POLSKIE FORUM PSYCHOLOGICZNE

2004 • tom 9 • numer 1

POLSKIE
FORUM
PSYCHOLOGICZNE

Prosimy o przygotowanie materiałów zgodnie z Instrukcją dla Autorów zamieszczoną na trzeciej stronie okładki i nadsyłanie ich pod adresem Redakcji

ADRES REDAKCJI
EDITORIAL ADDRESS

Instytut Psychologii
Akademia Bydgoska im. Kazimierza Wielkiego
85-867 Bydgoszcz, ul. Leopolda Staffa 1
tel. (52) 370 84 02, fax (52) 370 84 01
e-mail: tremjan@ab.edu.pl

Czasopismo ukazuje się dwa razy w roku (jeden tom w roku – dwa numery) i jest wydawane przez:

Wydawnictwo
Akademii Bydgoskiej
im. Kazimierza Wielkiego
(Członek Polskiej Izby Książki)
85-090 Bydgoszcz, ul. Powstańców Wlkp. 2
Poz. 1069

Komitet Redakcyjny
Wojciech Tomasik (przewodniczący),
Adam Boratyński, Stanisław Kowalik,
Józef Kubik, Jolanta Mędelska,
Dorota Skublicka (sekretarz)

Dystrybucja i promocja
Jarosław Nowak

Projekt okładki
Andrzej Tkaczyk

Skład
ArtStudio

ISSN 1642-1043

Zamówienia kierować pod adresem:
Wydawnictwo Akademii Bydgoskiej
im. Kazimierza Wielkiego
85-090 Bydgoszcz, ul. Powstańców Wlkp. 2
Numer konta:
Kredyt Bank S.A. O/Bydgoszcz
92 1500 1360 1213 6001 8602 0000

Redakcja
85-867 Bydgoszcz, ul. Leopolda Staffa 1
tel. (52) 370 84 02, fax (52) 370 84 01
e-mail: tremjan@ab.edu.pl

POLSKIE FORUM PSYCHOLOGICZNE

SPIS TREŚCI

SPECIAL ISSUE: MUSIC COGNITION AND THE BRAIN
(Guest Editor: Simone Dalla Bella)

KOMENTARZE
[Commentaries]

SPECIAL ISSUE: MUSIC COGNITION AND THE BRAIN

PREFACE

Music is everywhere. We can listen to it in a concert hall, in a music school, in a park, or simply while shopping or watching the TV. We typically like listening to our preferred style of music (e.g. Classical or Pop music) and sometimes we enjoy singing in a local choir, dancing or performing music. Some of us (i.e. professional musicians) are so deeply involved in music that they decided to make of it their main profession. Interestingly, music is not only ubiquitous in our Western culture. All known cultures, indeed, share some form of music as well as they share language. Why music is playing such a relevant role? What are the roots of such a universal activity? Which cognitive mechanisms and brain structures allow us to perceive and perform music? And ultimately, why do we have music?

Music Cognition and Neuropsychology of music address these questions. Their goal is to characterize the way we perceive, remember, create and perform music and to clarify which brain mechanisms underlie these functions. The origins of contemporary Music Cognition and Neuropsychology of music can be traced back to the middle of the 20th Century, with the pioneering works of Meyer (1956) and Francés (1958/1988). Nevertheless, it is only from the Eighties (e.g., see Critchley & Henson, 1977; Deutsch, 1982; Dowling & Harwood, 1986; Sloboda, 1985) that a remarkable increase of interest in the psychological processes and brain structures involved in music processing is observed. Accordingly, research works on the mechanisms underlying music perception and performance rapidly multiplied in several laboratories around the world (for reviews, see Deutsch, 1999; Peretz & Zatorre, 2003). Numerous phenomena in music perception and performance have been examined, for example, from melody recognition, music expectancies, emotional processes, to motor planning in piano performance. During the last decade particular attention has been devoted to the idea that music, as well as other important functions (e.g. language), may be rooted into biology (e.g. see Peretz & Zatorre, 2003; Zatorre & Peretz, 2001). This possibility is supported by converging evidence from Cognitive Psychology of music and Neuroscience.

In sum, during the Eighties Music Cognition and Neuropsychology of music were treated as infant disciplines and maybe regarded with a bit of scepticism ("Why should we be interested in music, given that other mental functions, such as language, are more useful and crucial for human beings?). However, nowadays the situation is radically different. Music Cognition and Neuropsychology of music are established disciplines with a precise and ambitious research program, established methods to study musical processes, and involving an increasing community of active researchers in several countries.

The main goal of this Special Issue of *Polskie Forum Psychologiczne*, aimed particularly to Polish scientists, but not only, is to illustrate some examples of contemporary research in Music Cognition and Neuropsychology of music. The authors who contributed to this Special Issue are well-known cognitive scientists interested in the cognitive processes and brain mechanisms underlying music perception, appreciation, and performance. The diversity of the research themes proposed is supposed to illustrate the variety and richness of topics that can be encountered in our field.

The first three contributions are review articles. The article by *Clément, Moroni, and Samson* describes the different types of sensory and short-term memory for properties of auditory stimuli (e.g. pitch and loudness), which also play a role in music perception. The authors show converging evidence from behavioural research and neuropsychology indicating that pitch, loudness, timbre, and modulated amplitude of a sound are held in separate sensory memory stores. This evidence supports a modular view of auditory sensory memory. *Racette, Hyde, and Peretz* thoroughly describe the work carried out mostly during the last fifteen years with amusic patients (with and without brain damage). The authors particularly focussed on music recognition processes (i.e. how we recognize and name well-known music). A modular model of music processing is proposed (see also Peretz & Coltheart, 2003). *Jungers'* article addresses parallels between memory for acoustic details in music and speech. Empirical evidence is reviewed showing that humans are able to remember global dimensions (e.g. tempo) as well as subtle relational aspects of music and speech (e.g. rhythm and intensity). The author showed how the memory of these acoustic details affects the production of music and speech.

The last two articles are detailed reports of recent experimental studies. *Thompson and Russo* addresses how the perceived happiness/sadness and meaningfulness of song lyrics is affected by whether or not they are presented in a musical context, and whether the music is familiar or unfamiliar to the listener. The reported findings reveal that the perceived happiness/sadness and meaningfulness of lyrics are influenced by some musical contexts, and that meaningfulness of lyrics varies as a function of familiarity with a song. The progressive exposure to a song leading to higher familiarity is hypothesized to increase the degree to which the music and the lyrics are integrated in memory. Finally, *Khalfa and Peretz* are interested in the psychophysiological measures of emotions evoked by music. The relation between the electrodermal response and emotional valence (i.e. pleasantness) is examined using consonant and dissonant stimuli in two groups of participants, controls and anhedonics (i.e. individuals who fail to experience sensorial pleasure). The results are consistent with the idea that the electrodermal response is sensitive to valence.

Simone Dalla Bella
Kazimierz Wielki University in Bydgoszcz, Poland
University of Finance and Management in Warsaw, Poland
October 1, 2004, Bydgoszcz

Correspondence concerning this special issue should be addressed to Simone Dalla Bella,
e-mail: sdallabella@vizja.pl

REFERENCES

Critchley, M., Henson, R. A. (1977). *Music and the Brain: Studies in the Neurology of Music*. London: Heinemann.

Deutsch, D. (1982). *The Psychology of music*. New York: Academic Press.

Deutsch, D. (1999). *The Psychology of music (2nd Edition)*. New York: Academic Press.

Dowling, W. J., Harwood, D. L. (1986). *Music cognition*. Orlando: Academic Press.

Francés, R. (1958/1988). La perception de la musique. Translated by W. J. Dowling as *The Perception of Music*. Hillsdale, N. J.: Lawrence Erlbaum Associates.

Meyer, L. B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Peretz, I., Zatorre, R. J. (2003). *The cognitive neuroscience of music*. Oxford: Oxford University Press.

Peretz, I., Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, 7, 688–691.

Sloboda, J. A. (1985). *The musical mind: the cognitive psychology of music*. Oxford: Oxford University Press.

Zatorre, R., Peretz, I. (2001). *The biological foundations of music* (vol. 930). Annals of the New York Academy of Sciences.

Polskie Forum Psychologiczne, 2004, tom 9, numer 1, s. 8–24

MODULARITY IN SENSORY AUDITORY MEMORY

Sylvain Clément, Christine Moroni
University of Lille 3, France
Séverine Samson
University of Lille 3, France
Epilepsy Unit, La Salpêtrière Hospital, France

The goal of this paper was to review various experimental and neuropsychological studies that support the modular conception of auditory sensory memory or auditory short-term memory. Based on initial findings demonstrating that verbal sensory memory system can be dissociated from a general auditory memory store at the functional and anatomical levels, we reported a series of studies that provided evidence in favor of multiple auditory sensory stores specialized in retaining either pitch, loudness, timbre or possibly modulated amplitude sounds. Finally, we also presented human data indicating the existence of a dissociation between auditory sensory memory for spatial and non spatial information. All these findings are consistent with multiple storage systems that are, to some extent, independent from one another.

Introduction

Auditory sensory memory is defined here as the retention of sensory features of sounds such as loudness, pitch, spectral composition, and duration. It must be distinguished from the retention of sound meaning ("a danger is approaching"). This particular memory system has generally been integrated in classical models of human memory as a "sensory store" providing the first representation of a just heard stimulus. This initial representation is transitory in nature (one or two seconds) and is considered as necessary for higher level encoding allowing extraction of word meaning to be stored in dedicated memory systems.

However, many studies demonstrated the existence of longer sensory traces. Using interference paradigm, Deutsch (1970) showed that participants are able to discriminate the frequency of two pure tones separated by a 5 sec delay. Based on this finding, she claimed that pure tone frequency can be remembered during 5 seconds in an auditory sensory store. In addition, this author as well as Massaro (1970) reported that discriminating the pitch of two tones becomes difficult when the time interval between them is filled by irrelevant distracting tones. This long trace lifetime (at least more than 5 sec) suggests that auditory sensory memory could not be restricted to the classical "sensory store" which can only held information for one or two seconds.

Correspondence concerning this article should be addressed to Séverine Samson, e-mail: severine.samson@univ-lille3.fr

Further evidence in favour of longer sensory traces has been reported by Harris (1952) and Wickelgren (1969). These authors demonstrated that the frequency of pure tones could be maintained in memory for more than 10 seconds. To support this claim, a discrimination task involving the presentation of two pure tones separated by a silent delay of different duration was used. The participants had to decide whether the two tones had the same frequency or not. By collecting performance for different delay durations, these authors have obtained a forgetting curve which describes the decay of performance when the delay increases. Even if performance decreased when the delay exceeded 1 sec, the frequency discrimination was still quite accurate for long delays. As an example, Harris (1952) showed that the mean Just Noticeable Difference (JND) when the two tones were separated by a 10 sec delay, was at least 0.79 % (7.9 Hz). All these findings suggest that auditory sensory memory can last several seconds.

In the last four decades, an alternative model emerged against this monolithic conception of auditory sensory memory. According to this model, auditory sensory memory is considered as a collection of memory registers with different characteristics. This modular conception of auditory memory is supported by several lines of evidence. The aim of this paper is to present the main arguments in favour of the modular representation of memory by reviewing psychophysical and neuropsychological studies based on "interference" and "forgetting curve" paradigms.

Psychophysical evidence

Forgetting curve studies

Testing modularity in auditory memory can be done by directly comparing the forgetting rate of different sound attributes. Assuming that two of those attributes (let's call them a_1 and a_2) are maintained in separate memory stores, it is then possible that the trace of a_1 does not fade away with time at the same rate as the trace of a_2. This can be tested using the above-mentioned "forgetting curve" paradigm to compare the resulting trace decay functions for a_1 and a_2. Such a comparison must be carried out between two equivalent experimental conditions.

Pitch and loudness. The decay of loudness traces as a function of time has been investigated by numerous authors (Berliner and Durlach, 1973; Berliner et al., 1977; Botte et al., 1992; Green et al., 1983; Kinchla and Smyzer, 1967; Lü et al., 1992). Similarly, the temporal decay of pitch traces has been examined in several studies (Wolfe, 1886; Harris, 1952; Bachem, 1954; Wickelgren, 1969; Rakowski, 1972). To verify if pitch and loudness are maintained in separate memory stores, the forgetting curves obtained for each sound attributes were compared with the same methodology.

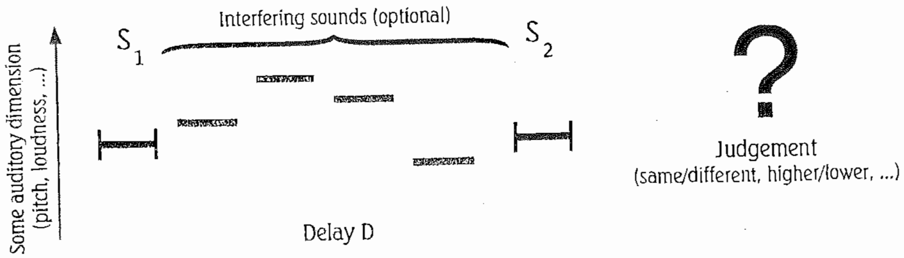

Figure 1. In typical auditory memory tasks with interference, participants have to compare two sounds (S_1 and S_2) separated by some delay (D) upon one sensory dimension (such as loudness, pitch, timbre, duration...). Interfering sounds (I) may also be inserted during the delay being generally ignored by the participants

As far as we know, Clément et al. (1999) reported the first study to compare the forgetting curve of pitch and loudness in the same experiment. However, an analysis of the previously cited studies led the authors to make the hypothesis that trace decay will be faster for loudness than for pitch. Each trial was composed of two test tones (S_1 and S_2) separated by a silent delay (from 0.5 to 10 seconds). S_2 was always different from S_1 with respect to either pitch or loudness, depending on the experimental condition. The participant had to decide whether there was an increment or a decrement of the considered attribute between the two test tones. Although four experimental conditions were used in this study, only two of them will be reported here. The intensity condition consists of pure tones (S_1 and S_2 of 1000 Hz) differing in intensity. The sound pressure level (SPL) of S_1 was randomly selected between 40 and 80 dB SPL to minimize context coding accuracy. The frequency (FREQ-PURE) condition consisted of pure tones differing in frequency, S_1 being chosen between 500 and 2000 Hz whereas the SPL was always 60 dB SPL. The physical changes used in this study were individually determined. They corresponded to 80% of correct responses with 0.5 sec delay measured with an adaptive procedure allowing comparison between different conditions. Therefore, similar performance was obtained in all the conditions at 0.5 sec delay providing a reference point for directly comparing the forgetting curves of intensity and frequency. The main result of this study showed that performance drops faster in the intensity than in the frequency condition when delay increases from 0.5 to 2 sec as illustrated in Figure 2.

The difference between pitch and loudness with respect to trace decay supports the sensory memory modularity hypothesis. This finding is also compatible with psychophysiological data (Giard et al., 1995). By studying scalp topographies of mismatch negativity (MMN) component of auditory event-related potentials (ERPs), the authors showed that a MMN is elicited when a deviant tone (differing from the standard by one or more features) is presented after some repetitions of a standard stimulus. The MMN is supposed to reflect a pre-attentive change detection and is considered as a physiological marker of auditory memory (see Schröger, 1997, for an extensive review). Giard and her colleagues also found that the MMN source location depends on the auditory feature that differs between the deviant tone and the standard one, at least for the three sound attributes examined in their study (duration, intensity and frequency). It seems then that even at the physiological level, one can find some arguments in favour of a separate storage of pitch and loudness sensation.

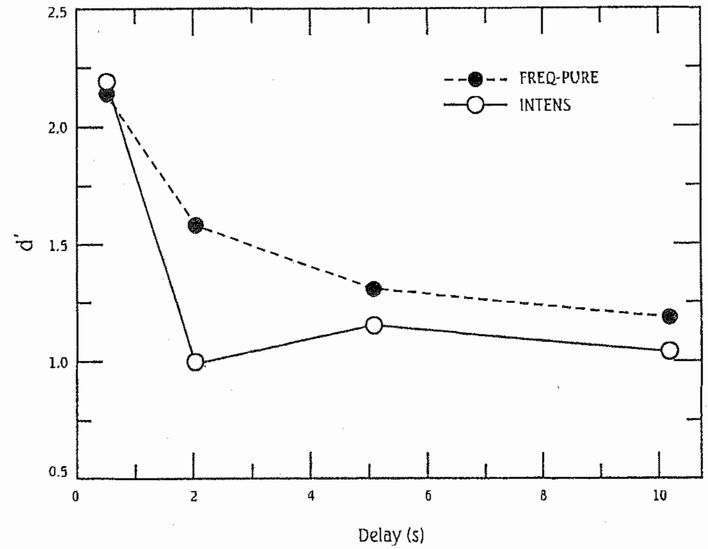

Figure 2. "Forgetting curves" for pitch (condition FREQ-PURE) and loudness (condition INTENS) reported in Clément et al. (1999)

However, Clément and his collaborators (1999) found that difference between intensity and frequency conditions disappears when delay was greater than 5 sec showing some kind of plateau. The authors argued that this plateau seems to be ascribed to a context-coding operation mode described by Braida and Durlach (1988). By using the forgetting curve paradigm, these authors made the distinction between two memory operation modes. In one mode, called the "trace mode", the sensation produced by S_2 was compared to the sensory trace left by S_1. The accuracy of this sensory trace was supposed (and found) to be high for short delay but to quickly decrease as a function of delay. In the other memory mode, called the "context-coding mode", the subjects compared the symbolic representation of two sounds with respect to the general context of the experimental session. This symbolic representation can be viewed as a categorical label attributed to each sound in relation with the general context. For example, a sound can be labelled as "one of the louder sound recently heard" or as "average intensity". This categorisation was based on a fuzzy long-term representation of the experimental context (the stimuli range) providing therefore low accuracy. Since the only thing to remember in this context-coding mode was a verbal label attributed to S_1, the accuracy would be constant for long delay. However, Braida and Durlach assumed that a constant number of categories could be extracted from the context. When the stimulus range (or the context) was widened, the categories were enlarged and the accuracy decreased. The authors did an impressive work in developing a highly detailed mathematical model predicting performance evolution as a function of delay and of size of the context.

Based on this prediction, Clément et al. computed the size of the context for the intensity and the frequency conditions in the following way. For the intensity condition, the mean threshold measured at 0.5 sec delay was of 1.8 dB and the intensity roving ranged from 40 to 80 dB SPL.

The size of the context was therefore relatively small corresponding to 22 units (i.e to 40÷1.8≈22). By contrast, the mean threshold of frequency measured at 0.5 sec delay was of 7.4 cents and the frequency roving ranged from 500 to 2000 Hz (corresponding to a difference of 2400 cents). The size of the context corresponded to 324 units (i.e 2400÷7.4≈324). Thus, the context was greatly wider in the frequency than in the intensity condition explaining the decrease of accuracy in the frequency task. It is therefore plausible that context coding mode may have limited the performance decay in the intensity but not in the frequency condition.

Another argument proposed later by Clément (2001) to understand the intensity and frequency forgetting curves was based on a mathematical model of the trace coding mechanism (Kinschla & Smyzer, 1967). According to this model, the accuracy of the sensory representation of some stimulus decreased with time as the result of a "random walk" process. This produced a linear increase of the variance of the representation as a function of time. In line with this model, performance in the frequency condition reflected only the trace coding accuracy whereas the performance drop in the intensity condition was limited by the context coding mode. In this case, the *d'* found in the frequency condition should be well predicted by the trace memory model whereas a stronger performance drop than the one observed for long delay should be predicted in the intensity condition. Computer simulations were run to compare the *d'* obtained by the four subjects examined by Clément and collaborators (1999) in the two conditions with the performance predicted by the random-walk model of sensory memory proposed by Kinshla and Smyzer. As displayed in Figure 3, the model made good predictions in the frequency condition. However, it predicted a greater trace decay than the one measured in the intensity condition emphasizing the initial interpretation in terms of context-coding mode previously reported (Clément et al., 1999).

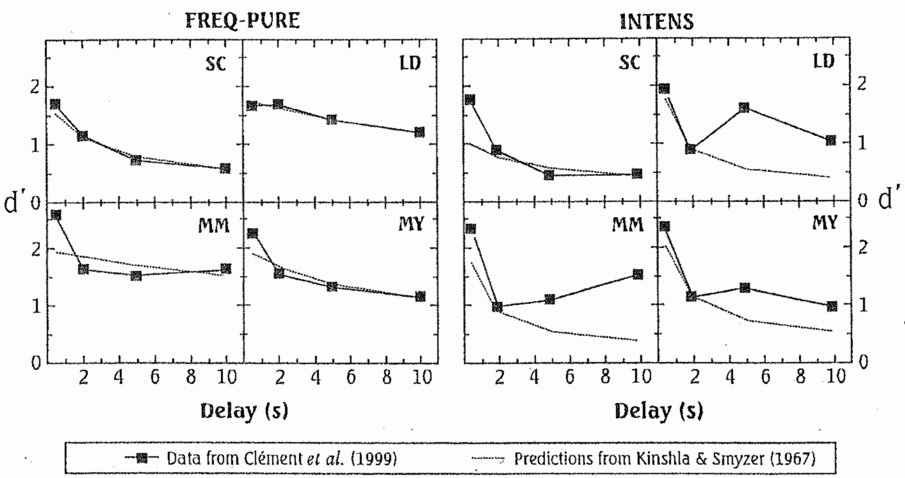

Figure 3. Comparison between the performances from Clément et al. (1999) and predictions made by the random-walk model of sensory memory proposed by Kinshla and Smyzer (1967)

One interesting question about loudness trace, is whether the observed decay of performance corresponds to a loss of accuracy of the remembered loudness or whether it corresponds to a systematic drift of the remembered loudness (e.g. the remembered loudness becoming softer with time). Koester (1945) addressed this issue and found no evidence supporting this last hypothesis. Similarly, the data from Clément et al. (1999) showed no systematic response bias in FREQ-PURE and INTENS conditions. Therefore the loudness as well as the pitch trace decays seem to correspond to a loss of accuracy of the trace. This is the original postulate proposed by Kinshla and Smyzer (1967) and by Braida and Durlach (1988) in their respective model.

The fact that loudness is forgotten faster than pitch suggests that pitch and loudness of sounds are held in different memory registers, the pitch register allowing a better trace maintenance over time than the loudness register.

Amplitude modulated sounds and intensity. Most studies investigating auditory sensory memory have used steady tones such as pure tones or harmonic complex tones. The intensity of these stimuli remained stable in time. However, the temporal envelope shown by the thick line in Figure 4 corresponds to the evolution of intensity with time and seems to be essential in music, speech or environmental sound perception. It is differentiated from the fine temporal structure illustrated by thin line which corresponds to the carrier of the waveform (Lorenzi et al., 2000). However, we still know very little about the sensory memory of the temporal envelope in amplitude modulated sounds.

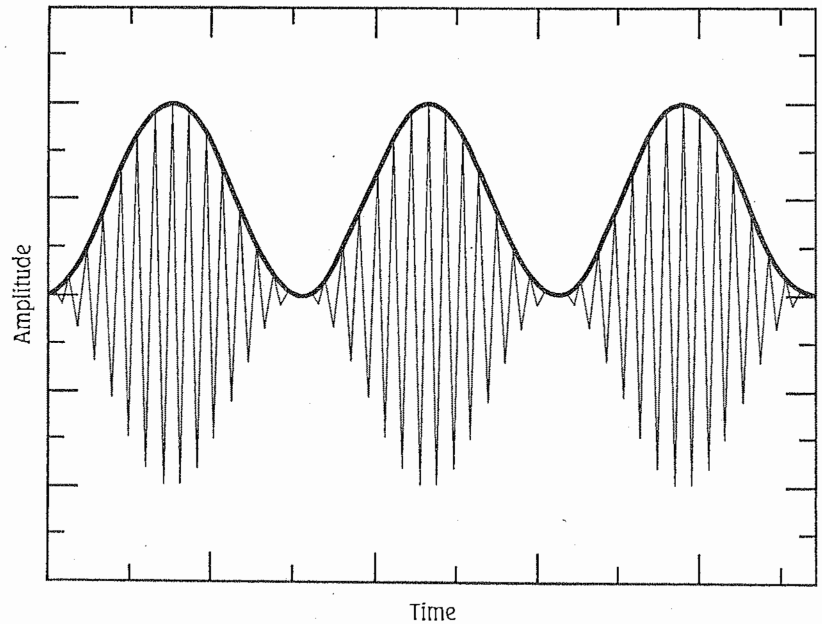

Figure 4. Waveform of a complex tone. The temporal fine structure of the sound is shown by the thin line and its temporal envelope is shown by the thick line

To address this question, Moroni, Demany and Lorenzi (2004) used sinusoidally amplitude modulated (SAM) white noises and compared the forgetting rate of various SAM white noises to the trace decay of intensity sound using a forced-choice discrimination paradigm. Four subjects without any known hearing deficit participated in this study. On each trial, two stimuli (1.25 seconds) separated by a delay were presented, the delay varying from 0.25 sec to 8 sec. Modulation frequency discrimination was assessed for an average frequency value of 8 Hz with a fixed modulation depth of either 12% or 100%. Intensity discrimination was assessed for an average level of 75 dB SPL using a modulation frequency of 8 Hz and a 100% depth. In each case, a roving procedure was employed to minimise the influence of context coding. Initially, frequency and intensity discrimination thresholds were measured for a delay of 0.25 sec. The corresponding physical differences (in dB) were presented at six different delays (0.25, 0.5, 1, 2, 4 and 8 sec). The results showed a significant effect of delay on performance but no effect of sound attributes (intensity, 100% or 12% modulation frequency) on performance. For both intensity and frequency discrimination using SAM noises, the sensitivity (d') decreased slowly as a function of delay but a significant forgetting was noted for 8 sec delay. Taken together, these results suggest that sensory memory for AM sounds and intensity resists very well to 4 sec delay. However, they did not allowed to demonstrate a dissociation between intensity and AM-frequency sensory memory using a forgetting curve paradigm. Further investigation using for example an interference paradigm will be necessary to try to clarify this point.

Interference studies

Pitch and timbre. Most studies examining auditory memory modularity have used the interference paradigm. In particular, Deutsch (1970) reported several experiments on pitch memory (for review, see Deutsch, 1999). For instance, she demonstrated that when the distracting material consists of spoken digits there is no decrement in performance (Deutsch, 1970) indicating that the frequency of sounds can be retained for more than 5 seconds. This finding as well as others (e.g., Massaro, 1970) suggest that retention of tonal information seems to be functionally distinct from the one responsible for processing auditory verbal information since discriminating pitch between two tones becomes more difficult when the time interval between them is filled by irrelevant distracting tones than by spoken digits.

Few years later, Deutsch (1972) showed that when the test tones (S_1 and S_2) and the interfering tones (I tones) were pure tones, the frequency distance between S_1 and S_2 influenced pitch discrimination. In a same/different task with S_1 differing from S_2 by 100 cents, the error rate was found to be highly dependant on the frequency distance between S_1 and I (S_1-I). The error rate increased from 2% for S_1-$I=0$ cent up to about 35% for S_1-$I=133$ cents, and then decreased to about 10% for the largest frequency distance used in the experiment (200 cents). This *selective interference* effect was explained by frequency selective lateral inhibition phenomena in the auditory store (Deutsch 1975, Deutsch & Feroe, 1975).

As mentioned above, the pitch comparison was greatly altered when the I tones were pure tones but it was not affected when I sounds were speech. This result can be explained in different ways. Firstly, speech sounds might be memorized in a separate memory system as proposed by Deutsch (1970). However, the effect of the I tones may also depend on physical similarities between I tones and test tones. Since I spoken sounds are quite different from the $S1$ and $S2$

pure tones (in term of spectral and temporal envelope), the interference effect should be minimal. Conversely, the deleterious effect of interference on pitch discrimination should increase if I tone timbre is very similar to $S1$ and $S2$ pure tones. Finally, it might also be possible that these interference effects are only determined by pitch similarities between $S1$ and the I tones whatever the overall physical similarity of these sounds is. Indeed, the two kinds of I tones (spoken sounds and pure tones) used by Deutsch (1970) not only differed by timbre but they also differed in pitch leading to different interference effects. In her study, the pitch of the pure tones ($S1$ and I tones) varied from 274 Hz to 517 Hz whereas the pitch of spoken I sounds ranged from 90 Hz (low-pitched male voice) up to 300 Hz (child voice) being therefore inferior to the pitch of pure tones. It seems therefore possible that the pitch distance between spoken I sounds and $S1$ was more important than the one between pure I and $S1$ tones. In all these cases, the error rate would be expected to be lower with spoken I sounds than with pure I tones, the sensory/verbal memory dissociation being thus not relevant.

Semal and Demany (1991) have tested this last interpretation. They tried to determine if pitch similarity could produce interference effects when I tones and test tones differ in timbre. In this study, untrained listeners were instructed to discriminate pairs of pure or of complex harmonic tones (S_1 and S_2) separated by a 4.3 sec delay during which I tones were presented. The pitch of S_1 was randomly chosen in a certain frequency range depending of the nature of the tests tones (pure or complex tones). In half of the trials, S_1 and S_2 differed in pitch by ±4%.

The similarity between the I tones and the test tones was investigated by manipulating 3 dimensions: pitch, timbre and spectral overlap. The pitch parameter was determined by the frequency of pure tones or by the fundamental frequency of harmonic tones (f_0). The timbre parameter was linked to the spectral envelope of the sounds (pure sounds, series of harmonics with eventually a missing fundamental). The spectral overlap parameter corresponded to the relation between the spectral envelopes of the I tones and the test tones; it was high when both I tones and test tones had energy in the same frequency region. One can assume that such an overlap could permit interactions between the two signals at a peripheral level producing backward masking in the cochlea.

The interference effect due to the I tones was found to be only determined by the pitch similarity parameter and to be independent from the timbre similarity and from the amount of spectral overlap. Semal and Demany (1991) concluded that this finding reflects a dissociation between pitch and timbre in the sensory auditory memory store. Based on these results, one can wonder if the results of Deutsch (1970) could be attributed to an uncontrolled pitch similarity problem as proposed Semal and Demany.

Two additional experiments testing this hypothesis were reported by Semal et al. (1996). Their participants were required to discriminate pitch of two test tones (S_1 and S_2) separated with a 6 seconds delay during which speech or nonspeech interfering stimuli were presented. The nonspeech I sounds were harmonic complex tones with a flat spectral profile. The speech I sounds were pitch-controlled monosyllabic spoken words. The two kinds of I sounds were matched in terms of pitch distance to the test tones; Δ-pitch could be either small (±100 cents), medium (±350–550 cents) or large (±800–1000 cents). In the first experiment, the test tones were composed of three equal-amplitude harmonics (rank 1 to 3). The nature of the I sounds (speech vs. nonspeech) was found to slightly affect discrimination but markedly less than did Δ-pitch. Performance was better when Δ-pitch was large than when it was small. The second experiment was quite similar to the first one except that S_1 and S_2 were also speech sounds differing in pitch in half of the trials. Again, discrimination was disrupted by small Δ-pitch but the nature of the I sounds produced no significant effect. All these results suggest that the lack of interference effect when spoken I tones were presented in Deutsch's study might be exclusively due to an uncontrolled, large Δ-pitch in such conditions. Making pitch comparisons may only involve a pitch-specific sensory store which seems "deaf" to timbre (Semal & Demany, 1991; Semal et al., 1996).

The existence of a sensory memory register specialised in the retention of pitch leads to another question. Is there any other feature-specific memory register? Starr and Pitt (1997) questioned the existence of an autonomous retention of some timbre dimension with a similar procedure. In their study, participants had to discriminate two harmonic tones (S_1 and S_2) separated by 5 sec delay that could differ in their harmonic composition. The tones were composed of the fundamental frequency (f_0) associated with 3 consecutive harmonics. For instance, a tone could be composed of several sinusoidal components (f_0, $7\times f_0$, $8\times f_0$ and $9\times f_0$). During the delay, six interfering tones with constant timbre could be played. These tones were synthesised in the same way as the test tones, but were chosen at a certain timbre distance from S_1. This distance depends on the harmonic ranks of the three harmonics added to the first spectral component of the I tones with respect those of S_1. Starr and Pitt have also studied the effect of pitch distance between the I and S_1 tones. The main result was that the effect of interference increased as timbre distance decreased between I tones and test tones.

When the pitch of the I tones was similar to the pitch of test tones, timbre comparison was slightly better than when they were different. However, when the pitch distance was small (but not null), timbre memory was not poorer than when it was large. All these results combined with Semal and Demany (1991) findings suggest that pitch and timbre are both held in independent sensory registers.

Pitch and loudness. At least two interference studies on loudness retention can provide clues about relations between pitch and loudness in sensory auditory memory. The first one found in Botte et al. (1992, experiment 3) used a loudness comparison task. Two test tones (S_1 and S_2) were separated with a 3.75 sec delay during which 6 interference tones could be inserted. There were eight conditions with different configurations for the I tones. These I tones could be softer or louder than the test tones with decreasing or increasing loudness. Interference effects were found when some I tones were louder than S_1 by at least 6 dB. No interference was found for I tones equal or softer than S_1. These findings suggest that loudness interference is not regulated by the same rules as pitch interference. As previously noted, pitch interference is dependant on pitch distance: a small Δ-pitch produce more interference than a large one. In the loudness domain, the deleterious effect of interference increased with loudness but was not dependent on loudness distance. This dissociation of interference operation mode between pitch and loudness can be understood if we assume that these two features of tones are held in different sensory memory registers.

Semal and Demany (1993) have also studied how interference effects are dependent on the loudness of the I tones. Once more, the participants were required to discriminate two pure tones separated by a 4.3 sec delay during which six I tones could be inserted. In the first experiment, the test tones had a sound pressure level (SPL) of 60 dB SPL and a frequency randomly chosen around 1000 Hz. The SPL of the I tones could be either 45, 54, 60, 66 or 75 dB depending on the condition (conditions 2 to 6, respectively). The frequency of each I tones was randomly selected

among four values differing by ±3% or ±6% from the frequency of S_1. The condition 1 served as "baseline" and no I tones were presented during the delay. As expected, the performance was significantly better in this first condition as compared to the other 5 conditions but no effect of condition was found for condition 2–6. As opposed to Botte et al. (1992), I tones generated the same amount of interference in all interfering conditions whatever the sound level was.

Semal and Demany also reported a second experiment because their negative results (no effect of I tone intensity) could be attributed to a floor effect, the performance being very poor in conditions 2–6. In this experiment, they used harmonic complex I tones. In the conditions 2–4, the I tones consisted of harmonics 1 to 5 of a fundamental frequency equal to 1/3, ±3 or 6% of the frequency of S_1. Thus, the median harmonic (the 3rd harmonic) of these tones was near S_1 but the pitch sensation was remotely lower. The SPL of the spectral components was 45, 60 and 75 dB in conditions 2, 3 and 4, respectively. In conditions 5–7, the I tones consisted of harmonics 28 to 32 of a fundamental frequency equal to 1/30th, ±3 or 6% of the frequency of S_1. The median harmonic (the 30th harmonic) was still close to S_1. Their fundamental frequency was always below 50 Hz i.e. in the *"infrapitch"* domain. Nonetheless, these tones had a very narrow spectral bandwidth (less than a critical band) so that they evoked a fuzzy pitch sensation near the frequency of S_1. More interference was found in conditions 5–7 than in conditions 2–4 which is consistent with previous findings (Semal & Demany, 1991). In conditions 5–7, the pitch evoked by the I tones was closer to the pitch of S_1 than in conditions 2–4. Furthermore, the same SPL values were used in conditions 5–7 and in conditions 2–4. The "loud" I tones (75 dB SPL) were found to produce less interference than "softer" I tones (45 dB SPL). Although this effect was small, it is surprising considering the results reported by Botte et al. (1992). Since these interferences seemed to be independent of loudness similarity, Semal and Demany concluded that pitch sensory memory may be independent from loudness sensory memory.

Neuropsychological evidence

Whereas experimental psychological studies provide arguments in favour of functional dissociation between different auditory sensory memory systems based on normal data, neuropsychological findings are used to reveal dissociation between performance obtained in brain damaged participants indicating that different memory systems depend on specific neural substrates. In the neuropsychological domain, the auditory sensory memory refers to the auditory short-term memory. However, relatively few studies attempted to probe the neural substrate for human auditory short-term memory. The goal of neuropsychological investigations is to identify the cerebral structures underlying auditory short-term memory.

Animal and human data suggest that a specialized neural mechanism exists for auditory short-term memory, and that it is linked to the function of the auditory association cortex in the superior temporal gyrus. Behavioral studies with monkeys have shown that lesions of the superior temporal cortex, sparing the primary auditory region, lead to impairments in the ability to retain auditory information in a delayed matching-to-sample task (Colombo et al., 1990; Stepien et al., 1960). This deficit appears to be selective to the auditory system, since such animals do not experience difficulty with similar visual discrimination abilites. Moreover, damage to this region does not lead to generalized auditory processing deficits because basic auditory discrimination ability is retained, even with large bilateral lesions of auditory cortex (Heffner

& Masterton, 1978; Jerison & Neff, 1953). In human literature, several lines of evidence suggest that the medial temporal lobe structures can also be involved in auditory short-term memory. In the context of this review, we will not give too much emphasis on anatomical structures but we will use the neuropsychological data to support the modularity of auditory short-term memory.

Memory traces

Very few studies investigated the forgetting curves in brain damaged patients. To our knowledge, no studies really compare the forgetting scores for different auditory attributes in patients with cerebral damage. Wickelgren (1968) examined the patient H.M., who underwent bilateral hippocampal resection, and who subsequently demonstrated very severe anterograde amnesia (Scoville and Milner, 1957). H.M.'s pitch retention function showed a normal short-term decay rate of about 5 sec, but the rate of decay drops after 40 sec or more.

Interference studies

Considering the well established predominance of left hemisphere structures for language, a deficit in verbal short-term memory is easily predicted after left temporal lobe lesion. This hypothesis found support in a study carried out in patients who had undergone anterior temporal removal to treat medically refractory epileptic seizures. Following lesions of the left medial temporal lobe structures, auditory verbal information appears to be rapidly lost when an interfering verbal task must be performed before recall (Corsi, 1972).

In neuropsychology, most brain lesion and functional imaging studies assessing auditory short-term memory for tonal information have used melody (Samson & Zatorre, 1988; Zatorre, 1985; Zatorre et al., 1994) or pitch (Alain et al., 2001; Johnsrude et al., 2001; Zatorre & Samson, 1991). In searching for the neural substrates involved in the maintenance of such non-verbal information in auditory short-term memory, these investigations clearly outlined the importance of the right anterior neocortical temporal and frontal lobe structures when melodic information was used. In this paper, we will only report studies using interference paradigm in patients who had undergone unilateral anterior temporal-lobe resection to control medically intractable epilepsy.

To examine directly the role of the temporal lobes in auditory retention of tonal materials, patients with unilateral temporal or frontal lobe lesion were tested in a pitch discrimination task with or without tonal interference (Zatorre & Samson, 1991). The retention of pitch was assessed in two conditions. In the experimental condition, which was very similar to the one used by Deutsch (1972), subjects had to compare the tonal pitch of two tones separated by a melodic interference pattern. In the control condition, there was a silent interval between the tones without any interference. In each condition, half of the trials were composed of same items, and the other half consisted of different items. When the tones differed, the comparison tone was always of a higher pitch than the test tone. Deficits in discriminating the pitch of two tones separated by distracting materials was predicted after damage to the temporal neocortex on the right side.

Results indicated that there was no significant deficit in any patient group on the control task. However, retention of pitch in the presence of interference was impaired in patients with right but not left temporal lobe lesion. The deficit was not related to the extent of the lateral neocortical damage since performance was not affected by the additional removal

of the primary auditory cortex. These data suggest that regions within the human right temporal lobe are important in maintaining auditory information in a short-term memory store. This finding is also supported by an animal study (Colombo et al., 1990) reporting auditory retention deficits in monkeys after bilateral lesions of the superior temporal gyrus sparing the primary auditory area. Results of patients also showed that lesions involving the right frontal lobe affect performance on this task suggesting that short-term auditory memory involves a distributed network in which connections between pre-frontal and temporal cortices serve to maintain information over filled delays (Chavis & Pandya, 1976; Perry et al., 1999). This later interpretation found support in a positron emission tomography study measuring cerebral blood flow in normal participants during the performance of a pitch judgement task (Zatorre et al., 1992). Activation of right frontal site during this task suggests that specific areas in the right prefrontal cortex play an important role in maintenance of pitch information, presumably through interaction with secondary auditory cortical regions in the superior temporal gyrus. The lateralised nature of the deficit obtained in this study is compatible with a large body of evidence suggesting a privileged role for the right cerebral hemisphere in certain aspect of processing musical sounds (Milner, 1962; Samson & Zatorre, 1988; Sidtis & Volpe, 1988; Zatorre, 1985). In this investigation, we also noted that the deficit was not exacerbated by extensive resection of the hippocampal region (or medial temporal lobe structures) in comparison with limited resection of this region. Therefore, it seems that hippocampal structures are not primarily involved in this memory task.

Based on these results, we predicted that the ability to retain tonal pitch in the context of a short-term memory task does not rely on the integrity of hippocampal regions. Evidence in favor of this hypothesis was provided by data obtained in epileptic patients with unilateral medial temporal lobe damage (Samson, 1999). Patients presenting unilateral hippocampal damage (atrophy) documented by MRI were tested at La Salpêtrière Hospital before brain surgery with the paradigm previously described. Results showed that performance of subjects with right or left medial temporal lobe damage (associated to hippocampal atrophy) did not differ from the normal data nor from the patients who had undergone a left temporal lobectomy in the previous study (Zatorre & Samson, 1991). This finding suggests that hippocampal dysfunction does not seem to affect *short-term retention of pitch*, which seems to depend mainly on the *anterior neocortical association areas of the right temporal lobe*.

Another dissociation between different short-term memory systems has also been reported in the literature. Recent studies with normal participants revealed a functional dissociation between short-term memory for auditory object and sound localization (Anourova et al., 1999; Clarke et al., 1998). In human neuropsychology, there is only one study (Lancelot et al., 2003) to our knowledge that investigated auditory short-term memory for spatial and non-spatial information in brain-damaged patients. The few studies reported in the literature examined auditory-spatial perception (Zatorre et al., 1995; Zatorre & Penhune, 2001). In Lancelot and al's study, two discrimination tasks were designed to compare short-term memory for sound localization with short-term memory for sound content with and without auditory interference. To insure that all participants were able to perceive sound localization and sound content, the exact same tasks with shorter inter-stimulus intervals (ISI) were used. For this purpose, the authors used bird songs as stimuli. The participants sat in a chair positioned in the center of a horizontal semicircular array with four speakers on the left side (-15°,-30°,-45°,-60°) and four on the right side (15°, 30°, 45°, 60°), body midline being aligned to a central position (0°). For each task, conditions with and without interference were contrasted using the same pairs of stimuli in a different order. Interferences (always different from target and comparison) consisted of a 2-sec sound made of the juxtaposition of four different bird song excerpts played backwards. *In the auditory object discrimination task*, the participant had to answer orally after each trial whether the two bird songs were identical or different. In the interference condition, the inter-stimulus interval was filled with distracting bird songs presented from the same loudspeaker. *In the sound localization discrimination task*, the participant had to answer whether the two locations were identical or different. Within a single trial, the exact same bird song was presented twice from either same or different loudspeakers. According to the predictions, patients with unilateral temporal-lobe lesions were expected to present a dissociation between spatial and non-spatial auditory short-term memory. More specifically, it was hypothesized that patients with right temporal-lobe damage will be impaired in non-spatial short-term memory involving bird songs whereas patients with either right or left temporal-lobe excision will be disturbed in short-term memory for sound location. Finally, interfering stimuli were expected to disrupt auditory short-term memory as compared to conditions without interference, particularly in patients with unilateral temporal-lobe lesions.

In keeping with our predictions, patients with right but not left temporal-lobe excision were inpaired in auditory object discrimination suggesting the predominant role of the right temporal-lobe structures in short-term memory for the auditory content. This result is supportive of previous work, suggesting the contribution of right temporal structures in short-term memory of tonal information, both in lesion (Johnsrude et al., 2001; Samson & Zatorre, 1988; Zatorre & Samson, 1991) and in imaging studies (Alain et al., 2001; Zatorre et al., 1994). Conversely, sound localization discrimination was affected by left and right temporal-lobe excision indicating that both temporal-lobe structures are important in auditory short-term memory for spatial information. This finding is in agreement with results of a previous lesion study that also demonstrated the role of both left and right temporal-lobe structures in auditory spatial processing (Zatorre & Penhune, 2001) therefore confirming our hypothesis. By contrast, the same patients were able to perform the tasks when short ISIs were used suggesting that the perception of sound content and sound localization does not depend on the integrity of the medial temporal lobe structures or of the temporal pole. Consequently, this study succeed in demonstrating the differential role of left and right temporal-lobe structures in the maintenance of spatial and non-spatial auditory information in short-term memory therefore revealing a dissociation between the side of the lesions and the nature of the auditory short-term memory deficits.

This result confirms the functional dissociation between short-term memory for auditory object and sound localization reported in studies with normal participants (Anourova et al., 1999; Clarke et al., 1998). It also complements previous neurophysiological reports indicating that these two types of auditory short-term memory depend on two anatomically distinct networks (Alain et al., 2001; Maeder et al., 2001). The differential influence of interference in spatial and non-spatial auditory short-term memory demonstrated in the present data further reinforces the presence of such a dissociation. Finally, our results revealed no reliable correlations between performance in our two discrimination tasks, emphasizing again the idea that spatial and non-spatial short-term memory depend on separate memory circuits. All of these findings are consistent with multiple storage systems that are, to some extent, independent from one another.

Conclusion

In addition to the well established dissociation between auditory short-term memory for verbal and non verbal information, the results reported in this paper provide psychological and psychophysical evidence suggesting that distinct sensory memory systems allow to retain pitch, loudness, timbre and possibly amplitude modulated sounds. Moreover, another dissociation between short-term memory for auditory object and sound localization has been recently demonstrated at the functional and anatomical levels. All these findings support a modular conception of sensory auditory memory (or short-term memory). Such a kind of modularity have also been proposed by Peretz and Coltheart (2003) in the musical domain. Our modular conception auditory sensory memory can be seen as an intermediate level between the acoustic analysis level and their musical processing components.

REFERENCES

Alain, C., Arnott, S. R., Hevenor, S., Graham, S., Grady, C. L. (2001). "What" and "where" in the human auditory system. *Proceeding of the National Academy of Sciences USA*, *98*, 12301–12306.

Anourova, I., Rama, P., Alho, K., Koivusalo, S., Kahnari, J., Carlson, S. (1999). Selective interference reveals dissociation between auditory memory for location and pitch. *Neuroreport*, *10*, 3543–3547.

Bachem, A. (1954). Time Factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America*, *26*, 751–753.

Berliner, J. E., Durlach, N. I. (1973). Intensity perception. IV. Resolution in roving-level discrimination. *Journal of the Acoustical Society of America*, *53*, 1270–1287.

Berliner, J. E., Durlach, N. I., Braida, L. D. (1977). Intensity perception. VII. Further data on roving-level discrimination and the resolution and bias edge effects. *Journal of the Acoustical Society of America*, *61*, 1577–1585.

Botte, M. C., Baruch, C., Mönikheim, S. (1992). Memory for Loudness: the Role of Loudness Contour. In Y. Cazals, L. Demany, and K. Horner (Eds.) *Auditory Physiology and Perception* (pp. 305–311). Oxford: Pergamon.

Braida, L. D., Durlach, N. I. (1988). Periphèral and central factors in intensity perception. In G. M. Edelman, W. E. Gall, W. M. Cowan (Eds.) *Auditory Function* (pp. 559–583). New York: John Wiley and Sons.

Chavis, D., Pandya, D. N. (1976). Further observations on corticofrontal pathways in the rhesus monkey. *Brain Research*, *117*, 369–386.

Clarke, S., Adriani, M., Bellmann, A. (1998). Distinct short-term memory systems for sound content and sound localization. *Neuroreport*, *9*, 3433–3437.

Clément, S. (2001). *La mémoire auditive humaine: Psychophysique et Neuroimagerie fonctionnelle*. Unpublished Doctoral Thesis, Université de Bordeaux 2, Bordeaux (France).

Clément, S., Demany, L., Semal, C. (1999). Memory for pitch versus memory for loudness. *J. Acoust. Soc. Am.*, *106*, 2805–2811.

Colombo, M., D'Amato, M. R., Rodman, H. R., Gross, C. G. (1990). Auditory association cortex lesions impair auditory short-term memory in monkeys. *Science*, *247*, 336–338.

Corsi, P. (1972). *Human memory and the medial temporal region of the brain.* Unpublished Doctoral Thesis, McGill University, Canada.

Deutsch, D. (1970). Tones and numbers: Specificity of interference in immediate memory. *Science*, *168*, 1604–1605.

Deutsch, D. (1972). Mapping of interaction in the pitch memory store. *Science*, *175*, 1020–1022.

Deutsch, D. (1975). The organization of short-term memory for a single acoustic attribute. In D. Deutsch, J. A. Deutsch (Eds.) *Short-Term memory* (pp. 107–151). New York: Academic Press.

Deutsch, D. (1999). The processing of pitch combinations. In D. Deutsch (Ed.) *The psychology of Music*. New York: Academic Press.

Deutsch, D., Feroe, J. (1975). Disinhibition in pitch memory. *Perception and Psychophysics*, *17*, 320–324.

Giard, M. H., Lavikainen, J., Reinikainen, K., Perrin, F., Bertrand, O., Pernier, J., et al. (1995). Separate Representation of stimulus frequency, intensity and duration in auditory sensory memory: An event-related potential and dipole-model analysis. *Journal of Cognitive Neuroscience*, *7*, 133–143.

Green, D. M., Kidd, G. J., Picardi, M. C. (1983). Successive versus simultaneous comparison in auditory intensity discrimination. *Journal of the Acoustical Society of America*, *73*, 639–643.

Harris, D. J. (1952). The decline of pitch discrimination with time. *Journal of Experimental Psychology*, *43*, 96–99.

Heffner, H. E., Masterton, B. (1978). Contribution of auditory cortex to hearing. In D. J. Chivers and J. Herbert (Eds.) *Recent Advances in Primatology*, Volume 1: *Behavior* (pp. 735–754). London: Academic Press.

Jerison, H. J., Neff W. D. (1953). Effect of cortical ablation in the monkey on discrimination on auditory patterns. *Federation Proceedings*, *12*, 73–74.

Johnsrudej, I. S., Penhune, V. B., Zatorre, R. J. (2000). Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain*, *123*, 155–163.

Kinchla, R. A., Smyzer, F. (1967). A diffusion model of perceptual memory. *Perception and Pyschophysics*, *2*, 219–229.

Koester, T. (1945). The time error and sensitivity in pitch and loudness discrimination as a function of time interval and stimulus level. *Archives of Psychology*, *297*, 3–9.

Lancelot, C., Ahad, P., Noulhiane, M., Hasboun, D., Baulac, M., Samson, S. (2003). Spatial and non-spatial auditory short-term memory in patients with temporal-lobe lesion. *Neuroreport*, *14*, 2203–2207.

Lorenzi, C., Wable, J., Moroni, C., Derobert, C., Frachet, B., Belin, C. (2000). Auditory temporal envelope processing in a patient with left-hemisphere damage. *NeuroCase*, *6*, 231–244.

Lü, S. L., Williamson, S. J., Kaufman, L. (1992). Behavioral Lifetime of Human Auditory Sensory Memory Predicted by Physiological Measures. *Science*, *258*, 1668–1670.

Maeder, P. P., Meuli, R. A., Adriani, M., Bellmann, A., Fornari, E., Thiran, J. P. (2001). Distinct pathways involved in sound recognition and localization: a human fMRI study. *NeuroImage*, *14*, 802–816.

Massaro, D. W. (1970). Preperceptual auditory images. *Journal of Experimental Psychology*, *85*, 411–417.

Milner, B. (1962). Laterality effects in audition. In V. B. Mountcastle (Ed.) *Interhemispheric Relations and Cerebral Dominance* (pp. 177–195). Baltimore: Johns Hopkins.

Moroni, C., Demany, L., Lorenzi, C. (2004). *Short-term memory for temporal envelopes.* Paper presented at the 7ème Congrès Français d'Acoustique (CFA/DAGA). Strasbourg (France).

Peretz, I. Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*, 688–691.

Perry, D. W., Zatorre, R. J., Petrides, M., Alivisatos, B., Meyer, E., et al. (1999). Localization of cerebral activity during simple singing. *Neuroreport*, *10*, 3979–3984.

Rakowski, A. (1972). *Direct comparison of absolute and relative pitch.* Symposium on Hearing Theory, IPO Eindhoven, Holland.

Samson, S. (1999). Musical function and temporal lobe structures: A review of brain lesion studies. *Journal of New Music Research*, *28*, 217–228.

Samson, S., Zatorre, R. J. (1988). Melodic and harmonic discrimination following unilateral cerebral excision. *Brain and Cognition*, *7*, 348–360.

Schröger, E. (1997). On the detection of auditory deviations: A pre-attentive activation model. *Psychophysiology*, *34*, 245–257.

Scoville, W. B, Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, and Psychiatry*, *20*, I, 11–21.

Semal, C., Demany, L. (1991). Dissociation of pitch from timbre in auditory short-term memory. *Journal of the Acoustical Society of America*, *89*, 2404–2410.

Semal, C., Demany, L. (1993). Further evidence for an autonomous processing of pitch in auditory short term memory. *Journal of the Acoustical Society of America*, *94*, 1315–1322.

Semal, C., Demany, L., Ueda, K., Hallé, P. A. (1996). Speech versus non-speech in pitch memory. *Journal of the Acoustical Society of America*, *100*, 1132–1140.

Sidtis, J. J., Volpe, B. T. (1988). Selective loss of complex-pitch or speech discrimination after unilateral lesion. *Brain and Language*, *34*, 235–245.

Starr, G. E., Pitt, M. A. (1997). Interference effects in short-term memory for timbre. *Journal of the Acoustical Society of America*, *102*, 486–494.

Stepien, L. S., Cordeau, J. P., Rasmussen, T. (1960). The effect of temporal lobe and hippocampal lesions on auditory and visual recent memory in monkeys. *Brain*, *83*, 470–489.

Wickelgren, W. A. (1968). Sparing of short-term memory in an amnesic patient: implications for strength theory of memory. *Neuropsychologia*, *6*, 235–244.

Wickelgren, W. A. (1969). Associative strength theory of recognition memory for pitch. *Journal of Mathematical Psychology*, *6*, 13–61.

Wolfe, H. K. (1886). Untersuschungen über das Tongedächtniss. *Philosophical Studies*, *3*, 524–571.

Zatorre, R. J. (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia*, *23*, 31–41.

Zatorre, R. J., Evans, A. C., Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, *14*, 1908–1919.

Zatorre, R. J., Evans, A. C., Meyer, E., Gjedde, A. (1992). Lateralization of phonetic and pitch processing in speech perception. *Science*, *256*, 846–849.

Zatorre, R. J., Ptito, A., Villemure, J. G. (1995). Preserved auditory spatial localization following cerebral hemispherectomy. *Brain*, *118*, 879–889.

Zatorre, R. J., Penhune, V. B. (2001). Spatial localization after excision of human auditory cortex. *Journal of Neuroscience*, *21*, 6321–6328.

Zatorre, R. J., Samson, S. (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain*, *114*, 2403–2417.

Polskie Forum Psychologiczne, 2004, tom 9, numer 1, s. 25–36

THE AMUSIAS

Amélie Racette, Krista L. Hyde, Isabelle Peretz
Department of Psychology, University of Montreal, Canada

It is quite recently that the study of the neural bases of music has become a rich and promising area of research. A key investigation in this field is the study of musical disorders, known as *amusias*. In the present review, we will first describe cases of acquired amusia in the presence or absence of an acquired language disorder. Next, we will present the congenital form of amusia. This will be followed by a description of a recent model of musical processing, including a discussion of the possible lateralization of the 'musical brain'. We will conclude by presenting the existing methods used to study the amusias.

Introduction

While it is obvious how having a language disorder might deeply affect one's life, it is perhaps less obvious to surmise how living with a problem in the musical domain may be difficult. However, music and language share many similar characteristics: both are universal across all cultures, and are defined by structured codes in auditory and motor modalities, as well as in written notation. Moreover, music is sub-served by specific and dedicated neural substrates (Peretz, 2001). In a world without music, one would lose an important affective and social means of communication. Despite the importance of these manifestations, musical disorders, which we refer to as amusias, have not nearly received as much attention as their language counterparts.

Historically speaking, the existence of amusia has been previously documented on several occasions. Bouillaud (1865), Proust (1872), Jellinek (1956) and Botez and Wertheim (1959) were among the first to describe amusic patients. However these studies suffered somewhat from lack of scientific rigor (see Marin & Perry, 1999 for a review). Just as there are multiple forms of aphasia, there exist various forms of amusia. For example, cases of receptive and expressive amusias, alexias, agraphias, impairments in musical memory as well as instrumental apraxias and rhythmic impairments have all been documented. Currently, amusia is a generic term which serves to describe both acquired and congenital disorders of music perception, memory, production, reading and writing of music, which cannot be accounted for by any auditory impairment, or a motor or intellectual deficit (Marin & Perry, 1999).

Correspondence concerning this article should be addressed to Isabelle Peretz,
e-mail: isabelle.peretz@umontreal.ca

The sparse literature on the study of music, as compared to language for instance, can be explained by certain inherent difficulties in this area. First, while it is accepted that most individuals can effectively communicate verbally, read and write, there exists an important variability in the general population in terms of level of musical experience. In an attempt to classify such a musical diversity (Grison, 1972), researchers in the field first concentrated on the study of musical impairments in musicians, while it is now known that non-musicians also possess musical abilities, without any prior musical training (see Tillmann, Bharucha, & Bigand, 2000). With the knowledge that it is possible to be a non-musician with amusia, there has been an increasing interest in the study of amusia. However, there is still no consensus as to the true definition of a 'musician': is an individual only a true musician when he/she has attained a famed status in the world of performance or composition, or should the definition also include the musician that may prefer simply to 'jam'?. The notion of musical expertise seems to be a fundamental factor in the study of music since an important degree of cerebral plasticity has been observed with differing levels of musical experience. For example, several studies have shown that certain brain areas known to be implicated in musical processing are larger in individuals with greater musical experience (Pantev et al., 1998; Schlaug, Jäncke, Huang, Staiger & Steinmetz, 1995; Schneider, Scherg, Dosch, Specht & Gutschalk, 2002; Pascual-Leone, 2001). Despite these issues, it is clear that the study of the amusias is important to understanding the neural bases of music. To this aim, one key avenue is the comparison of music and language.

Music and Language

A fundamental question when comparing music and language is whether music rests on a system originally designed to process language, or whether there exists a specialized system dedicated to music. The study of amusia has previously been limited to the study of aphasic patients (see Marin & Perry, 1999 for a review). In cases where amusia and aphasia are concomitant, the same impairments will most often be seen in both music and speech. For example, a patient that it is alexic in language may also have musical reading difficulties (Kawamura, Midorikawa & Kezuka, 2000). The question arises as to whether the same cerebral network is involved in both processes, or whether they correspond to distinct but adjacent networks. It has been shown that music reading can be selectively impaired (Cappelletti, Waley-Cohen, Butterworth & Kopelman, 2000) or spared (see Assal, 1973; Signoret, Van Eeckhout, Poncet & Castaigne, 1987 for an example in Braille). The presence of a double-dissociation between the reading of written language versus the reading of musical notation suggests that there exists an independent neural network for music reading.

In addition, cases in areas other than reading have been described (see Brust, 2001, for a review), such as cases of aphasia with no amusia (Basso & Capitani, 1985) or cases of amusia with no aphasia (Griffiths et al., 1997; Piccirilli, Sciarma & Luzzi, 2000). In the present chapter, we will pay particular attention to the latter case. Among these types of patients, are two classic ones, Chébaline and IR, each of who have a massive impairment in one of the two systems. Despite a cerebro-vascular accident resulting in aphasia, the Russian composer Chébaline was able to continue to compose highly provocative music (Luria, Tsvetkova & Futer, 1965). In contrast, IR (Peretz, Belleville & Fontaine, 1997) retained the ability to effectively communicate verbally, but lost the ability to recognize or produce musical songs that had once been familiar to her.

Thus, the musical system can be selectively impaired or spared with respect to the verbal language domain. In this case, the difference between the two domains seems to be in the nature, and not the level, of difficulty. If language and music were to lie on the same continuum of difficulty, a cerebral lesion should systematically affect the more complex domain. While this is not the case, it is possible that the two domains are independent at a certain level, but associated at another level. It is therefore essential to compare language and music at similar processing levels in order to determine which processes are dissociated and which are shared.

Songs represent a very important means for comparison between language and music, given that song is comprised of both lyrics and melody. The question arises as to whether songs might be encoded uniquely in memory, and thus that language and music may be processed under a unique code. This idea has motivated rehabilitation therapies of language disorders such as *Melodic Intonation Therapy*, which utilizes melodic and rhythmic exercises to improve speech production (Sparks, Albert & Helm, 1974).

In support of an integrated theory of song lyrics and melody (Serafine, Crowder and Repp, 1984; Serafine, Davidson, Crowder & Repp, 1986), it has been found that normal subjects have difficultly to access melody without the aid of any song lyrics. Such an integration of song lyrics and melody in song memory has also been found in brain-lesioned patients. Samson and Zatorre (1991) evaluated the ability of unilaterally temporally lesioned patients to recognize both song lyrics and song melodies. It was found that recognition of song lyrics was more difficult following a left lobectomy. Moreover, the authors also observed that melodic recognition was dependent on the lyrics with which the melodies were initially learned. These results led the authors to propose that there exist two distinct codes for song memory: a code where song lyrics and music are integrated, and a separate code for song lyrics alone.

Steinke, Cuddy and Jakobson (2000) have equally described a patient, KB, who demonstrates the same pattern of impairments, suggesting that there exists an integrated code in memory for songs. Consequent to a cerebral vascular accident (CVA) to the right hemisphere, KB was unable to recognize instrumental music, while retaining the ability to recognize previously learned song melodies with lyrics. The authors interpreted such a remarkable preservation of songs as the result of the retention of song lyrics in memory, which in turn, would facilitate musical recognition. This result is also consistent with Samson and Zatorre's proposition that there exists a type of memory for songs where lyrics and music are integrated, and another memory store for instrumental music. In this case, KB's problem would have stemmed from an impaired access to musical memory, while access to song memory was preserved.

Recent results from the study of song in aphasic patients provide further support for the idea of an integrated code for song lyrics and melodies (Hébert, Racette, Gagnon & Peretz, 2003; Peretz, Gagnon, Macoir & Hébert, 2004). Both patients had difficulty to produce song lyrics, whether by singing or speaking, however, they had no problem to hum the song melody. In this case, there would exist two separate codes for songs, one for the melody and one for lyrics. The latter code would be selectively impaired in the case of expressive aphasia, and would affect both song and speech.

In summary, the study of songs represents a unique way to compare music and language in both normal listeners and singers. It is unfortunate that such few studies have been undertaken to address these issues. It is important to note that all normal individuals possess the ability to sing. Song is not an exclusive ability to professional singers, and in future, the study of song should be addressed more systematically.

Congenital Amusia

In the previous sections, we have described cases of 'acquired' amusia, that result consequent to a brain lesion. These types of amusias are well known since they were discovered in patients who had undergone treatment through the health care system. The existence of an amusia that is present from birth has recently been termed congenital amusia (Peretz & Hyde, 2003). This type of amusia is not consequent to any cerebral damage and afflicted individuals have normal intelligence and a normal exposure to music. Congenital amusics are unable to recognize or hum familiar melodies, have no sensitivity to dissonance, an ability present in infants (Trainor & Heinmiller, 1998), and have great difficulty to detect wrong notes in melodies (Ayotte, Peretz & Hyde, 2002). This latter difficulty has been observed in a previous study of the British population (Kalmus & Fry, 1980). It has been estimated that about 4% of the population may be afflicted with congenital amusia.

It is possible that the origin of this amusic disorder may stem from a difficulty in the ability to discriminate fine changes in musical pitch (for example, at a semi-tone distance which corresponds to two adjacent notes on a keyboard) or smaller. This would explain why amusics have problems to perceive, and thus memorize, a melody that employs intervals of one semitone. Certain amusics are able to distinguish a question from a statement, which differ in terms of a final pitch change (Ayotte, Peretz & Hyde, 2002). It is important to note that speech employs pitch intervals that are much larger than in music, and thus explains why amusics show a deficit in musical pitch discrimination while sparing speech intonation. However, this pitch deficit is not specific to the musical domain, but rather 'musically-relevant', since music is composed of more subtle pitch variations as compared to speech (Peretz & Hyde, 2003).

Now that we have reviewed both the acquired and congenital forms of amusia, in the next section, we turn to describing a neuro-cognitive model of musical processing. This model presents the various components of musical processing, whereby the failure of one component (or box) or communication between components (an arrow) may result in an impairment in musical processing ability. In the case of a patient, a systematic evaluation will reveal at what point in the model an impairment may lie. In this way, pathologies may be very informative, in that their study enables the decomposition of a complex system into its elemental components (McCloskey, 2001).

A Modular Model of Musical Processing

In order to explain how we recognize a melody, Peretz (1993a) introduced a model where musical perception and memory are represented in a modular fashion. This model was primarily constructed by way of studies on musical recognition, a ubiquitous ability common to both musicians and non-musicians alike. Here we present the most recent version of this model (Peretz & Colheart, 2003), which has become a general model of musical processing.

Following an acoustic input, which refers to basic perceptual components including frequency, temporal duration, intensity and timbre of sound, we access a musical level of processing. This level is comprised of two principle pathways that will activate melodies stored in memory. These are the melodic pathway, which is defined by sequential pitch variations (the

'what' pathway), as well as the temporal pathway (the 'when' pathway). In perception, as well as in song and in reading, melody and rhythm can be selectively impaired, which demonstrates a certain independence (Peretz, 2001). The melodic pathway has a privileged access to the musical lexicon: through this pathway a song can be recognized with greater ease if the melody is played without rhythm, as opposed to the case where only rhythm is heard without melody (Hébert & Peretz, 1997). Moreover, a patient with an impaired melodic pathway is not able to compensate by way of the temporal pathway in order to recognize a melody (Peretz, 1994).

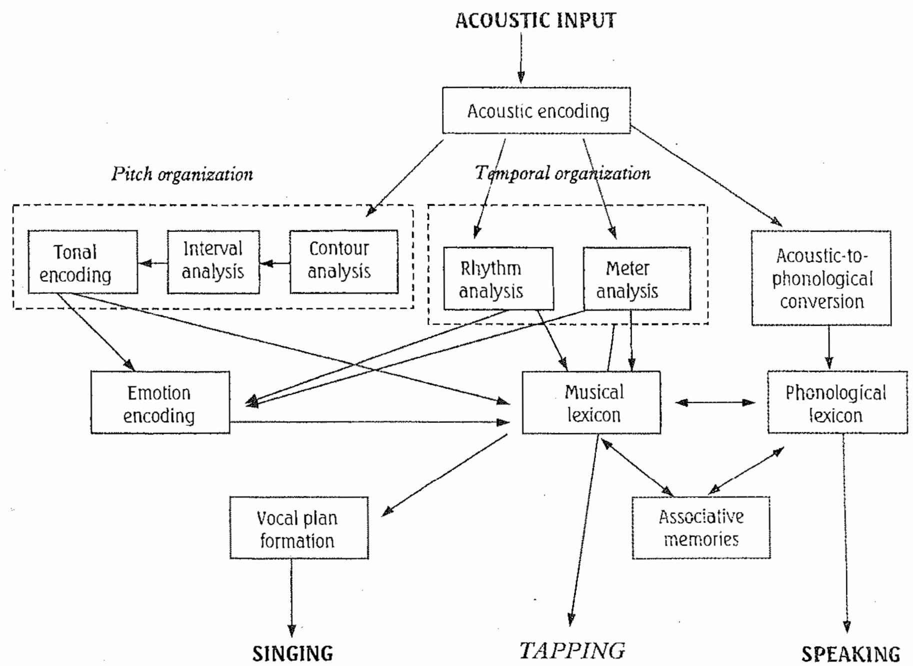

Figure 1. The modular musical processing model of Peretz & Coltheart (2003). An amusia may result consequent to damage to one component (a box) or the flow of information between components (an arrow)

The Melodic Pathway

The melodic pathway is comprised of three components: contour, intervals and tonality. Contour is described as the trajectory of the melodic line, in both ascending and descending directions. This pattern is stored in working memory as basic information, following the novel presentation of a melody (Dowling, 1982). Musical contour is similar to intonation, the melody of language. Several studies of amusic patients have explored to what degree the same process may play a role in music and language. It has been found that contour discrimination of music and intonation (with no verbal component) can both be impaired or spared, which suggests the existence of shared mechanisms between language and music (Nicholson et al., 2003; Patel, Peretz, Tramo & Labrecque, 1998). However, as we have previously seen in the case of congenital amusia, the processing of intonation contour may be spared in the presence of amusia. In this case, the dissociation would result form a mildly severe impairment in acoustic pitch discrimination common to both intonation and music.

Studies of brain-lesioned patients and neuroimaging studies of normals, have illustrated the role of the right superior temporal gyrus in the processing of sequential tonal pitch (Zatorre, 1985; Zatorre & Binder, 2000; Zatorre, Evans & Meyer, 1994). This region seems to be associated with more frontal structures in the short-term retention of melodic information (Penhune, Zatorre & Feindel, 1999; Zatorre & Samson, 1991). Since the right hemisphere is specialized in the processing of pitch, the 'melodic' amusias should be more frequently observed consequent to right hemisphere lesions. However, the most often reported amusic cases are seen in patients with bilateral lesions (Ayotte, Peretz, Rousseau, Bard & Bojanowski, 2000). Thus, it seems that a lateralization for music is less evident than for language, which is left lateralized (see Demonet & Thierry, 2001 for a review).

Other than absolute pitch and melodic contour, the critical information that allows for the melodic recognition is the pattern of melodic intervals, the second component of the melodic pathway. Intervals correspond to the exact distance between any two consecutive notes. Both contour and intervals have been associated with different musical processing modes, each respectively, the global and local approach. These two processing modes have been distinguished in brain-damaged patients. In left temporal lobe lesioned patients, only the processing of intervals is affected, while contour discrimination is spared (Liégeois-Chauvel, Peretz, Babaï, Laguitton & Chauvel, 1998; Peretz, 1990), while in right lesioned patients, contour discrimination is affected, and interval processing is also impaired. Thus, the global approach, would precede the local one. In other words, it seems that contour discrimination serves as an anchorage point for the processing of intervals.

Similarly, a classic study by Bever & Chiarello (1974), which explored melodic recognition by way of a dichotic listening task, showed that musicians have a right ear superiority (due to the left hemisphere), while non-musicians have a left ear advantage (right hemisphere). The authors suggested that this difference may be accounted for by a different type of melodic analysis, where musicians would use a local approach to contour analysis, while non-musicians would use a global approach. However, this distinction between processing modes and musical expertise is not absolute. Musicians are able to employ global contour (Peretz & Babaï, 1992) and non-musicians are able to extract local interval changes (Peretz & Morais, 1987).

The third component of the melodic pathway concerns the encoding of musical intervals by way of Western music 'tonal' rules. In music, pitch variations generate a determinate scale, a pattern of unequal-spaced pitches that are organized around 5 to 7 focal pitches. Scale tones are organized around a central tone, called the tonic, and typically a piece of music starts and ends on the tonic. Among the other scale tones, there is a hierarchy of importance or stability where the non-scale tones are the least related and often sound anomalous. This implicit tonal knowledge allows any given individual to detect when a musician strikes a wrong note for example. Indeed, sensitivity to musical tonality appears very early on in development, since babies have demonstrated preferences for scales with unequal steps, as in the tonal musical system (Trehub, Schellenberg & Kamenetsky, 1999). However, this widespread ability may, be lost or compromised as a consequence of brain damage.

I. Peretz (1993b) presents the case of GL, an amusic patient with an impaired ability to use this type of tonal knowledge, while both the use of contour and intervals as well as the temporal pathway are preserved. GL suffered from amusia consequent to a bilateral rupture of the middle anterior cerebral artery. Contrary to normal control subjects, GL had an impaired ability to detect a wrong note in a melody (Anomalous Pitch Detection), and did not show the typical preference for tonal excerpts versus atonal excerpts. The classic probe-tone method is another way to evaluate sensitivity to tonality. In this task, the participant has to indicate whether the final (probe) tone is congruous or not with the preceding tonal context. Next, a profile based on the context is generated as a function of each final tone judgement (Krumhansl, 1990). GL's profile was not as expected since he showed no preference when the final tone was indeed congruous. Another amusic case, MS presenting the reverse dissociation, serves to evidence the independence of tonal processing within the melodic route (Tramo et al., 1991). Despite impaired pitch perception consequent to a bilateral CVA, MS retained the ability to use the tonal system.

In summary, an amusia due to an anomaly in the melodic pathway can affect pitch processing at various levels (contour, intervals, tonality). In this way, music perception is then limited, as well as access to melodies stored in memory. The processing of pitch variations is associated primarily (but not exclusively) to right secondary auditory areas.

The Temporal Pathway

The temporal pathway has two levels of organization: meter and rhythm. Meter refers to the periodic alternation between strong and weak beats, whereas rhythm may be described as the organization of note durations. While Wilson and colleagues (2002) presented the case of an amusic patient who suffered from both metrical and rhythmic problems after sustaining a right temporoparietal infarct, there is also evidence to show that both rhythm and meter can be selectively impaired in brain-lesioned patients. For example, the posterior part of the superior temporal gyrus has been found to be implicated in rhythmic perception, whereas the anterior part is involved in the perception of meter (Liégeois-Chauvel et al., 1998). Thus both rhythm and meter are independent both anatomically and functionally (Liégeois-Chauvel et al., 1998; Peretz, 1990). However, as compared to the pitch domain, there is a relatively small literature on cases of patients with temporal deficits.

In production, rhythmic tasks often involve the reproduction of rhythmic sequences, while a typical metrical task would involve tapping the beat to different musical styles (e.g. disco, folklore, classical etc.) By way of these types of tasks, cases of dissociations between rhythm and meter have been described (Polk & Kertesz, 1993). Mavlov (1980) described the case of a professional musician who had rhythmic difficulties, where he was unable to recognize or reproduce rhythmic sequences. This type of rhythmic impairment, consequent to a left cerebro-vascular lesion, seems to result in a severe receptive and expressive amusia. Fries & Swihart (1990) described the case of a patient with metrical problems. Consequent to a right hemisphere lesion, this left-handed patient could no longer tap the beat along with melodies, but did retain the ability to reproduce rhythmic sequences.

In summary, impairments along the temporal pathway can selectively affect either rhythm or meter. The impact of these impairments are not necessarily specific to musical processing, but can also affect temporal processing in other modalities (Mavlov, 1980). This view is supported by a recent theory postulating that the left hemisphere would be preferentially sensitive to rapidly changing temporal information, as in speech, whereas the right hemisphere would be specialized for the processing of fine spectral changes as used in music (Zatorre, Belin & Penhune, 2002). The localization would no longer be dependent on the domain, but rather the required type of acoustic processing.

The Musical Lexicon

The melodic and temporal pathways give access to the musical lexicon, which comprises all of the melodies that have been previously heard and allow new representations to be stored in memory. The recognition of a melody is only possible if there is an adequate matching between the abstract representation made by both analysis pathways (with a greater importance on the melodic pathway in our Western musical system) and the representation stored in the musical lexicon. The musical lexicon is not limited to musical recognition. It is possible for the representations stored in the musical lexicon to activate perceptual pathways as well. This is a reverse pathway from recognition, and occurs when we hear a melody 'in our head'. This so called 'musical imagery', activates brain regions similar to those implicated in musical perception (Zatorre, Halpern, Perry, Meyer & Evans, 1996). The musical lexicon also plays a role in production, by activating processes responsible for humming or playing a familiar melody.

According to the model detailed above, an impaired musical recognition system may be due to a deficit in the access to the musical lexicon despite an intact lexicon (aperceptive amusia), or in the event of an actual damaged musical lexicon (associative amusia) as in the famous case of CN (Peretz, 1996). CN showed a selective impairment on tasks requiring the recognition of musical material such as familiar or non-familiar melodies, while she performed as controls in the recognition of non-musical material such as environmental sounds and song lyrics. In sum, the musical lexicon is considered the memory component of the modular model of music processing. Any alteration to this model will compromise the recognition and production (of memory) for familiar music.

Evaluation of the Amusias

Based on the model described above, and with the aim to screen for amusia, a battery of musical tests has been developed, the Montreal Battery of Evaluation of Amusia (Peretz, Champod & Hyde, 2003). The battery involves six tests which correspond to the components in Peretz (1993a) model. Three of these assess the ability to discriminate changes in melody (by pitch contour, scale, and interval size), while one tests rhythmic discrimination (by temporal grouping). Both the melodic and rhythmic tests use a "same-different" discrimination task, with the same set of novel but conventional sounding music. In a metric task, the subject is asked to decide whether a heard melody corresponds to a march (binary meter) or a waltz (ternary meter). The final test is an incidental memory task, where the subject must decide if a melody corresponds to one heard previously or not.

The MBEA has served as a diagnostic test for musical disorders in brain-lesioned patients (Ayotte et al., 2000; Liégeois-Chauvel et al., 1998; Peretz et al., 1997; Peretz, 1994; Steinke et al., 2001) as well as in congenital amusia (Ayotte et al., 2002; Peretz et al., 2002). For more details on the MBEA, it is possible to visit our laboratory website where norms are available (www.fas.umontreal.ca/psy/iperetz.html).

Conclusion

In conclusion, the systematic and relatively recent study of the amusias has revealed that there exists a neural network for music processing that is independent from that of language. As we have documented, this musical processing system is very complex, and is composed of multiple treatment components, each of which can be selectively impaired or spared. In this way, the study of the amusias constitutes one of the richest avenues to better understand the musical brain.

REFERENCES

Assal, G. (1973). Aphasie de Wernicke chez un pianiste. *Revue neurologique*, *29*, 251–255.

Ayotte, J., Peretz, I., Hyde, K. (2002). Congenital amusia: A group study of adults afflicted with a music-specific disorder. *Brain*, *125*, 238–251.

Ayotte, J., Peretz, I., Rousseau, I., Bard, C., Bojanowski, M. (2000). Patterns of music agnosia associated with middle cerebral artery infarcts. *Brain*, *123*, 1926–1938.

Basso, A., Capitani, E. (1985). Spared musical abilities in a conductor with global aphasia and ideomotor apraxia. *Journal of Neurology, Neurosurgery and Psychiatry*, *48*, 407–412.

Bever, T., Chiarello, R. (1974). Cerebral dominance in musicians and non-musicians. *Science*, *185*, 537–539.

Botez, M. I., Wertheim, N. (1959). Expressive aphasia and amusia following right frontal lesion in a right-handed man. *Brain*, *82*, 186–202.

Bouillaud, J. B. (1865). Sur la faculté du langage articulé. *Bulletin de l'Académie de Médecine*, *30*, 752–768.

Brust, J. (2001). Music and the neurologist: A historical perspective. *Annals of the New York Academy of Sciences*, *930*, 143–152.

Cappelletti, M., Waley-Cohen, H., Butterworth, B., Kopelman, M. (2000). A selective loss of the ability to read and write music. *Neurocase*, *6*, 332–341.

Demonet, J. F., Thierry, G. (2001). Language and brain: What is up? What is coming up? *Journal of Clinical and Experimental Neuropsychology*, *23* (1), 49–73.

Dowling, W. (1982). Melodic information processing and its development. In D. Deutsch (Ed.) *The psychology of music* (pp. 413–430). New York: Academic Press.

Fries, W., Swihart, A. (1990). Disturbance of rhythm sense following right hemisphere damage. *Neuropsychologia*, *28*, 1317–1323.

Griffiths, T., Rees, A., Witton, C., Cross, P., Shakir, R., Green, G. (1997). Spatial and temporal auditory processing deficits following right hemisphere infarction: A psychophysical study. *Brain*, *120*, 785–794.

Grison, B. (1972). *Étude sur les altérations musicales au cours des lésions hémisphériques.* Thèse de doctorat, Université de Paris.

Hébert, S., Peretz, I. (1997). Recognition of music in long-term memory : Are melodic and temporal patterns equal partners? *Memory & Cognition*, *25* (4), 518–533.

Hébert, S., Racette, A., Gagnon, L., Peretz, I. (2003). Revisiting the dissociation between singing and speaking in expressive aphasia. *Brain*, *126* (8), 1838–1850.

Jellinek, A. (1956). Amusia. *Folia Phoniatr*, *8*, 124–149.

Kalmus, H., Fry, D. B. (1980). On tune deafness (dysmelodia): frequency, development, genetics and musical background. *Annals of Human Genetics*, *43* (4), 369–382.

Kawamura, M., Midorikawa, A., Kezuka, M. (2000). Cerebral localization of the centre for reading and writing music. *NeuroReport*, *11* (14), 3299–3303.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Liégeois-Chauvel, C., Peretz, I., Babaï, M., Laguitton, V., Chauvel, P. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain*, *121*, 1853–1867.

Luria, A., Tsvetkova, L., Futer, J. (1965). Aphasia in a composer. *Journal of Neurological Science*, *2*, 288–292.

Marin, O. S., Perry, D. W. (1999). Neurological aspects of music perception and performance. In D. Deutsch (Ed.) *The psychology of music* (pp. 653–724). San Diego: Academic Press.

Mavlov, L. (1980). Amusia due to rhythmic agnosia in a musician with left hemisphere damage: A non-auditory supramodal defect. *Cortex*, *16*, 321–338.

McCloskey, M. (2001). The future of cognitive neuropsychology. In B. Rapp (Ed.) *The handbook of cognitive neuropsychology: What deficit reveal about the human mind* (pp. 593–610). Philadelphia: Psychology Press.

Nicholson, K. G., Baum, S., Kilgour, A., Koh, C. K., Munhall, K. G., Cuddy, L. L. (2003). Impaired processing of prosodic and musical patterns after right hemisphere damage. *Brain and Cognition*, *52*, 382–389.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L., Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature*, *392*, 811–814.

Pascual-Leone, A. (2001). The brain that plays music and is changed by it. *Ann N Y Acad Sci*, *930*, 315–29.

Patel, A. D., Peretz, I., Tramo, M., Labrecque, R. (1998). Processing prosodic and musical patterns: A neuropsychological investigation. *Brain and Language*, *61* (2), 123–144.

Penhune, V. B., Zatorre, R. J., Feindel, W. H. (1999). The role of auditory cortex in retention of rhythmic patterns as studied in patients with temporal lobe removals including Heschl's gyrus. *Neuropsychologia*, *37*, 315–331.

Peretz, I. (1990). Processing of local and global musical information in unilateral brain-damaged patients. *Brain*, *113*, 1185–1205.

Peretz, I. (1993a). Auditory agnosia: A functional analysis. In S. McAdams, E. Bigand (Eds.) *Thinking in Sound. The Cognitive Psychology of Human Audition* (pp.199–230). New York: Oxford University Press.

Peretz, I. (1993b). Auditory atonalia for melodies. *Cognitive Neuropsychology*, *10*, 21–56.

Peretz, I. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, *117*, 1283–1301.

Peretz, I. (1996). Can we lose memory for music? A case of music agnosia in a no musician. *Journal of Cognitive Neuroscience*, *8* (6), 481–496.

Peretz, I. (2001). Musical perception and recognition. In B. Rapp (Ed.) *The handbook of cognitive neuropsychology: what deficit reveal about the human mind* (pp. 519–540). Philadelphia: Psychology Press.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., Jutras, B. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron*, *33*, 185–191.

Peretz, I., Babaï, M. (1992). The role of contour and intervals in the recognition of melody parts: Evidence from cerebral asymmetries in musicians. *Neuropsychologia*, *30* (3), 277–292.

Peretz, I., Belleville, S., Fontaine, S. (1997). Dissociations entre musique et langage après atteinte cérébrale: un nouveau cas d'amusie sans aphasie. *Revue canadienne de psychologie expérimentale*, *51* (4), 354–367.

Peretz, I., Champod, A-S., Hyde, K. L. (2003). Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, *999*, 58–75.

Peretz, I., Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6* (7), 688–691.

Peretz, I., Gagnon, L., Macoir, J., Hébert, S. Singing in the brain: Insights from cognitive neuropsychology. *Music Perception*, submitted.

Peretz, I., Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Sciences*, *7* (8), 362–367.

Peretz, I., Morais, J. (1987). Analytic processing in the classification of melodies as same or different. *Neuropsychologia*, *25* (4), 645–652.

Piccirilli, M., Sciarma, T., Luzzi, S. (2000). Modularity of music : Evidence from a case of pure amusia. *Journal of Neurology, Neurosurgery and Psychiatry*, *69* (4), 541–545.

Polk, M., Kertesz, A. (1993). Music and language in degenerative disease of the brain. *Brain and Cognition*, *22*, 98–117.

Proust, A. (1872). De l'aphasie. *Archives générales de Médecine*, *1*, 147–166, 303–318, 653–685.

Samson, S., Zatorre, R. J. (1991). Recognition memory for text and melody of songs after unilateral temporal lobe lesion: Evidence for dual encoding. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *17* (4), 793–804.

Schlaug, G., Jäncke, L., Huang, Y., Staiger, J., Steinmetz, H. (1995). Increased corpus callosum size in musicians, *Neuropsychologia*, *33*, 1047–1055.

Schneider, P., Scherg, M., Dosch, G., Specht, H., Gutschalk, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nature Neuroscience*, *5* (7), 688–694.

Serafine, M. L., Crowder, R. G., Repp, B. (1984). Integration of melody and text in memory for songs. *Cognition*, *16*, 285–303.

Serafine, M. L., Davidson, J., Crowder, R. G., Repp, B. H. (1986). On the nature of melody-text integration in memory for songs. *Journal of Memory and Language*, *25*, 123–135.

Signoret, J. L., Van Eeckhout, P., Poncet, M., Castaigne, P. (1987). Aphasie sans amusie chez un organiste aveugle. *Revue neurologique*, *143*, 172–181.

Sparks, R., Helm, N., Albert, M. (1974). Aphasia rehabilitation resulting from Melodic Intonation Therapy. *Cortex*, *10*, 303–316.

Steinke, W. R., Cuddy, L. L., Jakobson, L. S. (2000). Dissociations among functional subsystems governing melody recognition after right-hemisphere damage. *Cognitive Neuropsychology*, *18* (5), 411–437.

Tillmann, B., Bharucha, J. J., Bigand, E. (2000). Implicit learning of tonality: a self-organizing approach. *Psychogical Review*, *107* (4), 885–913.

Trainor, L. J., Heinmiller, B. M. (1998). The development of evaluative responses to music: Infant prefer to listen to consonance over dissonance. *Infant Behavior and Development*, *21* (1), 77–88.

Tramo, M., Bharucha, J., Musiek, F. (1991). Music perception and cognition following bilateral lesions of auditory cortex. *Journal of Cognitive Neuroscience*, *2*, 195–212.

Trehub, S., Schellenberg, E., Kamenetsky, S. (1999). Infants' and adults' perception of scale structure. *Journal of Experimental Psychology: Human Perception & Performance*, *25* (4), 965–975.

Wilson, S. J., Pressing, J. L., Wales, R. J. (2002). Modelling rhythmic function in a musician post-stroke. *Neuropsychologia*, *40* (8), 1494–505

Zatorre, R. J. (1985). Discrimination and recognition of tonal melodies after unilateral cerebral excisions. *Neuropsychologia*, *23* (1), 31–41.

Zatorre, R. J., Belin, P., Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, *6* (1), 37–46.

Zatorre, R. J., Binder, J. R. (2000). Functional and structural imaging of the human auditory system. In A. W. Toga, J.C. Mazziotta (Eds.) *Brain Mapping: The Systems* (pp. 365–402). San Diego: Academic Press.

Zatorre, R. J., Evans, A., Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, *14*, 1908–1919.

Zatorre, R. J., Halpern, A., Perry, D., Meyer, E., Evans, A. (1996). Hearing in the mind's ear: A PET investigation of musical imagery and perception. *Journal of Cognitive Neuroscience*, *8*, 29–46.

Zatorre, R. J., Samson, S. (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain*, *114*, 2403–2417.

Polskie Forum Psychologiczne, 2004, tom 9, numer 1, s. 37–47

PERSISTENCE OF PERFORMANCE DETAILS IN MUSIC AND SPEECH

Melissa K. Jungers
Ohio State University, Newark, USA

What aspects of music and speech are retained in memory? How do remembered performance details influence future performances? This paper focuses on memory for performance details in music and speech and the influence of these elements from perception to performance. Listeners form a memory for a sentence or melody that includes timing and intensity details. These details then influence performance. Musicians persist in the tempo of a melody they have just heard. They also incorporate details of timing and intensity into their subsequent performances. Speakers persist in the rate of sentences when they produce similar sentences. As in music, this persistence extends beyond the global dimension of rate.

Introduction

The orchestra finishes triumphantly and the final notes reverberate in the concert hall. The audience members applaud and rise to their feet. What is left of that ephemeral sound? What part of the performance will the listeners keep? Will they walk away humming the tune? Research suggests listeners form a memory for the performance that includes not only the melody, but also the more subtle dynamics, timing, and nuances of the music. This ability to remember more than the basic categorical information is not specific to music: a similar ability to remember performance details is found in the domain of speech. In both domains, these acoustic details are part of the representation in memory for the melody or sentence. When producing a new sentence or melody, the listener is influenced by what was just heard. Thus, the music that seems to float away actually becomes part of the listener's memory and influences how the listener performs in the future.

The focus of this paper is on this influence of acoustic details from perception to performance. Pianists persist in the tempo of a melody they have just heard when they perform a similar melody. This persistence is not restricted to tempo. Musicians also retain information about meter and timing and incorporate these details into their subsequent performances. This persistence effect is also found in the domain of speech. Speakers persist in the rate of sentences when they produce similar sentences. As in music, this persistence extends beyond the global dimension of rate.

Correspondence concerning this article should be addressed to Melissa K. Jungers, e-mail: jungers.2@osu.edu

Musical details in performance and perception

Music is composed with specific notes and rhythms, but performers add acoustic features that are not in the notation. For example, performers may lengthen notes or increase accents. Such subtle acoustic variations are known as performance "expression" and they help listeners to differentiate two performances (Palmer, 1997). These musical nuances relate to the musical structure in an ordered way. Western tonal music is organized by meter and grouping principles (Cooper & Meyer, 1960, Lerdahl & Jackendoff, 1983). Meter refers to the alternation of strong and weak beats. For example, a march notated in 2/4 time will have an alternating strong-weak beat pattern while a waltz in 3/4 time will follow a strong-weak-weak beat pattern. Grouping refers to pitch relationships and rhythmic patterns (Lerdahl & Jackendoff, 1983; Cooper & Meyer, 1960). Meter and grouping are both arranged in a hierarchy, with smaller pitch or rhythmic events nested within larger events.

The fine-grained acoustic details in performance are often tied to this musical structure. For example, performers typically mark the ends of musical phrases by a decrease in tempo and dynamics (Henderson, 1936). This performance nuance is known as phrase-final lengthening and it highlights the importance of the phrase in the musical hierarchy (Lerdahl & Jackendoff, 1983; Palmer, 1996; Palmer, 1997). Performers also use performance features to mark meter. They perform with increased duration, more legato articulation, and louder accents on events that align with metrically strong beats than events that align with weak beats (Sloboda, 1983, 1985). Do these acoustic details interact in performance expression? In a study of musical accents associated with meter, rhythmic grouping, and melodic accents, meter and rhythm independently influenced performance expression (Drake & Palmer, 1993). However, the influence of melodic accent on performance expression depended on the context (Drake & Palmer, 1993).

The connection between performance expression and music structure also influences listeners' perceptions. In one study, listeners heard performances that contained one or more performance cues and judged the intended meter (Sloboda, 1985). The listeners relied on articulation cues and loudness to choose the intended meter, although some performers did not use loudness to indicate meter (Sloboda, 1985). Listeners base their judgments on musical expectations. For example, listeners had difficulty detecting a computer-lengthened event that occurred before a long duration in a simple rhythm (Drake, 1993). This perceptual error occurred in the same place in the music at which performers often lengthen events (Drake & Palmer, 1993). Also, when a computer-generated performance contained a lengthened event at a structurally-expected location, listeners were less likely to detect it (Repp, 1992). In another task, listeners judged how a probe beat fit into a metrical context. Their judgments revealed their implicit knowledge of metrical structure (Palmer & Krumhansl, 1990). Thus, both performers and perceivers relate fine-grained performance details to musical structure.

Musical details in memory

Performance details shape performance and perception, but are these details remembered by the listener? When the audience leaves the concert hall humming, what aspects of the performance have they retained in memory? Early perceptual research focused on listeners' abilities to recognize a tune as the same even when performed by a different instrument or at a different

tempo. This idea of perceptual constancy suggests that a basic version of the music is retained, stripped of the timbre, tempo, and other performance details (Dowling & Harwood, 1986; Large, Palmer, & Pollack, 1995). Thus, a normalized version of the melody remains (Large et al., 1995).

Perhaps listeners do not have memories for performance details because these features are used only to form the basic pitch and rhythmic categories, but are not retained in memory (Raffman, 1993). Raffman (1993) points to research by Siegel and Siegel (1977), that shows trained musicians do not accurately detect small pitch differences within categories. If trained musicians have perceptual difficulty with fine-grained pitch differences, surely the average listener does not retain acoustic details in memory.

However, there is evidence that listeners do remember performance details. In a study of memory for performance details, listeners with and without musical training were familiarized with one of two performances of the same short musical excerpt (Palmer, Jungers & Jusczyk, 2001). These performances contained the same pitches and rhythmic patterns, but differed in articulation, intensity, and interonset interval cues. At test, listeners heard both the original performance from familiarization as well as a different performance of the same melody. Listeners were required to identify the performance they heard at familiarization. Even though the pitch and rhythm categories in the two performances were the same, listeners recognized the performance from familiarization (Palmer et al., 2001).

The Palmer et al. (2001) study demonstrated memory for music performance details in adults. Listeners with and without formal music training could remember and differentiate performances based on fine acoustic details, but this result may be due partly to years of exposure to Western tonal music. To address whether musical acculturation is necessary for memory of musical details, Palmer et al. (2001) tested 10-month-old infants for performance memory using the same melodies. Infants were first familiarized with one performance of each melody. They were then tested with a head-turn preference procedure (Kemler Nelson et al., 1995) on the original and different performances of the same melodies. Infants oriented longer to the familiar performances during test, suggesting even infants retain acoustic cues for performances in memory (Palmer et al., 2001).

In addition to articulation details, musicians retain performance tempi for long periods of time. For example, musicians can perform an entire movement of a symphony at the same tempo as previous performances, with very little variability (Clynes & Walker, 1986; Collier & Collier, 1994). This ability to retain musical timing is not limited to those with musical training. In one study, nonmusicians reproduced popular songs from memory at tempi very close to the original tempo (Levitin & Cook, 1996). Also, the participants showed wide tempo variability for songs that did not have a standard original tempo (Levitin & Cook, 1996).

Speech details in production and perception

Music is not the only domain in which subtle performance variations are produced and perceived. The element of speech that includes these performance details is known as prosody. Informally, prosody is the way something is said. Prosody is both a structure that organizes sound and the suprasegmental features of speech including pitch, timing, and loudness (Cutler, Dahan, & von Donselaar, 1997). Prosody is also described as the "stress, rhythm, and intonation in spoken sentences" (Kjelgaard & Speer, 1999).

The prosody of a sentence, including word duration, timing, and intonation, can influence the listeners' interpretation of meaning. Prosody helps to disambiguate grammatically ambiguous sentences. In one experiment, listeners had to guess the meaning of ambiguous sentences read by four speakers (Lehiste, 1973). Listeners relied on timing and intonation cues and were better than chance for 10 of the 15 sentences (Lehiste, 1973). In another study, listeners heard syntactically ambiguous sentences with prosodic emphasis on different words, such as "They are FRYING chickens" and "They are frying CHICKENS" (Speer, Crowder, & Thomas, 1993). When listeners paraphrased the meaning for each sentence, their interpretations revealed the influence of the prosodic emphasis (Speer et al., 1993). The acoustic details associated with a speaker's voice can also aid sentence interpretation (Nygaard & Pisoni, 1998). Listeners were familiarized with isolated words produced by ten speakers. When they were later tested for intelligibility of novel words in noise, they better identified words spoken by a familiar voice than a new voice (Nygaard & Pisoni, 1998).

How does prosody relate to syntax: the grammatical rules for putting words together? In music, there is a clear connection between performance expression and musical structure, but the connection between prosody and syntax is not as straightforward. Although prosody often marks syntax, the relationship between these aspects of speech is not isomorphic (Cutler et al., 1997). Prosody has its own hierarchical structure (Beckman, 1996). Prosody helps to disambiguate syntax (Price et al., 1991). In one study, listeners judged the point at which recorded sentences switched from one ear to the other (Wingfield & Klein, 1971). The sentences contained a phrase that matched or did not match the intonation of the sentence. The listeners relied on both syntactic form and prosodic pronunciation to determine the switching point. Wingfield and Klein (1971) argued that syntax is the primary cue for sentence segmentation, although prosody helps to mark this syntax.

Although past research indicated that listeners use prosody to interpret syntactically ambiguous sentences (Lehiste, 1973, Lehiste et al., 1976), several recent papers questioned the generalizability of this effect outside of the laboratory. In one study, judges rated the intended meaning of syntactically ambiguous sentences produced by trained and untrained speakers (Allbritton, McKoon, & Ratcliff, 1996). The trained speakers had amateur or professional experience in acting or broadcasting. The speakers produced these sentences by reading them within two different passages that clarified the meaning. Untrained speakers and trained speakers who were unaware of the ambiguity did not disambiguate the sentences, according to the judges' ratings. Only the productions by trained speakers who were informed of the ambiguity were judged to disambiguate the meaning (Allbritton et al., 1996).

Another study of natural speech pitted passage context against sentence prosody (Fox Tree & Meijer, 2000). To create the stimuli, speakers read and memorized a short passage and then produced it. The middle sentence was then replaced with a sentence whose prosodic cues fit or did not fit the context of the passage. Listeners heard the recreated passages and chose the intended meaning of each passage. Listeners' choices were based on the context and not the prosody of the embedded sentence. The authors cited this as evidence that prosodic cues are not useful for syntactic disambiguation in a conversational context (Fox Tree & Meijer, 2000). However, there are several concerns with the experiment that make this conclusion less clear. For example, the stimuli were created by speakers whose task was to memorize and produce the passages verbatim. Although this production method is more natural than reading, the speakers may not have been using the full range of prosodic cues since their focus was to memorize and repeat the passage. Also, listeners could rely on the first sentence alone to interpret the passage.

Additionally, listeners were not instructed to use prosody (or even the middle sentence) to make their decisions. Although the debate about the use of prosody in syntactic disambiguation continues, there is evidence that prosody interacts with syntax (Wingfield & Klein, 1971; Lehiste, 1973).

Memory for prosody

Listeners have an amazing ability to understand speech under many conditions. They understand words spoken by children, men, and women, even though the vocal range is quite different for these groups. They understand speakers with unfamiliar accents. This human ability to understand spoken language with widely varying acoustic properties led early researchers to look for normalization processes. The idea behind normalization is that listeners form a representation of speech that lacks prosodic details (Pisoni, 1997). Thus, according to this view, timing and intonation are not part of the memory for a sentence.

More recent studies suggest that prosody is retained in our memory for language. Sentences that are presented with the same prosody at learning and test are recognized more accurately than sentences with different prosody (Speer et al., 1993). Also, listeners use prosodic cues to remember syntactically ambiguous sentences (Speer et al., 1993). Extralinguistic information, such as talker identity and talker rate, help listeners to identify words presented previously (Bradlow, Nygaard, & Pisoni, 1999). In addition, the presentation rate influences listeners' memory abilities for items produced by different speakers. Listeners more accurately recognize items presented at the same rate from familiarization to test than items presented at different rates (Nygaard, Sommers, & Pisoni, 1995). Prosody is incorporated into the memory for language.

Persistence in music

When musicians play in an ensemble and trade the melody from instrument to instrument, how does the performance of one player influence the performances of the others? Musicians include expressive nuances in their performances that are not notated in the musical score. Further, memory for music includes these details. How do these fine-grained performance details influence future performances?

One aspect of music that may persist from performer to performer is tempo. In one study, pianists were instructed to perform one melody at a particular tempo and then a second melody at either a slower or a faster tempo (Cathcart & Dawson, 1928). Pianists then played the original melody and tried to reproduce the original tempo, but their tempi drifted in the direction of the intervening performance tempo (Cathcart & Dawson, 1928). In a review of many studies with tasks as varied as color perception and weight lifting, Warren (1985) found a more general trend to explain this drift. Each domain showed a perceptual homeostasis so that perceivers' criteria shifted according to the current environmental conditions (Warren, 1985).

More recent research found evidence for persistence of musical tempo (Jungers, Palmer & Speer, 2002). Trained adult pianists first sight-read two melodies at their preferred rate. On each of the following trials, the pianists heard a computer-generated melody and then performed a similar melody. The pianists were not instructed to perform at a particular tempo. The computer-generated melodies (prime melodies) were blocked by fast (300 ms per eighth-note beat) or slow (600 ms per eighth-note beat) tempo. Pianists performed slower following the slow prime melodies than the fast prime melodies. However, their performances did not demonstrate simple imitation of the performance tempi they had just heard. Instead, the pianists' tempi reflected a drift away from their preferred tempo toward the prime melody tempo. Thus, the tempo of the prime melodies influenced the pianists' performances of the target melodies.

Although Jungers et al. (2002) demonstrated persistence of tempo in music, their study did not address other acoustic dimensions. In a set of experiments, Jungers (2003) examined whether pianists persist in the intensity or the articulation of what they have just heard. The intensity pattern of a musical performance often coordinates with the strong and weak beats in a given meter (Sloboda, 1983). Articulation represents the separation between note events and is measured as the offset time of one event minus the onset time of the next event, so negative values are staccato (separated) and positive values are legato (overlapping).

Computer-generated melodies with either binary or ternary intensity patterns served as stimuli. These prime melodies were produced with either a staccato or a legato articulation across all note events, unrelated to the intensity pattern. The notated musical stimuli, known as target melodies, were metrically ambiguous and contained no bar lines or articulation cues. They could be performed in either binary or ternary meter. The goal of the experiment was to test whether pianists persist in the performance cues that are structurally-related (intensity) or structurally-unrelated (articulation) (Jungers, 2003).

In the experiment, pianists first sight-read two of the notated melodies to assess any bias for performing in one meter or another. Then on each trial, pianists heard a computer-generated prime melody and performed a similar target melody (with a similar number of events and musical structure). The prime melodies were blocked by meter. Pianists were instructed to concentrate on the melodies for a later recognition task.

Pianists performed with a more separated style following the staccato than the legato prime melodies. Thus, pianists persisted in the metrically-unrelated cue of articulation. Pianists did not show a significant difference in intensity on metrically strong and weak events, although there was a trend for more intense events on metrically strong beats. However, pianists did incorporate the meter into their performances. The events that were expected to be more intense (if pianists persisted in the meter from the prime) were instead played with more length. Thus, instead of producing a strong-weak intensity pattern for a target melody that followed a binary prime melody, pianists produced a long-short articulation pattern. Pianists used articulation cues rather than intensity cues to produce a binary or ternary metrical interpretation. This means pianists perceived the meter and persisted in the meter, but they instantiated the meter with different performance cues they those they had heard in the prime melody. Thus, the pianists' performances revealed persistence of metrically-related and metrically-unrelated performance dimensions (Jungers, 2003).

Persistence in speech

When two people engage in conversation, how does the way one person speaks influence the speech of the other? Do conversation partners persist in the speech patterns of one another? One aspect of speech that persists is the syntactic form. Listeners who repeated a sentence they

had heard were more likely to use the same syntax from the sentence when describing a picture (Bock, 1986). For example, when subjects heard and repeated the passive sentence, "The referee was punched by one of the fans," they were more likely to describe a picture with a lightning bolt and church in the passive form as "The church is being struck by lightning" instead of "Lightning is striking the church" (Bock, 1986). Such structural priming lasts beyond one or two sentences (Bock & Griffin, 2000).

There is also evidence that speech rate persists. In one study, participants heard a recording of two male speakers: one at a fast rate and one at a slow rate (Kosslyn & Matt, 1977). Participants then read a paragraph they were told was written by one of the two speakers. The participants read the passage faster when they thought the fast-speaking person wrote it. However, the participants did not perfectly mimic the rate; their production rates were always slower than the rate of the fast speaker (Kosslyn & Matt, 1977).

Further evidence for rate persistence was demonstrated by Jungers, Palmer, and Speer (2002). Participants first read two sentences aloud as a measure of their preferred speech rate. Next, they heard a prime sentence and then read a written target sentence that was matched for number of syllables, lexical stress pattern, and syntactic structure. For example, participants heard, "She read the paper in a hurry" and then read aloud, "He smelled the coffee for a moment" (Jungers et al., 2002). The prime sentences were recorded by a naive female speaker at slow (750 ms or 80 bpm per accent) and fast (375 ms or 160 bpm per accent) rates. As in the music task, participants were instructed to attend carefully to the sentences for a later recognition task. The participants' rates showed an influence of both the prime rate and their preferred speaking rate. Although the speakers and the musicians in the parallel task were both influenced by the prime and preferred performance rates, the speakers were more influenced by the preferred rate and the musicians were more influenced by the prime rate.

Rate is a global prosodic cue that affects the entire utterance. Do speakers also persist in the fine-grained prosodic details of sentences they have heard? Jungers (2003) examined whether prosodic details such as phrase break location and pitch pattern persist in speech. Speech is produced in phrases, often with a pause at a phrase break location. These phrases make the sentence meaning clear and correlate with the sentence structure. The pitch pattern of a sentence can carry meaning, such as a rising pitch at the end of a sentence in English indicating a question. The pitch patterns used in this study occurred at the phrase break locations, but they did not independently add meaning to the sentence. Thus, the phrase break locations were structurally-related while the pitch patterns were structurally-unrelated.

The experimental stimuli were syntactically ambiguous sentences. For example, "Either Brett or Mike and Kay will come to babysit" can be produced with a prosodic phrase break after Brett, which implies that Brett alone or Mike and Kay together will come. This sentence could also be produced with a phrase break after Mike, which implies that Kay will come and one of the two men will also come (Jungers, 2003). A naïve female speaker recorded four versions of each sentence with an early or late phrase break and two different pitch patterns. The written version of each sentence was presented on a computer screen in capital letters without punctuation marks.

The participants first read three sentences aloud to assess their preferred prosodic production. Then participants listened to a prime sentence and produced a target sentence on each trial. Trials were blocked by phrase break location. Participants were instructed to pay careful attention to the sentences for a later recognition task.

Participants used similar phrase breaks in their target sentences as they had heard in the prime sentences. They incorporated the structurally-related phrasing cues into their own utterances. The musicians in the parallel task incorporated both structurally-related (metrical) and structurally-unrelated cues into their performance. Although listeners did not persist in the specific pitch pattern of the prime sentences, this may be partly because the participants did not use the full range of pitch patterns and phrasing. Current research is exploring whether listeners also incorporate pitch from perception to production.

Why persist?

Recent research demonstrates persistence of performance details in both speech and music. This persistence is not limited to tempo. Instead, it includes structurally-related details such as meter and phrasing, as well as structurally-unrelated details such as articulation. What is the advantage of persisting in these performance variations? One possibility is that persistence aids communication in speech and music. By producing similar prosodic patterns, conversation partners may be able to more quickly understand each other. Speakers adjust their utterances to aid listeners. For example, speakers add fillers such as "um" and "uh," that may help to pace the conversation (Clark, 2002). Also, special words and phrases such as "so," "now," "uh," and "um" serve to mark either a new turn or a continuing turn (Fox Tree, 2000). Prosodic persistence may be another way speakers adjust so listeners are prepared for the utterances and can understand easily.

A second possible advantage of persistence is that similar prosodic cues may be easier to produce because they are already primed in the speaker's memory. Syntactic persistence effects in speech have been explained as a type of implicit learning (Bock and Griffin, 2000). Perhaps a similar implicit learning explains persistence of performance cues in music and speech. Exposure to a longer stimulus could lead to a stronger memory representation or to greater implicit learning.

Music and language – common mechanisms?

Does the persistence of performance cues in speech and music stem from common mechanisms? In the study of rate persistence, both musicians and speakers were influenced by their preferred rate as well as the prime rate (Jungers, Palmer, Speer, 2002). However, there are differences in the degree of influence, with the musicians showing greater influence of the prime rate and speakers showing greater influence of their preferred rate. This difference may be due to differing performance expectations. Musicians traditionally perform in an ensemble where the goal is to synchronize with the conductor and fellow performers. Speakers, on the other hand, are more concerned with presenting a clear message and there is little pressure to speak ata particular rate.

The examination of prosodic persistence of structurally-related and structurally-unrelated cues also revealed a distinction between music and language (Jungers, 2003). The musicians showed persistence of metrically-unrelated cues, but they persisted with metrically-related cues to a lesser degree. The speakers persisted in structurally-related cues, but there was little evidence of persistence for structurally-unrelated cues. This difference may be due to the specific task or it may point to a more fundamental difference between the domains.

There is some evidence for a relationship between prosodic and musical patterns. One study examined the perception of two "amusic" subjects who had specific music perception impairments due to brain damage (Patel et al., 1998). The subjects performed a prosodic discrimination task as well as a parallel music discrimination task with stimuli derived from the language task. One subject performed well on both tasks while the other subject performed poorly on both, suggesting a common neural mechanism for interpreting linguistic and musical prosody (Patel et al., 1998). Another study showed that classical compositions by French and English composers differed in rhythmic patterns, paralleling the rhythmic differences between the two languages (Patel & Daniele, 2003). This study suggests linguistic prosody influences musical prosody. This connection between prosody in music and speech is unique, since many aspects of music processing are thought to be specific to music and may be organized modularly (Peretz & Coltheart, 2003).

The literature reviewed here suggests both commonalities and differences in the persistence of performance details across the domains of music and speech. Research continues to explore this persistence effect and the relationship between music and speech.

REFERENCES

Allbritton, D. W., McKoon, G., Ratcliff, R. (1996). Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 22*, 714–735.

Beckman, M. E. (1996). The parsing of prosody. *Language and Cognition Processses, 11*, 17–67.

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*, 355–387.

Bock, K., Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology: General, 129*, 177–192.

Bradlow, A. R., Nygaard, L. C., Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception & Psychophysics, 61*, 206–219.

Cathcart, E. P., Dawson, S. (1928). Persistence: A characteristic of remembering. *British Journal of Psychology, 18*, 262–275.

Clark, H. H. (2002). Speaking in time. *Speech Communication, 36*, 5–13.

Clynes, M., Walker, J. (1986). Music as time's measure. *Music Perception, 4*, 85–119.

Collier, G. L, Collier, J. L. (1994). An exploration of the use of tempo in jazz. *Music Perception, 11*, 219–242.

Cooper, G., Meyer, L. B. (1960). *The rhythmic structure of music*. Chicago: University of Chicago Press.

Cutler, A., Dahan, D., van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech, 40*, 141–201.

Dowling, W. J., Harwood, D. L. (1986). *Music cognition*. Orlando: Academic Press.

Drake, C. (1993). Perceptual and performed accents in musical sequences. *Bulletin of the Psychonomic Society, 31*, 107–110.

Drake, C., Palmer, C. (1993). Accent structures in music performance. *Music Perception, 10*, 343–378.

Fox Tree, J. E. (2000). Coordinating spontaneous talk. In L. Wheeldon (Ed.) *Aspects of language production* (pp. 375–406). Philadelphia: Psychology Press.

Fox Tree, J. E., Meijer, P. (2000). Untrained speakers' use of prosody in syntactic disambiguation and listeners' interpretations. *Psychological Research, 63*, 1–13.

Henderson, M. T. (1936). *Rhythmic organization in artistic piano performance* (pp.281–305). See CF. Seashore.

Jungers, M. K. (2003). *Prosodic persistence in music performance and language production.* Unpublished doctoral dissertation, Ohio State University, Columbus.

Jungers, M. K., Palmer, C. (2000). Episodic memory for music performance. *Abstracts of the Psychonomic Society, 5*, 105.

Jungers, M. K., Palmer, C., Speer, S. R. (2002). Time after time: The coordinating influence of tempo in music and speech. *Cognitive Processing, 2*, 21–35.

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., Gerken, L. A. (1995). The head-turn preference procedure for testing auditory perception. *Infant Behavior and Development, 18*, 111–116.

Kjelgaard, M. M., Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language, 40*, 153–194.

Kosslyn, S. M., Matt, A. M. (1977). If you speak slowly, do people read your prose slowly? Person-particular speech recoding during reading. *Bulletin of the Psychonomic Society, 9*, 250–252.

Large, E. W., Palmer, C. (2001). Perceiving temporal regularity in music. *Cognitive Science, 26*, 1–37.

Large, E. W., Palmer, C., Pollack, J. B. (1995). Reduced memory representations for music. *Cognitive Science, 19*, 53–96.

Lehiste, I. (1973). Phonetic disambiguation of syntactic ambiguity. *Glossa, 7*, 106–122.

Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics, 5*, 253–263.

Lehiste, I., Olive, J. P., Streeter, L. (1976). Role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America, 60*, 1199–1202.

Lerdahl, F., Jackendoff, R. (1983). *A generative theory of tonal music.* Cambridge, MA: MIT Press.

Levitin, D. J., Cook, P. R. (1996). Memory for musical tempo: Additional evidence that auditory memory is absolute. *Perception & Psychophysics, 58*, 927–935.

Nygaard, L. C., Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics, 60(3)*, 355–376.

Nygaard, L. C., Sommers, M. S., Pisoni, D. B. (1993). Effects of stimulus variability on perception and representation of spoken words in memory. *Perception & Psychophysics, 57*, 989–1001.

Palmer, C. (1996). Anatomy of a performance: sources of musical expression. *Music Perception, 13*, 433–454.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48*, 115–138.

Palmer, C., Jungers, M. K., Jusczyk, P. W. (2001). Memory for musical prosody. *Journal of Memory and Language, 45*, 526–545.

Palmer, C., Krumhansl, C. L. (1990). Mental representations for musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 728–741.

Patel, A. D., Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition, 87*, B35–B45.

Patel, A. D., Peretz, I., Tramo, M., Labreque, R. (1998). Processing Prosodic and Musical Patterns: A Neuropsychological Investigation. *Brain and Language, 61*, 123–144.

Peretz, I. Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*, 688–691.

Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson, J. W. Mullenix (Eds.) *Talker variability in speech processing* (pp. 9–32). San Diego: Academic Press.

Price, P., Ostendorf, M., Shattuck-Hufnagle, S., Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America, 90*, 723–735.

Raffman, D. (1993). *Language, music, and mind.* (Bradford Book, ed.). Cambridge, MA: MIT Press.

Repp, B. H. (1992). Probing the cognitive representation of musical time: structural constraints on the perception of timing perturbations. *Cognition, 44*, 241–281.

Seashore, C. E. (Ed.) (1936). *Objective Analysis of Musical Performance* (vol. 4). Iowa City: University of Iowa Press.

Siegel, J. A., Siegel, W. (1977). Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception and Psychophysics, 21*, 399–407.

Sloboda, J. A. (1983). The communication of music metre in piano performance. *Quarterly Journal of Experimental Psychology, 35*, 377–395.

Sloboda, J. A. (1985). Expressive skill in two pianists: metrical communication in real and simulated performances. *Canadian Journal of Psychology, 39*, 273–293.

Speer, S. R., Crowder, R. G., Thomas, L. M. (1993). Prosodic structure and sentence recognition. *Journal of Memory and Language, 32*, 336–358.

Warren, R. M. (1985). Criterion shift rule and perceptual homeostasis. *Psychological Review, 92*, 574–584.

Wingfield, A., Klein, J. F. (1971). Syntactic structure and acoustic pattern in speech perception. *Perception and Psychophysics, 9*, 23–25.

Polskie Forum Psychologiczne, 2004, tom 9, numer 1, s. 48–59

THE ATTRIBUTION OF EMOTION AND MEANING TO SONG LYRICS

William Forde Thompson, Frank A. Russo
University of Toronto, Canada

We examined the effect of music on the interpretation of song lyrics. Listeners were presented with sung lyrics, spoken lyrics, or written poetry, and judged the text for emotional valence and meaningfulness. Experiment 1 revealed that, for some songs, music influenced whether lyrics were interpreted as conveying a positive or negative message. Experiment 2 showed that for familiar music, sung lyrics were judged as more meaningful than the same lyrics presented as spoken text, suggesting that personal associations or other significance implied by familiar music are attributed to the accompanying lyrics. In Experiment 3, repeated exposure to unfamiliar songs led to an increase in the perceived meaningfulness of the lyrics. We raise the possibility that music and lyrics become represented in an increasingly integrated manner with increased exposure and familiarity, allowing greater cross-talk between the two media.

Introduction

Songs involve a strong and complex connection between melody and lyrics. Melody and lyrics are often coordinated for overall emotional meaning, points of stress, grouping, expectancy and closure. In some cases prosodic patterns associated with speech influence compositional choices that are made in creating the melodic and rhythmic aspects of a song. Music performers may also express emotions by manipulating the same acoustic variables that are manipulated in emotional speech (for a review, see Juslin & Laukka, 2003).

Although some lyrics are well written, they are not always highly meaningful outside of a musical context. Yet in the context of a song, they may take on unexpected significance. Music invites listeners to search for deeper meaning in lyrics, enhancing the process of semiosis. Whether silly love songs, indulgent songs about personal growth, or naïve political songs, messages conveyed by lyrics seem richer, more profound, more persuasive, and more emotional when embedded in a musical context than when read as straight text (Galizio & Hendrick, 1972; Iverson, Rees & Revlin, 1989; Stratton & Zalanowski, 1994).

Correspondence concerning this article should be addressed to William Forde Thompson, e-mail: b.thompson@utoronto.ca

This research was supported by a grant to the first author from the Natural Sciences and Engineering Research Council. We thank Aubrey Litvak for technical assistance.

For example, it is generally agreed that protest movements of the 60's were greatly influenced by folk and folk-rock music of the time. Yet the lyrics of that music tended to deal with political and social issues in simplistic ways. They conveyed a vague sentiment for political change but rarely if ever articulated a coherent political framework for bringing about such change. In an examination of this phenomenon, Galizio & Hendrick (1972) presented listeners with four folk songs, each of which had or did not have guitar accompaniment. Lyrics presented in the context of guitar accompaniment were perceived to be more emotionally arousing and persuasive than lyrics presented without guitar accompaniment. The authors proposed that because music listening is a pleasant activity, it creates a temporary mood of compliance toward any recommendations that may be available in an ongoing communication (see also Dabbs & Janis, 1965).

There are three views on the synchronization of words and song. The first is called *assimilation*. This position argues that lyrics subserve the music. Assimilationists point out that lyrics are rarely intended to lead an independent life but, rather, presuppose musical completion. According to Suzanne Langer (1957): *when words and music come together in song, music swallows words... song is not a compromise between poetry and music... song is music.* That is, song lyrics should be interpreted merely as one of the textures of the music.

Another view is *independence*. In this view, language and music make uneven partners because whereas language is a full-fledged semiotic system with a definable repertory of signs, it is not possible for music to be broken down into meaningful units (Benveniste, 1969). Because music has no identifiable relationship with the semiotics of language, the combination of words and music is destined to fail as an integrated medium.

A third view is *interaction*. In this view, music and lyrics are partially independent systems with their own rules, but there is enough overlap between them to allow interaction (Gorlée, 1997; Juslin & Laukka, 2003; Serafine, Crowder & Repp, 1984). More specifically, overlap in the structural characteristics of prosody and melody might increase the potential for one medium to influence the interpretation of the other medium. Areas of overlap include expectancy and closure, connections between prosody and melody, overlap in rhythm and grouping, and similarities in overall emotional meaning (for related discussions, see Peretz, Radeau & Arguin, 2004; Samson & Zatorre 1991; Besson et al., 1998).

Because music is temporally patterned, has syntax, creates expectancies, and has nested structure, it connotes meaning and significance, even though it typically makes no actual reference to objects or events in the world. This implied or free-floating meaning might feed into or spill over onto accompanying lyrics.

Figure 1 illustrates the idea schematically as a *trickle-down* effect. The Figure illustrates how music and lyrics have an expressive shape that unfolds over time. Structural attributes of music such as syntax, rhythm, expectancy, and grouping are often perceived to be meaningful, but their meaning is non-referential. In contrast, the meaning of lyrics relates to *definable referents*. Because music implies meaning but does not refer to specific objects or events, this free-floating meaning may readily become attached to accompanying media such as song lyrics, or film materials.

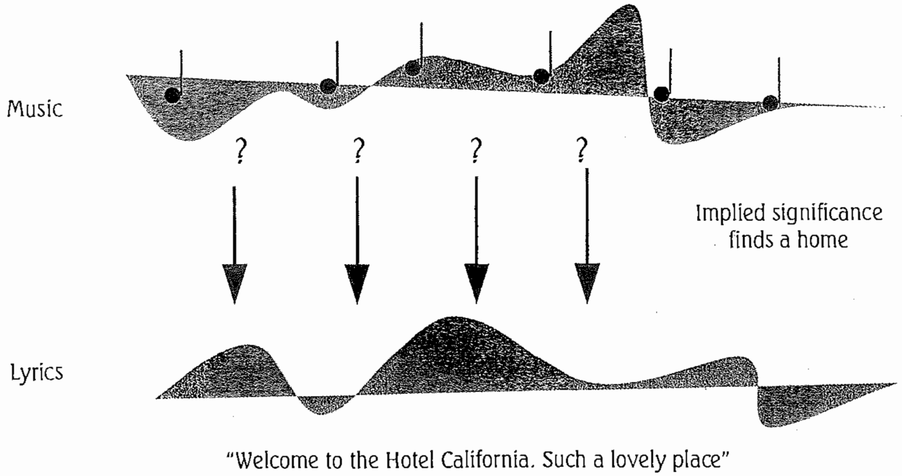

Figure 1. An illustration of the trickle-down effect, whereby free-floating meaning is attached to accompanying media such as song lyrics

In three experiments, we examined whether music influences responses to song lyrics. We first assessed whether music can manipulate interpretation of the emotional valence of lyrics. We then assessed whether song lyrics are more meaningful and significant when they are embedded in a musical context than when they are evaluated outside of a musical context. Finally, we examined the effects of repeated exposure to unfamiliar songs on the perceived significance of the lyrics.

Experiment 1: Judgments of emotional valence

The purpose of Experiment 1 was to examine whether the emotional connotation of spoken verbal information can be significantly altered when communicated in the context of a song. It is well known that music is often perceived to convey emotional meaning. For example, emotions conveyed in musical soundtracks significantly influence viewers' interpretation of films (Cohen, 2001; Thompson, Russo, & Sinclair, 1994). The question addressed in this experiment is whether emotional connotations of music also influence listeners' interpretations of the emotional meaning of accompanying verbal information.

Method

Participants. Thirty-four participants were recruited from the University of Toronto community. Participants were assigned to either a music presentation or spoken presentation condition. The age and musical training of participants was well matched across conditions. Participants in the music presentation group included 13 females and 4 males with a mean age of 19.58 years

(SD = 2.39) and an average of 2.15 years of musical training (SD = 3.23). Participants in the spoken presentation group included 13 females and 4 males with a mean age of 19.94 years (SD = 3.21) and an average of 2.09 years of musical training (SD = 3.17).

Stimuli. Four pop songs were selected that had been hits in the 1970's (all were in the Billboard Top-100 at some point) but were relatively unknown to our sample of participants in the year 2003. These songs were: 1. "Friend of the Devil" by the Grateful Dead (1970); 2. "Kodachrome" by Paul Simon (1973); 3. "Still the same" by Bob Seger (1978); and 4. "Sundown" by Gordon Lightfoot (1973). For each song, we selected a continuous excerpt of lyrics that included one verse and the chorus.

A professional musician/producer, Aubrey Litvak, recorded all lyrics in both the sung and spoken conditions. Performances of lyrics were recorded using a Rode NT-2 microphone routed through a Yamaha 01V microphone pre-amp into a PC computer running Steinberg Cubase SX. Recordings were digitized at 44.1 Khz using 24-bit ADA conversion. A rendering of the original musical accompaniment to the sung lyrics (e.g., drums, guitar, keyboard) was performed by the musician/producer and was recorded using the same setup. The performance of the sung and spoken versions was mindful of timing – specifically, the onset time of each word was roughly matched across sung and spoken versions of each lyric. In addition, the RMS levels of all recorded lyrics (singing with accompaniment and spoken) were matched to within 2dB of maximum and minimum readings. Four pieces of anonymous poetry from the public domain were selected for the text presentation of verbal material. The presentation of audio was realized over Sennheiser HD280 headphones. The presentation of visual text was realized over a computer monitor.

Procedure. Participants in the spoken group rated the valence of both spoken lyrics and written poetry. Participants in the music presentation group rated the valence of lyrics that were sung with musical accompaniment as well as the valence of poetry presented as text on the screen. The trials involving presentations of poetry were identical in spoken and music presentation conditions and served as an anchor so that ratings could be compared across the two conditions. The presentation of poetry and lyrics within conditions was fully randomized.

Participants were asked to rate the extent to which lyrics and poetry conveyed a positive or negative message. Responses were made on a 7-point scale where "1" represented "sad" and "7" represented "happy".

Results and Discussion

An alpha level of .05 was used for all statistical tests. The primary analysis involved comparing ratings of lyrics when they were *sung* with *music* to ratings of lyrics when they were *spoken*.

Figure 2 shows the effects of music on whether the lyrics were interpreted as conveying a positive or a negative message. The Figure illustrates mean ratings of sung lyrics, spoken lyrics, and written poetry, and illustrates that music influenced the affective valence of the lyrics in different ways depending on the song. In some cases the music increased the sense that the lyrics were conveying a positive message, as in Grateful Dead's Friend of the Devil (song 1) or Paul Simon's Kodachrome (song 2), but music had no effect for Bob Segar's Still the same (song 3), and the opposite effect for Gordon Lightfoot's Sundown (song 4).

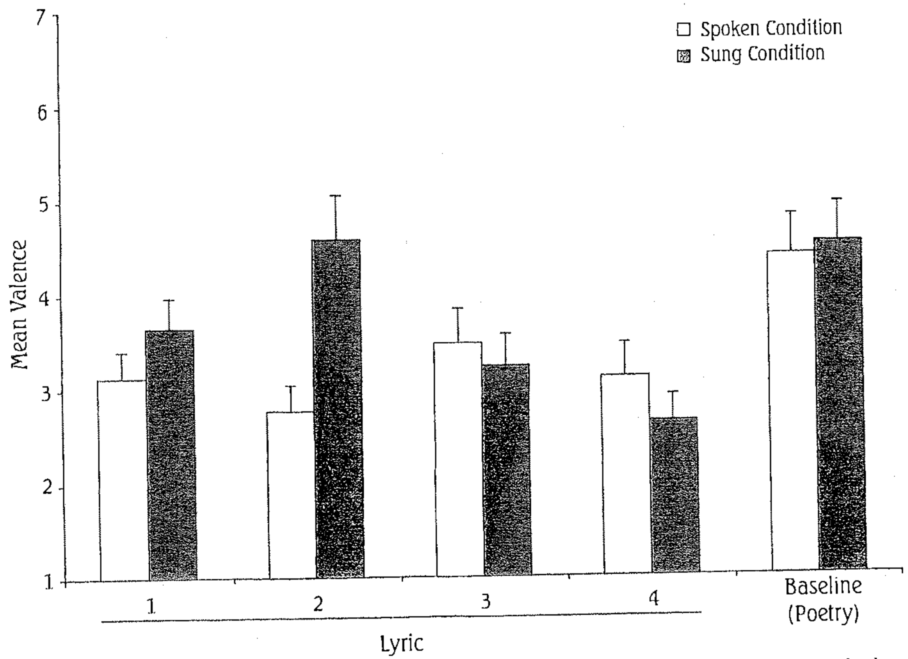

Figure 2. Valence ratings for lyrics and poetry (baseline) in the sung and spoken conditions. Standard errors are shown as error bars

Ratings were subjected to a mixed analysis of variance. The within-subjects variable was *lyrics* with four levels. The between-subjects variable was *presentation mode* with two levels (sung or spoken). The effects of lyrics was not significant, F (3, 96) = 1.67, *n.s.*; that is, the four lyrics did not differ significantly from each other with regard to perceived valence. There was a marginal effect of presentation mode such that sung lyrics were on the whole perceived to be slightly happier than spoken lyrics, F (1, 32) = 3.48, p = .07.

There was also a significant interaction between lyrics and presentation mode, F (3, 96) = 4.15, p < .01. As seen in Figure 2, whereas some lyrics were judged to be more positive when accompanied by music, other lyrics were judged to be less positive when accompanied by music. Paul Simon's Kodachrome is associated with the largest difference between the emotional message conveyed by lyrics & music (for Tukey's HSD, p<.05). The emotional message conveyed by the lyrics is generally negative, suggesting that there are limits on what can be achieved in life. This negative tone is reflected in the mean rating for the spoken lyrics (the lowest rating among the four spoken lyrics). In contrast, the music that accompanies these lyrics has a fast tempo (132 bpm), is in a major mode, has a syncopated rhythm, and is timbrally bright, thereby conveying a highly positive emotional message.

Experiment 2: Judgements of meaningfulness

Experiment 1 illustrated that music can sometimes affect a listener's interpretation of the emotional connotation of verbal information. Experiment 2 was conducted to examine whether music can also influence the perceived meaningfulness of accompanying verbal information.

Method

Participants. Seventy participants were recruited from the University of Toronto community. Participants were assigned to one of four lyric presentation conditions: 1) spoken / unfamiliar, 2) sung / unfamiliar, 3) spoken / familiar, and 4) sung / familiar. The age and musical training of participants was well matched across conditions. Participants in the spoken / unfamiliar group included 11 females and 6 males with a mean age of 19.8 years (SD = 2.49) and an average of 2.97 years of musical training (SD = 3.68). Participants in the sung / unfamiliar group included 9 females and 9 males with a mean age of 19.4 years (SD = 1.17) and an average of 1.44 years of musical training (SD = 2.58). Participants in the spoken / familiar group included 11 females and 6 males with a mean age of 20.0 years (SD = 1.37) and an average of 2.19 years of musical training (SD = 3.66). Participants in the sung / familiar group included 16 females and 2 males with a mean age of 19.8 years (SD = 1.64) and an average of 2.5 years of musical training (SD = 4.41).

Stimuli. Lyrics were selected, performed and recorded in the same manner as described in Experiment 1. Poetry was identical to that described in Experiment 1. In addition to the four pop songs used in Experiment 1, Experiment 2 included four additional songs from the same musical genre: 1) "Here Comes the Sun" by the Beatles (George Harrison, 1969); 2) "Hotel California" by the Eagles (Don Henley & Don Felder, 1976); 3) "Wild World" by Cat Stevens (1970); and 4) "Yesterday" by the Beatles (John Lennon & Paul McCartney, 1965). Unlike the songs used in Experiment 1, these other songs were familiar to listeners. Lyrics selected from the four songs used in Experiment 1 were referred to as unfamiliar and lyrics selected from the new songs introduced in Experiment 2 were referred to as familiar.

Procedure. Listeners provided ratings of meaningfulness. Judgments of meaningfulness were overall assessments of the significance of the lyrics, for example, whether the ideas presented are informative, artful, novel, generate strong and multiple associations, and are persuasive. Ratings were made on a 7-point scale where a rating of "1" represented not very meaningful and a rating of "7" represented very meaningful.

After making ratings of meaningfulness, participants in the sung / familiar and sung / unfamiliar conditions were asked to rate the familiarity of each song (i.e., familiarity prior to arriving at the test session). Familiarity ratings were made on a scale from 1 to 4, where a rating of 1 meant "I've never heard this song before," a rating of 2 meant, "I think I might have heard it once or twice before," a rating of 3 meant "I am somewhat familiar with this song," and a rating of 4 meant "I am very familiar with this song." Mean ratings for familiar songs (M = 3.13, SE = 0.19) were much higher than mean ratings for unfamiliar songs (M = 1.51, SE = 0.19), F (1, 34) = 36.79, p < .0001.

Results and Discussion

The primary analysis involved a 2 x 2 x 4 mixed design ANOVA with *familiarity* (familiar or unfamiliar) and *presentation mode* (sung with music or spoken) as between-subjects variables and *Lyrics* (4 levels) as the within subject variable. We were interested in whether ratings of lyrics when they were sung with music were higher than ratings of lyrics when they were spoken. Figure 3 illustrates the effect of music on the perceived meaningfulness of lyrics in familiar and unfamiliar songs. There was a main effect of familiarity on meaningfulness, F (1, 66) = 15.34, $p < .001$, with lyrics from familiar songs judged to be more meaningful (M = 4.05, SE = .27) than lyrics from unfamiliar songs (M = 3.02, SE = .27). The main effect of presentation mode was not significant, but there was a significant interaction between familiarity and presentation mode, F (1, 66) = 4.84, $p < .05$. As seen in Figure 3, music enhanced the perceived meaningfulness of lyrics only for familiar songs. Separate analyses of data for familiar and unfamiliar conditions confirmed that there was a significant effect of presentation mode for familiar songs, F (1, 33) = 6.08, $p < .05$, but not for unfamiliar songs, F (1, 33) < 1.0, ns.

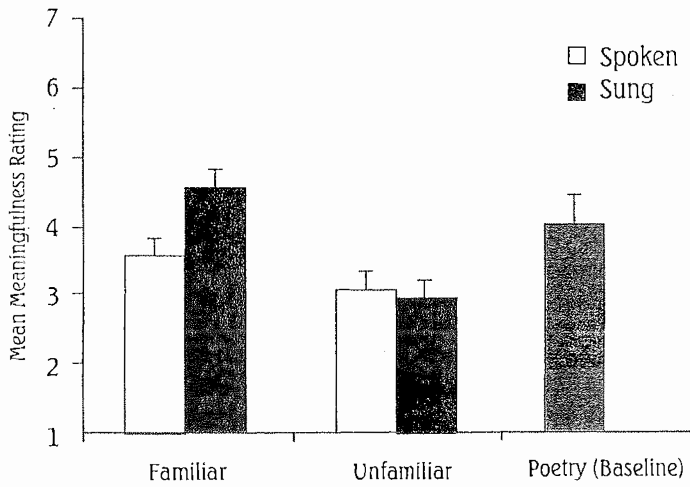

Figure 3. Meaningfulness ratings for lyrics and baseline poetry in all conditions of Experiment 2. Standard errors are shown as error bars

One interpretation of this interaction is that listeners had personal associations with the familiar music, for example, a particular summer vacation or party. Such personal associations with the music may have been conflated with the meaning of the lyrics themselves. A second possibility is that repeated exposure to music caused the music and lyrics to be represented in a more tightly integrated manner, which allowed greater cross-talk or mutual influence between the two media. A third possibility is that the familiar music was more predictable and therefore less distracting than the unfamiliar music. That is, selective attention to the meaning of lyrics may be easier for familiar music than for unfamiliar music.

If the first hypothesis is correct – that listeners confuse personal associations with the meaning of the lyrics – then *mere exposure* to unfamiliar songs in a neutral lab condition should have no or little effect on ratings of meaningfulness. Exposing listeners to music should increase familiarity, but it should not give rise to strong personal associations with that music. If the second or third hypothesis is correct, then increasing familiarity through mere exposure to songs should lead to an increase in ratings of meaningfulness. Experiment 3 was conducted to evaluate this possibility.

Experiment 3: Effects of repeated exposure

Experiment 2 indicated that verbal information is perceived to be more meaningful when it is expressed in song than when it is merely spoken, but this effect was limited to familiar songs. Experiment 3 was conducted to evaluate whether the perceived meaningfulness of song lyrics increases following increased exposure to the song.

Method

Participants. Thirty-seven participants were recruited from the University of Toronto community. Eighteen participants rated the lyrics of the unfamiliar song set after no previous exposure (no-exposure group); nineteen different participants rated the lyrics of the unfamiliar song set after repeated exposure to the songs (repeated-exposure group). Data for the no-exposure group were the same as those described for the unfamiliar sung condition in Experiment 2.

Participants in the no-exposure group are described in Experiment 2. Participants in the repeated-exposure group were well matched to the latter group, and included 15 females and 4 males with a mean age of 19.42 years (SD = 2.34) and an average of 2.45 years of musical training (SD = 3.24).

Procedure. Listeners provided ratings of meaningfulness, as in Experiment 2. In the exposure phase, listeners read magazines or books while the unfamiliar music samples were played repeatedly in blocked random order in the background. In total, each participant was exposed to five repetitions of each song prior to the test trials.

Stimuli. The lyrics presentation was identical to that described for the unfamiliar sung condition of Experiment 2. Visually presented poetry was also presented to both groups as an anchor for their other judgments.

Results and Discussion

The primary analysis involved comparing mean ratings of lyrics by the no-exposure and repeated-exposure groups. Figure 4 compares ratings of the meaningfulness of lyrics of unfamiliar music by the no-exposure group (M = 2.94, SE = 0.24) with ratings by the repeated-exposure group (M = 3.66, SE = 0.23). The Figure illustrates that repeated exposure to the songs led to an increase in the perceived meaningfulness of the lyrics, F (1, 35) = 4.66, $p < .05$. In other words, increasing familiarity through repeated exposure enhanced the perceived significance of lyrics.

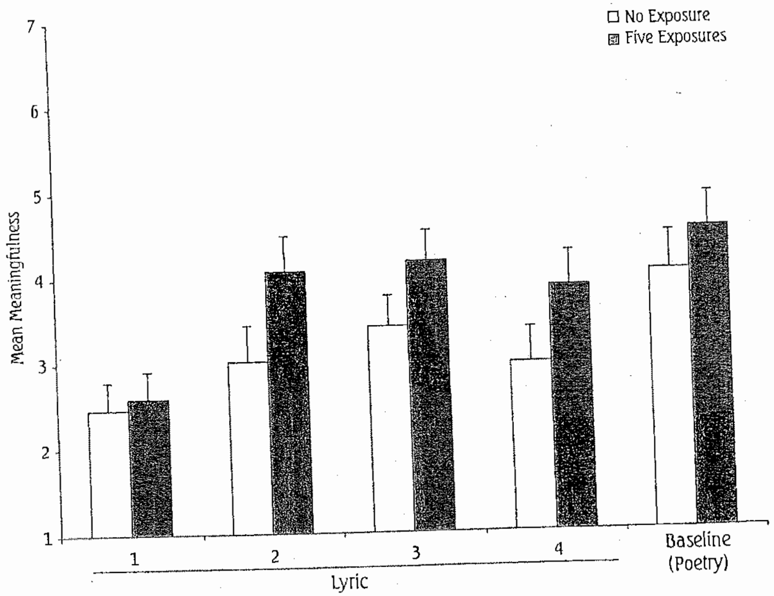

Figure 4. Ratings of the meaningfulness of the lyrics of unfamiliar music and poetry (baseline) after the first exposure and after repeated exposure. Standard errors are shown as error bars

General Discussion

The results of this study indicate that music influences the attribution of meaning and emotion to song lyrics. First, Experiment 1 indicated that a musical context can influence the affective valence attributed to accompanying lyrics. Thus, lyrics accompanying sad music are more likely to be interpreted as communicating a negative message, whereas lyrics accompanying joyful or happy music are more likely to be interpreted as communicating a positive message. Second, for familiar songs, listeners attributed greater overall significance and meaning to lyrics if they were sung with music than if they were merely spoken. Third, repeated exposure to songs increased the perceived meaningfulness of accompanying lyrics. This latter finding contradicts the view that the link between familiarity and perceived meaningfulness can be explained entirely by personal associations. For if this were the case, one would not expect repeated exposure to unfamiliar songs in laboratory conditions to influence the perceived meaningfulness of the accompanying lyrics.

Two other possibilities exist. First, repeated exposure may promote a more integrated representation of music and lyrics (Crowder, Serafine & Repp, 1990; Serafine, Crowder & Repp, 1984; Serafine, Davidson, Crowder & Repp, 1986). Stronger integration of music and lyrics, in turn, might allow greater transfer of implied significance from music to accompanying lyrics. Figure 5 illustrates the process. When songs are unfamiliar, musical and verbal textures are likely to be perceived as somewhat distinct, independent media. The degree of perceived independence will,

of course, depend on the extent to which the songwriter coordinated common elements of music and speech. On the whole, however, music and lyrics should be perceived as more distinct for unfamiliar songs than for familiar songs.

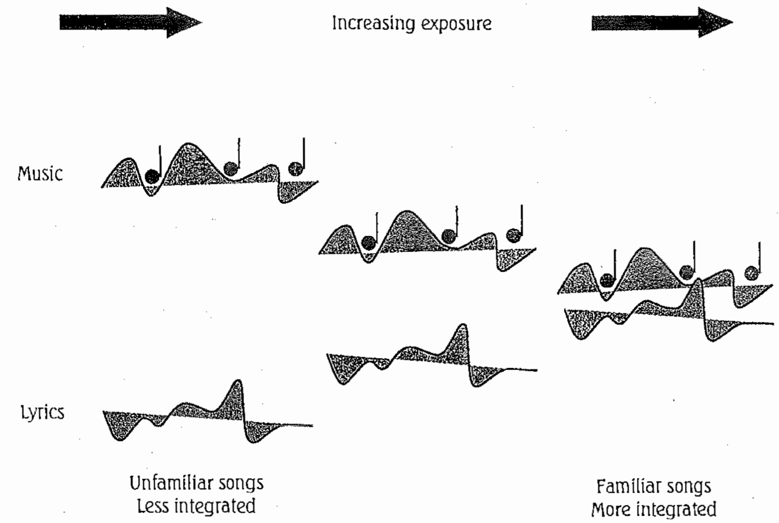

Figure 5. An illustration of the effect of repeated exposure on the degree of integration between music and lyrics, and the attribution of non-referential musical meaning to the accompanying lyrics

With repeated exposure, music and lyrics sound more and more like they "go together" and this increased integration increases the likelihood that free-floating (non referential) significance implied by the music will be attributed to the accompanying lyrics through a *trickle-down effect* (see also Figure 1). The process may be experienced in the following way: with greater familiarity, the musical texture is increasingly perceived as a fitting dramatic backdrop for the accompanying lyrics, aptly emphasizing, elaborating and clarifying those lyrics and thereby magnifying their apparent significance.

A second possibility is that ratings of meaningfulness are merely related to the amount of verbal information extracted from the lyrics. It can be difficult for listeners to understand lyrics in unfamiliar music because words are sometimes obscured in song, particularly during passages that involve a fast tempo, dense texture, or large changes in pitch. Repeated exposure to music may have allowed listeners to hear more of the words in the lyrics. Conversely, listeners may have found it easier to ignore the accompanying music when it was familiar because familiar music is more predictable and requires fewer cognitive resources than unfamiliar music. By increasing the amount of verbal information extracted from the lyrics (the "signal") and reducing the potential distraction arising from the accompanying music (the "noise"), it may have been easier for listeners to interpret the meaning of the lyrics. In other words, increases in familiarity may be associated with a higher functional signal (lyrics) to noise (music) ratio, leading to an enhanced ability to interpret the lyrics.

Although our study was not designed to disentangle these two hypotheses, the latter possibility implies that lyrics should be perceived as more meaningful when presented in isolation than when presented with potentially distracting music ("noise"). However, we observed the opposite effect in Experiment 2. Familiar lyrics were perceived as significantly more meaningful when they were sung in a musical context than when they were spoken in the absence of music. Thus, the effects of familiarity to songs on the perceived meaningfulness of accompanying lyrics cannot be explained solely in terms of the amount of verbal information extracted from the lyrics.

To conclude, music can influence whether lyrics convey a positive or negative message, it can affect the perceived meaningfulness of lyrics for familiar songs, and repeated exposure to a song increases the perceived meaningfulness of the lyrics. Such effects may occur through a trickle-down process that is enhanced through repeated exposure to the music.

REFERENCES

Benveniste, E. (1985). The semiology of language. In R. E. Innis (Ed.) *Semiotics: An Introductory Anthology*, Bloomington: Indiana University Press, 226–246.

Besson, M., Faïta, F., Peretz, I., Bonnel, A-M., Requin, J. (1998). Singing in the brain: independence of lyrics and tunes. *Psychological Science*, *9*, 494–498.

Cohen, A. (2001). Music as a source of emotion in film. In P. Juslin, J. Sloboda (Eds.) *Music and Emotion: Theory and Research* (pp. 249–274). Oxford: Oxford University Press.

Crowder, R., Serafine, M. L., Repp, B. (1990). Physical interaction and association by contiguity in memory for the words and melodies of songs. *Memory & Cognition*, *18*, 469–476.

Dabbs, J. M., Janis, I. L. (1965). Why does eating while reading facilitate opinion change? – An experimental inquiry. *Journal of Experimental Social Psychology*, *1*, 133–144.

Galizio, M., Hendrick, C. (1972). Effect of music accompaniment on attitude: The guitar as a prop for persuasion. *Journal of Applied Social Psychology*, *2*, 350–359.

Gorlée, D. L. (1997). Intercode translation: Words and music in opera. *Target*, *9*, 235–270.

Iverson, J., Rees, H., Revlin, R. (1989). The effect of music on the personal relevance of lyrics. *Psychology: A Journal of Human Behaviour*, *26*, 15–22.

Juslin, P. N., Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin*, *129*, 770–814.

Langer, S. (1957). *Philosophy in a New Key: A Study in the Symbolism of Reason, Rite, and Art*. Cambridge, MA: Harvard University Press.

Peretz, I., Radeau, M., Arguin, M. (2004). Two-way interactions between music and language: Evidence from priming recognition of tune and lyrics in familiar songs. *Memory & Cognition*, *32*, 142–152.

Samson, S., Zatorre, R. J. (1991). Recognition memory for text and melody of songs after unilateral temporal-lobe lesion: evidence for dual encoding. *Journal of Experimental Psychology: Learning, Memory & Cognition*, *17*, 793–804.

Serafine M. L., Crowder R. G., Repp, B. H. (1984). Integration of melody and text in memory for songs, *Cognition*, *16*, 285–303.

Serafine, M. L., Davidson, J., Crowder, R., Repp, B. (1986). On the nature of melody-text integration in memory for songs. *Journal of Memory and Language*, *25*, 123–135.

Stratton, V. N., Zalanowski, A. H. (1994). Affective impact of music vs. lyrics. *Empirical Studies of the Arts*, *12*, 173–184.

Thompson, W. F., Russo, F. A., Sinclair, D. (1994). Effects of underscoring on the perception of closure and intensity in film excerpts. *Psychomusicology*, *13*, 9–27.

Polskie Forum Psychologiczne, 2004, tom 9, numer 1, s. 60–69

ELECTRODERMAL RESPONSES TO DISSONANT AND CONSONANT MUSIC

Stéphanie Khalfa
Laboratoire de Neurophysiologie et Neuropsychologie
Université de la Mediterranée, France

Isabelle Peretz
Department of Psychology, University of Montreal, Canada

Owing to the close relation between music and emotions, autonomic responses to music have been often studied. In particular, we have previously shown that musical excerpts can elicit skin conductance responses (SCRs) depending on the level of arousal of the emotion as expressed by the music. In the present study, we extend the SCR sensitivity to valence, as conveyed by musical consonance/dissonance. Dissonant music is generally considered as unpleasant whereas consonant music is generally pleasant. Hence we tested the SCR to consonant and dissonant excerpts in university students. In order to assess whether the expected SCR sensitivity to consonance/dissonance is linked to valence rather than to differences in musical characteristics, rare individuals having no perceptual deficit but having difficulties in experiencing pleasure have been included in the experiment. These subjects who have no psychiatric disorders are known as non-clinical physical anhedonics. Anhedonics are expected to perceive the consonance/dissonance distinction as well as controls but to exhibit smaller SCR in comparison to normal listeners. The results are consistent with these predictions in showing larger SCRs to consonant than to dissonant musical excerpts, particularly so in controls. Thus, the results are in accordance with the notion that SCR is sensitive to valence, and not only to arousal in musical settings.

Introduction

Emotion is an integral part of music experience. Emotional responses appear to be the primary reason why people listen to music (Panksepp, 1995; Sloboda, 1991). Indeed, some emotions elicited by music seem able to modulate physiological responses (e.g., Krumhansl, 1997). In a recent study of ours (Khalfa et al., 2002), a measure of the sympathetic nervous system, the skin conductance response (SCR), was found to depend on the emotion expressed by the music. The SCRs were larger for the high-arousal stimuli (happy and scary music) than for the low-arousal stimuli (sad and peaceful music). In the present study, we focus on another emotional dimension, that of valence.

The valence (or pleasantness) of music is strongly related to its degree of consonance. Listeners often judge consonant music as pleasant whereas they find dissonant music unpleasant. Dissonance is related to the ratio in pitch frequencies between successive tones in a melody, or between simultaneous tones in a chord. When this ratio is complex (as in a minor second), the musical tones are considered dissonant (Schellenberg & Trehub, 1994). The preference for consonance over dissonance arises from mechanical constraints in the inner ear (often referred to as sensory dissonance; Plomp & Levelt, 1965), from psychological interpretation (referred to musical dissonance; Krumhansl, 1990), and from listener's experience (Cazden, 1945). This preference for consonance over dissonance is already present in infants (Trainor & Heinmiller, 1998; Zentner & Kagan, 1998) suggesting that this phenomenon may be innate (Trainor & Heinmiller, 1998).

At a neuroanatomical level, it has been shown in a PET study that dissonant and consonant musical excerpts may recruit brain regions such as prefrontal cortices and the subcallosal region, which have been previously associated with pleasant and unpleasant emotional states (Blood et al., 1999). This result confirms the ability of dissonant/consonant excerpts to evoke emotional states.

The degree of pleasantness conveyed by dissonant and consonant melodies can be assessed in subjects' reports. Its physiological impact is less easy to evaluate. In order to explore further the autonomic responses related to the emotional reaction to music, we assessed here whether SCRs are sensitive to the dissonance/consonance dimension in music (including both sensory and musical dissonance), that is to musical valence.

SCR represents the changes in electrical conductivity of a person's hand skin related to the amount of sweat within the eccrine sweat gland. The more sweat, the greater the skin conductance (Boucsein, 1992). The phasic changes of skin conductance, occurring between 1 and 4 sec after stimulus onset, are especially sensitive to emotions (Katkin, 1965). The electrodermal activity is controlled by the sympathetic branch of the autonomic nervous system, through the release of acetylcholine (Millington & Wilkinson, 1983). Via the sympathetic system, the SCR is under the control of cerebral structures that are involved in emotion processing, such as the prefrontal cortices (Papousek & Schulter, 2001), and the orbito-frontal cortex (Tranel & Damasio, 1994) that is known to participate in valence judgments (e.g., Anderson, 2003; Blood et al., 1999; Tranel & Damasio, 1994). It is because of these neural connections that we expect SCRs to vary according to musical valence.

A further demonstration that SRCs are elicited by emotions (including valence) is to show that these autonomic responses are attenuated in people who do not experience pleasure or emotions in general. To this aim, we tested rare individuals having no musical deficit but a specific disorder in experiencing emotions (Ribot, 1896). Those subjects are known as non-clinical physical anhedonics. Physical anhedonia is a failure in experiencing sensorial pleasure. In pathological cases, it may be a symptom of psychiatric disorders such as depression (Klein, 1974) and schizophrenia (Bernstein, 1987). In healthy subjects, anhedonia may appear as a personality trait identified using the Physical Anhedonia Scale (PAS) (Chapman et al., 1976). This scale measures the judgments of pleasantness for diverse sensory stimulations. Items include commonly pleasurable sights, sounds, smells, tastes, and tactile sensations. There is compelling evidence that anhedonics differ from controls in processing pleasant stimuli, and that self-described

Correspondence concerning this article should be addressed to Stéphanie Khalfa, e-mail: skhalfa@skhalfa.com

The research was supported by a grant from the Natural Science and Engineering Research Council to the second author, and by fellowships from the education government of Quebec, and from the Canadian Institutes of Health Research to the first author. We thank Aileen McGonigal for helpful comments regarding English editing and Simone Dalla Bella for insightful comments.

anhedonia is not a self-report response bias (Miller & Yee, 1994). This difficulty in experiencing pleasure is not due to perceptual disorders, such as deafness, hyperacusia or other similar deficits (Fernandes & Miller, 1995).

Given their impairment in experiencing pleasure, anhedonics should be less responsive to the emotional difference between consonant and dissonant music. That is, anhedonics' SRCs should be less sensitive to the pleasantness of the music than normal control individuals. To test this hypothesis, pleasantness judgment scores as well as the SCR were collected in both controls and healthy anhedonics while listening to pleasant and unpleasant musical stimuli, that were the consonant and dissonant versions of the same excerpts. The excerpts were taken from Western classical (consonant) music and modified so as to create dissonance by shifting the pitches of the leading voice by one semitone either upward or downward relative to its accompaniment, as used in Peretz et al. (2001).

Method

Participants

Participants were selected among 290 university science students, according to their score obtained on the Physical Anhedonia Scale (PAS) (Chapman et al., 1976). This scale has been validated in French (Loas, 1993), the native language of the participants. This questionnaire includes 61 items that must be rated as false or true. A response reflecting anhedonia gives 1 point and 0 otherwise. Hence, the higher the score, the more likely the participant suffers from physical anhedonia. According to the literature, a score of 29 and more on the PAS is sufficient to classify subjects as anhedonics (Loas, 1995; Lutzenberger et al., 1983). Control subjects were then selected if their scores were inferior to 10. Eleven anhedonic subjects accepted to participate in the experiment. However, despite the strong test-retest reliability of the PAS scores (r=0.94, $d.f.$=22, p<0.001), the score of one subject changed from 29 on the day of subjects' selection, to 22 on the day of the experiment. Since the mean scores of the two subjects' groups are calculated from the data obtained on the day of the experiment, this subject has been included as anhedonic in the study.

Four out of the 26 participants were excluded because they did not show detectable electrodermal responses (a minimum of 0.1 microSiemens (μS)) to any one stimulus. The remaining subjects were aged from 19 to 27 years. Eleven control subjects (7 males and 4 females) with a mean age of 22.9 years (SD= 2.5) and a mean anhedonic score of 4.5 (SD= 2.8) participated to the experiment as well as eleven anhedonics, all males, with a mean age of 22.7 years (SD= 2.7) and mean anhedonic score of 34.8 (SD= 5.72). None of subjects were taking medication.

The experiment has been approved by the institutional Ethics Committee of the University Institute of Geriatric Care of Montreal and was conducted with the fully informed consent of each subject.

Material

Twenty musical excerpts including 10 consonant and 10 dissonant versions of the same excerpts were used. They were taken from the set employed by Peretz et al. (2001) and shortened to the initial 7-sec duration. They were computer-generated via a synthesizer and delivered with

a piano timbre. Auditory examples are available at: www.fas.umontreal.ca/psy/iperetz.html. To ensure that the pleasant and unpleasant excerpts do not differ in level of arousal, 9 pilot university subjects (mean age= 21.7 years, SD= 2.4) judged each of the 20 musical clips, on a 10-point rating scale from 1 (relaxing) to 10 (stimulating). A paired t-test performed on the averaged score obtained for each excerpt showed no difference (t=0.72, $d.f.$=14, p=0.5) between the arousal ratings for the consonant (mean= 5.63, SD= 0.61) and the dissonant excerpts (mean= 6.08, SD= 0.38).

Electrodermal measures

Skin conductance was recorded from the right hand, with a pair of Ag-AgCl electrodes (0.8 cm diameter) attached on the palmar surface of the medial phalange of the index and middle fingers. These electrodes were filled with a 0.050 molar NaCl paste. All recording procedures followed the recommendations set forth by Fowles et al. (1981). Data acquisition, acoustic stimuli presentation and quantification were performed using a Coulbourn S71-23 Coupler, and a customized program created with InstEP Systems v3.3.

Procedure

The experiment was performed in a quiet room. After skin conductance electrodes were attached, the subject was instructed to rest quietly during the ensuing 5-min rest period.

Subjects were seated comfortably, asked to avoid moving as much as possible, and to listen attentively to the 20 musical selections presented through headphones. The excerpts were presented binaurally in an identical pseudo-randomized order for each participant, through MDR-v200 Sony headphones. Their task was to verbally determine the valence of the stimulus, on a 10-point scale ranging from 1 (désagréable/unpleasant) to 10 (agréable/pleasant). SCRs were recorded throughout the task. A microphone system allowed the experimenter to hear and then write the subjects' verbal answers.

Each stimulus presentation was followed by an interval of at least 20 seconds to allow for skin conductance recovery to baseline before the next stimulus was presented. This inter-trial interval lasted as long as necessary for skin conductance to regain stability, generally not more than 30 seconds. After SCR recording, subjects had to fill up the Physical Anhedonia Scale (Loas, 1993). The whole session lasted about 40 minutes.

Data quantification and analysis

The first three stimuli were considered as examples and were excluded from the analysis as they may have evoked strong orienting component in response to novelty rather than to consonance or dissonance. SCRs and ratings for each excerpt were only considered if control subjects gave a rating superior to 5.5 for consonant stimuli and inferior or equal to 5.5 for dissonant stimuli. From the 20 excerpts, only 14 matched these criteria. These 14 stimuli comprised 7 dissonant and 7 consonant stimuli. Moreover, given that the SCR decreases with time in most subjects due to habituation (Dawson, Schell & Filion, 1990), to compare results for the two categories of musical excerpts, it is necessary to take into account a putative order effect. As the average rank orders of presentation for the 7 pleasant and 7 unpleasant stimuli are almost similar, with 12.3 and 13 respectively, the habituation effect was controlled.

SCRs were considered in the analyses when they occurred within a latency window of 1-4 s following stimulus onset (Venables & Christie, 1980). Maximal amplitude of the SCR elicited by each stimulus was the electrodermal parameter measured. SCR magnitudes and valence ratings for the 7 melodies of each category were averaged for each participant. Statistical analyses on SCRs were performed using non-parametric tests such as the Mann-Whitney, since the Kolmogorov-Smirnov normality test failed to reach significance. In contrast, ratings were normally distributed. Hence, mean ratings obtained for each excerpt in term of pleasantness were analyzed with a 2 X 2 analysis of variance (ANOVA) using Population (anhedonics vs controls) as between-subjects factor and Category (Dissonant vs Consonant) as within-subjects factor. Pairwise comparisons (Bonferroni test) were performed when the ANOVA tests reached significance (p<0.05).

Results

Valence ratings

The ANOVA yielded a significant main effect of Valence ($F(1, 20)$= 38.6, p<0.001) and Group ($F(1, 20)$= 916, p<0.001) on pleasantness ratings, and the interaction between these two factors almost reached significance ($F(3, 40)$= 3.9, p=0.06). As can be seen in Figure 1, consonant excerpts were judged as more pleasant than dissonant ones, and controls judged all excerpts as more pleasant than anhedonics did. This group difference tends to be especially pronounced for consonant excerpts.

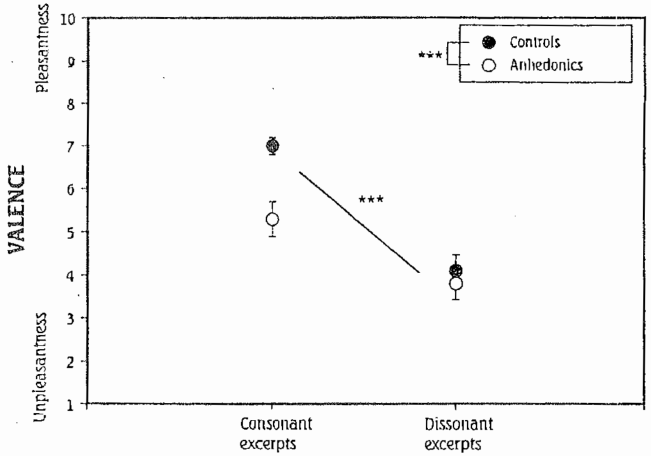

Figure 1. Means and standard error bars of pleasantness judgements for consonant and dissonant musical excerpts, in both anhedonic and control subjects. *** p<0.001

Valence and group effect on SCR

The representation of SCR in Figure 2 shows that the SCRs average magnitudes corresponding to the consonant clips were lower in anhedonics than in controls (U=30, p<0.05) whereas no significant difference was demonstrated between the two populations for the dissonant melodies. It is also shown in this figure that SCRs magnitudes were larger for consonant than for dissonant excerpts, in control subjects (z=-2.3, p<0.05) but this difference did not reach significance in anhedonic subjects.

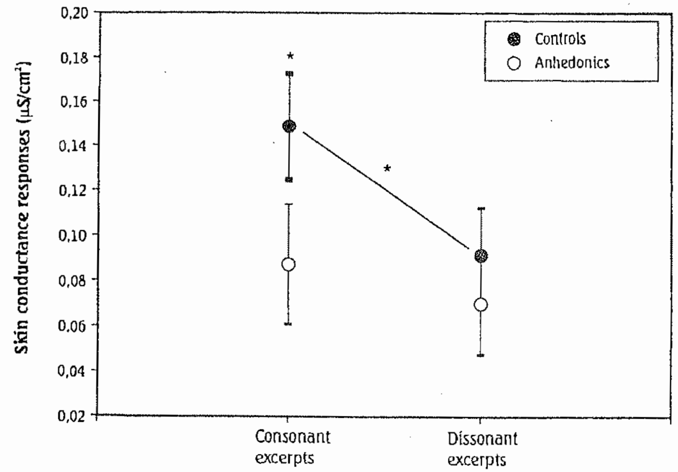

Figure 2. Means and standard error bars of SCR magnitudes while listening to consonant and dissonant musical excerpts, in both anhedonic and control subjects. * p<0.05

Relationships between SCRs, pleasantness judgments and PAS scores

Pleasantness judgments and SCR magnitudes were significantly correlated across the two groups (r=0.36, $d.f.$=44, p<.05). This correlation is not due to extreme SCR values (superior to 0.30 $\mu S/cm^2$); when these are discarded, the correlation is still significant (r=0.34, $d.f.$=42, p<.05). The higher the rating score (i.e. the more pleasant the excerpt), the greater the SCR magnitude (see Figure 3). No significant correlation was found between the SCR magnitudes and the PAS scores. However, as can be seen in Figure 4, when considering only the responses obtained for the consonant excerpts, pleasantness judgments did correlate with PAS scores (r=-0.75, p<0.001, N=22). That is, the higher the PAS score, the lower the valence rating for the consonant excerpts.

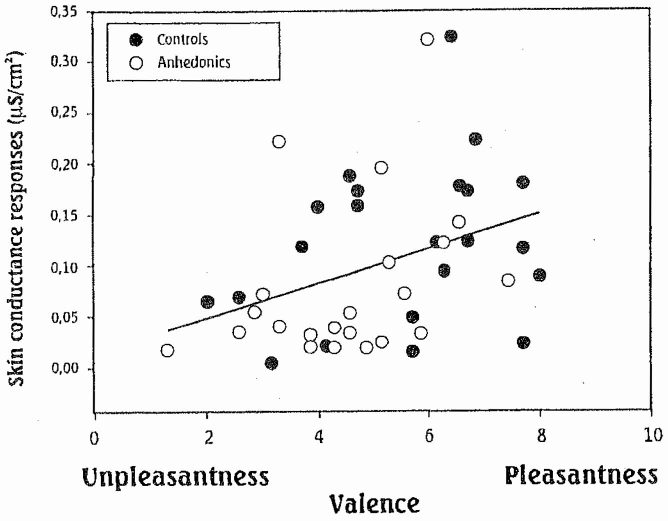

Figure 3. Relationships between valence ratings and SCRs

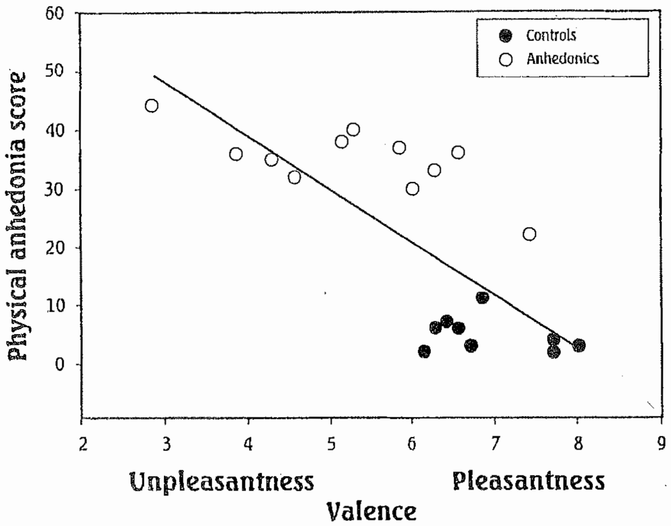

Figure 4. Relationships between physical anhedonia scores and valence ratings for the consonant excerpts

Discussion

The first important result of this experiment is the SCR sensitivity to differences between consonant and dissonant music. This distinction cannot be attributed to arousal differences between consonant and dissonant excerpts, since we controlled for this factor. Rather, the difference arises from the differential valence of consonant and dissonant music, as shown by the pleasantness judgments given by normal subjects. Therefore, SCR amplitude when listening to consonant as compared to dissonant music seems to reflect a valence effect. This finding contrasts with the literature where little support is given in favor of a link between SCR and valence (Bradley & Lang, 2000). One possibility is that the arousal dimension is generally more determinant than valence to elicit SCRs. However, when arousal does not differentiate the stimuli, as it is the case here, then valence may modulate SCRs. This is all the more likely since pleasantness judgments were found to be positively correlated with SCR magnitudes in the two groups of listeners tested here.

In support of the notion that SCRs are sensitive to valence is the observation of lower SCRs to consonant excerpts in anhedonics compared to controls. This result is in accordance with a previous study on anhedonia showing a decrease in SCR to pleasant stimuli (Rockstroh et al., 1982) as well as another one showing a trend for SCR to be similar or lower in anhedonia than in normal subjects (see review of Fernandes & Miller, 1995). As expressed in their ratings of valence, anhedonics tend to find consonant excerpts considerably less pleasant than controls did. In contrast, both anhedonics and controls judged the dissonant music to be unpleasant. Thus, when pleasure is somewhat attenuated as in anhedonia, SCRs are also decreased. The SCR sensitivity to dissonance/consonance would be the consequence of the affective appraisal of the music. This hypothesis is reinforced by the existence of a positive correlation between valence judgments and SCRs. However, the correlation only explains 13% of the variance. Other mechanisms may therefore be implicated in the SCR such as lower level emotional processing.

To conclude, our results support the idea that measuring SCR to musical excerpts is an appropriate means to assess emotional processing elicited by music. In the present experiment, this autonomic measure proved to vary according to the emotional content of the excerpts, especially the valence dimension which depends on the musical characteristics of dissonance/consonance. Further experiments on musical emotions would benefit from the use of SCR measures in conjunction with brain imaging techniques to deepen our understanding of emotion processing.

REFERENCES

Anderson, A. K., Christoff, K., Stappen, I., Panitz, D., Ghahremani, D. G., Glover, G., Gabrieli, J. D., Sobel., N. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, *6*(2), 196–202.

Bernstein, A. S. (1987). Orienting response research in schizophrenia: where we have gone and where we might go. *Schizophrenia Bulletin*, *13*, 623–641.

Blood, A. J., Zatorre, R. J., Bemudez, P., Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nature Neuroscience*, *2*(4), 382–387.

Boucsein, W. (1992). *Electrodermal activity*. New-York and London: Plenum Press.

Bradley, M. M., Lang, P. J. (2000). Affective reactions to acoustic stimuli. *Psychophysiology, 37*, 204–215.

Cazden, N. C. (1945). Musical consonance and dissonance: a cultural criterion. *The Journal of Aesthetics & Art Criticism, 4*, 3–11.

Chapman, L. J., Chapman, J. P., Raulin, M. L. (1976). Scales for physical and social anhedonia. *Journal of Abnormal Psychology, 85*, 374–382.

Dawson, M. E., Schell, A. M., Filion, D. L (1990). The electrodermal system. In J. T. Cacioppo, L. G. Tassinary (Eds.) *Principles of psychophysiology* (pp. 295–324). Cambridge University Press.

Fernandes L. O. L., Miller G. A. (1995). Compromised performance and abnormal psychophysiology associated with the Wisconsin Scales of Psychosis Proneness. In G.A. Miller (Eds.) *The behavioral high-risk paradigm in Psychopathology* (pp. 47–87). Springer.

Fowles, D. C., Christie, M. J., Grings, W. W., Lykken, D. T., Venables, P. H. (1981). Publication recommendations for electrodermal measurements. *Psychophysiology, 18*(3), 232–239.

Katkin, E. S. (1965). Relationship between manifest anxiety and two indices of autonomic response to stress. *Journal of Personality & Social Psychology, 2*, 324–333.

Khalfa, S., Peretz, I., Blondin, J-P., Robert, M. (2002). Event-related skin conductance responses to musical emotions. *Neuroscience Letters, 328*, 145–149.

Klein, D. F. (1974). Endogenomorphic depression: a conceptual and terminological revision. *Archives of General Psychiatry, 31*, 447–454.

Krumhansl, C. L. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology, 51*(4), 336–352.

Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Loas, G. (1993) Adaptation et validation française de l'échelle d'anhédonie physique de Chapman et Chapman (PAS). *L'Encéphale, XIX*, 639–644.

Loas, G. (1995). L'évaluation de l'anhédonie en psychopathologie. In J. D.Gulefi, V. Gaillac, R. Dardennes (Eds.) *Psychopathologie Quantitative*. Paris: Masson.

Lutzenberger, W., Birbaumer, N., Rockstroh, B., Elbert, T. (1983). Evaluation of contingencies and conditional probabilities. A psychological approach to anhedonia. *Archives of Psychiatry Nervenkr, 233*, 471–488.

Miller, G. A., Yee, C. M. (1994). Risk for severe psychopathology: psychometric screening and psychophysiological assessment. In J. R. Jennings, P. K. Ackles, M. G. H. Coles (Eds.) *Advances in Psychophysiology* (vol. 5, pp. 1–54). Jessica Kingsley Publishers Ltd.

Millington, P. F., Wilkinson, R. (1983). *Skin*. Cambridge: University Press.

Panksepp, J. (1995). The emotional sources of "chills" induced by music. *Music Perception, 13*, 171–208.

Papousek, I., Schulter, G. (2001). Associations between EEG asymmetries and electrodermal lability in low vs. high depressive and anxious normal individuals. *International Journal of Psychophysiology, 41*(2), 105–117.

Peretz, I., Blood, A., Penhune, V., Zatorre, R. (2001). Cortical deafness to dissonance. *Brain, 124*, 928–940.

Plomp, R., Levelt, W. J. M. (1965). Tonal consonance and critical bandwith. *Journal of the Acoustical Society of America, 38*, 518–560.

Ribot, T. (1896). *La psychologie des sentiments*. Vol. 1, Alcan 1896. 16th edition (1939), PUF, Bibliothèque de Philosophie Contemporaine.

Rockstroh, B., Elbert, T., Birbaumer, N., Lutzenberger, W. (1982). *Slow brain and behavior*. Baltimore: Urban & Schwarzenberg.

Schellenberg, E. G., Trehub, S. E. (1994). Frequency ratios and the perception of tone patterns. *Psychonomic Bulletin & Review, 1*, 191–201.

Sloboda, J. A. (1991). Musical structure and emotional response: Some empirical findings. *Psychology of Music, 19*, 110–120.

Trainor, L. J., Heinmiller, B. M. (1998). The development of evaluative responses to music: Infants prefer to listen to consonance over dissonance. *Infant Behavior & Development, 21*(1), 77–88.

Tranel, D., Damasio, H. (1994). Neuroanatomical correlates of electrodermal skin conductance responses. *Psychophysiology, 31*, 427–438.

Venables, P. H., Christie, M. J. (1980). Electrodermal activity. In I. Martin, P. H. Venables (Eds.) *Techniques in psychophysiology* (pp. 3–67). New York, NY: John Wiley & Sons, Ltd.

Zentner, M. R., Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behavior & Development, 21*(3), 483–492.

Polskie Forum Psychologiczne, 2004, tom 9, numer 1, s. 70-73

KOMENTARZE

AKTUALNY PEJZAŻ PSYCHOLOGII MUZYKI W POLSCE

Ilona Laskowska, Maciej Michalak
Instytut Psychologii
Akademia Bydgoska im. Kazimierza Wielkiego

Psychologia muzyki w Polsce ma już dosyć długą tradycję. Od samego początku (okres międzywojenny) rozwijała się niezależnie od muzykologii. W przeciwieństwie do badaczy historii muzyki czy etnomuzykologii pierwsi psychologowie muzyki koncentrowali swoją uwagę na zagadnieniach dotyczących zdolności muzycznych i psychologicznych wyznaczników determinujących umiejętność słuchania i przeżywania muzyki. Obszar zainteresowań badawczych psychologów muzyki w decydujący sposób rzutował na przedmiot badań i dobór metod badawczych. W polskiej psychologii muzyki generalnie dominował eksperymentalny nurt badawczy, którego korzenie sięgają psychofizyki Fechnera (1860, 1870) i psychoakustyki Hemholtza (1863) (Manturzewska, Kotarska, 1990). Uwagę badaczy pracujących w tym paradygmacie zajmowały zwłaszcza problemy związane z percepcją poszczególnych właściwości dźwięku i struktur akustyczno-muzycznych, a także zagadnienia dotyczące słuchu muzycznego (np. słuch absolutny), uzdolnień i osiągnięć muzycznych. Badanie uzdolnień muzycznych pociągało za sobą konieczność poszukiwania czynników warunkujących zróżnicowanie indywidualne w tym zakresie. Progresywny rozwój badań w ramach nurtu eksperymentalnego doprowadził do wydzielenia dwóch niezależnych dyscyplin naukowych: muzycznej psychometrii i psychoakustyki muzycznej (ibidem). Celem tej pierwszej jest konstruowanie i standaryzowanie, a w niektórych przypadkach adaptowanie do warunków polskich, obiektywnych i rzetelnych testów uzdolnień i osiągnięć muzycznych, drugiej natomiast prowadzenie między innymi badań w zakresie percepcji dźwięków, słuchu absolutnego i metodyki ocen słuchowych.

Do grona badaczy, którzy wywarli największy wpływ na kształt współczesnej psychologii muzyki w Polsce, zaliczyć należy: S. Szumana, M. Manturzewską, J. Horbulewicza, J. Wierszyłowskiego, H. Kotarską, A. Rakowskiego i K. Miklaszewskiego (Miklaszewski, Meyer-Borysewicz, 1991).

Adres do korespondencji: Ilona Laskowska, e-mail: morfeusz@ab-byd.edu.pl

Celem niniejszego opracowania nie jest szczegółowe omówienie wszystkich dokonań psychologii muzyki w Polsce, dlatego też skoncentrujemy się na jej najbardziej znaczących osiągnięciach, które stały się źródłem inspiracji dla współczesnych badaczy. W gronie osób szczególnie zasłużonych podkreślić należy działalność Marii Manturzewskiej, inicjatorki wielu szeroko zakrojonych programów badawczych poświęconych problematyce rozwoju muzycznego człowieka, różnic indywidualnych w tym zakresie, uzdolnień muzycznych oraz poszukiwania wyznaczników powodzenia w nauce muzyki i zawodzie muzyka (Manturzewska, 1966, 1969, 1973, 1974, 1987). Pod koniec lat pięćdziesiątych pod kierownictwem M. Manturzewskiej powstał międzyuczelniany zespół do badań nad zdolnościami muzycznymi i metodami ich pomiaru. Wynikiem działań tego zespołu było powołanie pierwszych poradni psychologicznych dla szkół muzycznych w Warszawie i Poznaniu oraz stworzenie narzędzia przydatnego podczas rekrutacji kandydatów do szkół muzycznych. Ponadto M. Manturzewska znacząco przyczyniła się do wdrożenia w warunkach polskich uznanych na świecie testów zdolności osiągnięć muzycznych. Działalność M. Manturzewskiej i jej zespołu lokuje się zwłaszcza w obszarze psychometrii muzycznej. Warto tu jednak nadmienić, że w swoich licznych badaniach niejednokrotnie sięgała do perspektywy rozwojowej.

Kontynuatorką idei Marii Manturzewskiej jest Barbara Kamińska, która obecnie kieruje Międzywydziałową Katedrą Psychologii Muzyki w Akademii Muzycznej im. F. Chopina w Warszawie. W latach 80. pracowała nad adaptacją polskiej wersji australijskiego testu osiągnięć muzycznych dla kandydatów na studia muzyczne (Kamińska, 1985b). Jednocześnie zajmowała się przystosowaniem do warunków polskich dwóch testów muzycznych dla małych dzieci, pierwszy autorstwa E. Gordona, drugi A. Zenatti. Od lat 90. prowadzi badania nad rozwojem oraz uwarunkowaniami psychologicznymi poprawnej intonacji w śpiewie dzieci (Kamińska, 1990). Ostatnie prace B. Kamińskiej we współpracy z M. Manturzewską dotyczą rozwoju muzycznego człowieka jako przedmiotu badań psychologicznych i pedagogicznych (Manturzewska, Kamińska, 1996). W obszarze ich zainteresowań mieści się również problematyka dotycząca postaw i oczekiwań młodzieży w odniesieniu do muzyki w szkole (Manturzewska, Kamińska, 1999).

Wśród przedstawicieli drugiego ze wspomnianych przez nas nurtów badawczych, na podkreślenie zasługują dokonania Andrzeja Rakowskiego. Jako jeden z pierwszych badaczy inicjował i projektował w Polsce badania w zakresie psychoakustyki. Jego zainteresowania badawcze koncentrowały się wokół percepcji wysokości dźwięków, organizacji materiału wysokościowego muzyki, wyznaczenia kryterium słuchu absolutnego. Późniejsze badania prowadzone pod kierownictwem A. Rakowskiego dotyczyły następujących problemów badawczych: 1) kształtowania dźwięków podczas wykonywania muzyki (por. Rakowski, Kiuila, Rogowski, 2002); 2) wysokości dźwięku w percepcji muzyki (por. Miśkiewicz, Rakowski, Rogoziński, Kocańda, 2002); 3) barwy i brzmienia dźwięku (por. Miśkiewicz, 2002); 4) koincydencji wrażeń muzycznych i wrażeń wzrokowych (por. Fyk, Graban, Rakowski, 2002); 5) zagrożenia słuchu towarzyszącego percepcji silnych dźwięków muzycznych (por. Rogowski, Rakowski, Jaroszewski, 2002). Od 2001 roku kierownictwo nad Katedrą Akustyki Muzycznej przejął Andrzej Miśkiewicz. Pod jego przewodnictwem nadal prowadzone są liczne badania z zakresu psychoakustyki muzycznej, metodyki ocen słuchowych, wybranych zagadnień z dziedziny akustyki pomieszczeń i akustyki instrumentów muzycznych oraz teorii słyszenia i problematyki zagrożeń słuchu hałasem.

Podsumowując można powiedzieć, że najważniejsze zdobycze psychologii muzyki w Polsce dotyczą nadal nurtu psychometrii muzycznej i psychoakustyki. Poznawcza psychologia muzyki i neuropsychologia muzyki, tak powszechna na Zachodzie, do tej pory nie znalazła odzwierciedlenia w badaniach polskich. Takiego stanu rzeczy należy doszukiwać się w historii psychologii muzyki w Polsce, w której mniej miejsca poświęcono kognitywistyce niż nurtom społecznym czy rozwojowym. Biorąc pod uwagę, że zachodnia psychologia poznawcza i neuropsychologia muzyki wyrosła na gruncie nauk poznawczych, prawdopodobnym jest, że w Polsce nie znalazła ona odpowiedniego kontekstu.

Wydaje się, że współcześnie istnieje potrzeba sprowadzenia na grunt polskiej psychologii nowoczesnych teorii i metod dotyczących psychologii i neuropsychologii muzyki. Nowe projekty badawcze powinny ewoluować w kierunku poszukiwania funkcjonalnych i neuronalnych podstaw procesów percepcji, komponowania i wykonywania muzyki, czyli tematyki badawczej aktualnej dla zachodnich przedstawicieli poznawczych i neuropsychologicznych podejść do psychologii muzyki.

LITERATURA CYTOWANA

Fyk, J., Graban B., Rakowski, A. (2002). Słuchowe wrażenie wysokości dźwięku a wzrokowe wrażenie barwy. W: A. Rakowski (red.) *Kształtowanie i percepcja sekwencji dźwięków muzycznych* (s. 273–290). Warszawa: Akademia Muzyczna im. Fryderyka Chopina.

Jordan-Szymańska, A. (1991). Psychologia muzyki w Polsce. W: K. Miklaszewski, M. Meyer-Borysewicz (red.) *Psychologia muzyki. Problemy – Zadania – Perspektywy* (s. 83–95). Warszawa: Akademia Muzyczna im. Fryderyka Chopina.

Kamińska, B. (1985b). *Doreen Bridges – Australijski Test Osiągnięć Muzycznych dla Kandydatów na Studia Muzyczne*. Warszawa: Akademia Muzyczna im. Fryderyka Chopina.

Kamińska, B. (1990). Studia wybranych czynników warunkujących rozwój poprawnej intonacji w śpiewie dzieci i młodzieży. W: K. Miklaszewski, M. Meyer-Borysewicz (red.) *Psychologia muzyki. Problemy – Zadania – Perspektywy* (s. 372–381). Warszawa: Akademia Muzyczna im. Fryderyka Chopina.

Manturzewska, M. (1966). Diagnostyczność testu uzdolnień muzycznych R. Drake'a. *Zeszyty Naukowe PWSM, 1.*

Manturzewska, M. (1969). Badania nad warunkami powodzenia w zawodzie pianisty. *Zeszyty Naukowe PWSM, 4.*

Manturzewska, M. (1973). Metodologiczne problemy prognozowania osiągnięć szkolnych. W: M. Manturzewska (red.) *Rekrutacja młodzieży na studia wyższe*. Warszawa: Państwowe Wydawnictwo Naukowe.

Manturzewska, M. (1974). *Psychologiczne wyznaczniki powodzenia w studiach muzycznych*. Warszawa: Centralny Ośrodek Pedagogiczny Szkolnictwa Artystycznego.

Manturzewska, M. (1987). Dotychczasowe badania nad rozwojem muzycznym człowieka: W: Z. Burowska, E. Głowacka (red.) *Z zagadnień zdolności, percepcji i kształcenia muzycznego w akademiach muzycznych*. Kraków: Akademia Muzyczna.

Manturzewska, M. (1990). Psychologia muzyki – kierunki i etapy rozwoju. W: M. Manturzewska, H. Kotarska (red.) *Wybrane zagadnienia z psychologii muzyki* (s. 13–27). Warszawa: Wydawnictwa Szkolne i Pedagogiczne.

Manturzewska, M., Kamińska, B. (1996). Rozwój muzyczny człowieka jako przedmiot badań psychologicznych refleksji pedagogicznych. W: J. Rostowski, T. Rostowska (red.) *Psychospołeczne aspekty rozwoju człowieka*. Łódź: Wydawnictwa Uniwersytetu Łódzkiego.

Manturzewska, M., Kamińska, B. (1999). Określenie przesłanek psychologicznych i socjologicznych postaw i oczekiwań młodzieży w odniesieniu do muzyki w szkole. W: J. Kurcz (red.) *Kształcenie nauczycieli muzyki, stan obecny i perspektywy*. Kraków: Akademia Muzyczna.

Miśkiewicz, A. (2002). Brzmienie i przestrzeń jako podstawowe cechy wielowymiarowe obrazu słuchowego. W: A. Rakowski (red.) *Kształtowanie i percepcja sekwencji dźwięków muzycznych* (s. 194–199). Warszawa: Akademia Muzyczna im. Fryderyka Chopina.

Miśkiewicz, A., Rakowski, A., Rogoziński, P., Kocańda, M. (2002). Progi różnicowe częstości i progi różnicowe wysokości dźwięków muzycznych. W: A. Rakowski (red.) *Kształtowanie i percepcja sekwencji dźwięków muzycznych* (s. 125–154). Warszawa: Akademia Muzyczna im. Fryderyka Chopina.

Rakowski, A., Kisiła, I., Rogowski, P. (2002). Analiza wysokości pięciu dźwięków skrzypcowych granych sposobem wibrato. W: A. Rakowski (red.) *Kształtowanie i percepcja sekwencji dźwięków muzycznych* (s. 15–23). Warszawa: Akademia Muzyczna im. Fryderyka Chopina.

Rogowski, P., Rakowski, A., Jaroszewski, A. (2002). Osobliwe ubytki słuchu u młodych grających na instrumentach perkusyjnych dętych blaszanych. W: A. Rakowski (red.) *Kształtowanie i percepcja sekwencji dźwięków muzycznych* (s. 301–314). Warszawa: Akademia Muzyczna im. Fryderyka Chopina.

Polskie Forum Psychologiczne, 2004, tom 9, numer 1, s. 74–78

ZABURZENIA W ZAKRESIE ROZPOZNAWANIA MUZYKI

Magdalena Berkowska, Ilona Laskowska, Maciej Michalak
Instytut Psychologii
Akademia Bydgoska im. Kazimierza Wielkiego

Wstęp

Czy człowiek może mieć trudności z rozpoznawaniem muzyki? Pytanie, z pozoru oczywiste, nastręcza wiele problemów definicyjnych. Należałoby zatem wyjaśnić, co rozumie się pod nazwą zaburzeń w zakresie odbioru treści muzycznych. Czy niezdolność odróżnienia oberka od mazurka, symfonii od chorału, rocka od jazzu możemy zakwalifikować do kategorii zaburzeń? Czy może problem dotyczy wyłącznie rozpoznawania przebiegów melodycznych?

Kiedyś powszechnie uważano, że brak zdolności do rozpoznawania muzyki związany jest wyłącznie z niesprawnym narządem słuchu. Takie założenie nie pociągało za sobą konieczności klasyfikowania rodzajów zaburzeń. A jak jest dziś?

Obecnie przyjmuje się, że trudności z rozpoznawaniem muzyki mogą występować bez stwierdzonego audiometrycznie upośledzenia słuchu oraz bez deficytów mowy i problemów w różnicowaniu dźwięków niewerbalnych. Niezdolność rozpoznawania muzyki ogranicza się do braku umiejętności identyfikowania powszechnie znanych utworów muzycznych, takich jak na przykład *Sto lat*. Taka dysfunkcja (tzw. agnozja muzyki) jest rodzajem amuzji.

Głównym celem niniejszej publikacji jest zapoznanie czytelnika z problemem amuzji, rozumianym, jako swoiste zaburzenie w zakresie percepcji i produkcji muzyki. Prezentowany artykuł nawiązuje do współczesnej koncepcji amuzji, opisanej przez A. Racette i in. (zamieszczony w tym numerze, artykuł 2). Tematem rozważań uczyniono wyjaśnienie kluczowego pojęcia, jakim jest termin „amuzja", przedstawienie klinicznego przypadku obrazującego zespół symptomów charakterystyczny dla amuzji, oraz modelu teoretycznego wyjaśniającego omawiane zaburzenie, a także wskazanie narzędzi służących jego diagnozie.

Istota amuzji

Amuzja jest specyficznym zaburzeniem, odnoszącym się wyłącznie do dysfunkcji w zakresie percepcji, pamięci, produkcji, czytania i pisania muzyki, które nie jest spowodowane uszkodzeniem słuchu ani deficytami intelektualnymi i motorycznymi (Racette i in., 2004). Dotknięte amuzją jednostki posiadają normalny iloraz inteligencji. Nie są natomiast zdolne do rozpoznawania, czy nucenia znanych im melodii oraz nie posiadają wrażliwości muzycznej potrzebnej do identyfikowania źle brzmiących nut w melodiach.

Adres do korespondencji: Ilona Laskowska, e-mail: morfeusz@ab-byd.edu.pl

Wyróżnia się dwa typy amuzji: nabytą i wrodzoną. Pierwszy typ jest spowodowany uszkodzeniami mózgu związanymi z wylewami, czy wypadkami, których następstwem są uszkodzenia zwłaszcza płata skroniowego. Amuzja nabyta może pojawić się na każdym etapie życia człowieka, podczas gdy wrodzona towarzyszy mu od narodzin, nie będąc jednocześnie konsekwencją uszkodzenia mechanicznego mózgu. A. Racette i in. (2004) wyjaśniają, że początek wrodzonych nieprawidłowości muzycznych może wywodzić się z trudności w zdolnościach potrzebnych do wykrywania zmian w wysokościach dźwięków muzycznych w melodii. Ilustracją jednego z rodzajów amuzji jest studium przypadku opisywane przez I. Peretz i in. (2002). Przypadek Moniki jest przykładem amuzji wrodzonej. Monika, z pochodzenia Kanadyjka, podczas badań miała około czterdziestu lat. Posiada wykształcenie wyższe. Wcześniej przez wiele lat pracowała jako pielęgniarka. Nigdy u niej nie stwierdzono żadnej choroby psychicznej lub neurologicznej. Rezonans magnetyczny nie wykazał widocznych nieprawidłowości anatomicznych. Ponadto nie wykazano zaburzeń w zakresie pamięci werbalnej i niewerbalnej (Test Wechslera: *Digit Span Test*), natomiast zdiagnozowano niezdolność do rozpoznawania lub odbierania muzyki, śpiewu i tańca. U Moniki słuchanie muzyki wywoływało stres. Taki stan rzeczy podyktowany był interpretacją bodźców muzycznych jako doznań o nieprzyjemnym, hałaśliwym charakterze.

Monikę poddano badaniom służącym diagnozie nieprawidłowości związanych z muzyką. Do tego celu użyto baterii testów (*The Montreal Battery of Evaluation of Amusia*, Peretz i in., 2003) mierzących zdolność detekcji zmian w nieznanych melodiach. Zadaniem badanej było określenie czy w prezentowanych parach melodii pojawia się modyfikacja (zmieniony jeden dźwięk, rytm itd.), czy też nie (szczegółowy opis baterii poniżej). Wyniki Moniki w większości zadań miały charakter bardziej przypadkowy niż intencjonalny oraz jasno wskazywały na jej niezdolność do różnicowania melodii na podłożu ogólnego konturu (kierunek melodii) oraz szczegółowego wzoru interwałów (odległości między wysokościami dźwięków). Należy zaznaczyć, że omawiane uszkodzenie nie mogło być wyjaśniane deficytami słuchu lub problemami w rozpoznawaniu dźwięków niemuzycznych (głosy, mówione słowa, dźwięki środowiska) (Peretz i in., 2002).

Model rozpoznawania muzyki I. Peretz i M. Colthearta (2003)

Rezultatem prowadzonych przez ponad dziesięć lat badań u pacjentów z uszkodzeniem mózgu jest model rozpoznawania muzyki zaproponowany przez I. Peretz i M. Colthearta (2003; patrz również Peretz, 2001). Model dotyczy mechanizmów zaangażowanych w proces percepcji i rozpoznawania muzyki i służy do diagnozowania nieprawidłowości związanych z odbiorem i rozpoznawaniem muzyki. Reprezentacja graficzna modelu znajduje się w artykule A. Racette i wsp. w tym numerze.

Model złożony jest z dwóch komponentów przetwarzania informacji: melodycznej i rytmicznej. Komponent melodyczny zawiera trzy główne elementy: kontur, interwały i tonalność. Kontur definiowany jest jako kierunek melodii, na przykład wznoszący lub opadający (analogicznie do intonacji w mowie). Stwierdzono, że mechanizmy odpowiedzialne za przetwarzanie konturu muzyki i intonacji języka są wspólne (Patel, Peretz, 1997). Drugim elementem komponentu melodycznego są interwały określające odległości pomiędzy dwoma dźwiękami. Są one niezbędne podczas czynności rozpoznawania melodii. Trzeci element stanowi tonalne odkodowywanie interwałów na podstawie tonacji utworu. Swoista wrażliwość na tonalność pojawia się już we wczesnym dzieciństwie.

Drugi komponent dotyczy przetwarzania informacji rytmicznej, składającej się z rytmu i metrum. Rytm jest organizacją nut w czasie. Natomiast metrum czynnikiem porządkującym przebieg rytmiczny utworu przez jego podział na takty za pomocą odpowiednio rozłożonych regularnych akcentów.

Kiedy przetwarzanie informacji na poziomie komponentu melodycznego oraz rytmicznego zostaje zakończone, aktywizuje się reprezentacja utworu muzycznego. Aktywizacja tej reprezentacji jest równoznaczna z rozpoznaniem melodii. Jednak ten proces nie jest wystarczający, żeby podać tytuł utworu. Niezbędne jest uruchomienie procesów językowych (np. dotyczących reprezentacji fonologicznej tytułu) w celu jego werbalizacji. Zaburzenie jednego chociażby modułu może spowodować deficyty w rozpoznawaniu muzyki.

Wykorzystanie tego modelu w kontekście diagnozy amuzji nabytej i wrodzonej możliwe jest przy zastosowaniu baterii testów proponowanej przez autorki (*The Montreal Battery of Evaluation of Amusia*, Peretz i in., 2003). Bateria złożona jest z sześciu testów pozwalających na przeanalizowanie każdego z komponentów muzycznych wchodzących w skład wcześniej nakreślonego modelu poznawczego (Peretz i in., 2003). Wszystkie testy dotyczą kolejno organizacji melodycznej (konturu, interwałów, skali), czasowej (rytmu, metrum) oraz sprawności pamięciowej.

Pierwsza część składa się z trzech typów manipulacji zastosowanych do zmiany tego samego tonu. Jedna manipulacja zakłada naruszenie stopni skali określonej melodii przy jednoczesnym zachowaniu oryginalnego konturu melodycznego. Druga manipulacja dotyczy zmiany konturu stworzonej melodii, ale bez dokonywania modyfikacji w obrębie samej skali. Natomiast trzecia manipulacja ma na celu naruszenie interwałów melodii, w której wcześniej dokonano przemiany konturu i skali. Poszczególne typy melodycznych modyfikacji testowane są w dwóch próbach ćwiczeniowych i trzydziestu eksperymentalnych. Każdy z nich obejmuje melodię właściwą i porównawczą oddzieloną od siebie dwusekundową przerwą i poprzedzony jest dźwiękiem próbnym służącym ogniskowaniu uwagi. Pierwszy rodzaj testów (zmodyfikowana skala) skonstruowano w taki sposób, że w skład piętnastu prób wchodzą identyczne melodie, a piętnaście kolejnych pojawia się w zmienionej skali. Drugi i trzeci rodzaj (zmodyfikowany kontur i interwał) są podobne do pierwszego, ponieważ wykorzystują tę samą melodię właściwą, natomiast różnią się tym, że każda porównywana melodia ulega zmianie w wyniku naruszenia bądź pozostawienia bez zmian kierunku konturu. Pary melodii prezentuje się w przypadkowej kolejności. Uczestnicy są proszeni o wykonanie zadania testowego określanego jako „taki sam-różny" przebieg melodyczny. W każdej próbie badani mają zdecydować, czy sekwencja właściwa i porównawcza różnią się od siebie, czy pozostają bez zmian.

Druga część dotyczy organizacji czasowej, na którą składają się testy metryczne i rytmiczne. W testach rytmicznych używa się tych samych bodźców, analogicznie do organizacji melodycznej. W celu stworzenia różnych, porównywalnych wzorów dokonano zmiany wartości czasu trwania dwóch sąsiadujących tonów, poprzez przetworzenie struktury rytmicznej, zachowując jednocześnie to samo metrum i całkowitą liczbę dźwięków. Można to było osiągnąć w wyniku

dodania kropki ćwierćnucie i ósemce lub zmieniając porządek dwóch kolejnych, ale różnych wartości czasowych (np. półnuta, po której następuje ćwierćnuta staje się ćwierćnutą, po której następnie pojawia się półnuta). Seryjne pozycje takich zmian różnią się w kolejnych przykładach. W tym teście stosowane są również dwie próby ćwiczeniowe i trzydzieści eksperymentalnych. Zadaniem badanych jest określenie różnic i podobieństw w prezentowanych fragmentach muzycznych. W teście metrycznym używa się dwufrazowych sekwencji w przeciwieństwie do jednofrazowych występujących w poprzednich zadaniach. Ponadto prezentowane melodie posiadają zharmonizowaną wersję. Połowa sekwencji napisana jest w dwudzielnym (jak w marszu), a połowa w trójdzielnym (jak w walcu) metrum. Utwory są nagrane w przypadkowej kolejności. Badanych informuje się, że usłyszą serię melodii, które będą musieli rozróżnić jako mające formę walca lub marsza.

Ostatnią część baterii muzycznej stanowi rozpoznawanie pamięciowe. Materiałem testowym wykorzystywanym na tym etapie jest zbiór piętnastu losowo wybranych melodii z poprzednich części testu (organizacja melodyczna i czasowa), które zestawia się z piętnastoma nowymi melodiami (niesłyszanymi wcześniej przez badanego), w celu ich wzajemnego porównania. Należy podkreślić, że prezentowane melodie pojawiają się w przypadkowej kolejności. Każda jest eksponowana badanemu najwyżej pięć razy w tym samym formacie (włączając jedną osadzoną w dwufrazowo zharmonizowanej sekwencji). Badani są proszeni o odpowiadanie „tak", jeżeli rozpoznają melodię lub „nie", jeżeli uznają ją za obcą.

Prezentowana bateria jest przydatna do diagnozowania obecności zaburzeń muzycznych w percepcji i pamięci, w związku z tym można jej używać do wykrywania przypadków amuzji wrodzonej w ogólnej populacji (Peretz i in., 2003).

Podsumowanie

Przedstawiony powyżej model rozpoznawania muzyki jest jednym z obowiązujących na świecie sposobów opisu funkcji zaangażowanych w proces percepcji i produkcji muzyki. Coraz częściej jest on wykorzystywany w praktyce klinicznej zaburzeń neuropsychologicznych. Jak już wspomniano niezbędnym uzupełnieniem prezentowanego modelu jest bateria testów – *The Montreal Battery of Evaluation of Amusia*, będąca praktycznym narzędziem umożliwiającym skuteczne określenie charakteru zaburzeń. Odniesienie deficytów do poszczególnych modułów funkcjonalnych modelu pozwala na precyzyjne wyjaśnienie poziomu, na którym dochodzi do wystąpienia zaburzeń. W Polsce nie wykonuje się standardowo badań diagnozujących amuzję. Zaadaptowanie baterii (*The Montreal Battery of Evaluation of Amusia*) do warunków polskich wydaje się zatem konieczne. W tym celu zespół badawczy pod kierunkiem Simone Dalla Bella zainicjował pierwsze prace dotyczące normalizacji testu w Polsce. Przewiduje się zbadanie stu osób w pięciu grupach wiekowych (18–25, 26–35, 36–45, 46–55, od 56), po dwadzieścia w każdej. Należy nadmienić, że badania kanadyjskie nie wykazały istotnego zróżnicowania wyników w poszczególnych grupach wiekowych. Na tej podstawie wnioskujemy, że zmienna niezależna, jaką jest wiek nie powinna wpływać na zróżnicowanie uzyskiwanych wyników. W celu weryfikacji tej hipotezy i kontrolowania wpływu czynnika kulturowego wyodrębniono pięć grup wiekowych (patrz powyżej). Za kryterium doboru przyjęto preferencje muzyczne określone na podstawie kwestionariusza muzycznego w badanych kohortach. Osoby badane będą testowane pod kątem trzech rodzajów umiejętności muzycznych: analizy melodycznej, czasowej i zdolności pamięciowych. W tym celu zostaną zaprezentowane zadania testowe dotyczące kolejno: konturu, interwałów, skali, rytmu, metrum i pamięci muzycznej. Podstawowe sześć testów zachowa wyjściowy materiał muzyczny z wyjątkiem jednego z dodanych podtestów związanych bezpośrednio z pamięcią – *Testu Pamięci Muzycznej*. W teście tym w miejsce kanadyjskich popularnych piosenek wprowadzone zostaną tradycyjne, polskie, powszechnie znane melodie, na przykład *Stary niedźwiedź mocno śpi*. Prowadzone w tym zakresie działania zaowocują pierwszą polską adaptacją testu do badania amuzji nabytej i wrodzonej. W przyszłości możemy się spodziewać wprowadzenia polskiej wersji testu do poradni psychologicznych.

LITERATURA CYTOWANA

Patel, A., Peretz, I. (1997). Is music autonomous from language? A neuropsychological appraisal. Perception and cognition of music. W: I. Deliège, J. Sloboda (red.) *Perception and cognition of music* (s. 191–215). Hove, UK: Psychology Press.

Peretz, I. (2001). Music perception and recognition. W: B. Rapp (red.) *The handbook of cognitive neuropsychology* (s. 519–540). Hove, UK: Psychology Press.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., Jutras, B. (2002). Congenital Amusia: A Disorder of Fine- Grained Pitch Discrimination. *Neuron, 33*, 185–191.

Peretz, I., Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*(7), 688–691.

Peretz, I., Champod, A. S., Hyde, K. (2003). Varietes of musical disorders. The Montreal Battery of Evaluation of Amusia. W: G. Avanzini, C. Faienza, D. Minciacchi, L. Lopez, M. Majno (red.) *The neurosciences of music* (t. 999, s. 58–75). Annals of the New York: Academy of Science.

Racette, A., Hyde, K. L., Peretz, I. (2004). The Amusias. *Polskie Forum Psychologiczne, 9* (1).

SPIS TREŚCI

SPECIAL ISSUE: MUSIC COGNITION AND THE BRAIN
(Guest Editor: Simone Dalla Bella)

KOMENTARZE
[Commentaries]