Learning lyrics: To sing or not to sing?

AMÉLIE RACETTE AND ISABELLE PERETZ
Université de Montréal, Montréal, Québec, Canada

According to common practice and oral tradition, learning verbal materials through song should facilitate word recall. In the present study, we provide evidence against this belief. In Experiment 1, 36 university students, half of them musicians, learned an unfamiliar song in three conditions. In the sung–sung condition, the song to be learned was sung, and the response was sung too. In the sung–spoken condition, the response was spoken. In the divided–spoken condition, the presented lyrics (accompanied by music) and the response were both spoken. Superior word recall in the sung–sung condition was predicted. However, fewer words were recalled when singing than when speaking. Furthermore, the mode of presentation, whether sung or spoken, had no influence on lyric recall, in either short- or long-term recall. In Experiment 2, singing was assessed with and without words. Altogether, the results indicate that the text and the melody of a song have separate representations in memory, making singing a dual task to perform, at least in the first steps of learning. Interestingly, musical training had little impact on performance, suggesting that vocal learning is a basic and widespread skill.

The notion that music may serve as a mnemonic technique for learning verbal material has a long history. Minstrels transmitted stories through songs (Calvert & Tart, 1993; Rubin, 1995), and this practice is still influential today. Among the most familiar experiences of musical learning are jingles for brand names and the alphabet song children learn. Other examples that have been described consist of learning the laws of physics through karaoke (Dickson & Grant, 2003) and learning English as a second language via songs (Medina, 1993). The goal of the present study was to contribute to the understanding of this phenomenon from both empirical and theoretical perspectives.

Indeed, it is not obvious why music should facilitate word recall, since there is more to learn in a song than in a text. To our surprise, this simple notion has not been properly assessed. Song learning is typically assessed through written recall (Kilgour, Jakobson, & Cuddy, 2000; McElhinney & Annett, 1996; Wallace, 1994). This change in format between perception and performance introduces a bias in word recall in favor of the spoken version, because extracting words from the sung version requires filtering out the music component. Moreover, written recall requires participants to perform a task that is not familiar to them. Lyrics are typically learned to be sung, not to be written. Thus, a putative advantage of singing over reciting words should be assessed with a vocal response. To our knowledge, this procedure has been used only once (Jellison & Miller, 1982), and the results were negative: Music was found to interfere with digit recall and had no effect on word recall. However, in this experiment the words were unrelated and probably were not optimally aligned to the music, hence introducing an additional difficulty.

Thus, to properly test the idea that music may serve as a mnemotechnique for recalling words, one must not only examine oral responses, but also select material in which the words are appropriately set to the music (Gingold & Abravanel, 1987); in short, one must use real songs. This was done in the present study.

Although an adequate test of the idea that music facilitates text recall requires consideration of both input and output factors, the influence of music on word recall starts at the encoding stage. Thus, all prior studies that used written recall but looked at input factors may shed light on the idea that sung words are easier to encode than spoken words. Support for this notion is mixed. In several studies, participants recalled as many sung as spoken words (Gingold & Abravanel, 1987; Wolfe & Hom, 1993) or even did worse on sung material (Calvert & Billingsley, 1998). Yet, in many other studies, an advantage of sung over spoken presentation has been shown (Calvert & Tart, 1993; Chazin & Neuschatz, 1990; Kilgour et al., 2000; McElhinney & Annett, 1996; Rainey & Larsen, 2002; Wallace, 1994; Wolfe & Hom, 1993).

This advantage of sung over spoken text at encoding has been attributed to speed (Kilgour et al., 2000) and to melody simplicity (Wallace, 1994). In effect, words are pronounced more slowly when singing than when speaking. When the sung version of a text is compressed to match its spoken duration, there is no longer a difference in recall, suggesting that the slower rate of singing in comparison with speaking is a key variable in song learnability (Kilgour et al., 2000). Similarly, in order for a sung text to be recalled better than one that is recited, it has to be presented on a simple and repeated melody, as typically found in songs. Lyrics that are sung to a complex and

I. Peretz, isabelle.peretz@umontreal.ca

changing melody can be more difficult to remember than their spoken version (Wallace, 1994).

Songs also possess structural characteristics that may assist text recall. For instance, the metrical structure of music and the number of musical notes in a line can cue word recall. Similarly, song lyrics are usually constrained by both semantics (a story underlies the words, generally through a schema or a script) and sound patterns (e.g., rhymes, alliteration), which may again limit the possibilities. Indeed, when errors occur in song recall, the changes usually preserve the rhyme (Rubin, 1995) and the number of syllables in the line (Wallace, 1994).

Nevertheless, as mentioned previously, texts of real songs are not systematically memorized better when sung than when recited (see, e.g., Gingold & Abravanel, 1987, a study with children; Wallace, 1994, Experiment 3; Wolfe & Hom, 1993). This lack of consistency might be related to the mode of response, as has been pointed out previously. Writing down or reciting the words requires filtering them out from the music. This filtering process might be difficult, especially when words are sung at high pitches (Scotto Di Carlo & Germain, 1985). One way to control for this perceptual disparity between sung and spoken presentations is to present the spoken lyrics accompanied by music. We refer to this situation as the "divided song" condition. By adding the musical background, this condition also maintains the presence of the melody at encoding. None of the prior studies that aimed at testing the effect of music on word recall have included such a control condition. Finally, in order to promote the use of musical cues as a structural aid in the retrieval process, one needs to assess sung recall.

Consideration of all these factors is not solely motivated by experimental elegance, since the contribution of music to verbal memory is a theoretically important question. As alluded to previously, text and melody are aligned in songs in such a way that they promote binding of speech and musical sounds at multiple levels of processing. These tight relations may enhance memory for relatively distinct representations of text and melody in songs by linking elements of words and tones in rich, multiple-linked representations (Peretz, Radeau, & Arguin, 2004). Alternatively, the text and melody of songs might be integrated in a unitary representation, especially when singing is required. Central to the distinction between these two positions is a difference in the way recall is assumed to operate. If integrated, a part of the song's representation will reinstate the whole—namely, singing the melody will reinstate the text. If separate, a part (the melody) may or may not connect with the other part (the text), depending on the strength of the links. Thus, the integrated view of song memory would predict superior text recall in singing over speaking, whereas a separate-memory view of song components would not.

The idea that melody and text might be represented in a unitary memory trace has been relatively neglected in performance, but it has been studied in perception and memory. The prevailing paradigm in the field involves the recognition of unrelated song lines (Crowder, Serafine, & Repp, 1990; Morrongiello & Roes, 1990; Peretz, Radeau,

& Arguin, 2004; Samson & Zatorre, 1991; Serafine, Crowder, & Repp, 1984; Serafine, Davidson, Crowder, & Repp, 1986). In the recognition of song lines, melody and text appear to be highly associated, even after a single hearing, suggesting that lyrics and melody representations are integrated in memory for songs (Serafine et al., 1984; Serafine et al., 1986). However, there is increasing evidence that the music and language components of songs maintain autonomy in both perception (Besson, Faïta, Peretz, Bonnel, & Requin, 1998; Bonnel, Faïta, Peretz, & Besson, 2001) and memory (Crowder et al., 1990, Experiment 3; Peretz, 1996). Very recently, we extended these conclusions to singing by studying brain-damaged patients who suffered from a severe speech disorder without a concomitant musical disorder (Hébert, Racette, Gagnon, & Peretz, 2003; Peretz, Gagnon, Hébert, & Macoir, 2004; Racette, Bard, & Peretz, 2006). The results indicate that verbal production, be it sung or spoken, is mediated by the same (impaired) language output system and that this speech route is distinct from the (spared) melodic route. These neuropsychological findings strongly suggest that singing taps into distinct codes for melody and text. Thus, the present study should help us to shed further light on this issue by testing singing in the normal population.

The general population is musically untrained. However, we also considered a group of professional musicians because these individuals might exploit musical cues more effectively than nonmusicians, and therefore might benefit more from the presence of music on text recall. Moreover, musicians seem to have better verbal memory than nonmusicians (Chan, Ho, & Cheung, 1998; Jellison & Miller, 1982; Kilgour et al., 2000), apparently from childhood (Ho, Cheung, & Chan, 2003). Thus, it is possible that musical training strengthens auditory temporal processing, which would mediate verbal recall (Jakobson, Cuddy, & Kilgour, 2003; Jellison & Miller, 1982). These results in turn suggest that music may assist in text recall, but only in those individuals who regularly use the two codes.

Therefore, in the present study 36 students, half with musical expertise, had to learn novel songs in three different conditions. As illustrated in Table 1, the text to be learned was either sung or spoken. When spoken, its corresponding melody was sung on /la/ in the background. Recall of text was either sung (on the melody) or spoken (lyrics alone). We predicted that word recall would be superior in the sung–sung condition, especially for musicians, simply because singing is slowed down relative to normal speech. The sung–spoken condition was expected to be the most difficult, because in this condition the text needs to be extracted from the song. In the divided–spoken condition, there would be no cost of extracting the words,

Table 1
Modes of Presentation and Recall
in the Three Conditions of Experiment 1

Presentation of the Song	Recall of the Lyrics
Sung	Sung
Sung	Spoken
Spoken (divided)	Spoken

but neither would there be an advantage from hearing them sung (at a slow speed).

EXPERIMENT 1

Method

Participants

Thirty-six French-speaking university students (mean age, 25; range, 20–37) who felt comfortable singing were recruited. Half of the participants (14 women, 4 men) were considered nonmusicians, with a mean of 2 years of music training before the end of high school. The other half (14 women, 4 men) were students in the music faculty or professional musicians. This second group had on average 13.6 years of musical training: 5 were singers, 5 had singing as their second instrument, and 8 had no formal training in singing besides *solfeggio*. Participants were paid for their participation.

Materials

Unfamiliar songs were chosen from the repertoire of a popular French-Canadian folksinger, author, and composer (Claude Gauthier). Six songs with few word or melodic line repetitions were selected. From these, eight-line excerpts were chosen for the learning task (see the Appendix for an example). Each line carried, on average, 6 words and 8 notes. Thus, on average, a song contained 49 words (range, 45–57) and 68 notes (range, 64–74). An additional eight-line excerpt from an unfamiliar choir song by Johann Steuerlein served as a training song.

The six song excerpts from Claude Gauthier were considered to be "good" songs, as assessed by 7 pilot participants who were unfamiliar with the singer. The judges were presented with the six song excerpts in their original version, and these were randomly mixed with excerpts of six hit songs by the same folksinger. Each excerpt was presented twice in a random order. For each song excerpt, the judges rated its musicality, simplicity, and potential to be a hit on three 6-point scales, in which 1 meant *poor* and 6 *excellent*. Very similar ratings were obtained for the hits and the experimental songs on each dimension (3.7 and 3.8, respectively, for musicality, 3.4 and 3.5 for simplicity, and 3.2 and 3.4 for hit potential, with first and second ratings pooled together), supporting the idea that the selected material corresponded to well-formed songs.

The six songs and the training song were produced *a cappella* (without instrumental accompaniment) by a female singer, who learned the songs beforehand. The same singer also sang each song on /la/ and pronounced the lyrics with a natural intonation. The best performance of each song in each version was recorded on a DAT Sony via a Shure 565SD microphone, and then transferred into a computer and edited with the Cool Edit program (Syntrillium Software, 1996). The three versions of the same song served to create two types of stimuli, the sung songs and the "divided" songs. The latter were created by coupling each spoken line with its corresponding melody sung on /la/. In these divided songs, the intensity of the melody had to be decreased by 32%, on average, in order to make the spoken version intelligible. The intelligibility of songs' lyrics in the sung and the divided versions was verified by measuring the number of errors made by the participants when repeating the lines in a spoken mode immediately after hearing them. Mean correct repetitions were 94.8% and 95.3% for the sung and divided presentations, respectively [$t(35) = 0.40$, $SE = 1.46$, n.s.], suggesting that understanding the words from the sung and divided versions of the songs was equally easy.

As expected, the duration of the spoken version was about 43% shorter than the sung version [$M = 2.51$ and 4.40 sec per line, respectively; $t(47) = 9.60$, $SE = 0.20$, $p < .01$]. Because the divided condition combined both the spoken and the "sung on /la/" versions ($M = 4.42$ sec per line), divided and sung presentations had equivalent durations ($M = 4.46$ and $M = 4.40$ sec per line, respectively) [$t(47) = 1.58$, $SE = 0.04$, n.s.]. In the divided condition, the shorter spoken line was placed in the middle of the sung melody, so that it was preceded and followed by equivalent durations of the melody.

Song Analysis

The six songs used in the present study had a theme, most often referring to love. The words used were thus predictable (e.g., *amour* "love," *fleurs* "flowers," *cœur* "heart"), but they were different enough across songs to prevent confusions. Themes diverging from love were related to music (one song) and to patriotism (one song). Frequency of usage of the song words was high. The 155 different words used in the songs had a mean frequency of 2,057 per million, based on a French lexical database (New, Pallier, Ferrand, & Matos, 2001): 72% were highly frequent, with a frequency of usage higher than 50 per million. Only 13% of the words had a low frequency, corresponding to less than 15 per million. Rhymes were present in all of the six songs except one, which only rhymed in the second half (see the Appendix). In all songs, there was a one-to-one mapping between syllables and tones, with each syllable coupled to a single note. There were no ties (see the Appendix). However, notes outnumbered words [$t(47) = 13.49$, $SE = 0.17$, $p < .001$], because even if most of the words (69%) were monosyllabic, 25% were disyllabic, and 6% were trisyllabic.

Regarding musical structure, the six songs had stable and standard meters. Four of the six songs had a duple meter (4/4), and the other two, a triple meter (3/4 and 6/8). All of the songs were in major mode and were written in the key of Ab, C, D, G, or F, and two songs contained a single modulation. Even if melodies were chosen for their diversity, the melody parts were highly coherent within a song (see the Appendix). Lines respected the grouping preference rules proposed by Lerdahl and Jackendoff (1983), which are mainly based on gestalt principles, such as symmetry, proximity, similarities, and parallelism. Thus, the melodic lines did not contain the same absolute pitches or the same exact rhythm, but they were structurally similar in a given song.

The alignment between the prosody of the text and the rhythm of the melody conformed to the rules of French songs (Dell, 1989). The last accentuated syllable of the verse coincided with a strong beat. This constraint was present in each of the eight lines of the songs. Note that, in French, there is no prescriptive rule regarding the alternation between strongly and weakly stressed syllables. Hence, whether the prosodic accents in French songs coincide with musical metrical accents, as found in English songs, remains debatable (Palmer & Kelly, 1992).

Procedure

Because pilot observations suggested high variability in the number of lines participants were able to recall, an adaptive procedure was used. This procedure is represented in Table 2. Participants first heard the whole song excerpt once, in order to familiarize themselves with it. Afterward, the first line was presented, and they had to repeat it. Lines were then repeated two by two. Once the first four lines were repeated, participants were asked to recall the four lines from the beginning without hearing them again. If more than 80% of the words were recalled, Lines 5 and 6 were presented and repeated before participants were again asked to recall every line starting at the beginning. If more than 80% of the words were again recalled, Lines 7 and 8 were presented, and participants made a last recall of the entire excerpt. Number of words correctly produced was calculated online by the experimenter so as to decide whether the procedure was to be continued or stopped.

Presentation of the song lines was either sung or spoken with the melody in the background. Repetition was either sung or spoken (see Table 1). In the sung–sung condition, participants listened to the sung version of the lyrics and sung them back. In the sung–spoken condition, participants listened to the sung version of the lines and were asked to recall only the lyrics and recite them in a natural way. In the divided–spoken condition, they listened to the divided version of the lines and recalled only the lyrics in a similarly natural manner.

Each participant learned one song in each condition, for a total of three different songs. The order of the conditions was counterbalanced across participants according to a Latin square. Care was taken to test each song in each condition across participants in each group.

Table 2
Illustration of the Adaptive Learning Procedure

Lyrics Presented	Lyrics Repeated	Lyrics to Be Recalled
1 Dans cette petite boîte vide	1 Dans cette petite boîte vide	1 Dans cette petite boîte vide
		2 Avec un ruban de velours
1 Dans cette petite boîte vide	1 Dans cette petite boîte vide	3 Il y a tout mon cœur et mes rides
2 Avec un ruban de velours	2 Avec un ruban de velours	4 Mon sourire et tout mon amour
3 Il y a tout mon cœur et mes rides	3 Il y a tout mon cœur et mes rides	
4 Mon sourire et tout mon amour	4 Mon sourire et tout mon amour	If less than 80% of words recalled, stop. If 80% or more words recalled, continue.
5 Il n'y a pas d'argent qui remplace	5 Il n'y a pas d'argent qui remplace	1 Dans cette petite boîte vide
6 Tout le temps que l'on peut donner	6 Tout le temps que l'on peut donner	2 Avec un ruban de velours
		3 Il y a tout mon cœur et mes rides
		4 Mon sourire et tout mon amour
		5 Il n'y a pas d'argent qui remplace
		6 Tout le temps que l'on peut donner
		If less than 80% of words recalled, stop. If 80% or more words recalled, continue.
7 À tous ceux que l'on aime hélas	7 À tous ceux que l'on aime hélas	1 Dans cette petite boîte vide
8 Trop souvent qu'on oublie d'aimer	8 Trop souvent qu'on oublie d'aimer	2 Avec un ruban de velours
		3 Il y a tout mon cœur et mes rides
		4 Mon sourire et tout mon amour
		5 Il n'y a pas d'argent qui remplace
		6 Tout le temps que l'on peut donner
		7 À tous ceux que l'on aime hélas
		8 Trop souvent qu'on oublie d'aimer

The practice song was learned before each condition in the corresponding version. Participants were asked to do their best to recall the exact words and, if they did not remember a part, to report whatever came to mind. The participants listened to digital recordings through speakers, and their performance was recorded on a Sony DAT.

In order to assess verbal memory independently from song memory, the Rey Auditory Verbal Learning Test (RAVLT; Rey, 1964) was administered after participants had learned the three songs. In this task, they had five trials to recall a list of 15 unrelated words. The RAVLT also served as a distraction task. Afterward, participants were asked to make a written recall[1] of the three songs they had previously learned (excluding the practice song). The time elapsed between recalls was approximately 20 min. This delayed recall came as a surprise test, because participants were not warned in advance that their memory would be assessed one more time. Since music may help long-term memory (Rainey & Larsen, 2002), 25 of the 36 participants were contacted 7 months later (2–10 months after the first administration) and asked again for vocal song recall, which was recorded on tape.

Data Scoring

For text recall, words were considered correct or incorrect, irrespective of their pitch and duration when sung. Words were chosen over syllables as the criterion because number of syllables sometimes differed across conditions; mute vowels are often sung but not pronounced. The words had to be produced in the correct order to obtain a point. Omissions and substitutions received no points. A point was lost when words were added, and half-points were subtracted when words were mispronounced but recognizable. In cases in which participants made an error because they misperceived a word and did not repeat it correctly when first heard, the repeated version of the word was considered correct in recall. Finally, a point was lost if the correct word was spoken instead of sung, and vice versa. The total numbers of words correctly reproduced in the last immediate recall and in the delayed recall were then divided by the number of words contained in the lines to be recalled by a given participant, thus taking into consideration the differing number of lines that participants were recalling. This ratio was multiplied by 100 to obtain a percentage.

The raw number of words recalled and the number of lines (4, 6, or 8) attempted in immediate recall were also taken into account. Number of hesitations, defined as a marked pause or a corrected attempt (the participant tried something and then changed her/his answer), was also noted. Finally, the locations of breaths were recorded.

In the sung mode of recall, the musical notes in each final recall were transcribed by two independent musicians. The agreement between the judges was very low for rhythm, and therefore rhythm was not considered in the present study. Instead, pitch intervals and directions were analyzed. The number of correct notes was defined as the number of notes both judges gave a point to. When there was a disagreement (in 15% of the cases—i.e., for 228 out of 1,559 notes produced), the note was discarded. Thus, the score corresponded to the number of correct pitches divided by the total number of possible notes minus the notes both raters disagreed upon, multiplied by 100.

Results

Performance in the immediate recall of the song was first examined by considering the percentages of words that were correctly sung and spoken after presentation of sung and recited songs. The number of lines completed, the total number of words recalled, the position of the forgotten lines in the song, the types of errors made by the participants, and pitch accuracy were also analyzed. Word recall was also examined as a function of learning condition both after a delay of 20 min and after several months. Finally, performance in lyrics learning was compared with performance in the Rey Auditory Verbal Learning Test.

Immediate Recall

Correct words. An initial repeated measures ANOVA, with both group (musician, nonmusician) and order of

Table 3
Means (and Standard Errors) Obtained in Each Condition
on Immediate Recall in Experiment 1

Group	Dependent Variable	Sung–Sung M	Sung–Sung SE	Sung–Spoken M	Sung–Spoken SE	Divided–Spoken M	Divided–Spoken SE
Nonmusicians	%	63.9	4.8	73.0	3.5	74.0	3.2
	Words	20.4	2.9	22.7	2.3	27.9	2.5
	Lines	5.0	0.4	5.0	0.4	6.0	0.4
Musicians	%	56.1	4.3	68.9	3.0	73.9	3.6
	Words	18.7	2.3	23.7	2.3	27.1	3.0
	Lines	5.2	0.4	5.6	0.4	5.8	0.4
Mean	%	60.0	3.2	70.9	2.3	74.0	2.4
	Words	19.6	1.8	23.2	1.6	27.5	1.9
	Lines	5.1	0.4	5.3	0.4	5.9	0.4

presentation (1, 2, 3) as between-subjects variables and condition (sung–sung, sung–spoken, divided–spoken) as a within-subjects variable, was performed on the percentage of words recalled. Since there was neither an effect of order [$F(2,30) = 1.19$, $MS_e = 175.84$, $p > .05$] nor an interaction between order and the other factors, order was not considered in the following analyses.

In Table 3, performance is expressed in percentage of words recalled, as well as in terms of the total number of words recalled and the number of lines attempted. As can be seen, recall appears more difficult when participants have to sing, regardless of their musical background. This was supported by an ANOVA performed on percentage of words recalled, with condition (sung–sung, sung–spoken, divided–spoken) as the within-subjects variable and group (musician, nonmusician) as the between-subjects variable. The ANOVA revealed a main effect of condition [$F(2,68) = 11.78$, $MS_e = 165.78$, $p < .001$] and no group effect ($F < 1$) or interaction between condition and group ($F < 1$). Post hoc Tukey tests revealed that recall did not differ in the two spoken-recall conditions ($p > .05$) and was significantly better than sung recall ($p < .01$).

The superiority of spoken recall was also apparent when the other measures were considered. When the raw number of correct words was considered as the dependent variable, the main effect of condition [$F(2,68) = 7.34$, $MS_e = 78.14$, $p < .01$] also reached significance. There was no group effect nor an interaction with condition ($Fs < 1$). When considering the number of attempted lines (4, 6, or 8 lines), a similar trend was observed, since in this respect performance in the spoken conditions was also superior to that in the sung condition [$F(2,68) = 2.83$, $MS_e = 2.14$, $p = .07$]. Moreover, as shown in Figure 1, nonmusicians could learn as much of the songs as musicians when singing.

An aspect that is worth examining is serial recall. The beginning of a song acts as an anchor point for the whole song. This refers to the fact that the beginning of a sequence is a determinant for the recall of the sequence in question (Peretz, Radeau, & Arguin, 2004). Therefore, recall of the first song lines should be best, and recall should decline as the song progresses. Because the first line was presented twice to participants, whereas the subsequent lines were presented only once, we considered the recall of the second line, which was forgotten by only 17% of

the participants, and compared it with recall of the last line. The last line was not recalled by 42% of the participants while singing, but by only 18% while reciting. This difference was significant [$Q(2) = 8.78$, $p < .05$, using Cochran's test]. Thus, serial position of the line appears to be more important in singing than in speaking. Moreover, forgetting an entire line was more frequent in sung recall (24% of the lines) than in the spoken recalls (11%) [$F(2,70) = 8.25$, $MS_e = 0.02$, $p < .01$]. When a line was omitted in singing, the next line was omitted in 71% of the cases. In contrast, when a line was omitted in reciting, only 55% of the following lines were missed. This suggests that text recall in singing is more strictly sequential, because it appears to be more dependent on the serial order of information than is reciting.

Word errors. Types of errors are useful in determining the nature of the memory code used by participants. For example, one word can be substituted for another in order to preserve the song line structure, and this type of error would be expected to occur more often while singing than while speaking. Indeed, words were often replaced by a word with the same number of syllables (e.g., "Je t'écris *cet' lettr'* par amitié" for "Je t'écris *ces mots* par amitié").
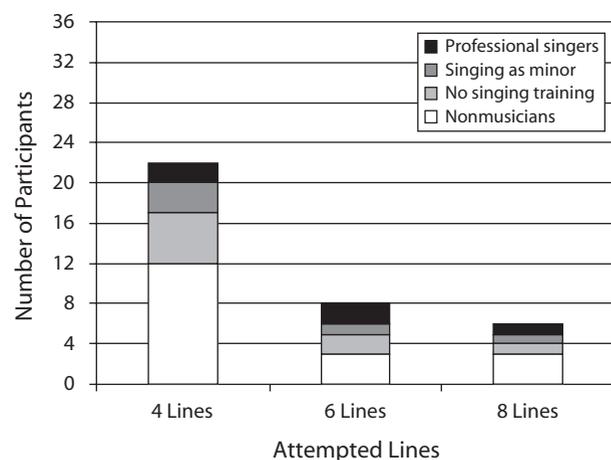

Figure 1. Number of nonmusicians and musicians reaching each level of song line recall in Experiment 1, as a function of their singing experience. Nonmusicians are represented in white and musicians in gray shades.

Table 4
Mean Percentages of Errors (and Standard Errors) in Each Condition
As a Function of Line Structure

Type	Line Structure	Sung–Sung		Sung–Spoken		Divided–Spoken	
		M	*SE*	*M*	*SE*	*M*	*SE*
Omissions	Preserved	13	4	11	5	9	4
	Altered	10	3	21	4	24	5
Substitutions	Preserved	20	4	27	4	22	4
	Altered	10	3	12	3	20	4

Similarly, in singing, when a word was omitted, participants could replace it by a meaningless syllable (/na/) in order to preserve line structure. These omissions and substitution errors were assessed with respect to the number of syllables in the line. If a match was found, the line structure was considered preserved. When the number of syllables did not match, the line structure was considered altered. The results of this analysis are presented in Table 4. Other types of errors, such as the addition of words (2% of total errors) or pronunciation errors (0.2%) were too rare to be examined.

As can be seen in Table 4, errors tended to preserve the line structure, especially in singing. An ANOVA with condition (sung–sung, sung–spoken, divided–spoken), type of errors (omission, substitution), and line structure (preserved, altered) as within-subjects variables and group (musician, nonmusician) as a between-subjects variable yielded an interaction between type of error and line structure [$F(1,34) = 8.96$, $MS_e = 0.08$, $p < .01$]. As expected, the omission of words tends to alter the line structure [18% vs. 11%; $t(35) = 1.90$, $SE = 0.04$, $p = .07$], whereas substitutions more often preserved it [23% vs. 14%; $t(35) = 2.95$, $SE = 0.03$, $p < .01$]. This pattern was not significantly affected by singing, since the interaction with condition was not significant ($F < 1$). There was no group effect, nor any interaction between group and any other variables. In addition, the substituted words were semantically related to the target (67% of the words), thus keeping the gist of the line (e.g., "si jamais vous *trouvez* cet *homme*"/"if you ever *find* this *man*" instead of "si jamais vous *tenez* cet *âme*"/"if you ever *hold* this *soul*"). Thus, participants tried to respect both the number of syllables and the meaning of words in their recall of lyrics, regardless of the mode of vocal reproduction.

Another factor that is known to enhance memory of lyrics is the presence of rhymes at the end of lines. In order to assess the contribution of rhyme, we examined word errors as a function of their serial position in the line. The final words of each line—that is, those bearing the rhyme—were incorrectly reproduced in only 15% of the lines (with 19% and 12.5% in singing and reciting, respectively). This error rate was smaller than the one observed for any prior position in the line (e.g., the error rate was 20% for the initial word of the line; $t(595) = 2.92$, $SE = 0.02$, $p < .01$). Moreover, when the last word was replaced by another word, it respected the rhyme in 39% of the cases (e.g., *tour* for *jour*).

In order to assess fluency, the number of hesitations per line was examined in each condition (see Table 5). As can

be seen, the amount of hesitations was equal for musicians in the singing and speaking conditions, but nonmusicians clearly made fewer hesitations when singing. The interaction between condition and group was close to significance [$F(2,68) = 2.88$, $MS_e = 0.03$, $p = .06$].

Finally, participants generally took a breath between lines (75%) instead of during a line. While singing, 47% of them took a breath after each line. While reciting, breaths were often taken after two or three lines. Indeed, more spoken than sung words can be produced in a single breath.

Notes. In the sung–sung condition, nonmusicians correctly sang 36% of the notes ($SE = 7.8$) and 65% of the words ($SE = 4.8$), whereas musicians correctly sang 48% of the notes ($SE = 7.2$) and 56% of the words ($SE = 4.3$). An ANOVA with material (word, note) as a within-subjects variable and group (musician, nonmusician) as a between-subjects variable revealed an interaction between material and group [$F(1,34) = 4.61$, $MS_e = 413.43$, $p < .05$]. Whereas nonmusicians recalled more words than notes ($p < .01$ using a post hoc Tukey test), musicians did not. Interestingly, musicians also did not reproduce more correct pitches than did nonmusicians (n.s.). When the total numbers of correct notes ($M = 13.1$, $SE = 2.83$, for nonmusicians and $M = 19.6$, $SE = 4.17$, for musicians) and words were examined instead of the proportions of correct notes and words (see Table 3), there was no effect of material [$F(1,34) = 1.85$, $MS_e = 97.63$, $p > .05$] nor any group effect ($F < 1$), but the interaction was again close to significance [$F(1,34) = 3.03$, $p = .09$].

Furthermore, there was no significant correlation between note and word recall, either in nonmusicians [$r(16) = .38$, n.s.] or in musicians [$r(16) = .27$, n.s.].[2]

Delayed recall and long-term retention. Recall after a 20-min delay is presented in Table 6. As can be seen, performance dropped by half. Moreover, word recall appeared to persist longer after a divided–spoken presentation. However, this trend was not significant, as revealed by an ANOVA with condition (sung–sung, sung–spoken,

Table 5
Mean Percentages of Hesitations per Line (and Standard
Errors) in Each Condition for Each Group

Group	Sung–Sung		Sung–Spoken		Divided–Spoken	
	M	*SE*	*M*	*SE*	*M*	*SE*
Nonmusicians	6	3	26	5	21	4
Musicians	14	3	17	3	15	5
Mean	9	3	21	4	18	5

Table 6
Mean Percentages of Correctly Recalled Words (and Standard Errors) in the Three Conditions After a 20-Min Delay and After Several Months (in Italics)

Group	Sung–Sung		Sung–Spoken		Divided–Spoken	
	M	SE	M	SE	M	SE
Nonmusicians	36.7	6.0	26.0	7.2	38.1	5.5
	2		*12*		*13*	
Musicians	25.8	6.7	32.3	6.6	42.9	7.9
	14		*5*		*32*	
Mean	31.2	4.5	29.1	4.9	40.5	4.7

divided–spoken) as the within-subjects variable and group (musician, nonmusician) as the between-subjects variable. The effect of condition failed to reach significance [$F(2,68) = 1.88$, $MS_e = 701.75$, n.s.], and there was no group effect ($F < 1$) or interaction [$F(2,68) = 1.16$, n.s.].

As can be seen in Table 6, very few song lyrics were remembered several months after they had been learned. The practice song, which was repeated three times during the initial procedure, was the most frequently recalled (by 12 of the 25 participants contacted), but it was sung by only 3 of the 12 participants who recalled it.

RAVLT. In this standard auditory memory test, musicians recalled as many words as nonmusicians (see Figure 2; $F < 1$), and both groups improved across the five trials [$F(4,136) = 161.28$, $MS_e = 1.17$, $p < .001$]. Furthermore, no significant correlation was found between performance on the RAVLT and word recall from songs in any of the learning conditions [all rs(34) $< .29$, $p > .05$].

Discussion

Contrary to expectations, music was found to interfere with rather than facilitate text recall. Participants recalled fewer words when singing than when reciting, despite the fact that words are articulated more slowly when sung. Moreover, music did not help the recovery of lines after a memory blank. On the contrary, sung recall seems to be

Figure 2. Mean number of recalled words (and standard errors) on each trial of the Rey Auditory Verbal Learning Test in nonmusicians and musicians.

more strictly sequential than spoken recall, since recall of a line was more dependent on the recall of the previous line in singing than in reciting. Yet there were fewer hesitations when singing. Hence, singing can give an impression of fluency, not because lyrics are better retrieved from memory, but because the flow is more continuous. This observation has also been reported in cases of speech disorders (Hébert et al., 2003). This apparent fluency was limited to the musically untrained; musicians did not hesitate less while singing than while reciting.

The fact that music did not help text recall cannot be ascribed to the fact that ordinary students are not used to singing. Musicians, including singers, did not perform differently from, and recalled as many lyrics as, nonmusicians. Hence, the results suggest that oral recall of lyrics is a widespread ability.

Music not only slightly impaired vocal production, it also had little impact at presentation. There was no difference in word recall between the sung and the divided presentations. Furthermore, there was not the slightest indication that music helped in the long run, since the addition of music during song presentation or during participants' responses had no effect on long-term recall either.

Yet, all the characteristics that qualify memory in oral tradition also govern performance in the present study. Recall of lyrics respected line structure, semantics, rhymes, and front anchoring. This form of memory is commonly used for stories and poems (Rubin, 1995). Contrary to expectations, music does not seem to add much to these constraints. Because the focus of the task was on text recall, it is possible that participants treated music as a secondary task, and hence treated it as an additional demand rather than as an aid to memory. The goal of Experiment 2 was to examine this possibility by asking participants to focus on the musical component.

EXPERIMENT 2

In this experiment, we compared the recall of text and melody alone and in combination. We reasoned that, if there was an advantage of text recall over music because the task demands focused on words in Experiment 1, then asking participants to pay attention to music should improve singing in general, particularly for musicians. In contrast, if singing is a dual task, producing the melody with the text should require more resources than producing either the text or the melody alone.

Method

Six musicians (3 women, 3 men) and 6 nonmusicians (5 women, 1 man) who had participated in Experiment 1, for a total of 12 participants (mean age, 23.3; range, 20–26), came back for an additional session 11 months later (range, 5–13 months). This subgroup was selected on the basis of their availability. No professional singers participated in this second experiment.

The participants were presented with the sung–sung and divided–spoken versions of the same six songs that were used in Experiment 1. However, care was taken to present each participant with the three songs that the participant had not learned in Experiment 1. Recall was tested with an adaptive procedure, as in Experiment 1. Each participant once again learned each of the three songs in a
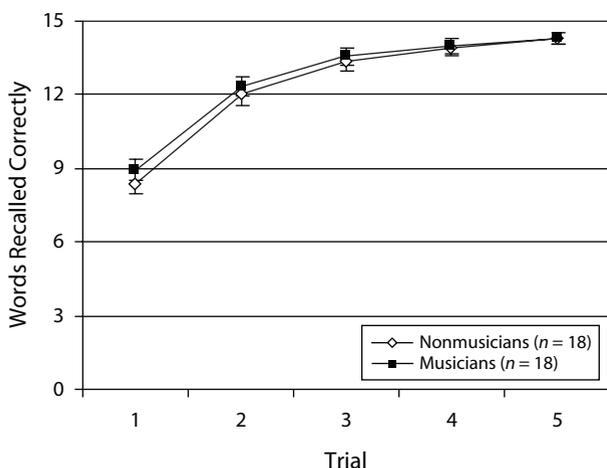
different condition: the sung–sung condition, in which the participant sang the lyrics after hearing them sung; the divided–spoken condition, in which the participant recited lyrics presented spoken along with the melody sung on /la/; and the divided–sung condition, in which the participant only repeated the melody sung on /la/. In the latter condition, participants were asked to recall the full eight lines of the song, because accuracy of note production could not be judged online by the experimenter. There were no other tasks, nor was there a delayed recall.

Data Scoring

The same scoring procedure was used as in Experiment 1. For the sung production, raters agreed on 83% of the 562 notes produced. For the sung melodies on /la/, if the number of notes correctly produced in the first recall (Lines 1–4) was less than 80%, the score was computed over these four lines. If 80% or more notes were recalled, the second recall (Lines 1–6) was rated. If more than 80% of those notes were recalled, scores were based on the last recall (Lines 1–8). The musical note score was computed on the basis of the number of correct notes both raters agreed upon—that is, 89% of the 565 notes produced.

Results and Comments

Because of the small number of participants, nonparametric tests were first performed on the data in order to assess whether nonmusicians' performance differed from musicians' performance. There were no differences between the two groups on word recall (sung–sung and divided–spoken conditions) or note recall (sung–sung and divided–sung conditions; all $ps > .05$ by Mann–Whitney tests). Even singing on /la/ did not significantly differentiate musicians from nonmusicians (see Figure 3; Mann–Whitney test, $Z = .96$, $p > .05$). These two groups did not differ in performance in Experiment 1, either ($p > .05$). Hence, all 12 participants were considered in a single group and parametric analyses were applied.

In Table 7, the *combined* condition refers to the sung production of both words and notes (sung–sung condition), and the *alone* condition refers to the production of words (divided–spoken) or of musical notes (divided–sung on /la/) only. As can be seen in Table 7, text recall

Table 7
Means (and Standard Errors) Obtained for Words and Notes in Experiment 2

| Material | Dependent Variable | Production | | | |
| | | Combined | | Alone | |
		M	SE	M	SE
Words	%	70.6	3.2	77.3	3.6
	Number	25.8	2.8	32.5	2.8
	Lines	5.5	0.5	7.0	0.4
Notes	%	58.8	9.5	50.9	6.4
	Number	23.0	4.1	20.8	4.3
	Lines	5.5	0.5	5.0	0.5

was again worse when combined with music (sung) than when spoken. However, the trend did not reach significance [$t(11) = 1.34$, $SE = 5.00$, $p > .05$]. Participants also tended to learn fewer words and fewer lines in singing than in speaking [$t(11) = 1.37$, $SE = 4.91$, n.s., and $t(11) = 2.02$, $SE = 0.37$, $p = .07$, respectively]. In order to assess the effect of familiarity with the task, the results obtained by the same participants in Experiment 1 were compared with their results obtained here. Task repetition improved singing [$t(11) = 2.81$, $SE = 4.69$, $p < .05$] but not reciting [$t(11) = 1.12$, $SE = 6.70$, n.s.]. Thus, participants seemed more comfortable with the task than in Experiment 1. However, this improvement was not sufficient to bring word recall to a higher level in singing than in speaking. Again, music does not seem to facilitate word recall.

Melody recall was more variable than word recall, both in singing with words and in singing on /la/ (see Table 7). The percentages of notes correctly recalled in singing with and without words did not differ [$t(11) = 1.01$, $SE = 7.76$, n.s.]. In fact, note recall with words was not better than in Experiment 1 ($M = 50.2$, $SE = 9.0$; $t(11) = 0.63$, n.s.). Thus, melody recall generally appears to be poor, whether the focus is on text or music. Participants do not seem to have much flexibility in the quantity of resources they can allocate to the musical component.

The majority of participants (9) did not go farther than the fourth line in singing on /la/ (see Figure 3), whereas the majority (7) reached the end of the song when reciting lyrics. The performances in word and note recall were not significantly correlated, whether produced together in singing [$r(10) = .16$, $p > .05$] or produced alone [$r(10) = .30$, $p > .05$].

As in Experiment 1, there were more words than notes correctly recalled in singing. In an ANOVA with material (word, note) and production (combined, alone) as within-subjects variables and percentage as a dependent variable, a main effect of material was obtained, with more words produced than notes [$F(1,11) = 6.11$, $MS_e = 720.18$, $p < .05$]. There was no effect of the mode of production [$F(1,11) = 3.89$, $MS_e = 284.39$, $p > .05$], nor was there an interaction [$F(1,11) = 2.79$, $MS_e = 227.23$, $p > .05$]. The same effects were obtained with the total number of words and notes used as the dependent variables.

In summary, the results are similar to those obtained previously, indicating a slight advantage of speaking over singing in text recall. This advantage of text over music
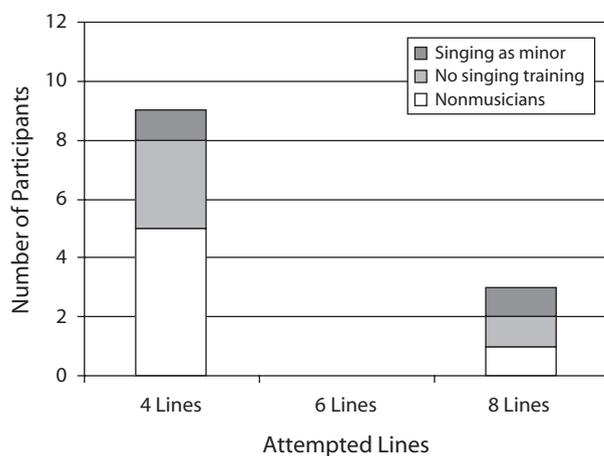

Figure 3. Number of nonmusicians and musicians reaching each level of melody line recall, as a function of their singing experience, in Experiment 2.

does not seem related to a trade-off between the two components. Accuracy in singing the melody was similar whether it carried lyrics or not. Furthermore, there was no correlation between words and notes recalled, suggesting that these two components are supported by separate memory representations.

GENERAL DISCUSSION

The present findings suggest that the best strategy for learning song lyrics is to ignore the melody. The melody seems to interfere rather than facilitate word recall in songs in both musically trained and untrained learners. Music was found to be of little help for text recall in either encoding or response. Hearing the lyrics embedded in the melody (i.e., sung) or spoken with the melody in the background did not affect word recall, even after a time delay (Experiment 1) and after task familiarization (Experiment 2). The same conclusion applies to the mode of expression: Performance at reproducing both the lyrics and the melody while singing was either impaired (Experiment 1) or slightly inferior (Experiment 2) to the recall of the text alone. Melody recall was generally less precise than word recall, whether it was sung with the lyrics (Experiments 1 and 2) or on /la/ (Experiment 2), in both musicians and nonmusicians. Thus, the results suggest that in the first steps of learning a new song, melody and lyrics are remembered separately, making singing a dual task.

The cost of singing was reflected by a 14% word loss (Experiment 1), but it was associated with an 8% increase in the recall of notes (Experiment 2); the cost was reliable, but the benefit was not. This cost–benefit analysis is more compatible with the view that the melody and lyrics of songs are processed independently (Besson et al., 1998; Bonnel et al., 2001; Hébert & Peretz, 2001; Peretz, 1996) rather than treated as an integrated unit (see, e.g., Serafine et al., 1984). Thus, the present results extend to singing what has been found in the normal functioning of perception and memory (see Peretz, Radeau, & Arguin, 2004, for a recent review and discussion).

However, separate production of melody and lyrics does not entail interference, unless attention to one component adversely affects the other. In the present case, it seems that lyric recall was either prioritized or much easier than note recall. Such a discrepancy between the processing of words and notes has repeatedly been found in the literature pertaining to perception of songs, with words always being more salient than musical notes (Hébert & Peretz, 2001; Peretz, Radeau, & Arguin, 2004). There are several factors that can account for this advantage of lyrics over melody. First, the lyrics were organized like a poem, and hence their memorability benefited from the use of several language constraints that are known to help remembering (Rubin, 1995). Semantics, rhymes, and line structure were all found to affect recall, whether recited or sung. In contrast, the melody had no semantics or rhymes, but has rhythm, line structure, and pitch accents. These musical characteristics were instrumental in decreasing hesitations, making singing more fluent, but were not sufficient to give additional assistance to lyric recall. On the contrary, it was observed that when a line was forgotten, participants were usually unable to continue singing, but that they could continue reciting after a break. This might be a drawback of the strictly sequential nature of singing, in which melodic lines are represented in connected strings with front anchoring.

Nevertheless, one important cue for auditory–vocal remembering that is common to both music and poems is rhythm. The regular organization of stresses, mostly alternating between strong and weak beats/syllables, is supposed to limit the words that are compatible with it, and thereby constrains word selection. At least in English, the rhythmic similarity between the prosodic accent structure of spoken words and the metric structure of the melody is striking and has long been noted by linguists (see, e.g., Hayes & Kaun, 1996) and music theorists (Lerdahl & Jackendoff, 1983). Moreover, Palmer and Kelly (1992) have shown that linguistic accent structure and musical meter are generally aligned in Western songs. Hence, rhythmic structure, as determined by the number of syllables (or notes) and the location of primary stress, may serve as a compatible format for setting words to tones. By this account, recalling a particular stress pattern in a melody (or spoken text) activates a metrical grid that constrains the type of text (melody) that is compatible with it. A common metrical grid is typically used throughout a song. Therefore, metric structure provides a means by which lines of an entire song are organized in a common hierarchical structure, thereby relating nonadjacent song components and helping memory.

A limitation of the present study is that we were unable to assess the specific contribution of rhythm to memory. First, the raters failed to provide consistent judgments for the rhythmic aspect of the productions. Second, French is not a stress-based language, so it is possible that musical meter (and rhythm in general) is not as efficient a memory aid for French lyrics as it is for English lyrics. Yet, as mentioned in the introduction, support for the contribution of music to lyric recall in English is scant (Kilgour et al., 2000, Experiment 1 but not 2; Wallace, 1994, Experiments 1 and 2 but not 3). There are also many negative reports of this contribution, even in English (Calvert & Billingsley, 1998; Jellison & Miller, 1982). Therefore, and even though the contribution of rhythm to lyric recall has not been established yet in French, musical constraints appear to be of limited help for lyric recall in general.

This conclusion raises the question of why music is believed to be so important for verbal memory, not only in oral tradition but also in everyday life. We believe this is due to a misunderstanding of the utility of music. Music is not at the service of language. In songs, music contributes to the creation of a general mood that is shared with others (Bowra, 1962; see also Thompson & Russo, 2004, for empirical support). As Booth (1981) writes, a singer tells people "nothing they need to decode or learn. He evokes in them ways of seeing life that they already have" (p. 28). In fact, oral transmission of text is rarely word for word (verbatim) in singing. Although singers believe that they sing the text exactly as heard, they never do so (see Rubin, 1995, for a review). This applies to music recall as well.

Singers, with and without musical training, never recall note for note what they have been presented (Sloboda & Parker, 1985). Rather, singers memorize a schema in which the surface detail is not retained. Recall involves processes akin to improvisation that fill in structurally important events according to general constraints. Learning a new song for faithful reproduction is thus a laborious task that requires hours of practice.

It is interesting to note that when musicians with different expertise *spontaneously* learn an opera song, words and melodies are practiced independently before they are practiced together (Ginsborg, 2002). Moreover, expert singers do not take more time than novice singers practicing words and melodies together when learning a new song. Rather, they use more variable modes of learning. This is probably the best procedure to create detailed memory representations in which words and notes are tightly connected. Hence, we may predict that in a follow-up study, a stronger association between words and melodies may emerge with further training of the same songs under variable modes of recall, as has been observed for highly familiar songs (Peretz, Radeau, & Arguin, 2004). Hence, prolonged practice may confer an advantage to singing over reciting, but this would require considerable time and effort, with probably little payoff for the nonmusician.

Without much practice, however, experts and novices perform quantitatively and qualitatively in a similar manner. This was a rather pleasant unexpected finding, because it suggests that everyone is able to sing fairly well, even in the laboratory, and that song learning is a basic, though difficult, skill. That is, singing appears to be a musical ability that is shared by musicians and nonmusicians of the same culture. This result provides further support for the notion that everyone (unless tone deaf) is equipped to become musically proficient, although only a minority will become experts, usually through extensive practice and explicit tutoring (see Bigand & Poulin-Charronnat, 2006, and Peretz & Hyde, 2003, for recent reviews). Perhaps our expert singers were at a disadvantage here because they did not have the musical score to refer to, which, for them, is the normal procedure for learning. In the procedure used here, musicians had to draw on a common auditory–vocal code that is exercised from childhood for learning popular songs. This widespread mode of vocal learning is a basic mechanism by which humans learn not only to sing, but also to speak. This capacity might very well be shaped by innate mechanisms. The capacity to adjust vocal output so as to imitate an auditory model is a remarkable ability that is rare in the animal kingdom (Merker, 2004). Humans are vocal learners, as are a few bird species, whales, and bats, whereas our closest ancestors, the chimpanzees, are not (Janik & Slater, 1997). This confers to vocal learning a privileged role in the study of the most sophisticated human-specific traits—namely, music and speech.

AUTHOR NOTE

REFERENCES

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

Bigand, E., & Poulin-Charronnat, B. (2006). Are we experienced listeners? A review of the musical capacities that do not depend on formal musical training. *Cognition*, **100**, 100-130.

Bonnel, A.-M., Faïta, F., Peretz, I., & Besson, M. (2001). Divided attention between lyrics and tunes of operatic songs: Evidence for independent processing. *Perception & Psychophysics*, **63**, 1201-1213.

Booth, M. W. (1981). *The experience of songs*. New Haven, CT: Yale University Press.

Bowra, C. M. (1962). *Primitive song*. London: Weidenfeld & Nicolson.

Calvert, S. L., & Billingsley, R. L. (1998). Young children's recitation and comprehension of information presented by songs. *Journal of Applied Developmental Psychology*, **19**, 97-108.

Calvert, S. L., & Tart, M. (1993). Song versus verbal forms for very-long-term, long-term, and short-term verbatim recall. *Journal of Applied Developmental Psychology*, **14**, 245-260.

Chan, A. S., Ho, Y.-C., & Cheung, M.-C. (1998). Music training improves verbal memory. *Nature*, **396**, 128.

Chazin, S., & Neuschatz, J. S. (1990). Using a mnemonic to aid in the recall of unfamiliar information. *Perceptual & Motor Skills*, **71**, 1067-1071.

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

Dell, F. (1989). Concordances rythmiques entre la musique et les paroles dans le chant: L'accent de l'e muet dans la chanson française. In M. Dominicy (Ed.), *Le souci des apparences* (pp. 121-136). Bruxelles: Université de Bruxelles.

Dickson, D., & Grant, L. (2003). Physics karaoke: Why not? *Physics Education*, **4**, 320-323.

Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: Relations between planning and temporal control. *Cognition*, **74**, 1-32.

Gingold, H., & Abravanel, E. (1987). Music as a mnemonic: The effects of good- and bad-music settings on verbatim recall of short passages by young children. *Psychomusicology*, **7**, 25-39.

Ginsborg, J. (2002). Classical singers learning and memorising a new song: An observational study. *Psychology of Music*, **30**, 58-101.

Hayes, N., & Kaun, A. (1996). The role of phonological phrasing in sung and chanted verse. *Linguistic Review*, **13**, 243-303.

Hébert, S., & Peretz, I. (2001). Are text and tune of familiar songs separable by brain damage? *Brain & Cognition*, **46**, 169-175.

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

Ho, Y.-C., Cheung, M.-C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: Cross-sectional and longitudinal explorations in children. *Neuropsychology*, **17**, 439-450.

Jakobson, L. S., Cuddy, L. L., & Kilgour, A. R. (2003). Time tagging: A key to musicians' superior memory. *Music Perception*, **20**, 307-313.

Janik, V. M., & Slater, P. J. B. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, **26**, 56-99.

Jellison, J. A., & Miller, N. L. (1982). Recall of digit and word sequences by musicians and nonmusicians as a function of spoken or sung input and task. *Journal of Music Therapy*, **19**, 194-209.

Kilgour, A. R., Jakobson, L. S., & Cuddy, L. L. (2000). Music training and rate of presentation as mediators of text and song recall. *Memory & Cognition*, **28**, 700-710.

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

McElhinney, M., & Annett, J. M. (1996). Pattern of efficacy of a mu-

sical mnemonic on recall of familiar words over several presentations. *Perceptual & Motor Skills*, **82**, 395-400.

MEDINA, S. L. (1993). The effect of music on second language vocabulary acquisition. *National Network for Early Language Learning*, **6**, 1-8.

MERKER, B. (2004). *Vocal learning for song: Key constraint on the path to language*. Unpublished manuscript.

MORRONGIELLO, B. A., & ROES, C. L. (1990). Children's memory for new songs: Integration or independent storage of words and tunes? *Journal of Experimental Child Psychology*, **50**, 25-38.

NEW, B., PALLIER, C., FERRAND, L., & MATOS, R. (2001). Une base de données lexicales du français contemporain sur Internet: Lexique. *L'Année Psychologique*, **101**, 447-462.

PALMER, C., & KELLY, M. H. (1992). Linguistic prosody and musical meter in song. *Journal of Memory & Language*, **31**, 525-542.

PERETZ, I. (1996). Can we lose memories for music? The case of music agnosia in a nonmusician. *Journal of Cognitive Neurosciences*, **8**, 481-496.

PERETZ, I., GAGNON, L., HÉBERT, S., & MACOIR, J. (2004). Singing in the brain: Insights from cognitive neuropsychology. *Music Perception*, **21**, 373-390.

PERETZ, I., & HYDE, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Sciences*, **7**, 362-367.

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

RACETTE, A., BARD, C., & PERETZ, I. (2006). Making non-fluent aphasics speak: Sing along! *Brain*, **129**, 2571-2584.

RAINEY, D. W., & LARSEN, J. D. (2002). The effects of familiar melodies on initial learning and long-term memory for unconnected text. *Music Perception*, **20**, 173-186.

REY, A. (1964). *L'examen clinique en psychologie*. Paris: Presses Universitaires de France.

RUBIN, D. C. (1995). *Memory in oral traditions: The cognitive psychology of counting-out rhymes, ballads, and epics*. New York: Oxford University Press.

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

SCOTTO DI CARLO, N., & GERMAIN, A. (1985). A perceptual study of the influence of pitch on the intelligibility of sung vowels. *Phonetica*, **42**, 188-197.

SERAFINE, M. L., CROWDER, R. G., & REPP, B. H. (1984). Integration of melody and text in memory for songs. *Cognition*, **16**, 285-303.

SERAFINE, M. L., DAVIDSON, J., CROWDER, R. G., & REPP, B. H. (1986). On the nature of melody–text integration in memory for songs. *Journal of Memory & Language*, **25**, 123-135.

SLOBODA, J. A., & PARKER, D. H. H. (1985). Immediate recall of melodies. In P. Howell, I. Cross, & R. West (Eds.), *Musical structure and cognition* (pp. 143-168). London: Academic Press.

THOMPSON, W. F., & RUSSO, F. A. (2004). The attribution of emotion and meaning to song lyrics. *Polskie Forum Psychologiczne*, **9**, 51-62.

WALLACE, W. T. (1994). Memory for music: Effect of melody on recall of text. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 1471-1485.

WOLFE, D. E., & HOM, C. (1993). Use of melodies as structural prompts for learning and retention of sequential verbal information by preschool students. *Journal of Music Therapy*, **30**, 100-118.

NOTES

1. Written delayed recall was used instead of vocal recall in order to avoid confounding the mode of production and its associated set of lyrics. This procedure also provided an opportunity to assess the effect of music on text encoding, as has been done in most prior studies.

2. This lack of correlation between verbal and musical recall suggests that these two components are produced independently. To test for independence between word and pitch errors, we would need to compute the probability of joint word/pitch errors on the basis of the error rates for the separately occurring word (W) and pitch (P) errors (prob-W multiplied by prob-P, as applied by Drake & Palmer, 2000, to the pitch and time errors obtained in piano performance). However, because 67% of the participants had no errors on the words only and/or on the notes only, the probability of joint errors was mostly zero, so independence could not be assessed properly.

APPENDIX

CACHEZ-LA

Claude Gauthier

C'EST TELLE - MENT   TELLE-MENT DIF - FI - CI - LE___

DE   TROU - VER___   QUEL-QU'UN QUI VOUS AIME___

SI   JA - MAIS___   VOUS TE - NEZ CETTE   ÂME___

CET - TE PER - LE   RARE___   CE BOU-QUET DE FLEURS

CA - CHEZ - LE___   DANS___ VO - TRE   COEUR

CA - CHEZ - LA___   À L'A - BRI   DU   TEMPS

AI - MEZ   LE___   DE TOU VO - TRE   COEUR

AI - MEZ - LA___   DE TOUT VO - TRE   TEMPS