Neuropsychologia 51 (2013) 2740–2746

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Electrophysiological correlates of the retention of tones differing in timbre in auditory short-term memory

Sophie Nolden [a,b,*], Patrick Bermudez [a,b], Kristelle Alunni-Menichini [a,b], Christine Lefebvre [a,b], Stephan Grimault [a,b,c], Pierre Jolicoeur [a,b]

[a] CERNEC, Université de Montréal, Département de Psychologie, C.P. 6128, succursale Centre ville, Montreal, QC, Canada H3C 3J7
[b] BRAMS, Montréal, QC, Canada
[c] CNRS, Marseille, France

ARTICLE INFO

Article history:
Received 4 March 2013
Received in revised form
3 September 2013
Accepted 4 September 2013
Available online 11 September 2013

Keywords:
Audition
Short-term memory
Working memory
Timbre
Electroencephalography (EEG)
Event-related potentials (ERP)
Sustained anterior negativity (SAN)

ABSTRACT

We examined the electrophysiological correlates of retention in auditory short-term memory (ASTM) for sequences of one, two, or three tones differing in timbre but having the same pitch. We focused on event-related potentials (ERPs) during the retention interval and revealed a sustained fronto-central ERP component (most likely a sustained anterior negativity; SAN) that became more negative as memory load increased. Our results are consistent with recent ERP studies on the retention of pitch and suggest that the SAN reflects brain activity mediating the low-level retention of basic acoustic features in ASTM. The present work shows that the retention of timbre shares common features with the retention of pitch, hence supporting the notion that the retention of basic sensory features is an active process that recruits modality-specific brain areas.

1. Introduction

Short-term memory, the capacity to retain information from the environment for a short period of time, is a key capacity for human cognition and behaviour. To participate in an interesting conversation or for enjoying an artful piece of music, we need to represent the information that we have just heard for a certain time after it is no longer present in the environment. In this study, we used event-related potentials (ERPs) as indices of auditory short-term memory (ASTM). More precisely, we targeted a very basic underlying mechanism—the simple retention of tones differing in timbre. We thus aimed to isolate brain activity that directly reflects representations of these tones as they are actively held in short-term memory.

Research on short-term memory has inspired the development of various theoretical models. An influential model proposed by Baddeley and colleagues emphasises that information is not only stored for a short period of time but can also be manipulated and transformed (Baddeley, 2000; Baddeley & Hitch, 1974). The authors proposed a central executive that controls several subsidiary systems that process information in various ways. According to the model, a visuo-spatial sketch pad and a phonological loop store and manipulate modality-specific information (visual and phonological, respectively) whereas an episodic buffer provides multimodal codes to represent events or episodes. Other models of short-term memory that incorporate neurological findings also propose that simple maintenance in short-term memory would be modality-specific, taking place in sensory areas responsible for the perception of the features, though more elaborate processes (e. g. monitoring and manipulation of information held in short-term memory) might engage executive functions and their subservient areas. Activity patterns arising from perception would thereby overlap with those that contribute to retention (Cowan, 2008; D'Esposito, 2007; D'Esposito, Postle, & Rypma, 2000; Goldman-Rakic, 1987; Petrides, 1991, 2005; Postle, 2006; Postle, Berger, & D'Esposito, 1999; Ruchkin, Grafman, Cameron, & Berndt, 2003).

Given that the models described above emphasise that short-term memory is at least partly modality-specific, it seems reasonable to use stimuli that target specific senses. Regarding ASTM, much of the research to date has used language-related material, usually in the form of aurally presented words (Crottaz-Herbette, Anagnoson, & Menon, 2004; Koelsch et al., 2009) or digits (Pelosi, Hayward, & Blumhardt, 1998). Despite the value of this approach,

* Corresponding author at: Université de Montréal, Département de Psychologie, C.P. 6128, succursale Centre-ville, Montreal, QC, Canada H3C 3J7.
Tel.: +1 514 343 6111 2631; fax: +1 514 343 2285.
E-mail address: sophie.nolden@umontreal.ca (S. Nolden).

0028-3932/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neuropsychologia.2013.09.010

the results might not uniquely reflect fundamental processes of ASTM. When we retain language-related material, it is very likely that the brain activity we observe also reflects the use our explicit and/or implicit knowledge of language and its related codes (phonological, orthographic, etc.) to accomplish the task (Penney, 1989; Schumacher et al., 1996). The use of musical stimuli is similarly problematic given our implicit and (in a minority of the population) explicit knowledge of musical syntax.

Despite a preponderance of work using verbal material, research on the electrophysiological correlates of ASTM has gone some way to dissociate language and non-language related as well as sensory-specific electrophysiological signatures of short-term memory. For example, an interesting effect of memory load has been found during probe presentation (Pratt, Michalewski, Barrett, & Starr, 1989; see also Pratt, Michalewski, Patterson, & Starr, 1989a, 1989b). Participants had to retain visually or aurally presented digits or musical notes while brain activity was measured with electroencephalography (EEG). The number of items to be retained (memory load) was varied. The authors observed a fronto-central positive component after probe presentation that was less pronounced for higher memory loads and generally larger for the visually-presented digits than for aurally-presented digits or musical notes. Furthermore, a later, parieto-central component was characterized by an interaction between stimulus type and memory load. Additionally, ERPs differed between the memory load conditions, thus reflecting different degrees of memory demand. Note, however, that the data were analyzed during a relatively long interval during and after probe presentation and may consequently have been contaminated with a variety of processes other than memory, per se, such as perception, decision making, and response selection. Nonetheless, the study provided useful data concerning variations in memory load and indications that short-term memory for different types of information – visual and auditory short-term memory – may rely on at least partly distinct structures that produce dissociable brain activity, consistent with theoretical notions on the modality-specificity of short-term memory.

Another study focused on brain activity during the retention interval (Lang, Starr, Lang, Lindinger, & Deecke, 1992), participants had to retain visually or aurally presented digits or musical notes. ERPs were assessed during the retention interval, thus focusing on activity related to short-term memory and not to other processes such as encoding, retrieval, comparison, decision-making, or response selection. A frontal negative shift was observed that was more pronounced for aurally-presented digits than for visually-presented digits. Furthermore, this negative shift was more pronounced for musical notes than for aurally-presented digits. Over posterior temporal regions, on the other hand, a negative shift occurred that was more pronounced for visual than auditory stimuli. The authors interpreted this to mean that short-term memory relies on modality-specific processes and processes specific for verbal material. Although no memory-load manipulation was used in this study, the results suggest that a sustained negative frontal ERP might reflect the retention of aurally presented material and be sensitive to whether or not the stimuli are linguistic in nature (see also Pelosi et al., 1998; Ruchkin et al., 1997).

Other ERP studies (Alunni-Menichini, Guimond, Bermudez, Nolden, Lefebvre, & Jolicoeur, 2013; Guimond et al., 2011; Lefebvre et al., 2013) have introduced a paradigm that used several strategies to minimize possible confounds between brain activity related to ASTM and other cognitive processes (see also Grimault et al., 2013; Nolden et al., 2013). First, the stimuli were pure tones differing in pitch that were not very likely to be verbally encoded. Second, the relations of frequencies (pitches) between these tones did not correspond to those of known musical scales and therefore tone sequences would be less likely to match, and activate, our implicit knowledge of music (Trehub, Schellenberg, & Kamenetsky, 1999). Third, memory load was varied while other parameters of the task remained constant, thus making it possible to measure activity related to changes in short-term memory load and minimize the effect of activity related to the task. Fourth, ERPs were assessed during the retention interval (between stimulation and probe), thereby avoiding contamination of brain activity related to perception, encoding, retrieval, comparison, decision-making, or response selection. A fronto-central ERP, predominantly at electrode site AFz, was isolated during the retention interval and named sustained anterior negativity (SAN). Importantly, this ERP became more negative as memory load increased. Additionally, the growth of the SAN with memory load was larger for participants with higher memory capacity (Alunni-Menichini et al., 2013; Lefebvre et al., 2013).

Lefebvre et al. (2013) provided evidence for modality-specific electrophysiological indices of STM by comparing the correlates of the retention of tones and visual items. The SAN was only observed for the retention of tones whereas only the visual items gave rise to a sustained posterior contralateral negativity (SPCN, Klaver, Talsma, Wijers, Heinze, & Mulder, 1999; Robitaille et al., 2010; Vogel & Machizawa, 2004).

The electrophysiological studies of ASTM described above have only used tones varying in multiple tonal characteristics or pure tones varying in pitch. The first approach made difficult the isolation of low-level processes related to the retention of tones and, though the latter approach permitted to focus on the low-level retention of very simple acoustic material, the interpretation of the results of these studies was strictly speaking only valid for the retention of pitch and not for ASTM in general, including the retention of other characteristics of sounds.

Sounds have different characteristics such as pitch, duration, loudness, spatial location, and timbre. One approach to study simple tone retention is to vary one sound characteristic at a time. In this study, we chose to focus on timbre. Timbre can be described as the quality of a sound that distinguishes different musical instruments or differences in voice quality across speakers, as for example when we compare a note played on a violoncello with the same note played on a saxophone, or a Les Paul.[1] Timbre is a multidimensional characteristic that is determined by frequency and time, with attack and the relative intensity of harmonics often emphasized as having particular importance. (Caclin et al., 2006; Caclin, McAdams, Smith, & Winsberg, 2005; Grey, 1977; Marozeau, de Cheveigné, McAdams, & Winsberg, 2003; McAdams, Winsberg, Donnadieu, Soete, & Krimphoff, 1995; Plomb, 1970; Samson, Zatorre, & Ramsay, 1997). As will soon be described, we mainly made use of the variation of the relative intensity of the harmonics in the design of our stimuli.

Timbre and pitch are distinct tonal characteristics and a behavioral study has shown that timbre and pitch are dissociated in auditory short-term memory (Semal & Demany, 1991). There are nonetheless small interactions between the perception of pitch and timbre (for example Singh & Hirsh, 1992; Vurma, Raju, & Kuuda, 2010; Warrier & Zatorre, 2002). Russo and Thompson (2005) demonstrated that timbre can also influence the perceived size of a pitch interval. In their study, two tones were played consecutively with an interval of either six or seven semitones. Pitch could either be ascending or descending. The brightness of the timbre was either congruent (a dull timbre for the low pitch tone and a bright timbre for the high pitch tone), incongruent (a dull timbre for the high pitch tone and a bright timbre for the low pitch tone), or identical for both tones (both tones either dull or bright). Participants judged the size of the tone interval. Six-semitone intervals with congruent timbre were judged to be larger

[1] A Les Paul is an electric guitar with a very characteristic timbre.