.406 | 0.083 |
| SCG/EHA | SCG | 38.035 | 0.569 | ≤ 0.0001 | 8.398 | 0.431 | 0.002 |
| | EHA | 18.373 | 0.602 | ≤ 0.0001 | 23.869 | 0.537 | ≤ 0.0001 |

Significant effects are highlighted in bold. GG=Greenhouse–Geisser.

For the *Attack/Even-Harmonic Attenuation (EHA)* pair, the effect of the Condition factor on RT for Attack judgments was due to slower RTs in the Filtering condition relative to the others ($p < 0.0002$), and slower RTs in Correlation 1 than in Correlation 2 ($p = 0.04$). None of the other comparisons were significant ($p > 0.08$). For EHA judgments, RTs were slower in the Filtering condition relative to all four others ($p < 0.003$), with no other significant comparisons ($p > 0.58$).

Finally, for the *SCG/EHA* pair, slower RTs were observed in the Filtering condition relative to all four others ($p < 0.0002$) for SCG judgments, with no other significant comparison ($p > 0.08$). Similarly for EHA judgments Filtering RTs were slower than RTs in other conditions ($p < 0.02$), with additionally slower RTs in the Correlation 2 condition relative to both Baselines ($p = 0.02$ and $p = 0.08$) and Correlation 1 ($p = 0.0004$). All other comparisons were not significant ($p > 0.36$).

In summary, for all three pairs of dimensions, RT varied according to the experimental condition with, in particular, a systematic increase in RT for the Filtering condition (see Fig. 3b). This was always true for both dimensions, and we can thus conclude that all three pairs of dimensions are *non-separable* but rather exhibit *bi-directional crosstalk* during their processing. Analysis of congruency effects will allow us to further characterize the interaction patterns for these three pairs of timbre dimensions.

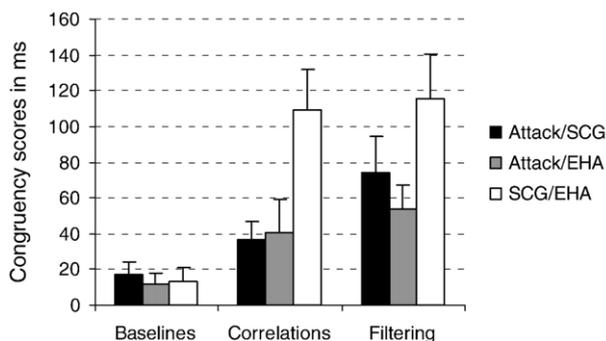


Fig. 4 – Congruency effects. Congruency scores (RT difference between the two diagonals of the stimulus squares: Incongruent diagonal – Congruent diagonal) are plotted separately for each condition and each pair of dimensions, after averaging the data for the two dimensions of a pair. Error bars represent the standard error of the mean.

2.2. Congruency effects

Congruency effects were also examined in the RT data. For each criterial dimension, mean RTs were computed for A0B1 and A1B0 (first diagonal) and for A0B0 and A1B1 (second diagonal) in the combined Baselines, combined Correlations, and Filtering conditions (notations from Fig. 1a). For each pair of dimensions, congruency was analyzed using a three-way ANOVA with Dimension (2 levels), Congruency (2 levels corresponding to the two diagonals), and Condition (3 levels: Baseline, Correlation, and Filtering) as within-subject factors. A significant effect of the Congruency factor would reveal congruency effects as discussed in the Introduction section. As such effects were observed, the diagonals (and the correlation conditions) were subsequently labeled ‘congruent’ and ‘incongruent’ according to mean RTs (‘congruent’ being the one with shorter RTs). Congruency scores (RT difference between the two diagonals of the stimulus square; see Experimental procedures) for the different pairs of dimensions are illustrated Fig. 4.

The results of the three-way Congruency \times Condition \times Dimension ANOVAs on RTs for the three pairs of dimensions are presented in Table 3. Significant congruency effects were observed for all three pairs of dimensions. As expected from the previous RT analyses, the effect of the Condition factor was significant for all three pairs of dimensions and will not be explored further here. The effect of the Dimension factor was not significant for any pair, so within a given pair of dimensions RTs did not depend on the criterial dimension. Interestingly, for the *Attack/SCG* and *SCG/EHA* pairs there was a significant interaction between Congruency and Condition. This interaction was also marginally significant for the *Attack/EHA* pair. This result implies that congruency effects vary as a function of the experimental condition (i.e. the stimulus context). To explore this interaction, we computed congruency scores (i.e. RT for the Incongruent diagonal minus RT for the Congruent diagonal) for each dimension in each of the three pairs, for the combined Baselines, combined Correlations, and Filtering conditions. Mathematically, stating that there is a significant Congruency \times Condition interaction in the RT data is equivalent to stating that there is a significant effect of Condition on congruency scores. The effects of Condition on congruency scores were explored using Tukey–Kramer HSD post-hoc tests.

For the *Attack/SCG* pair, the effect of Condition on congruency scores is mainly due to higher congruency scores in

Table 3 – Results of the ANOVAs exploring Congruency for all three pairs of dimensions

| Factor (df) | Attack/SCG | | | Attack/EHA | | | SCG/EHA | | |
|---|------------|-----------------|----------------|------------|-----------------|----------------|---------|-----------------|----------------|
| | F | ϵ (GG) | p | F | ϵ (GG) | p | F | ϵ (GG) | p |
| Congruency (1,19) | 20.090 | 1 | 0.0003 | 15.065 | 1 | 0.001 | 36.939 | 1 | ≤0.0001 |
| Condition (2,38) | 30.141 | 0.984 | ≤0.0001 | 47.225 | 0.742 | ≤0.0001 | 87.098 | 0.689 | ≤0.0001 |
| Dimension (1,19) | 3.100 | 1 | 0.09 | 0.524 | 1 | 0.48 | 2.195 | 1 | 0.15 |
| Congruency × Condition (2,38) | 6.148 | 0.796 | 0.009 | 2.841 | 0.766 | 0.09 | 9.298 | 0.982 | 0.0006 |
| Congruency × Dimension (1,19) | 2.129 | 1 | 0.16 | 0.980 | 1 | 0.33 | 1.431 | 1 | 0.25 |
| Condition × Dimension (2,38) | 1.465 | 0.663 | 0.25 | 0.154 | 0.721 | 0.79 | 0.428 | 0.915 | 0.64 |
| Congruency × Condition × Dimension (2,38) | 2.190 | 0.903 | 0.13 | 2.003 | 0.823 | 0.16 | 4.363 | 0.884 | 0.025 |

Significant effects are highlighted in bold. GG=Greenhouse–Geisser.

the Filtering condition compared to the Baseline condition ($p=0.004$) and to the Correlation condition ($p=0.07$, marginal). Conversely there was no difference between the Baseline and Correlation conditions ($p=0.47$). For the *Attack/EHA* pair, the marginal effect of the Condition factor was mostly due to a marginal difference between Filtering and Baseline conditions ($p=0.06$). Other comparisons were not significant ($p>0.25$). Finally for the *SCG/EHA* pair, the effect of the Condition factor arose mostly from a difference between the Baseline and the other conditions ($p=0.002$ for Correlation and $p=0.001$ for Filtering). There was no difference between Correlation and Filtering conditions ($p=0.98$). For this last pair of dimensions, a significant three-way Congruency × Condition × Dimension interaction was observed in the ANOVA on RTs (see Table 3). This effect seemed to arise from a difference in congruency scores between Baseline and Correlation conditions for EHA judgments only ($p=0.0002$ for EHA judgments and $p=0.33$ for SCG judgments).

To summarize the congruency analysis, we found that all three pairs of timbre dimensions investigated here exhibit congruency effects. These congruency effects depended on the experimental conditions (with, in particular, a difference between Baseline and Filtering conditions). Further, this dependency has a different pattern for the *SCG/EHA* pair compared with the other two pairs. Direct comparison of the RT data and congruency scores between pairs of dimensions will allow us to clarify this point.

2.3. Comparison between pairs of dimensions

All three pairs of dimensions display common characteristics in the data patterns: Filtering has longer RTs than Baseline and Correlation conditions, and RTs in the Correlation conditions are faster than or equal to those in the Baseline conditions with the exception of the Incongruent Correlation in the *SCG/EHA* pair. Further, all three pairs of dimensions exhibit congruency effects, which depend on the experimental condition (Baseline, Correlation, or Filtering).

2.3.1. RT analysis

In order to compare directly the results for the three pairs of dimensions, we combined for each subject the data from the two criterial dimensions by averaging the RTs for i) the four Baseline conditions, ii) the two Congruent Correlation conditions (Correlation 1 for *ATT/SCG* and *SCG/EHA*, Correlation 2 for *ATT/EHA*, see Fig. 3), iii) the two Incongruent Correlation

conditions (Correlation 2 for *ATT/SCG* and *SCG/EHA*, Correlation 1 for *ATT/EHA*, see Fig. 3), and iv) the two Filtering conditions. This approach allowed us to analyze RTs for all three pairs of dimensions using an ANOVA with Pair-of-dimensions (3 levels) and Condition (4 levels: Baseline, Congruent Correlation, Incongruent Correlation, and Filtering) as factors (see Experimental procedures). In this Pair-of-dimensions × Condition ANOVA, we found a significant effect of the Condition factor ($F(3,171)=109.837$, $\epsilon=0.857$, $p\leq 0.0001$), a marginal effect of the Pair-of-dimensions factor ($F(2,57)=2.968$, $p=0.06$), and a significant interaction between these two factors ($F(3,171)=11.327$, $\epsilon=0.857$, $p\leq 0.0001$). A Tukey–Kramer HSD post-hoc test revealed that for the *SCG/EHA* pair only, there was a significant difference in RTs between the Incongruent Correlation and Baseline ($p=0.01$) and between the Incongruent Correlation and the Congruent Correlation ($p\leq 0.0001$) (for the other two pairs, $p>0.23$).

To explore further the differences between the pairs of dimensions, we computed for each subject the RT difference between Baseline (mean of the four Baseline conditions, as described above) and the other three conditions and analyzed this RT difference for each condition in a one-way ANOVA with Pair-of-dimensions as a between-subjects factor. Baseline RTs were virtually equal among all pairs of dimensions as a consequence of stimulus construction (this was checked in a one-way ANOVA with Pair-of-dimensions as a between-subjects factors on mean Baseline RTs: $F(2,57)=0.211$, $p=0.81$). The effect of the Pair-of-dimensions factor was significant for the Incongruent Correlation minus Baseline ($F(2,57)=5.951$, $p=0.0045$) and for the Filtering minus Baseline ($F(2,57)=23.351$, $p\leq 0.0001$) cases, but not for Congruent Correlation minus Baseline case ($F(2,57)=0.928$, $p=0.40$). The Pair-of-dimensions factor effects on these RT differences arose from a difference between the *SCG/EHA* pair and the other two pairs, for both Incongruent Correlation ($p=0.004$ with *Attack/SCG* and $p=0.06$ with *Attack/EHA*) and Filtering ($p\leq 0.0001$ with *Attack/SCG* and *Attack/EHA*). To summarize, there is a greater cost of Filtering (relative to Baseline) for the *SCG/EHA* pair than for the other two pairs, and for the *SCG/EHA* pair only there is an increase in RT in Incongruent Correlation (primarily due to EHA judgments) relative to Baseline.

2.3.2. Congruency analysis

Finally, to compare congruency effects between the three groups of subjects, we averaged for each subject the congruency scores for the two criterial dimensions in each of the

three conditions (combined Baselines, combined Correlations, and Filtering). An ANOVA on congruency scores with Pair-of-dimensions (3 levels) and Condition (3 levels) as factors led to a significant effect of the Pair-of-dimensions factor ($F(2,57)=4.957$, $p=0.01$) and of the Condition factor ($F(2,114)=16.736$, $\epsilon=0.976$, $p\leq 0.0001$). The interaction between these two factors was also marginally significant ($F(2,114)=2.323$, $\epsilon=0.976$, $p=0.06$). The effect of Pair-of-dimensions was due to a difference between the SCG/EHA pair and the other two pairs ($p<0.05$) and the effect of Condition to a difference between Baseline and the other two conditions ($p<0.0005$).

We also analyzed congruency scores separately for each condition in a one-way ANOVA with Pair-of-dimensions as a between-subjects factor. The effect of this factor was significant for Correlation ($F(2,57)=5.262$, $p=0.008$) and also marginally significant for Filtering ($F(2,57)=2.487$, $p=0.09$). It was not significant for Baseline ($F(2,57)=0.155$, $p=0.86$). The effect observed for Correlation was due to a difference between SCG/EHA and the other two pairs ($p<0.02$). Overall, congruency scores were greater for the SCG/EHA pair, especially in the Correlation conditions.

2.4. Error analysis

Additionally, error rate data were analyzed using the same procedure as for the RT data. The results of the error analysis were similar overall to those obtained with the RT data, arguing for the absence of a speed–accuracy trade-off in the experiment, and are therefore only briefly summarized here. For the two criterial dimensions of each of the three pairs of dimensions investigated, we explored the error rates using an ANOVA with Condition as a five-level within-subject factor (as for RTs in section 2.1). There was a significant effect of the Condition factor on error rates (see Table 2) for both dimensions in all pairs except EHA in the Attack/EHA pair where this effect was only marginally significant. The highest error rates were observed in the Filtering conditions. Further for each pair of dimensions we found significant Congruency effects in the error data, which, as observed for RT data, depended on the experimental condition, with the largest congruency effects in the Filtering condition.

The comparison of error rates for the three pairs of timbre dimensions tested revealed larger error rates for the SCG/EHA pair than for the other two pairs in the Filtering condition. Congruency scores computed with error rates were also higher for the SCG/EHA pair than for the other two pairs, particularly in the Filtering condition.

3. Discussion

3.1. Processing of timbre dimensions

The present study highlights the existence of interactions in the processing of different timbre dimensions. The results confirm our hypothesis that timbre dimensions should exhibit Garner interference. For all three pairs of dimensions tested, we have indeed observed that RTs and error rates varied as a function of the experimental condition, with in particular an increase of RT and error rate in the Filtering condition relative

to Baseline in all cases, i.e. a filtering cost. The exact pattern of interactions departed from that of *integral* dimensions, because we observed redundancy gains in only one of the two Correlation conditions relative to Baseline (Correlation 1 for Attack/SCG and SCG/EHA, Correlation 2 for Attack/SCG, see Table 1 and Fig. 3), and often only as numerical trends, whereas redundancy gains are typically observed in both correlation conditions for *integral* dimensions (e.g. Garner and Felfoldy, 1970). A related finding was that of significant congruency effects for all three pairs of timbre dimensions. These congruency effects were coherent across the three pairs of dimensions: knowing which diagonal of the stimulus square was the congruent one for two pairs of dimensions (e.g. sharp attack/high SCG and slow Attack/low SCG for Attack/SCG; sharp attack/no EHA and slow attack/high EHA for Attack/EHA) allows us to guess correctly which diagonal will be congruent for the third pair (i.e. high SCG/no EHA, corresponding both to ‘sharp attack’, and low SCG/high EHA, corresponding to ‘slow attack’; see Melara and Marks, 1990b, for an example where it is not the case). Congruency effects thus allow us to distinguish two sets of congruent attributes: sharp attack, brilliant (high SCG), and neutral (no EHA), on the one hand, and slow attack, dull (low SCG), and hollow (high EHA), on the other hand. Taken together these results argue for timbre dimensions to be *correspondent* or *configural* dimensions, that is interacting dimensions exhibiting congruency effects (see Fig. 1b).

Following Macmillan and Kingston (1995), we computed the mean RT in the Correlation conditions and compared it to the mean RT in the Baseline conditions for each pair of dimensions, in order to distinguish between *correspondent* and *configural* dimensions (mean RT in the Correlation conditions is shorter than in the Baseline conditions for *correspondent* dimensions, and the reverse is true for *configural* dimensions, see Fig. 1b). According to this criterion, Attack/SCG (mean correlation RT=450 ms < mean baseline RT=484 ms) and Attack/EHA (mean correlation RT=471 ms < mean baseline RT=482 ms) were classified as *correspondent*, and SCG/EHA (mean correlation RT=511 ms > mean baseline RT=504 ms) was classified as *configural*. Interestingly, this criterion distinguishes between Temporal-dimension/Spectral-dimension pairs (i.e. Attack/SCG and Attack/EHA) and the pair with two spectral dimensions (SCG/EHA), which fits rather well with the results of the direct statistical comparison between pairs of dimensions. This analysis did not reveal any difference between Attack/SCG and Attack/EHA, but allowed us to conclude that the SCG/EHA pair differed from the other two pairs both in RT and error rate profiles, in particular the highest filtering costs were observed for the SCG/EHA pair, and with respect to congruency effects, the highest congruency scores were observed for the SCG/EHA pair.

As mentioned in the Introduction, it is not a simple question to determine at which levels of processing the interactions between dimensions exhibiting Garner interference occur. *Correspondence* is often considered as a sign of interactions at a semantic level, but in the literature it has mostly involved linguistic dimensions (e.g. Melara and Marks, 1990d). Such an interpretation does not seem very likely here given the nature of the dimensions manipulated (basic auditory dimensions). Interactions at a perceptual level of

processing seem more likely (see Melara and Marks, 1990b, for similar results and interpretation with other auditory attributes). Nevertheless, it remains possible that the congruency effects observed for the three pairs of timbre dimensions tested could be partly mediated by semantic interactions between the verbal labels provided to the participants when explaining the task (see Experimental procedures). *Configurality* is described in the literature as the consequence of variations of an uncontrolled emergent dimension. It might be that varying spectral dimensions of timbre (SCG and EHA) gives rise to variations in yet unknown spectral timbre dimensions. This possibility is however unlikely given the results of a dissimilarity rating study we conducted using tones varying along the same three acoustic parameters as here. Multidimensional scaling (MDS) of these dissimilarity ratings led to a three-dimensional solution, where each of the dimensions recovered by the MDS closely corresponded to one of the three acoustical parameters manipulated (Caclin et al., 2005, Exp. 3). Any emergent timbre dimension obtained when manipulating attack time, SCG, and EHA should have been recovered in this study.

So the three pairs of timbre dimensions tested exhibited an interaction profile typical of what has come to be termed in the literature *correspondent* or *configural* dimensions, but it is unlikely that the mechanisms of interaction generally invoked to account for such interaction profiles (interactions at a semantic level, emergence of uncontrolled dimensions) could entirely explain the present results. A related question is whether the Condition effects (i.e. Filtering costs and redundancy gains in Correlation) and the Congruency effects are the consequence of the same interactive mechanisms or arise from different sources. On psychophysical grounds, it has been argued that the latter is more likely (Melara and Mounds, 1993; Ben-Artzi and Marks, 1995). These two types of mechanisms would nevertheless have to be linked in some way, as Congruency effects arise mostly in Correlation and Filtering conditions, but not in Baseline conditions in this study and others (see Melara and Mounds, 1993, for similar results; Ben-Artzi and Marks, 1995). Congruency effects could reflect a facilitation of (or a cost in) the processing of one dimension caused by another dimension, which manifests itself only when this second dimension also varies in the sound sequence. We tentatively propose that the interactions between timbre dimensions observed in the present study originate at different levels of processing, including early perceptual processing stages, and that filtering costs, redundancy gains, and congruency effects might reflect different mechanisms of interaction. Additional work is needed to delineate more precisely the processing levels at which interactions take place and whether Condition and Congruency effects rely on interactions at similar or different processing levels.

The patterns of interactions found for the Attack/Spectral-dimension pairs are very similar to those reported between other auditory attributes such as pitch and loudness (e.g. Melara and Marks, 1990c), with a cost of the Filtering condition relative to Baseline, small redundancy gains in Correlation, and moderate Congruency effects. This suggests rather similar mechanisms of interaction in all these cases. In contrast, the two spectral dimensions of

timbre exhibited a different pattern of interactions, with a greater cost of Filtering and larger Congruency effects (especially in the Correlation conditions) than the other pairs, as well as an increase of RT in Incongruent Correlation relative to Baseline. Future work should aim at clarifying whether this is the consequence of different mechanisms of interaction, or if the same types of mechanisms are involved, but to different degrees for all pairs of auditory dimensions.

3.2. Implications for the functional organization of the auditory system

In the present study using the Garner interference paradigm, we found evidence for an interactive processing of timbre dimensions. Yet we have recently reported separate representations in auditory sensory memory for the same timbre dimensions using an electrophysiological measure, the MMN (Caclin et al., 2006). Concerning the other basic auditory attributes (e.g. pitch and loudness), evidence for their interactive processing has also been found using Garner's paradigm (see Introduction). Furthermore, separate sensory memory representations have been found using behavioral and electrophysiological measures for pitch, loudness, and duration (Giard et al., 1995; Frodl-Bauch et al., 1997; Clément et al., 1999; Rosburg, 2003), and for pitch and timbre (Semal and Demany, 1991; Krumhansl and Iverson, 1992). The apparent contradiction between the studies using Garner's paradigm (suggesting interactive processing of auditory dimensions) and sensory memory studies (suggesting separate processing of the very same dimensions) is likely to arise from the fact that the two types of studies tackle different levels of processing. A general rule therefore emerges: auditory attributes (pitch, loudness, duration, position, timbre dimensions) would be initially processed interactively, as revealed by Garner interference studies, but are nevertheless represented in a fairly separate fashion in sensory memory. In addition, feature (attribute) conjunctions can also be stored in auditory sensory memory (Gomes et al., 1997; Sussman et al., 1998). This set of data is compatible with a model postulating separate channels of processing for different auditory attributes (as revealed by sensory memory studies), which can interact at various stages of processing (as revealed by Garner interference studies). Alternatively, it may be that the segregation into specialized processing channels occurs progressively along the auditory pathway. Grouping mechanisms that bind together the different attributes pertaining to a single object would operate on (and/or interact with) these different processing channels. To incorporate the present findings about timbre dimensions into these models, we can postulate that there is more crosstalk between the processing channels for the spectral dimensions of timbre than for other pairs of auditory dimensions, as manifested by larger filtering costs and larger congruency effects. If one hypothesizes a progressive segregation of processing channels, these more pronounced behavioral effects could imply that the separation of channels for the spectral dimensions of timbre occur later in the auditory pathway than between the temporal and spectral dimensions of timbre.

3.3. Conclusion

When analyzing the relationships between the processing of timbre dimensions, we found evidence for an interactive processing of these dimensions (present report), yet with fairly separate representations of the very same dimensions in auditory sensory memory (Caclin et al., 2006), supporting a model with separate processing channels for these dimensions, and crosstalk between channels. The precise level(s) of processing at which crosstalk occurs remain to be determined, as well as potential differences in the levels involved for different pairs of timbre dimensions. Exploring the diversity of possible mechanisms of interactions between timbre dimensions (and more generally auditory dimensions) sheds light on the functional architecture of the auditory system.

4. Experimental procedures

4.1. Participants

Sixty-three listeners participated in a 1-h behavioral experiment (35 females, 57 right-handed, aged 19–56). Thirty-three of them reported regular musical practice. Participants gave written informed consent prior to the experiment. They were naïve as to the purpose of the experiment and were paid for their participation. They were randomly assigned to one of three groups of 21 participants, and each group was tested on a particular pair of dimensions.

4.2. Stimuli

Stimuli were created by additive synthesis of 20 harmonics with a 311-Hz fundamental frequency (E_{b4}). The amplitude envelope of the sounds was composed of a linear attack (15 or 100 ms, labeled t_0 and t_1 , respectively), followed by a plateau and an exponential decay (see Fig. 2). The duration of the plateau was adjusted to keep perceived sound duration constant. The global shape of the spectrum was manipulated through a power-function relation between harmonic amplitude and harmonic rank, which determined the value of SCG. SCG was fixed at 3 (c_0) or 4.5 (c_1) in harmonic rank units, which corresponded to values of -1.67 and -1.23 , respectively, for the exponent of the power function determining harmonic amplitude for each harmonic rank. The local shape of the spectrum was controlled by selectively attenuating even harmonics by 0 (h_0) or 10 dB (h_1) relative to odd ones. Global levels were adjusted to ensure constant loudness across stimuli. The two values taken by each of the three dimensions were selected on the basis of preliminary testing in order to obtain equal performance in all baseline conditions. Further details about stimulus synthesis can be found in Caclin et al. (2005).

4.3. Procedure

Experimental routines were programmed using PsiExp (Smith, 1995). Experiments took place in a sound-proof booth, and sounds were presented over Sennheiser HD 520 II headphones after digital-to-analog conversion and amplification (Yamaha

P2075 power amplifier). Response times were recorded using an external two-button box. Each participant was tested on a single pair of dimensions. The experiment started with a presentation of the four tones to be classified. Verbal labels associated with the two values along each dimension were used to explain the categorization task. The approximate English translation of these French terms would be ‘sharp’ (*rapide*) (t_0 , see labels in Table 1) vs. ‘slow’ (*lente*) (t_1) attack, ‘dull’ (*mat*) (c_0) vs. ‘brilliant’ (*brillant*) (c_1), and ‘neutral’ (*neutre*) (h_0) vs. ‘hollow’ (*creux*) (h_1). The experiment was divided into two halves in which participants were requested to categorize sounds according to each of the two criterial dimensions successively. The order of the two dimensions was balanced across participants.

For each dimension, participants started with a training phase in the Baseline conditions and then were to perform the classification task in the five experimental conditions (see Table 1). There were 48 sounds in each condition. The order of these conditions was balanced across participants using a latin square design. Each of the two values along the dimension to be classified was associated with one of the buttons of the box. This association varied randomly across participants. The key assignment was apparent on the computer screen during the entire experiment. Participants were asked to respond as quickly and accurately as possible, with emphasis on speed. Feedback was given after each response.

4.4. Statistical analyses

RTs were averaged for each participant for each of the 10 blocks in which he/she participated (2 dimensions \times 5 experimental conditions). RTs exceeding 2 s were excluded from averaging. Error rates were also computed for each block. For the three pairs of dimensions, RTs were analyzed separately for each criterial dimension using one-way ANOVAs with Condition as a five-level within-subject factor. Congruency effects were also examined in the RT data. For each criterial dimension, mean RTs were computed for A0B1 and A1B0 (first diagonal) and for A0B0 and A1B1 (second diagonal) in the Baseline, Correlation, and Filtering conditions (notations from Fig. 1a). Notice that for the Correlation conditions, this does not imply sorting the raw data again because stimuli are already paired according to the diagonals of the stimulus square. For each pair of dimensions, congruency was analyzed using a three-way ANOVA with Dimension (2 levels), Congruency (2 levels corresponding to the two diagonals) and Condition (3 levels: Baseline, Correlation, and Filtering) as within-subject factors. As we observed significant effects of the Congruency factor, the diagonals (and the correlation conditions) were labeled ‘congruent’ and ‘incongruent’ according to mean RTs (‘congruent’ being the one with shorter RTs).

Finally, to compare between pairs of dimensions, for each participant, data from the four ‘Baseline’ conditions were averaged, as were those from the two ‘Congruent Correlations’, the two ‘Incongruent Correlations’ and the two ‘Filtering’ conditions. Mean RTs were then compared in a two-way ANOVA with Pair-of-dimensions (3 levels) as a between-subjects factor and Condition (4 levels: Baseline, Congruent

Correlation, Incongruent Correlation, and Filtering) as a within-subject factor. *Congruency effects* were also compared across pairs of dimensions. Congruency scores were obtained for each participant in each condition as the mean RT difference for the two diagonals of the stimulus square (Incongruent diagonal–Congruent diagonal). We then performed a two-way ANOVA on congruency scores (averaged across the two criterial dimensions) with Pair-of-dimensions (3 levels) as a between-subjects factor and Condition (3 levels: Baseline, Correlation, and Filtering) as a within-subject factor.

Additionally, *error rates* were analyzed using the same procedure as for RT data.

ANOVA results were corrected with the Greenhouse–Geisser procedure where applicable, and epsilon and corrected *p* are reported. Significant effects (alpha was set at 0.05) were further analyzed using Tukey–Kramer HSD post-hoc tests.

Acknowledgment

This work was supported by a CNRS (Centre National de la Recherche Scientifique) grant (CTI01-28).

REFERENCES

- ANSI, 1973. American National Standard Psychoacoustical Terminology. American National Standards Institute, Inc.
- Ben-Artzi, E., Marks, L.E., 1995. Congruence effects in classifying auditory stimuli: a review and a model. In: Possamai, C.A. (Ed.), *Fechner Day 95, Cassis, France*. International Society for Psychophysics, pp. 145–150.
- Berger, K.W., 1964. Some factors in the recognition of timbre. *J. Acoust. Soc. Am.* 36, 1888–1891.
- Caclin, A., McAdams, S., Smith, B.K., Winsberg, S., 2005. Acoustic correlates of timbre space dimensions: a confirmatory study using synthetic tones. *J. Acoust. Soc. Am.* 118, 471–482.
- Caclin, A., Brattico, E., Tervaniemi, M., Näätänen, R., Morlet, D., Giard, M.H., McAdams, S., 2006. Separate neural processing of timbre dimensions in auditory sensory memory. *J. Cogn. Neurosci.* 18, 1959–1972.
- Clément, S., Demany, L., Semal, C., 1999. Memory for pitch versus memory for loudness. *J. Acoust. Soc. Am.* 106, 2805–2811.
- Frodl-Bauch, T., Kathmann, N., Moller, H.J., Hegerl, U., 1997. Dipole localization and test–retest reliability of frequency and duration mismatch negativity generator processes. *Brain Topogr.* 10, 3–8.
- Garner, W.R., 1974. *The Processing of Information and Structure*. Erlbaum, Potomac, Wiley, New York.
- Garner, W.R., Felfoldy, G.L., 1970. Integrality of stimulus dimensions in various types of information processing. *Cognit. Psychol.* 1, 225–241.
- Giard, M.H., Lavikainen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., Näätänen, R., 1995. Separate representation of stimulus frequency, intensity, and duration in auditory sensory memory: an event-related potential and dipole-model analysis. *J. Cogn. Neurosci.* 7, 133–143.
- Gomes, H., Bernstein, R., Ritter, W., Vaughan Jr., H.G., Miller, J., 1997. Storage of feature conjunctions in transient auditory memory. *Psychophysiology* 34, 712–716.
- Grau, J.W., Kemler-Nelson, D.G., 1988. The distinction between integral and separable dimensions: evidence for integrality of pitch and loudness. *J. Exp. Psychol. Gen.* 117, 347–370.
- Grey, J.M., 1977. Multidimensional perceptual scaling of musical timbres. *J. Acoust. Soc. Am.* 61, 1270–1277.
- Hajda, J. M., 1999. The effect of time-variant acoustical properties on orchestral instrument timbres. Unpublished PhD thesis, University of California, Los Angeles.
- Hajda, J.M., Kendall, R.A., Carterette, E.C., Harschberger, M.L., 1997. Methodological issues in timbre research. In: Deliège, I., Sloboda, J. (Eds.), *Perception and Cognition of Music*. Psychology Press, Hove, pp. 253–307.
- Handel, S., 1995. Timbre perception and auditory object identification. In: Moore, B.C.J. (Ed.), *Hearing*. Academic Press, San Diego, pp. 425–461.
- Kendall, R.A., Carterette, E.C., 1993a. Verbal attributes of simultaneous wind instruments timbres: I. von Bismarck's adjectives. *Music Percept.* 10, 445–468.
- Kendall, R.A., Carterette, E.C., 1993b. Verbal attributes of simultaneous wind instruments timbres: II. Adjectives induced from Piston's *Orchestration*. *Music Percept.* 10, 469–502.
- Krimphoff, J., McAdams, S., Winsberg, S., 1994. Caractérisation du timbre des sons complexes. II. Analyses acoustiques et quantification psychophysique (Characterization of the timbre of complex sounds. II. Acoustical analyses and psychophysical quantification). *J. Phys. IV C5*, 625–628.
- Krumhansl, C.L., 1989. Why is musical timbre so hard to understand? In: Nielsen, S., Olsson, O. (Eds.), *Structure and Perception of Electroacoustic Sound and Music*. Elsevier, Amsterdam, pp. 43–53.
- Krumhansl, C.L., Iverson, P., 1992. Perceptual interactions between musical pitch and timbre. *J. Exp. Psychol. Hum. Percept. Perform.* 18, 739–751.
- Lakatos, S., 2000. A common perceptual space for harmonic and percussive timbres. *Percept. Psychophys.* 62, 1426–1439.
- Lockhead, G.R., King, M.C., 1977. Classifying integral stimuli. *J. Exp. Psychol. Hum. Percept. Perform.* 3, 746–755.
- Macmillan, N.A., Kingston, J., 1995. Integrality, correspondence and configurality represent different degrees of perceptual interaction, not different types. In: Possamai, C.A. (Ed.), *Fechner Day 95, Cassis, France*. International Society for Psychophysics, pp. 243–248.
- Marozeau, J., de Cheveigné, A., McAdams, S., Winsberg, S., 2003. The dependency of timbre on fundamental frequency. *J. Acoust. Soc. Am.* 114, 2946–2957.
- McAdams, S., 1993. Recognition of sound sources and events. In: McAdams, S., Bigand, E. (Eds.), *Thinking in Sound: The Cognitive Psychology of Human Audition*. Oxford Univ. Press, Oxford, pp. 146–198.
- McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., Krimphoff, J., 1995. Perceptual scaling of synthesized musical timbres: common dimensions, specificities and latent subject classes. *Psychol. Res.* 58, 177–192.
- Melara, R.D., Marks, L.E., 1990a. Dimensional interactions in language processing: investigating directions and levels of crosstalk. *J. Exp. Psychol. Learn. Mem. Cogn.* 16, 539–554.
- Melara, R.D., Marks, L.E., 1990b. Interaction among auditory dimensions: timbre, pitch and loudness. *Percept. Psychophys.* 48, 169–178.
- Melara, R.D., Marks, L.E., 1990c. Perceptual primacy of dimensions: support for a model of dimensional interaction. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 398–414.
- Melara, R.D., Marks, L.E., 1990d. Processes underlying dimensional interactions: correspondences between linguistic and nonlinguistic dimensions. *Mem. Cogn.* 18, 477–495.
- Melara, R.D., Mounts, J.R.W., 1993. Selective attention to Stroop dimensions: effects of baseline discriminability, response mode and practice. *Mem. Cogn.* 21, 627–645.
- Melara, R.D., Mounts, J.R.W., 1994. Contextual influences on interactive processing: effects of discriminability, quantity, and uncertainty. *Percept. Psychophys.* 56, 73–90.
- Melara, R.D., O'Brien, T.P., 1987. Interaction between

- synesthetically corresponding dimensions. *J. Exp. Psychol. Gen.* 116, 323–336.
- Melara, R.D., Marks, L.E., Lesko, K.E., 1992. Optional processes in similarity judgments. *Percept. Psychophys.* 51, 123–133.
- Menon, V., Levitin, D.J., Smith, B.K., Lembke, A., Krasnow, B.D., Glazer, D., Glover, G.H., McAdams, S., 2002. Neural correlates of timbre change in harmonic sounds. *NeuroImage* 17, 1742–1754.
- Näätänen, R., Winkler, I., 1999. The concept of auditory stimulus representation in cognitive neuroscience. *Psychol. Bull.* 125, 826–859.
- Plomp, R., 1970. Timbre as a multidimensional attribute of complex tones. In: Plomp, R., Smoorenburg, G.F. (Eds.), *Frequency Analysis and Periodicity Detection in Hearing*. A.W. Sijthoff, Leiden, pp. 397–414.
- Pomerantz, J.R., Garner, W.R., 1973. Stimulus configurations in selective attention tasks. *Percept. Psychophys.* 14, 565–569.
- Ritter, W., Deacon, D., Gomes, H., Javitt, D.C., Vaughan Jr., H.G., 1995. The mismatch negativity of event-related potentials as a probe of transient auditory memory: a review. *Ear Hear.* 16, 52–67.
- Rosburg, T., 2003. Left hemispheric dipole locations of the neuromagnetic mismatch negativity to frequency, intensity and duration deviants. *Brain Res. Cogn. Brain Res.* 16, 83–90.
- Saldanha, E.L., Corso, J.F., 1964. Timbre cues and the identification of musical instruments. *J. Acoust. Soc. Am.* 36, 2021–2026.
- Samson, S., Zatorre, R.J., Ramsay, J.O., 1997. Multidimensional scaling of synthetic musical timbre: perception of spectral and temporal characteristics. *Can. J. Exp. Psychol.* 51, 307–315.
- Semal, C., Demany, L., 1991. Dissociation of pitch from timbre in auditory short-term memory. *J. Acoust. Soc. Am.* 89, 2404–2410.
- Smith, B., 1995. *PsiExp: an environment for psychoacoustic experimentation using the IRCAM musical workstation*. Paper presented at the Society for Music Perception and Cognition, Berkeley.
- Sussman, E., Gomes, H., Noursak, J.M., Ritter, W., Vaughan Jr., H.G., 1998. Feature conjunction and auditory sensory memory. *Brain Res.* 793, 95–102.
- Wedin, L., Goude, G., 1972. Dimension analysis of the perception of instrumental timbre. *Scand. J. Psychol.* 13, 228–240.