Does This Ring a Bell? Music-cued Retrieval of Semantic Knowledge and Metamemory Judgments

Maya Zuckerman[1], Daniel A. Levy[2], Roni Tibon[3], Niv Reggev[1], and Anat Maril[1]

Abstract

■ Failed knowledge recall attempts are sometimes accompanied by a strong feeling of imminent success, giving rise to a "tip-of-the-tongue" (TOT) experience. Similar to successful retrieval (i.e., the Know state, K), a TOT commences with strong cue familiarity but involves only partial retrieval of related information. We sought to characterize the cognitive processes and temporal dynamics of these retrieval states and to extend the applicability of previous findings about TOT to the auditory modality. Participants heard 3-sec initial segments of popular songs and were asked to recall their names. EEG was recorded while participants indicated their retrieval state via button press. Stimulus-locked analyses revealed a significant early left fronto-central difference between TOT and K, at 300–550 msec postcue onset. Post hoc analysis revealed that, in this time window, TOT also differed from DK (Don't Know) responses, which themselves were similar to the K responses. This finding indicates that neural processes, which may reflect strategy selection, ease of semantic processing, familiarity-related processes, or conflict monitoring, are indicative of the fate of our knowledge judgments long before we actually execute them. ■

INTRODUCTION

Metacognition refers to individuals' knowledge, regulation, and control of their own cognitive systems. Within the scope of metacognition, *metamemory* is defined as individuals' knowledge about their own memory and strategies that can aid memory (Koriat, Metcalfe, & Shimamura, 1994). One metamemorial function that has been extensively investigated is the positive prediction judgment. Given upon unsuccessful cued recall, this judgment reflects individuals' positive predictions about their ability to retrieve the answer at a later time (Schwartz, Benjamin, & Bjork, 1997; Schwartz, 1994).

A number of theoretical accounts have been proposed to explain such positive predictions. Hart (1965) put forward the direct access account, proposing that a cognitive monitoring mechanism can detect whether target information that is not immediately accessible is nevertheless available in memory, resulting in an output in the form of positive predictions. Alongside this account stand two leading inferential theoretical accounts: the cue-familiarity account and the accessibility account. The cue-familiarity account (Metcalfe, Schwartz, & Joaquim, 1993; Reder, 1987) suggests that positive predictions stem from an overall assessment of cue familiarity, not from the target's actual level of retrievability. Therefore, if a cue is highly familiar, the positive prediction will be strong, and if a cue is not familiar, the positive prediction will be weak or absent. The accessibility account suggests that positive predictions are based on the overall accessibility of partial information relating to the target rather than the sheer familiarity of the cue (Koriat, 1993). More recently, Koriat and Levy-Sadot (2001) proposed the interactive hypothesis, suggesting an interplay between the cue-familiarity and accessibility accounts. This theory suggests that familiarity is assessed first, followed by the activation of accessibility mechanisms when familiarity is sufficiently high. Accessibility mechanisms, if activated, lead to further interrogation of the memory store in search of the potential target (Koriat & Levy-Sadot, 2001). Importantly, a rapid initial cue-related stage is proposed to occur before memory search, followed by a slower process of a retrieval attempt that may produce partial information. Rapid positive predictions may be formed on the basis of the initial cue-dependent stage (Reder & Ritter, 1992), whereas slower positive predictions may develop as the quantity and accessibility of information produced during the retrieval attempt are assessed.

The focus of the current study is on the following unique retrieval state: Despite failing to recall sought-after knowledge, people may nevertheless (a) have very high confidence that they know the answer and (b) feel that recovery of the answer is imminent. This retrieval state is characterized by a strong feeling of knowing, often to the point that the answer is felt to be "on the tip of the tongue" (TOT; Brown & McNeill, 1966). During a TOT state, people are unable to retrieve a specific bit of information, but at the same time they experience an intense

---

[1]The Hebrew University of Jerusalem, [2]The Interdisciplinary Center, Herzliya, Israel, [3]Bar-Ilan University, Ramat Gan, Israel

subjective feeling that they know the item and are on the verge of recovering it (for a review, see Brown, 1991). Although originally characterized as a voluntary, spontaneously occurring retrieval state (James, 1890), cognitive and neurocognitive research typically characterizes a retrieval experience as TOT if, when failing to retrieve an answer from memory, people are both very confident that they do in fact know the answer, and they have the feeling that successful recovery of the elusive answer is imminent[1] (e.g., Schwartz & Metcalfe, 2011; Schwartz, 1999; Smith, 1994).

According to the cue-familiarity account (Metcalfe et al., 1993), both TOT and successful retrieval (i.e., Know state, K) are hypothesized to commence with high levels of cue familiarity, eliciting a strong feeling that the target may be retrieved. However, these two states have different eventualities; whereas in K the sought-for target is found, in TOT it is not, despite the rememberer's conviction that it is represented in memory (Schwartz, 1998, 1999). This account suggests that, although cognitive processes of TOT and K judgments and their underlying physiological substrates differ in relatively late stages of the retrieval attempt, they will be similar at the initial, familiarity assessment stage and possibly even at the intermediate stages of the retrieval attempt. Data from fMRI studies of metamemory predictions (e.g., Maril, Simons, Weaver, & Schacter, 2005; Kikyo, Ohki, Ishiura, & Sekihara, 2001; Maril, Wagner, & Schacter, 2001) have provided firm evidence for neural differences between a TOT state and a K state, as well as evidence about the spatial localization of the neural circuits involved in the genesis of the TOT state. Within experimental variation, various areas in bilateral and medial frontal regions have been observed to differentiate TOT trials from all other trial types. However, being of limited temporal resolution, these studies could not detect the point in time at which the two states start to diverge. Electrophysiological studies of memory predictions have begun to investigate this divergence and its temporal characterization.

Differentiating Signals of Familiarity Judgments

Focusing on these cue-familiarity assessment stages of positive metamemory predictions, Reder (1987) postulated that, when strategizing about how to answer a question, individuals might quantify the familiarity of the question itself to decide whether the answer is likely to be known, and consequently, whether a memory search should be initiated. A further examination of the initial prediction heuristic led Reder and Ritter (1992) to conduct a study in which participants were presented with unfamiliar math problems, some of which were repeated many times over the course of the experiment, resulting in increased familiarity for the repeated items. Participants were required to rapidly decide (i.e., in less than 850 msec) whether they could quickly retrieve the answer from memory (retrieve trials) or if they would have to calculate the answer (calculate trials). Results revealed that individuals can rapidly assess whether they are likely to know an answer before actually attempting to retrieve it. Furthermore, positive judgments were better predicted by the frequency of presentation of the problem parts than by knowledge of the answer itself, suggesting that an initial positive prediction is based on a feeling of familiarity of the relevant cue (Reder & Ritter, 1992). Employing the rapid positive prediction task, a recent electrophysiological study by Paynter, Reder, and Kieffaber (2009) showed that high-familiarity accurate retrieval trials were associated with greater positivity in an early frontal P2 component (180–280 msec) and a fronto-central P3 component (300–550 msec), in comparison with high-familiarity calculate trials. These results suggest that knowledge predictions may emerge early in the retrieval process (i.e., within 200 msec of stimulus onset) and are influenced strongly by cue familiarity.

Further data about these differences come from a number of ERP studies that have specifically investigated the brain correlates of the TOT state arising during a face-naming task (Lindin & Diaz, 2010; Galdo-Alvarez, Lindin, & Diaz, 2009; Diaz, Lindin, Galdo-Alvarez, Facal, & Juncos-Rabadan, 2007). In one study (Diaz et al., 2007), participants were asked to name famous people shown in photographs, and the ERP correlates of successful naming (K) and TOT states were characterized. Results revealed no ERP differences between TOT and K conditions before 550 msec poststimulus onset. Significant differences in amplitude at posterior locations were observed between the two conditions at a later time interval (550–750 msec). In a later study, Galdo-Alvarez et al. (2009) replicated these results in another group of young adults, and in a follow-up study, Lindin and Diaz (2010) applied an extended version of the same task (Diaz et al., 2007) and found that the latency of ERP component correlates of TOT responses was longer than those of K responses in the range of the early P3 and the N450 components at posterior locations. Recently, Lindin, Diaz, Capilla, Ortiz, and Maestu (2010) aimed to better characterize the spatio-temporal course of brain activation in successful naming and TOT states by utilizing their original face-naming paradigm in a magneto-encephalographic (MEG) study. Results revealed that early in the process (310–520 msec poststimulus onset), the TOT state showed lower activation than the K state at temporal and prefrontal areas, predominantly in the left hemisphere.

In summary, extant findings pertaining to the timing of differences between TOT and K are mixed. Whereas two studies found difference at relatively late time intervals (550–750 msec; Galdo-Alvarez et al., 2009; Diaz et al., 2007) and one study demonstrated latency differences at posterior sites only (Lindin & Diaz, 2010), there is some evidence from both EEG and MEG studies for the existence of a difference related to earlier stages of processing (250–550 msec). Such an early difference distinguishing between predictions of future retrieval has been found in response to rapid predictions elicited by mathematical problems (Paynter et al., 2009) and in a face-naming paradigm comparing TOT and K (Lindin et al., 2010).

In all of the aforementioned studies, the retrieval cues were presented visually. However, studies of patients with various forms of brain damage provide evidence suggesting that the sensory modality in which a cue is presented can impact memory processes (e.g., Straube, Schulz, Geipel, Mentzel, & Miltner, 2008; Wilson, Parsons, & Reutens, 2006; Cuddy & Duffin, 2005; Peretz, Gagnon, Hebert, & Macoir, 2004; Hebert, Racette, Gagnon, & Peretz, 2003; Halpern & O'Connor, 2000; Samson & Zatorre, 1991). Therefore, we were interested in exploring whether these initial, metamemory-related ERP differences are cross-modal or specific to a particular cue modality. Note that most of the previous ERP studies used not only the visual modality, but specifically face stimuli as cues (with the exception of Paynter et al., 2009, who used numbers). Employing musical cues as means of accessing semantic knowledge enables examination of the generalizability of previous results, namely, do previously observed differences between K and TOT reflect cue (face) specific processing or could they be related to cue-independent metamemory processes?

In addition, we were interested in a finer characterization of the differences observed between TOT and successful retrieval (e.g., Lindin et al., 2010). To achieve these goals, we modified the research paradigms used in previous studies in two ways. First, we examined retrieval processes elicited by musical cues; that is, in response to cues given in the auditory modality. Second, whereas previous studies employed three response options, with only one response option reflecting a subjective feeling state between knowing and not knowing (i.e., K, TOT, and Don't Know [DK]), our paradigm included an additional option, which we term "Weak Positive Prediction" (WPP), to indicate all other positive predictions regarding future successful retrieval (i.e., positive predictions not accompanied by high confidence nor a feeling of imminence regarding future retrieval). Including the WPP response option ensured that the TOT response bin would comprise only trials in which participants had a strong feeling of imminent retrieval and not a weak or an intermediate feeling of possible future retrieval (for which participants were instructed to respond with a WPP judgment). In addition, the WPP response provided for a meaningful DK condition. Previous researchers (Lindin et al., 2010; Galdo-Alvarez et al., 2009) did not include DK responses in their analyses because they reasoned that the DK condition in their experiments included both WPPs and real DK responses. Thus, by providing participants with the WPP response option, we hoped to achieve a DK condition that included only negative predictions (as opposed to a mixture of weak positive and negative predictions). Although our main analysis directly compared K and TOT, by comparing those conditions to DK in a post hoc analysis we were able to constrain possible interpretations of previous and current results with regard to the differences between K and TOT (see Discussion).

Musical Stimuli

For the purpose of this study, auditory knowledge cues must potentially elicit a feeling of familiarity and effectively entrain access to semantic information. Verbal presentations of a math problem or a general knowledge question pose methodological challenges because stimulus duration would be relatively long and thus processing span alone might obscure any early familiarity-related signal. Therefore, we sought a stimulus that could elicit familiarity rapidly and, within a short time frame, "pose a question" to which the participant might or might not retrieve the answer. Accordingly, in this experiment, we used short musical excerpts as cues. The properties relating to the appropriateness of musical excerpts were delineated on the basis of previous research. In two recent studies, Daltrozzo and Schon (2009a, 2009b) examined the semantic relatedness of music to verbal information as revealed by the N400 component (see Koelsch et al., 2004, for further explanation). The aforementioned relatedness task was tested both explicitly (Daltrozzo & Schon, 2009a) and implicitly (Daltrozzo & Schon, 2009b). The results suggested that music may elicit conceptual knowledge as early as 250-msec poststimulus onset. In a later study, Daltrozzo, Tillmann, Platel, and Schon (2010) presented participants with highly familiar and less familiar melodies, testing musical recognition with a gating paradigm (Dalla Bella, Peretz, & Aronoff, 2003). In this paradigm, a musical excerpt is presented several times, adding a tone to each repetition. Dalla Bella and colleagues (2003) defined the "familiarity emergence point" (FEP) as the number of tones required by participants to consider the stimulus familiar. Daltrozzo and colleagues (2010) instructed participants to judge a melody as either familiar or unfamiliar, using repeated presentations with additional tones at each repetition. ERPs time-locked to the FEP showed larger fronto-central negativity for highly familiar melodies compared with moderately familiar melodies at 200–500 msec (after the FEP tone). These findings provide further support for the ability of music to evoke patterns stored in memory relatively rapidly. The findings additionally imply that the greater the familiarity of the music, the more concepts are conveyed, as revealed by the increased N400-like component (see also Filipic, Tillmann, & Bigand, 2010, for additional supporting behavioral findings). Thus, the aforementioned studies present evidence that musical stimuli are meaningful enough to convey semantic concepts and to concurrently elicit a rapid familiarity-based response related to prior knowledge. Moreover, both processes occur within a very short time frame, relative to the length of the entire retrieval process.

The Current Study

In brief, the aim of this study was to compare brain activity accompanying TOT and K responses to explore questions regarding semantic knowledge activated via brief musical cues. Our focus was on the differences in neural activity associated with the initial stages of the retrieval process that might distinguish between these retrieval states. In addition to extending previous findings from the visual domain to retrieval in response to auditory cues, we also wished to focus on TOT specifically (as opposed to positive metamemory predictions in general). We conducted an EEG study in which names of commonly known popular songs were queried using short initial song segments. Participants indicated whether they spontaneously recalled the name of the song (K response), experienced a TOT state, had a weaker level of feeling of knowing regarding future knowledge (WPP), or did not know the answer (DK response). We were specifically interested in characterizing the ERP deflections elicited by musical cues, as indications of the temporal dynamics of the two types of judgment: the point at which processing paths between the TOT and K begin to differ as well as the interval during which differences are manifested.

METHODS

Participants

Twenty healthy nonmusician volunteers participated in the experiment. All participants were native Hebrew speakers who were self-reportedly free from neurological or psychological illness, had normal or corrected-to-normal vision, and had no auditory impairments. All participants were either paid (~$5 per hour) or given course credit for their participation in the experiment. Two participants were excluded from analyses because of excessively noisy recordings. The remaining 18 participants (10 men and 8 women) were 20–30 years old (mean age = 24.3 years).

The experiment was approved by the ethics committee of the Department of Psychology at the Hebrew University of Jerusalem, and informed consent was obtained after the experimental procedures were explained to the participants.

Stimuli

A separate behavioral norming study was conducted to determine the selection of musical segments to be used in the EEG experiment. Our aim was to reach a large enough corpus with a distribution that would provide enough trials in each analysis bin. The initial list included 400 segments. Fifty-seven participants who did not participate in the EEG study gave their responses to all 400 segments. Songs that were recognized by all participants or were not recognized by any of the participants were removed from the list. A list of 300 musical segments (150 Hebrew popular music and 150 English popular music) was selected. All segments were taken from original music recordings. All stimuli were edited using Cool Edit Pro audio editing software (Adobe Systems Incorporated, San Jose, CA) at 44 kHz, 16-bit resolution, stereo mode. Short instrumental segments (3 sec long) were sampled from candidate songs. Musical segments were generally initial segments without lyrics (in eight cases they were drawn from prominent non-initial purely instrumental passages). All musical segments were such that in the original song there are no lyrics attached to them (e.g., the last harmonica section in Neil Young's "Heart of Gold" or the guitar section about 4 min into Guns N' Roses' "November Rain"). Thus, none of the segments were susceptible to verbal contamination (i.e., participants were not able to "sing" lyrics along with the music that they heard).

All auditory stimuli were presented binaurally at a comfortable hearing level through in-ear dynamic earphones driven by a computer soundcard. Stimulus presentation was controlled by a PC running Presentation software (Neurobehavioral Systems, Inc., Albany, CA). Responses were collected from an external four-button USB port response box connected to the PC.

Procedure and Cognitive Task

Participants were tested individually in a quiet room. Each experimental session began with instructions about the general aim of the experiment and instructions on EEG artifact reduction by relaxation and attempting to avoid eye movements and blinks. Following EEG headcap preparation, participants were seated in front of a computer monitor. Their index and middle fingers of both hands rested on the four response buttons. During the course of the presentation of the musical stimuli, participants watched a computer screen to minimize eye movements. A colored circle was presented in the center of the screen, changing in color between blocks, and participants were instructed to keep their eyes focused on it.

Each block began with a 10-sec fixation cross in the center of the screen, followed by two filler trials that were discarded from the analysis. The main experimental trials consisted of auditory presentation of a song segment, for which participants were asked to indicate their knowledge of the song title or lyrics. Four response options were available:

(1) Know (K): This is the successful retrieval of the name/lyrics of the song.
(2) Tip-of-the-Tongue (TOT): Given a little more time or a few more notes of the melody, the participant would definitely retrieve the name/lyrics of the song and would be able to later recognize the answer given alternatives on a forced-choice recognition test. Instructions also emphasized that when choosing this response option, participants should feel that successful recovery of the target was imminent.
(3) Weak Positive Predictions (WPP): There is no feeling of imminent retrieval of the target information, but a feeling that with more time the name/lyrics might be retrieved.

(4) Don't Know (DK): Negative prediction—Even given more time or cues, the participant would not be able to retrieve the name/lyrics of the song.

All trials were 4000 msec long (3000 msec of stimuli presentation and 1000 msec of silence). Participants were instructed to respond as quickly as possible at any point during the trial interval. The experimental session was preceded by a short practice session, which was also used for sound-level adjustment, and provision of detailed instructions, including examples to clarify what constitutes TOT and WPP states and the difference between them. The 300 stimuli were fully randomized across participants and presented in six blocks, with a short rest break between blocks.

Participants' actual recognition of the songs was subsequently confirmed off-line. During this recognition test, each song segment appeared with its corresponding (correct) name or line of lyrics from the song, and participants were required to say whether they matched or not. Participants had the option to write the name of a different song if they thought it to be the correct answer. This assessment method, rather than an alternative forced-choice, was employed because targets in this study were names of popular songs, and relevant distractors from the same artists or genres would necessarily have a high contamination potential (see Discussion). Following the recognition test, participants were debriefed. None of the participants exhibited knowledge that all of the song–name pairs presented in the recognition session were actually correct pairs. Moreover, the proportion of accurately recognized songs did not differ between different parts of the recognition test (see Results), indicating that participants did not "discover" that valid song–name pairings were always used.

EEG Methods: Data Recording and Analysis

EEG Acquisition and Preprocessing

The EEG was recorded using the Active II system (BioSemi, Amsterdam, The Netherlands) from 64 electrodes mounted in an elastic cap according to the extended 10–20 system, with the addition of four external electrodes. EOG were recorded using external electrodes located at the outer canthi of the right and left eyes and below the center of the left eye. A reference electrode was placed on the tip of the nose. The ground function during recording was provided by common mode signal and direct right leg electrodes forming a feedback loop, placed over parieto-occipital scalp. The on-line filter settings of the EEG amplifiers were 0.16–100 Hz. Both EEG and EOG were continuously sampled at 512 Hz and stored for off-line analysis.

Data Processing

Data were analyzed using BrainVision Analyzer (Brain Products GmbH; www.brainproducts.com). Raw EEG data were 0.5 Hz high-pass filtered (24 dB/octave) with a notch filter at 50 Hz and were referenced off-line to the tip of the nose. Ocular artifacts were removed using the independent component analysis (ICA) method (Jung et al., 2000; as implemented in BrainVision Analyzer). Following the independent component analysis performed on the unsegmented data, we detected and nullified the blink-related component based on its typical scalp topography and on its time course, which had to match the observable blink artifacts in the raw EEG. Segments contaminated by other artifacts were discarded (rejection criteria: more than 150 μV absolute difference between samples within segments of 200 msec; absolute amplitude > 100 μV). A minimum of 89% of the segments remained after this procedure for each of the participants. In addition, the practice blocks and the filler trials of each block were omitted from the analysis. The remaining EEG data were parsed into 2250-msec segments, beginning 250 msec before stimulus onset. The segments were averaged separately for the possible metamemory ratings. Potentials were measured relative to a −250 to 0 msec prestimulus baseline period. The average waveforms were low-pass filtered with a cutoff of 30 Hz, and the baseline was adjusted by subtracting the mean amplitude of the prestimulus period (250 msec) of each ERP from all the data points in the segment.

ERP Analysis

The waveforms of our main conditions of interest (i.e., K and TOT) and of the DK and WPP conditions were visually inspected to identify time-windows of interest in which there were well-characterized deflections. This inspection confirmed a priori windows of comparison for acoustic processing: an N1 component in the range between 90 and 120 msec poststimulus onset over temporal lobe recording sites and a P2 component in the range between 180 and 230 msec poststimulus onset over fronto-central recording sites. Peaks were detected individually for each participant for each response type and for each component for the purpose of latency measures. The dependent measures in the ERP analyses were the mean amplitudes during the entire a priori time windows and peak latencies of the ERP components within those time ranges. The amplitudes and latencies were compared using repeated-measures ANOVAs with response types and electrodes as factors.

Because the experiment employed a novel paradigm with complex musical stimuli, we were interested in further examination of the spatio-temporal distribution (including all channels and all time points) of differences between experimental conditions. Therefore, in an initial exploratory analysis, we tested the statistical significance of observed differences between the main conditions of interest (K and TOT) at each time point for each electrode. To address the problem of multiple comparisons, we used a correction procedure based on the method introduced by Maris and Oostenveld (2007) to control Type I error rate (implemented using an in-house Matlab script). In this procedure, the statistical significance of the comparison between conditions is determined for a limited number of clusters of adjacent time points, separately for each electrode. First, we identified time points whose *t* statistic for the comparison of the two conditions exceeded a critical threshold ($p < .05$, two-tailed), and we computed a cluster-level statistic as the sum of all *t* values for each of the time points in a cluster of adjacent time points. Next, the distribution of this statistic under the null hypothesis of no difference between conditions was obtained by randomly swapping the conditions across participants 1000 times and calculating the cluster-level statistics for each of these permutations. Then, for each cluster of time points from the original data, a Monte Carlo estimate of the permutation *p* value was calculated as the proportion of random partition values in the distribution of the observed maximal cluster-level statistics exceeding the actual cluster-level statistic of this cluster drawn from the permutation distribution. The 95th quantile of this randomization distribution of maximal cluster-level statistics was used as a critical value to retain or reject the null hypothesis of no differences between conditions. That is, a *p* value smaller than .05 indicates that the two conditions differ significantly at that specific cluster of time points and electrode.

This exploratory analysis revealed a group of electrodes with a strong left fronto-central concentration, in which the conditions of interest differed during the 300–550 msec time windows, as described in the Results. Accordingly, for the following analysis step, we defined a "region of interest" of 12 contiguous electrodes, covering the left-fronto-central area. Differences in mean voltage between the conditions in our "time-window of interest" (300–550 msec) and "region of interest" were then examined using repeated-measures ANOVA and post hoc analyses using the Tukey post hoc criterion for significance ($p < .05$). The Greenhouse–Geisser correction was applied as appropriate.

RESULTS

Behavioral Results

For each participant, the 300 trials were sorted into four response-based bins: K, TOT, WPP, and DK. The average number of trials per condition is presented in Table 1. A TOT response was given for about 18% of all trials, a WPP response was given for 19% of all trials, and the remaining responses were divided fairly evenly between the K and DK bins (about 30% each). These results are generally consistent with previous research (for reviews, see Schwartz, 1994, 1999; Brown, 1991).

On the off-line recognition test following the recording session (see Methods), accuracy rates increased monotonically from the low predictions (i.e., DK and WPP) to the high predictions (i.e., TOT and K). Percentage of correctly recognized items in each response condition was entered as the dependent variable to a one-way repeated-measure ANOVA. Overall, the effect of Response Type was significant, $F(3, 51) = 131.4$, $p < .001$. Post hoc analyses using the Tukey post hoc criterion for significance indicated significant differences between accuracies observed in each response type and all other response types: DK and WPP: $F(1, 17) = 57.7$, WPP and TOT: $F(1, 17) = 53.4$, TOT and K: $F(1, 17) = 41.4$ (all ps $< .001$). This confirmed the validity of the participants' predictions for future retrieval or failure to retrieve (see Table 1).

The off-line recognition test in this study used only valid pairings of songs and titles. To rule out the consequent potential concern that participants may have become aware of the fact that all answers were correct, thereby invalidating this measure of their recognition accuracy, each participant's recognition test was divided into the first, second, and third parts. The proportions of correctly identified segments from each part were compared across all participants in a repeated-measure ANOVA, which did not reveal any influence of order on Recognition Rates ($M_{\text{first}} = 63.4\%$, $M_{\text{second}} = 62.7\%$, $M_{\text{third}} = 61.8\%$, $F(1, 17) < 1$, $p = .6$), indicating that participants were not more likely to accurately identify segments in different parts of the test, confirming the validity of the off-line confirmation procedure.

ERP Results

It was first necessary to establish that the auditory evoked responses reflecting perceptual as opposed to meta-memory aspects of processing were the same across the

Table 1. Average Number of Trials (±*SD*) and Minimum–Maximum Range during the EEG Session and Percentage of Correct Responses on Recognition Test (±*SD*) as a Function of Subjective Rating Response

Music Semantic Task	*DK*	*WPP*	*TOT*	*K*
Number of trials	88.8 (52.8)	57.9 (14.7)	54.6 (19.1)	93.4 (56.6)
Minimum	25[a]	32	23	22[b]
Maximum	191	93	89	203
% of correctly recognized	0.31 (0.17)	0.54 (0.18)	0.75 (0.17)	0.94 (0.06)

[a] With the exception of one participant providing 11 responses in the DK bin.

[b] With the exception of one participant providing 16 responses in the K bin.

different response conditions. Therefore, the N1 and P2 components were analyzed to examine whether differences affecting later recognized effects of interest are found in these acoustic processing windows according to previous literature (e.g., N1: Woods, Alain, Covarrubias, & Zaidel, 1995; Näätanen & Picton, 1987; Wolpaw & Penry, 1975; P2: Rugg & Nieto-Vegas, 1999; Doyle, Rugg, & Wells, 1996; Rugg & Nagy, 1987). As such, the time window defined for the N1 component was 90–120 msec, and the time window defined for the P2 component was 180–230 msec.

N1 Component

The N1 component analysis included bilateral temporoparietal sites (T7, TP7, P7 on the left; T8, TP8, P8 on the right) and midline electrodes (Cz, CPz, and Pz) because previous research has shown that these electrode sites best represent activity of the auditory N1 component (e.g., Woods et al., 1995; Näätanen & Picton, 1987; Wolpaw & Penry, 1975). Repeated measure ANOVAs were performed with electrode and response condition (K, TOT, WPP, and DK) as factors. As expected, there was no significant difference between the four response conditions in the mean amplitude of this component (K, −0.19 μV; TOT, −0.48 μV; WPP, −0.59 μV; DK, −0.87 μV), $F(3, 51) = 0.57, p > .6, \mathrm{MS_e} = 24.97$. In addition, there was no significant difference in mean latency between the four response conditions (K, 107 msec; TOT, 103 msec; WPP, 102 msec; DK, 102 msec), $F(3, 51) = 1.43, p > .2, \mathrm{MS_e} = 709.88$.

P2 Component

Visual inspection of the waveforms showed that the P2 component was centered at the frontal-central region. Therefore, the P2 analyses included the FCz electrode and the four surrounding electrodes (Fz, FC1, FC2, and Cz). Repeated-measure ANOVAs were performed with Electrode and Response Condition as factors. There was no significant difference in the amplitude of the P2 component between the four response conditions (K, 5.92 μV; TOT, 5.81 μV; WPP, 6.27 μV; DK, 6.05 μV), $F(3, 51) = 0.137, p > .9, \mathrm{MS_e} = 22.86$. In addition, there was no significant difference in mean latency between the four response conditions (K, 208 msec; TOT, 199 msec; WPP, 204 msec; DK, 206 msec), $F(3, 51) = 1.93, p > .1, \mathrm{MS_e} = 860.92$.

The N1 and P2 components are believed to reflect early stimulus processing phases. The N1 component has been shown in previous research to be an auditory onset component (e.g., Woods et al., 1995; Näätanen & Picton, 1987; Wolpaw & Penry, 1975). The P2 component has generally been associated with perceptual processing of stimuli (e.g., Rugg & Nieto-Vegas, 1999; Doyle et al., 1996; Rugg & Nagy, 1987). The difference between K, TOT, and DK reported below is not likely to be attributable to early perceptual processes.

Whole-scalp Analysis

Having established that there were no significant electrophysiological differences between response conditions in earlier ERP components linked with perceptual processes, we proceeded to investigate the differences between conditions in the subsequent epochs. We began with the implementation of the whole-scalp analysis using the correction procedure of Maris and Oostenveld (2007) described in the Methods section. We examined K and TOT, the two conditions of primary theoretical interest, to identify the time windows of interest for further analysis. This procedure yielded significant differences between K and TOT at approximately 300–550 msec poststimulus onset in eight fronto-central electrodes: F5, C3, C5, FC1, FC3, FC5, AFz, and AF8 ($p < .05$, two-tailed for all after 1000 iterations; the averaged waveforms of the eight significant electrodes are presented in Figure 1). Electrodes approaching significance (i.e., $p < .07$, two-tailed for all after 1000 iterations) at the same time interval were FP1, F3, C1, CP3, CP5, P5, FPz, Cz, and FC2. Thus, differences between K and TOT were most pronounced in electrodes placed over left fronto-central scalp (see Figure 1). To round out the analysis, we also checked for differences between the other two conditions—DK and WPP (and their differences from the conditions of interest—K and TOT). No coherent cluster of electrodes was identified where differences between these two conditions or between them and K or TOT were observed at any time window (for a list of the electrodes/time windows in which any difference related to these conditions was observed, see Appendix B).

To test whether the differences observed in this group of left-lateralized electrodes reflected a true laterality effect, we conducted an additional repeated-measures ANOVA on the mean voltage of the difference wave between K and TOT in the time window of interest (300–550 msec), with Hemisphere[2] and Electrodes (F5, C3, C5, FC1, FC3, FC5 vs. F6, C4, C6, FC2, FC4, FC6) as factors. A trend toward a laterality effect was found, $F(1, 17) = 3.43, p = .08$, with greater differences in the left hemisphere, which corroborated the above-reported identification of the left fronto-central regions as the area in which the metamemorial processes exhibit divergent activity.

To examine an alternative theoretical account of the data (see Discussion for further explanation), post hoc analyses were carried out including the grand averaged amplitude of the DK response condition. Additionally, we included the grand averaged amplitude of the WPP response condition for the sake of comprehensive data analysis (although we believe that the WPP condition in our study is less well theoretically characterized than the other conditions; see Discussion).

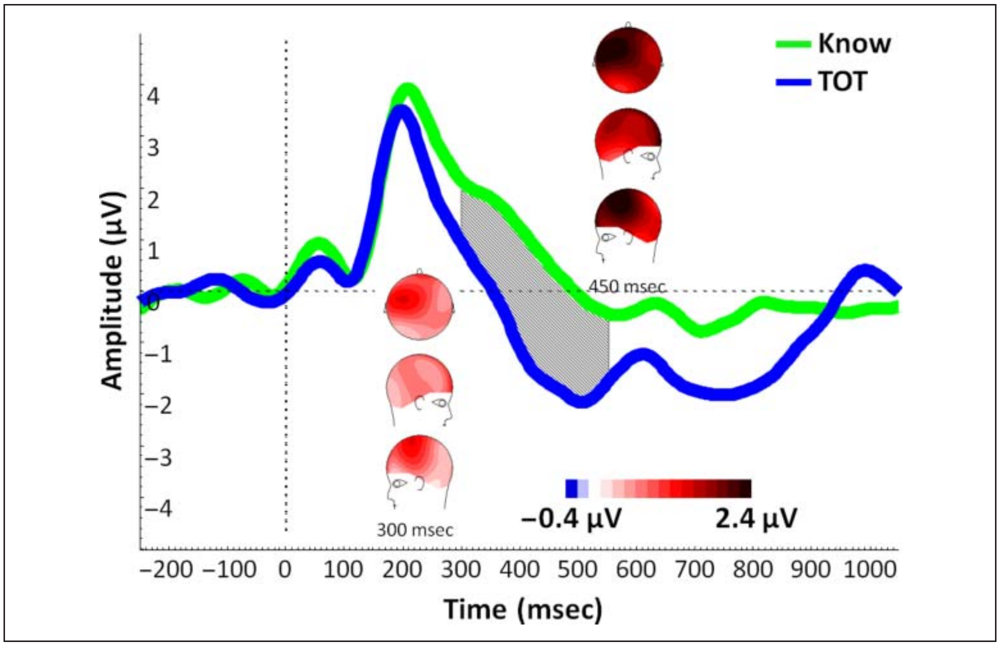

Figure 1. Stimulus-locked grand averaged ERP waveforms for the K and TOT response conditions calculated as the average of the observed eight significant electrodes (F5, C3, C5, FC1, FC3, FC5, AFz, AF8) after Monte Carlo correction. The average waveforms are presented along with the averaged scalp distribution of the difference wave between K and TOT at 300 and 450 msec. The time window of interest (300–550 msec) is marked.

Given the indication that the time window of 300–550 msec poststimulus onset included distinctions of importance in left frontal regions, we demarcated a contiguous cluster of left fronto-central electrodes including but not limited to those exhibiting the above-mentioned significant differences revealed between K and TOT.[3] We conducted a repeated-measures ANOVA with factors of Response Condition (K, TOT, WPP, and DK) and Electrode (12 fronto-central electrodes: F1, F3, F5, FC1, FC3, FC5, C1, C3, C5, CP1, CP3, and CP5; see Figure 2). There was a significant main effect of Response Condition, $F(3, 54) = 3.47$, $p < .05$, $MS_e = 50.99$, $\epsilon = 0.756$, but no effect of Electrode, $F(11, 198) = 2.51$, $p > .07$, $MS_e = 7.87$, $\epsilon = 0.255$, and no interaction effect between Response Condition and Electrode, $F(33, 594) = 0.84$, $p > .5$, $MS_e = 5.18$, $\epsilon = 0.137$. Post hoc pairwise comparisons using the Tukey post hoc criterion for significance indicated that the main effect of Response Condition stemmed from significant differences between the K and TOT ($p = .003$) conditions and between the DK and TOT conditions ($p = .042$). There were no significant differences between the WPP condition and the other conditions (K, TOT, and DK). Thus, TOT differed from both K and DK conditions, which did not differ from each other, whereas the WPP response condition was not distinguished from any of the other conditions (see Figure 3 for the grand averaged waveforms of the four conditions).

Finally, we ran two additional analyses to rule out alternative accounts of the data. First, some of the segments carried more diagnostic information in the first beat than did other segments because of the nature of the instrument or sound quality of the first beat. This could have led to the patterns observed resulting from a stimulus effect rather than from judgment differences. If this stimulus effect were indeed the case, many of the same segments would be sorted into the K, TOT, and so forth, categories across different participants. We therefore conducted a careful item analysis, in which we examined the proportion of participants that gave the same (K or TOT) response to any given segment. For K responses, the average proportion was 0.31 and the median was 0.28; for TOT responses, the average was 0.18 and the median 0.17. This analysis renders highly unlikely the hypothesis that the EEG differences observed here stem from rudimentary physical differences of TOT musical sequences relative to all others. Second, it is theoretically possible that participants might have sometimes been more attentive or more effortful in their identification attempts, leading to enhanced N400 processing independent of the knowledge of the stimulus that was presented. If state effect was indeed the case, subjective reports of K/TOT, and so forth, would be expected to cluster together in time more than the chance distribution expected given the random order of stimulus presentation. To examine this alternative, we first estimated the distribution of responses expected by chance alone by running a simulation, separately for each participant and based on his or her individual proportions of each response type (K, TOT, WPP, and DK), whereby 10,000 sequences of responses were generated. The maximum number of consecutive responses of the same type observed in this simulation at a $p = .05$ cutoff was four. Five participants had such a string in one of the relevant conditions (K and TOT; an additional participant had a string of five WPP responses). However, (a) excluding these participants from the group analysis revealed the same results as the whole 18-participant analysis, with the exception that the significant differences between the K

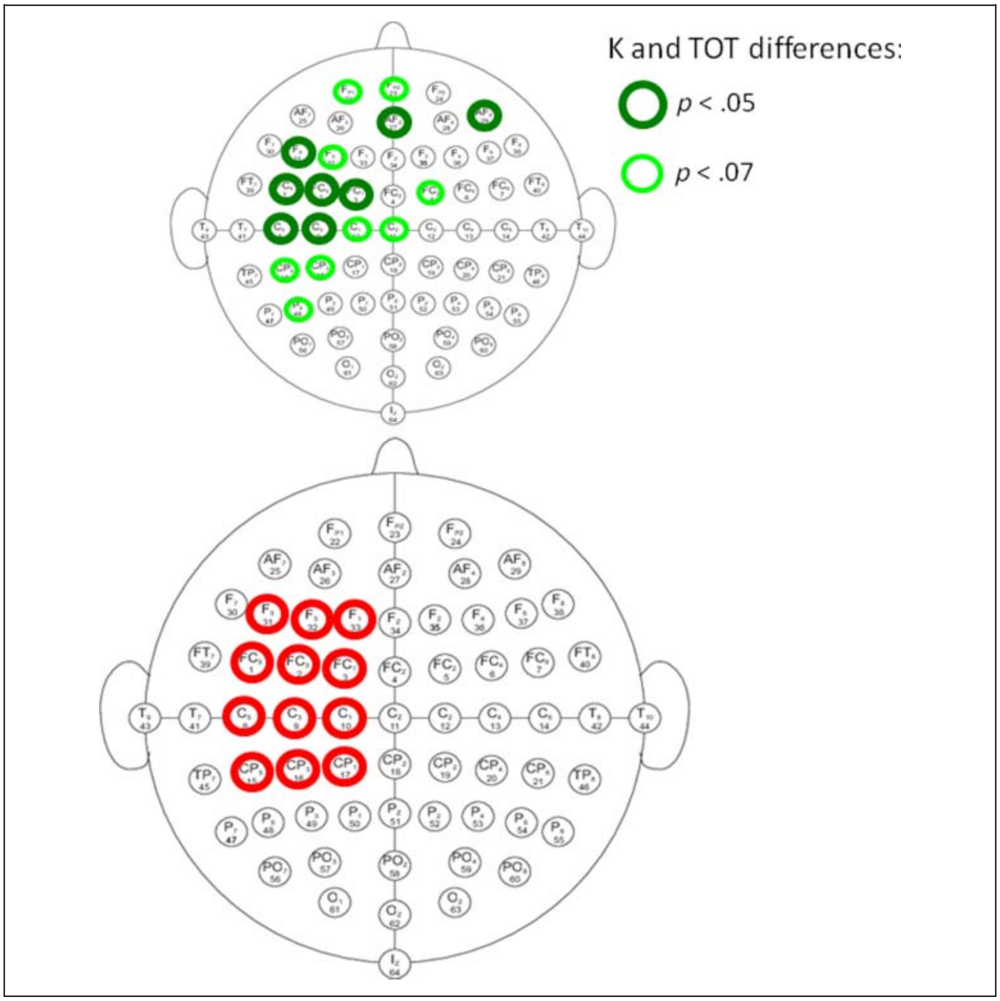

Figure 2. Top: Electrode array with significant electrodes (dark green) and approaching significant electrodes (light green) marked. Bottom: Electrode array with electrodes of the ROI marked in red.

and TOT conditions found in the same time window and in the same electrodes did not pass the (rather strict) permutations test correction (expected in light of the reduction on power—almost a third of the participants was removed), and (b) even within these five participants, only one such "cluster" was found, and only in one of the two conditions, K or TOT. No participant was identified who exhibited what would be expected under the state

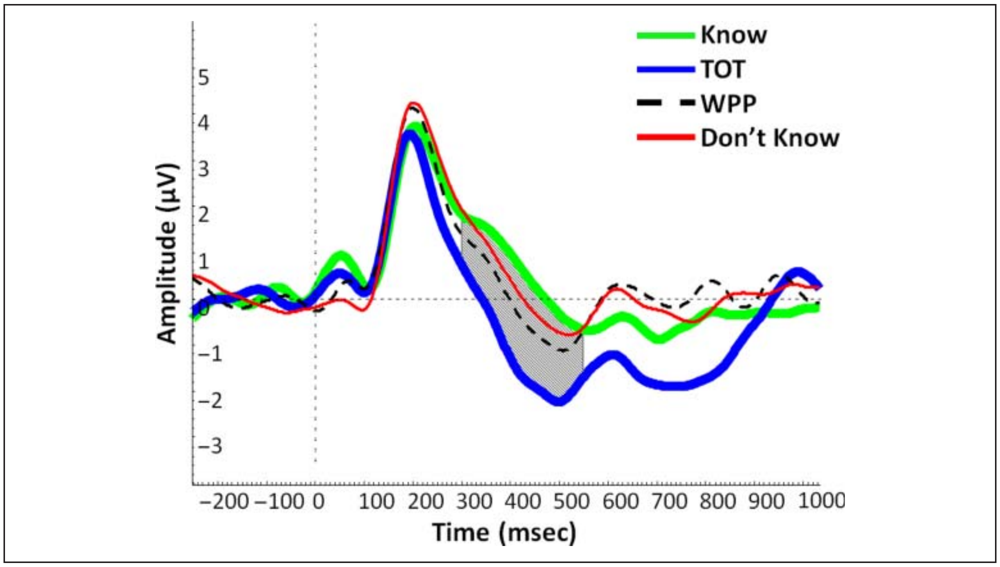

Figure 3. Stimulus-locked grand averaged ERP waveforms for each response condition (K, TOT, WPP, and DK) calculated as the average of the 12 electrodes of the defined ROI (F1, F3, F5, C1, C3, C5, FC1, FC3, FC5, CP1, CP3, CP5). The time window of interest (300–550 msec) is marked.

effect (clustering) hypothesis, namely, repeated clusters of Ks and TOTs.

DISCUSSION

The objectives of the current study were (1) to probe for initial stages of processing differentiating between the TOT state and successful retrieval (K) in semantic metamemory judgments using a more comprehensive study design than those used in previous studies and (2) to extend the findings of previous studies to a different cue stimulus modality. These aims were achieved by investigating the electrophysiological correlates of the retrieval states elicited in response to musical cues and by including an additional, weaker positive prediction response condition (WPP).

We identified a cluster of electrodes in the left frontocentral region in which TOT responses elicited a significantly more negative-going deflection than K responses at a very early time interval (i.e., 300–550 msec poststimulus onset) relative to the actual behavioral response (on average, 2–3 sec poststimulus onset). These electrophysiological results suggest a distinction between a strong positive retrieval prediction (TOT) and actual retrieval (K), occurring long before response selection.

Our results converge temporally with two previous studies (MEG, Lindin et al., 2010; EEG, Paynter et al., 2009) and spatially with a recent MEG study using visually presented faces (Lindin et al., 2010). Beyond replicating the finding that K and TOT response conditions differentiate already at initial stages of retrieval, the current study extended those findings to cues presented in the auditory modality, suggesting that results previously obtained are not cue specific but are related to the domain of target information, that is, semantic knowledge.

Initially, the WPP condition was intended to collect responses in which there was a weak positive feeling of knowing and to differentiate them from our condition of interest, the TOT response, allowing us to examine a "purer" TOT response condition than in previous studies employing three response conditions (K, TOT, and DK). However, this response option also provides for a meaningful DK condition. Whereas previous researchers (Lindin et al., 2010; Galdo-Alvarez et al., 2009) did not include DK responses in their analyses because they reasoned that the DK condition in their experiments included both WPPs and real DK responses, we were able to use the DK condition in an informative way. Although our main interest in characterizing the spatial and temporal differences between TOT and K dictated our primary analyses, the availability of an informative DK condition was useful in restricting the possible interpretations of the observed differences, as we discuss below.

Our results call into question the interpretation suggested by Lindin and colleagues (2010), who suggested that the differences between K and TOT in the 300–500 msec time window might be associated with different levels of access to the memory stores of information (both semantic and lexical) about the target (in line with the model offered by Valentine, Brennen, & Bredart, 1996). In contrast, we have shown that in the time window during which TOT differed from K, it also differed from DK, which itself was similar to K. This pattern of results is not easily accommodated within the semantic/lexical interpretation offered by Lindin et al. (2010). Their interpretation of our data would require that, in trials in which the participant did not know the answer, the level of access to semantic/lexical target information is the same as in trials where the answer was known and further that the level of access for DK trials was higher than for TOT trials. Clearly, these are unlikely scenarios.

TOT and (F)N400

The cognitive theoretical model of metamemorial processes proposed by Koriat and Levy-Sadot (2001) reflects a hierarchical process whereby a cue-familiarity assessment precedes retrieval processes (and the accessibility-related processes they elicit). Therefore, it could be the case that our results reflect early cue familiarity-based processes rather than semantic accessibility mechanisms. Accordingly, the waveform characteristics of the TOT condition could be related to a similar deflection (in time window and spatial location) of a difference often reported in EEG studies of episodic memory (reviewed in Rugg & Curran, 2007; Rugg & Yonelinas, 2003). Familiarity-based differences between old and new items on an episodic recognition test have been marked by brain potentials in the time frame of 300–500 msec designated as the FN400 or midfrontal old/new effects, exhibiting greater negativity for new items than for old studied items (for reviews, see Voss, Lucas, & Paller, 2010; Rugg & Curran, 2007; Curran, DeBuse, Woroch, & Hirshman, 2006; Mecklinger, 2006; for a different view, see Voss & Paller, 2006, 2007, 2008, 2009). In accordance with results from the ERP episodic memory literature (e.g., Curran, Schacter, Johnson, & Spinks, 2001; Curran, 2000), our results could be seen as reflecting (semantic) cue familiarity-related processes eliciting the FN400 potentials. This interpretation would suggest that, although K and TOT are both high-familiarity states, with the phenomenological difference between them being the time to resolution, differences in familiarity between them do exist.

To assess the FN400 interpretation, we took advantage of the DK response condition trials in our analysis. DK trials should reflect, on average, a lower degree of familiarity compared with both K and TOT trials. If our results reflect the FN400 component and FN400 indexes familiarity strength of the cue (and not exclusively in an episodic recognition context), we would expect an increased negative amplitude for the average waveform of the DK response compared with the average waveform of the TOT response at our time window of interest (300–550 msec): the lower the familiarity, the greater the negativity. However, our analysis revealed that the electrophysiological traces of the TOT state differed significantly from both K and DK at the 300–550 msec time window, whereas no significant differences between K and DK were found (see Figure 3). These results are not consistent with this component being the familiarity-based FN400 potential as reported in episodic memory literature.

Given these reasons to reject an FN400 familiarity interpretation of the TOT-elicited ERP deflection, we suggest that our observed differences may reflect an initial stage of the metamemorial monitoring process, distinguishing between the definite state of knowing (K) or not knowing (DK) and the indefinite TOT state. We propose that rather than reflecting familiarity strength, the signal observed in our study may instead reflect the degree of task-related informativeness of familiar stimuli. Whereas both the K and DK states reflect high levels of informativeness (i.e., extremely familiar and extremely unfamiliar cues are both highly informative regarding the prospect of their ability to elicit target information), the TOT state may be related to a less definite degree of informativeness. Therefore, it is possible that the observed difference, while related to familiarity-based processes, does not reflect the simple familiarity strength of a given cue, but rather its task-related information value.

Recent research has questioned the distinction between the FN400 and the N400 potentials (Voss & Paller, 2009; Boldini, Algarabel, Ibanez, & Bajo, 2008; Meyer, Mecklinger, & Friederici, 2007; Nessler, Mecklinger, & Penney, 2005; Curran & Cleary, 2003), a well-known marker of semantic processing (Schon, Ystad, Kronland-Martinet, & Besson, 2009; Orgs, Lange, Dombrowski, & Heil, 2006, 2007; Federmeier & Kutas, 2001; Castle, Van Toller, & Miligan, 2000; Kutas & Federmeier, 2000; Kutas & Van Petten, 1994; Bentin, McCarthy, & Wood, 1985; Kutas & Hillyard, 1980). This proposal provides a possible explanation for the results of our study. As noted above, processing differences between metamemorial states reflected in EEG amplitude in the 300–550 msec epoch might index task-related informativeness of the cue rather than its familiarity strength. High informativeness entrains limited subsequent semantic processing, irrespective of the prediction's direction (i.e., K, positive prediction; DK, negative prediction). In contrast, low informativeness yielding the TOT state initiates an extensive and more difficult search process. Thus, the negative deflection associated with the TOT condition seems most likely to belong to the N400 family of components, which have been characterized as delineating "a temporal interval in which unimodal sensory analysis gives way to multimodal associations in a manner that makes use of—and has consequences for—(semantic) long-term memory" (Kutas & Federmeier, 2011, p. 639). The early expression of differences between trials which later culminate in decisive success or failure of knowledge retrieval, and those in which doubt lingers regarding the possible availability of the target information indicates that, as in other N400-indexed cognitive processes, semantic access may have been initiated before recognition was complete. Regarding the question of interest of our study, the implication is that metamemory processes may begin even before the retrieval cue is fully processed. Although N400 is typically observed in posterior sites, when elicited by auditory stimuli it has been shown to have a more frontal distribution than following visual presentation (e.g., Holcomb & Neville, 1990), and so the TOT waveform may certainly be related to that family of components.

Concluding Remarks

In our ROI, the ERP deflection elicited in the WPP condition either lies between the TOT condition and the K (or DK) conditions or along the DK condition (see Figure 3); in neither case, it is significantly differentiated from the other three conditions. This result is expected, because in the current design, the WPP response condition includes a range of intermediate-to-weaker positive predictions (see e.g., Schmitter-Edgecombe & Anderson, 2007; Koriat, 1993; Costermans, Lories, & Ansay, 1992; Gruneberg & Monks, 1974; Hart, 1965). Moreover, evidence from neuroimaging studies reveals that specific brain regions show gradation in activity correlated with different levels of positive ratings other than TOT, implying that these different states are correlated with different neural activity (Schnyer, Nicholls, & Verfaellie, 2005; Kikyo & Miyashita, 2004). Therefore, we believe that the WPP condition in our study is theoretically undercharacterized, as it includes a scale of intermediate levels of predicted knowledge between DK and TOT. Possibly a finer division of this condition would separate the WPP response condition into ERPs that are similar to the TOT condition for the higher predictions and those that are similar to the DK condition for the lower predictions, but this remains to be demonstrated in future research.

The scalp distribution of the difference wave between K and TOT suggests that the difference between the two states, both elicited by a semantic musical cue, was relatively left lateralized. This may seem to contrast with literature, suggesting that music is processed by the right hemisphere (for a review, see Peretz & Zatorre, 2005; although see Levitin & Tirovolas, 2009 for a different view). However, in this study musical segments served as cues whereas the sought-for targets consisted of verbal information. Therefore, the observed left-lateralized patterns may reflect modality-general memory retrieval processes. Additionally, the left-lateralized patterns further extend previous fMRI studies revealing right-lateralized frontal differences between K and TOT (Maril et al., 2001, 2005), suggesting that it is possible that the cognitive processes reflected in the right pFC activation as observed with fMRI actually originate with early processes supported by the left pFC, the latter process being too fleeting to be observed within the time resolution of fMRI.

The off-line recognition test in this study used only valid pairings of songs and titles. Although there is no single perfect way to assess participants' accuracy of prediction (Nelson, Gerler, & Narens, 1984), this type of no-distractor recognition may be a relatively weak test of participants' knowledge, as it may raise the potential concern that participants may have become aware of the fact that all answers were correct, thereby invalidating this measure of their recognition accuracy. However, as mentioned earlier, participants' debriefing revealed no sign of them "discovering" that the name provided was always the correct one. More importantly, had participants been aware of this fact, their accuracy should have been at ceiling, at least in the latter part of the test (after they have had a chance to make this "discovery"). Analysis of the behavioral data, comparing participants' accuracy in the first, second, and third parts of the recognition test, revealed that there was no difference between participants' accuracy when responding to items located in the test's first, second, or third parts, reinforcing our claim that participants were not aware of the fact that only valid pairs were presented.

In this study, the electrophysiological differences revealed between response types were observed much earlier than actual response execution, probably before the actual formation of the retrieval judgment response. The nature of the relationship between these neural processes and our actual eventual response remains unknown. Whether these relatively early differences actually determine the eventual future response, reflect initial preretrieval processes such as strategy selection and conflict monitoring, are related to ease of semantic processing or task-relevant informativeness remains a subject for future research.

APPENDIX A. List of 150 Musical Excerpts from English Popular Music

1. 1973 (James Blunt)
2. A Forest (The Cure)
3. Africa (Toto)
4. Alive (Pearl Jam)
5. All I Wanna Do (Sheryl Crow)
6. Ayo Technology (50 Cent)
7. Babooshka (Kate Bush)
8. Bachelorette (Bjork)
9. Beverly Hills 90210 theme
10. Big In Japan (Alphaville)
11. Bigmouth Strikes Again (The Smiths)
12. Blowin' in the Wind (Bob Dylan)
13. Bohemian Raphsody (Queen)
14. Born Slippy (Underworld)
15. Boys Don't Cry (The Cure)
16. Brothers In Arms (Dire Straits)
17. Call Me (Blondie)

APPENDIX A. *(continued)*

18. Can You Feel the Love Tonight (Elton John)
19. Close To Me (The Cure)
20. Coffee and TV (Blur)
21. Come As You Are (Nirvana)
22. Come on Eileen (Dexys Midnight Runners)
23. Come Undone (Duran Duran)
24. Creep (Radiohead)
25. Dancing Queen (ABBA)
26. Day Tripper (The Beatles)
27. Daysleeper (REM)
28. Dazed and Confused (Led Zeppelin)
29. De Do Do Do, De Da Da Da (The Police)
30. Don't Know Why (Norah Jones)
31. Don't Let Me Be Misunderstood (The Animals)
32. Don't Look Back In Anger (Oasis)
33. Don't Stop (Fleetwood Mac)
34. Down Under (Men At Work)
35. Dream On (Depeche Mode)
36. Drive (REM)
37. Englishman In New York (Sting)
38. Enjoy the Silence (Depeche Mode)
39. ET theme
40. Every Breath You Take (Sting & The Police)
41. Everybody Hurts (REM)
42. Everywhere (Fleetwood Mac)
43. Eye of the Tiger (Survivor)
44. Free Your Mind (En Vogue)
45. Friday I'm in Love (The Cure)
46. Friends theme
47. Funky Town (Lipps Inc)
48. Girls and Boys (Blur)
49. Hand in My Pocket (Alanis Morissette)
50. Heart of Gold (Neil Young)
51. Hello I Love You (The Doors)
52. House of the Rising Sun (The Animals)
53. I Don't Want To Wait (Paula Cole)
54. I Feel Fine (The Beatles)
55. I Get Around (Beach Boys)
56. Imitation of Life (REM)
57. In Bloom (Nirvana)

APPENDIX A. *(continued)*

58. In the Middle of the Night (Billy Joel)
59. Independent Women (Destiny's Child)
60. Invisible Touch (Genesis)
61. James Bond 007 theme
62. Jeremy (Pearl Jam)
63. Jesus He Knows Me (Genesis)
64. Johnny B. Goode (Chuck Berry)
65. Jump Around (House of Pain)
66. Just Can't Get Enough (Depeche Mode)
67. Karma Police (Radiohead)
68. Kill Bill theme
69. Lady Madonna (The Beatles)
70. Le Freak (Chic)
71. Lemon Tree (Fool's Garden)
72. Let It Be (The Beatles)
73. Let's Dance (David Bowie)
74. Light My Fire (The Doors)
75. Love In This Club (Usher)
76. Love Me Do (The Beatles)
77. Love Song (The Cure)
78. Lullaby (The Cure)
79. Mamma Mia (ABBA)
80. Man On the Moon (REM)
81. Matrix theme
82. Moonlight Shadow (Mike Oldfield)
83. More Than A Woman (Bee Gees)
84. Mr. Jones (Counting Crows)
85. Mr. Tambourine Man (Bob Dylan)
86. My Lovin' (En Vogue)
87. Mysterious Ways (U2)
88. No Surprises (Radiohead)
89. November Rain (Guns N' Roses)
90. One Of Us (Joan Osborne)
91. Only Love Can Break Your Heart (Neil Young)
92. Ordinary World (Red)
93. Our House (Madness)
94. Paranoid Android (Radiohead)
95. People Are Strange (The Doors)
96. Perfect Day (Lou Reed)
97. Personal Jesus (Depeche Mode)

APPENDIX A. *(continued)*

98. Piano Man (Billy Joel)
99. Pink (Aerosmith)
100. Precious (Depeche Mode)
101. Respect (Aretha Franklin)
102. Rocky theme – Gonna Fly Now
103. Romeo and Juliet (The Killers)
104. Roxanne (The Police)
105. Satisfaction (The Rolling Stones)
106. Seinfield theme
107. Seven Nation Army (The White Stripes)
108. She's a Rainbow (The Rolling Stones)
109. She's Always a Woman To Me (Billy Joel)
110. She's Like the Wind (Patrick Swayze)
111. Shiny Happy People (REM)
112. Smells Like Teen Spirit (Nirvana)
113. Smoke On the Water (Deep Purple)
114. So Far Away (Dire Straits)
115. Solsbury Hill (Peter Gabriel)
116. Something (The Beatles)
117. Song2 (Blur)
118. Spread a Little Happiness (Sting)
119. Stayin' Alive (Bee Gees)
120. Sultans of Swing (Dire Straits)
121. Surfin' USA (Beach Boys)
122. Suzanne (Leonard Cohen)
123. Sweet Dreams (Eurythmics)
124. Sweet Home Alabama (Lynyrd Skynrd)
125. Tears In Heaven (Eric Clapton)
126. The Flinstones theme
127. The Man Who Sold the World (Nirvana)
128. The One I Love (REM)
129. The Simpsons theme
130. The Unforgiven (Metallica)
131. The Universal (Blur)
132. The Winner Takes It All (ABBA)
133. Time of Your Life (Green Day)
134. To the Left (Beyonce)
135. Touch Me (The Doors)
136. Tragedy (Bee Gees)
137. Trouble (Coldplay)

APPENDIX A. *(continued)*

138. Unbreak My Heart (Toni Braxton)
139. Under the Bridge (Red Hot Chili Peppers)
140. Upside Down (Diana Ross)
141. Walk On the Wild Side (Lou Reed)
142. We Be Burnin' (Sean Paul)
143. We Didn't Start the Fire (Billy Joel)
144. What a Feeling (Irene Cara)
145. When I Come Around (Green Day)
146. Wonderwall (Oasis)
147. Yeah (Usher)
148. Yellow (Coldplay)
149. Ziggy Stardust (David Bowie)
150. Zombie (The Cranberries)

APPENDIX B. Electrodes and Time Window in which a Significant Difference Was Found between DK or WPP and Any of the Other Conditions

Comparison	*Electrode*	Time Window (msec)
TOT and DK	FC4	735–990
TOT and WPP	P8	670–950
TOT and WPP	P10	690–950

Acknowledgments

We thank Assaf Breska and Shani Shalgi for their assistance in the EEG data analysis. This work was supported by The Israel Science Foundation (in part; Grant No. 350/10 to A. M. and Grant No. 611/09 to D. A. L.).

Reprint requests should be sent to Anat Maril, Department of Psychology, Social Sciences Building, The Hebrew University of Jerusalem, Mt. Scopus, Jerusalem 91905, Israel, or via e-mail: marila@pluto.huji.ac.il.

Notes

1. Using the term TOT to characterize these trials should not be taken to imply that we take a theoretical position on the question concerning whether the differences between strong feeling of knowing and TOT are quantitative or qualitative in nature (e.g., Schwartz, 2008).
2. Laterality was also examined in light of the literature suggesting that music is processed by the right hemisphere (see Discussion).
3. All presented effects withstand when limiting the analysis to the electrodes revealing significant differences between K and TOT.

REFERENCES

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology, 60,* 343–355.

Boldini, A., Algarabel, S., Ibanez, A., & Bajo, M. T. (2008). Perceptual and semantic familiarity in recognition memory: An event-related potential study. *NeuroReport, 19,* 305–308.

Brown, A. S. (1991). A review of the tip-of-the-tongue experience. *Psychological Bulletin, 109,* 204–223.

Brown, R., & McNeill, D. (1966). Tip of tongue phenomenon. *Journal of Verbal Learning and Verbal Behavior, 5,* 325–337.

Castle, P. C., Van Toller, S., & Miligan, G. J. (2000). The effect of odor priming on cortical EEG and visual ERP responses. *International Journal of Psychophysiology, 36,* 123–131.

Costermans, J., Lories, G., & Ansay, C. (1992). Confidence level and feeling of knowing in question answering: The weight of inferential processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 142–150.

Cuddy, L. L., & Duffin, J. (2005). Music, memory, and Alzheimer's disease: Is music recognition spared in dementia, and how can it be assessed? *Medical Hypotheses, 64,* 229–235.

Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & Cognition, 28,* 923–938.

Curran, T., & Cleary, A. M. (2003). Using ERPs to dissociate recollection from familiarity in picture recognition. *Cognitive Brain Research, 15,* 191–205.

Curran, T., DeBuse, C., Woroch, B., & Hirshman, E. (2006). Combined pharmacological and electrophysiological dissociation of familiarity and recollection. *Journal of Neuroscience, 26,* 1979–1985.

Curran, T., Schacter, D. L., Johnson, M. K., & Spinks, R. (2001). Brain potentials reflect behavioral differences in true and false recognition. *Journal of Cognitive Neuroscience, 13,* 201–216.

Dalla Bella, S., Peretz, I., & Aronoff, N. (2003). Time course of melody recognition: A gating paradigm study. *Perception & Psychophysics, 65,* 1019–1028.

Daltrozzo, J., & Schon, D. (2009a). Conceptual processing in music as revealed by N400 effects on words and musical targets. *Journal of Cognitive Neuroscience, 21,* 1882–1892.

Daltrozzo, J., & Schon, D. (2009b). Is conceptual processing in music automatic? An electrophysiological approach. *Brain Research, 1270,* 88–94.

Daltrozzo, J., Tillmann, B., Platel, H., & Schon, D. (2010). Temporal aspects of the feeling of familiarity for music and the emergence of conceptual processing. *Journal of Cognitive Neuroscience, 22,* 1754–1769.

Diaz, F., Lindin, M., Galdo-Alvarez, S., Facal, D., & Juncos-Rabadan, O. (2007). An event-related potentials study of face identification and naming: The tip-of-the-tongue state. *Psychophysiology, 44,* 50–68.

Doyle, M. C., Rugg, M. D., & Wells, T. (1996). A comparison of the electrophysiological effects of formal and repetition priming. *Psychophysiology, 33,* 132–147.

Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 202–224.

Filipic, S., Tillmann, B., & Bigand, E. (2010). Judging familiarity and emotion from very brief musical excerpts. *Psychonomic Bulletin & Review, 17,* 335–341.

Galdo-Alvarez, S., Lindin, M., & Diaz, F. (2009). The effect of age on event-related potentials (ERP) associated with face naming and with the tip-of-the-tongue (TOT) state. *Biological Psychology, 81,* 14–23.

Gruneberg, M. M., & Monks, J. (1974). "Feeling of knowing" and cued recall. *Acta Psychologica, 38,* 257–265.

Halpern, A. R., & O'Connor, M. G. (2000). Implicit memory for music in Alzheimer's disease. *Neuropsychology, 14,* 391–397.

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology, 56,* 208–216.

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

Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processing, 5,* 281–312.

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Holt.

Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., et al. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology, 37,* 163–178.

Kikyo, H., & Miyashita, Y. (2004). Temporal lobe activations of "feeling-of-knowing" induced by face-name associations. *Neuroimage, 23,* 1348–1357.

Kikyo, H., Ohki, K., Ishiura, H., & Sekihara, K. (2001). Temporal characterization of memory retrieval processes: A fMRI study of "tip of the tongue" phenomenon. *Neuroimage, 13,* S693.

Koelsch, S., Kasper, E., Sammler, D., Schulze, K., Gunter, T., & Friederici, A. D. (2004). Music, language and meaning: Brain signatures of semantic processing. *Nature Neuroscience, 7,* 302–307.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100,* 609–639.

Koriat, A., & Levy-Sadot, R. (2001). The combined contributions of the cue-familiarity and accessibility heuristics to feelings of knowing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 34–53.

Koriat, A., Metcalfe, J., & Shimamura, A. P. (1994). Memory's knowledge of its own knowledge: The accessibility account of the feeling of knowing. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 115–135). Cambridge, MA: MIT Press.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4,* 463–470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62,* 621–647.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 204,* 203–205.

Kutas, M., & Van Petten, C. (1994). Psycholinguistics electrified. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 83–143). San Diego, CA: Academic Press.

Levitin, D. J., & Tirovolas, A. K. (2009). Current advances in the cognitive neuroscience of music. *Year in Cognitive Neuroscience 2009: Annals of the New York Academy of Sciences, 1156,* 211–231.

Lindin, M., & Diaz, F. (2010). Event-related potentials in face naming and tip-of-the-tongue state: Further results. *International Journal of Psychophysiology, 77,* 53–58.

Lindin, M., Diaz, F., Capilla, A., Ortiz, T., & Maestu, F. (2010). On the characterization of the spatio-temporal profiles of brain activity associated with face naming and the tip-of-the-tongue state: A magnetoencephalographic (MEG) study. *Neuropsychologia, 48,* 1757–1766.

Maril, A., Simons, J. S., Weaver, J. J., & Schacter, D. L. (2005). Graded recall success: An event-related fMRI comparison of tip of the tongue and feeling of knowing. *Neuroimage, 24,* 1130–1138.

Maril, A., Wagner, A. D., & Schacter, D. L. (2001). On the tip of the tongue: An event-related fMRI study of semantic retrieval failure and cognitive conflict. *Neuron, 31,* 653–660.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164,* 177–190.

Mecklinger, A. (2006). Electrophysiological measures of familiarity memory. *Clinical EEG and Neuroscience, 37,* 292–299.

Metcalfe, J., Schwartz, B. L., & Joaquim, S. G. (1993). The cue-familiarity heuristic in metacognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 851–864.

Meyer, P., Mecklinger, A., & Friederici, A. D. (2007). Bridging the gap between the semantic N400 and the early old/new memory effect. *NeuroReport, 18,* 1009–1013.

Näätanen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound—A review and an analysis of the component structure. *Psychophysiology, 24,* 375–425.

Nelson, T. O., Gerler, D., & Narens, L. (1984). Accuracy of feeling-of-knowing judgments for predicting perceptual identification and relearning. *Journal of Experimental Psychology: General, 113,* 282–300.

Nessler, D., Mecklinger, A., & Penney, T. B. (2005). Perceptual fluency, semantic familiarity and recognition-related familiarity: An electrophysiological exploration. *Cognitive Brain Research, 22,* 265–288.

Orgs, G., Lange, K., Dombrowski, J., & Heil, M. (2006). Conceptual priming for environmental sounds and words: An ERP study. *Brain and Cognition, 62,* 267–272.

Orgs, G., Lange, K., Dombrowski, J., & Heil, M. (2007). Is conceptual priming for environmental sounds obligatory? *International Journal of Psychophysiology, 65,* 162–166.

Paynter, C. A., Reder, L. M., & Kieffaber, P. D. (2009). Knowing we know before we know: ERP correlates of initial feeling-of-knowing. *Neuropsychologia, 47,* 796–803.

Peretz, I., Gagnon, L., Hebert, S., & Macoir, J. (2004). Singing in the brain: Insights from cognitive neuropsychology. *Music Perception, 21,* 373–390.

Peretz, I., & Zatorre, R. J. (2005). Brain organization for music processing. *Annual Review of Psychology, 56,* 89–114.

Reder, L. M. (1987). Strategy selection in question answering. *Cognitive Psychology, 19,* 90–138.

Reder, L. M., & Ritter, F. E. (1992). What determines initial feeling of knowing—Familiarity with question terms, not with the answer. *Journal of Experimental Psychology: Learning Memory and Cognition, 18,* 435–451.

Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences, 11,* 251–257.

Rugg, M. D., & Nagy, M. E. (1987). Lexical contribution to nonword-repetition effects—Evidence from event-related potentials. *Memory & Cognition, 15,* 473–481.

Rugg, M. D., & Nieto-Vegas, M. (1999). Modality-specific effects of immediate word repetition: Electrophysiological evidence. *NeuroReport, 10,* 2661–2664.

Rugg, M. D., & Yonelinas, A. P. (2003). Human recognition memory: A cognitive neuroscience perspective. *Trends in Cognitive Sciences, 7,* 313–319.

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,* 793–804.

Schmitter-Edgecombe, M., & Anderson, J. W. (2007). Feeling of knowing in episodic memory following moderate to severe closed-head injury. *Neuropsychology, 21,* 224–234.

Schnyer, D. M., Nicholls, L., & Verfaellie, M. (2005). The role of VMPC in metamemorial judgments of content retrievability. *Journal of Cognitive Neuroscience, 17,* 832–846.

Schon, D., Ystad, Y., Kronland-Martinet, R., & Besson, M. (2009). The evocative power of sounds: Conceptual priming between words and nonverbal sounds. *Journal of Cognitive Neuroscience, 22,* 1026–1035.

Schwartz, B. L. (1994). Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychonomic Bulletin & Review, 1,* 357–375.

Schwartz, B. L. (1998). Illusory tip-of-the-tongue states. *Memory, 6,* 623–642.

Schwartz, B. L. (1999). Sparkling at the end of the tongue: The etiology of tip-of-the-tongue phenomenology. *Psychonomic Bulletin & Review, 6,* 379–393.

Schwartz, B. L. (2008). Working memory load differentially affects tip-of-the-tongue states and feeling-of-knowing judgment. *Memory & Cognition, 36,* 9–19.

Schwartz, B. L., Benjamin, A. S., & Bjork, R. A. (1997). The inferential and experiential bases of metamemory. *Current Directions in Psychological Science, 6,* 132–137.

Schwartz, B. L., & Metcalfe, J. (2011). Tip-of-the-tongue (TOT) states: Retrieval, behavior, and experience. *Memory & Cognition, 39,* 737–749.

Smith, S. M. (1994). Frustrated feelings of imminent recall: On the tip of the tongue. In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 27–46). Cambridge, MA: MIT Press.

Straube, T., Schulz, A., Geipel, K., Mentzel, H. J., & Miltner, W. H. R. (2008). Dissociation between singing and speaking in expressive aphasia: The role of song familiarity. *Neuropsychologia, 46,* 1505–1512.

Valentine, T., Brennen, T., & Bredart, S. (1996). *The cognitive psychology of proper names: On the importance of being Ernest*. London: Routledge.

Voss, J. L., Lucas, H. D., & Paller, K. A. (2010). Conceptual priming and familiarity: Different expressions of memory during recognition testing with distinct neurophysiological correlates. *Journal of Cognitive Neuroscience, 22,* 2638–2651.

Voss, J. L., & Paller, K. A. (2006). Fluent conceptual processing and explicit memory for faces are electrophysiologically distinct. *Journal of Neuroscience, 26,* 926–933.

Voss, J. L., & Paller, K. A. (2007). Distinguishing explicit memory and conceptual implicit memory electrophysiologically. *Psychophysiology, 44,* S4.

Voss, J. L., & Paller, K. A. (2008). Brain substrates of implicit and explicit memory: The importance of concurrently acquired neural signals of both memory types. *Neuropsychologia, 46,* 3021–3029.

Voss, J. L., & Paller, K. A. (2009). Remembering and knowing: Electrophysiological distinctions at encoding but not retrieval. *Neuroimage, 46,* 280–289.

Wilson, S. J., Parsons, K., & Reutens, D. C. (2006). Preserved singing in aphasia: A case study of the efficacy of melodic intonation therapy. *Music Perception, 24,* 23–35.

Wolpaw, J. R., & Penry, J. K. (1975). Temporal component of auditory evoked-response. *Electroencephalography and Clinical Neurophysiology, 39,* 609–620.

Woods, D. L., Alain, C., Covarrubias, D., & Zaidel, O. (1995). Middle latency auditory-evoked potentials to tones of different frequency. *Hearing Research, 85,* 69–75.