

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

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

lifetime (at least more than 5 sec) suggests that auditory sensory memory could not be restricted to the classical "sensory store" which can only hold information for one or two seconds.

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.

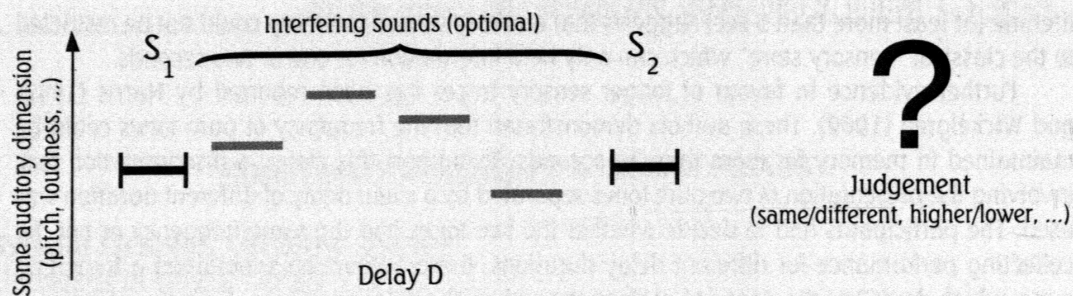


Figure 1. In typical auditory memory tasks with interference, participants have to compare two sounds (S_1 and S_2) separated by some delay (D) upon one sensory dimension (such as loudness, pitch, timbre, duration...). Interfering sounds (I) may also be inserted during the delay being generally ignored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

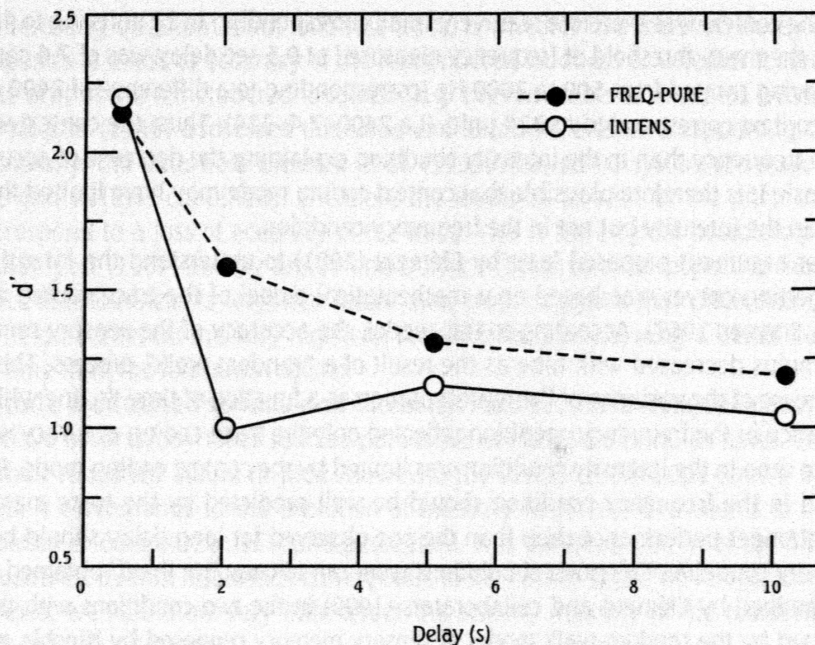


Figure 2. “Forgetting curves” for pitch (condition FREQ-PURE) and loudness (condition INTENS) reported in Clément et al. (1999)

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 $40 \div 1.8 \approx 22$). 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 $2400 \div 7.4 \approx 324$). 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).

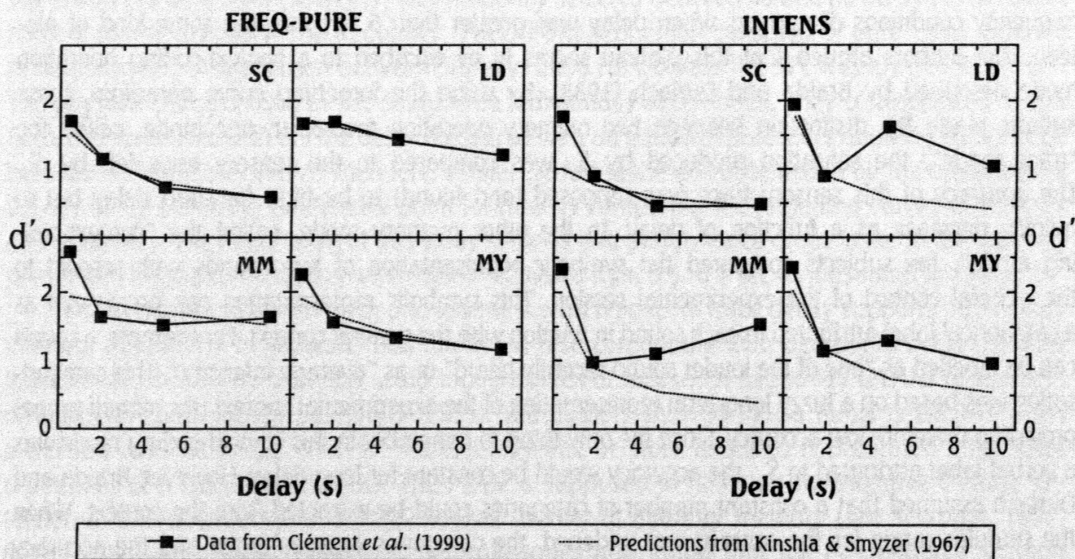


Figure 3. Comparison between the performances from Clément et al. (1999) and predictions made by the random-walk model of sensory memory proposed by Kinshla and Smyzer (1967)

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.

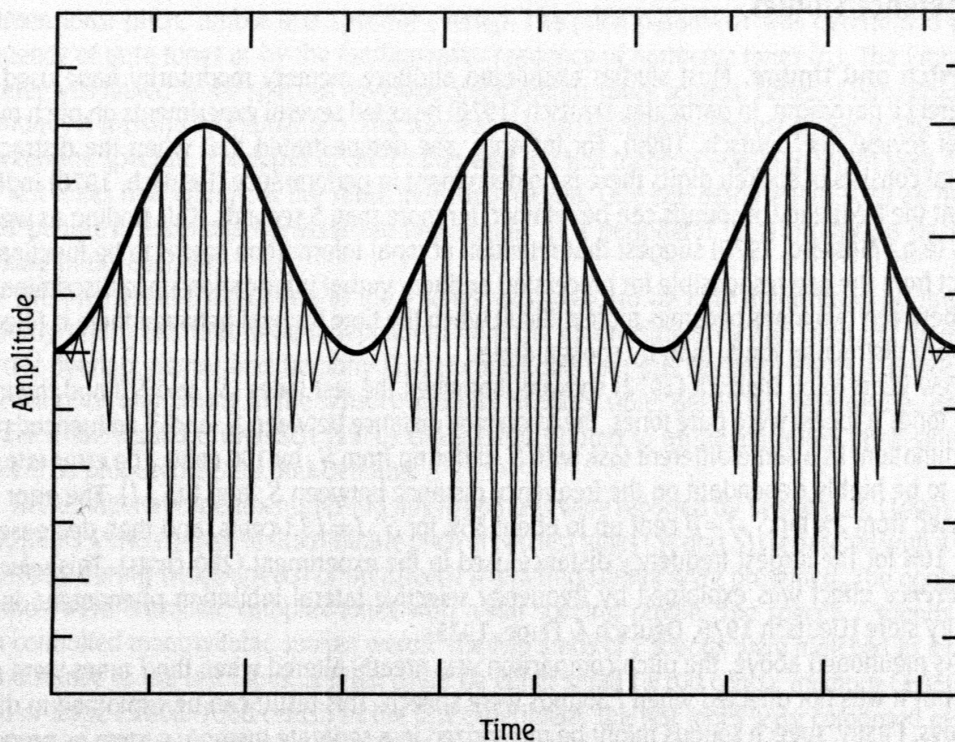


Figure 4. Waveform of a complex tone. The temporal fine structure of the sound is shown by the thin line and its temporal envelope is shown by the thick line

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 1 tones and test tones differ in timbre. In this study, untrained listeners were instructed to discriminate pairs of pure or of complex harmonic tones (*S₁* and *S₂*) separated by a 4.3 sec delay during which *I* tones were presented. The pitch of *S₁* 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, *S₁* and *S₂* differed in pitch by $\pm 4\%$.

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 (*S₁* and *S₂*) 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 $\pm 3\%$ or $\pm 6\%$ from the frequency of S_j . 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_j . Thus, the median harmonic (the 3rd harmonic) of these tones was near S_j 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 $1/30^{\text{th}}$, ± 3 or 6% of the frequency of S_j . The median harmonic (the 30th harmonic) was still close to S_j . 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_j . 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_j 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 abilities. 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 impaired 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 succeeded 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). Peripheral 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 Psychophysics*, 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, 1, 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). Untersuchungen 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.