2012

Semantic Priming of Familiar Songs

Sarah K. Johnson

Andrea R. Halpern
Bucknell University, ahalpern@bucknell.edu

Semantic priming of familiar songs

Sarah K. Johnson · Andrea R. Halpern

Abstract We explored the functional organization of semantic memory for music by comparing priming across familiar songs both within modalities (Experiment 1, tune to tune; Experiment 3, category label to lyrics) and across modalities (Experiment 2, category label to tune; Experiment 4, tune to lyrics). Participants judged whether or not the target tune or lyrics were real (akin to lexical decision tasks). We found significant priming, analogous to linguistic associative-priming effects, in reaction times for related primes as compared to unrelated primes, but primarily for within-modality comparisons. Reaction times to tunes (e.g., "Silent Night") were faster following related tunes ("Deck the Hall") than following unrelated tunes ("God Bless America"). However, a category label (e.g., *Christmas*) did not prime tunes from within that category. Lyrics were primed by a related category label, but not by a related tune. These results support the conceptual organization of music in semantic memory, but with potentially weaker associations across modalities.

Keywords Music · Priming · Semantic memory · Categorization

Music is a highly compelling and pervasive aspect of the lives of most individuals. Not surprisingly, music also represents a well-ingrained portion of our semantic memory—most people can attest to having a store of countless songs in their memory,

S. K. Johnson (✉)
Department of Psychology, Moravian College,
1200 Main Street,
Bethlehem, PA 18018-6650, USA
e-mail: skjohnson@moravian.edu

A. R. Halpern
Department of Psychology, Bucknell University,
Lewisburg, PA, USA

most of which are without clear episodic origin (Bartlett & Snelus, 1980). Despite this fact, few studies have explored the organization of musical semantic memory. Here we explored the conceptual organization of music using a priming paradigm, with the expectation that priming would occur across songs from within a shared musical category (e.g., Christmas songs). In a series of four experiments, we explored across-song priming of tunes and lyrics, in relation to each other and to their categorical labels.

The expectation of priming across songs was motivated by early work that revealed the robustness with which people cluster songs into thematic groupings. Halpern (1984a) investigated two possible ways that people may categorize songs in semantic memory: one based on musical similarity, the other based on nonmusical similarity. Participants were asked to sort the titles of familiar tunes according to first one, then the other, type of similarity. Clustering results from both methods yielded solutions based more on thematic or conceptual dimensions than on melodic aspects. Thus, the organization of memory for songs can be described in a way that parallels descriptions of other aspects of semantic memory.

Other evidence strengthens this general conclusion; for example, Schulkind (2004) noted a high rate of genre confusions in his study of identification of familiar melodies. These thematic-based errors were nearly as common as phrasing/meter errors, although the latter were substantially more relevant to the participants' task. As Schulkind pointed out, the large number of phrasing/meter errors implicates a role for musical organization in semantic memory—that is, for organization based on aspects of melody such as key, rhythm, and meter. However, the musical aspects of the semantic organization of songs may be less prevalent in situations in which the task does not immediately demand judgments based on those features. Under less constrained

task conditions, the conceptual features may govern the way that we use our semantic memory for songs.

The presence of errors of a semantic nature, even when it was not necessary to process this dimension of the music in order to complete the task (Schulkind, 2004), is suggestive of automatic semantic processing. A more direct way to explore the nature and automaticity of the associations between songs, as well as between the different components of songs (lyrics, melodies, genres, etc.), is via priming. For example, Peretz, Radeau, and Arguin (2004) used priming to explore the automatic activation between the musical and lyrical components of songs. Their focus was on within-song priming: That is, did portions of lyrics from a particular song prime the corresponding melody from that song, and vice versa, when participants made familiar/unfamiliar song judgments? Their results revealed priming on the order of ~100 ms (ranging from 50–250 ms) for targets related to the prime (i.e., from the same song) versus those unrelated to the prime (from a different song). The results from their Experiments 2 and 3 suggested that responses to lyrics were affected by the match in prime–target domain more than were responses to melodies; that is, melody-to-lyric priming (cross-domain) was weaker than lyric-to-lyric priming (same-domain). Melody targets were affected equally by melody (same-domain) or lyric (cross-domain) primes. Interestingly, these domain effects did not occur in the final experiment in that study, in which the authors controlled for recognition difficulty and stimulus duration between the two stimulus types. Nonetheless, Peretz et al.'s results support the presence of priming between components of songs. However, no studies to date have looked at priming across songs.

Priming across songs could potentially follow musical relationships, as in similar pitch contours or rhythmic patterns, or—of more interest here—conceptual ties. For example, shared membership in a superordinate category might promote priming across songs (e.g., priming between "My Country 'tis of Thee" and "The Star Spangled Banner," both belonging to the category of [American] Patriotic songs). One reason to expect conceptually based priming is Halpern's (1984a) finding regarding errors in a title–tune matching task. When title and tune did not correspond (mismatch trials), error rates were twice as high when the tune was from the same category as the referent of the title as when the presented tune and title were not from the same category. The error rates supported a role of conceptual relationships between songs. Reaction times (RTs) did not show the effect, but RT differences might emerge with other types of tasks more equivalent to the tasks typically used in the priming literature (e.g., lexical decision tasks).

Priming from one song to another could occur via spreading activation through associative links formed via coexposure (e.g., hearing Christmas songs played in succession on the radio during the holiday season), through the higher-level categorical relationships of songs from a particular genre (i.e., mediated priming), or via shared features relevant to categorical relationships, but not explicitly mediated through them (such as when two fruits are associated via shared features like juiciness and growing on a tree, as in proximity-based distributed network models; see McNamara, 2005, for a review). Word-priming research has contrasted priming of items from the same semantic category (without any other associative relationship) versus priming via associations. Several reviews of the literature on this distinction have suggested that pure categorical priming, albeit weaker than associative priming, does occur on lexical decision tasks, as well as on other tasks (McNamara, 2005; Neely, 1991). For example, McRae and Boisvert (1998) compared priming for a target (*chicken*) following a high-similarity prime with a shared categorical relationship (*duck*), a low-similarity prime with a shared categorical relationship (*pony*), and a dissimilar prime (*mittens*) using a semantic decision task and a lexical decision task. For both tasks, they found significantly faster RTs for targets following high-similarity primes than for those following dissimilar primes, regardless of the stimulus onset asynchrony (SOA). Importantly, low-similarity primes also led to faster RTs, but only with a long SOA. According to traditional spreading activation theories of priming (e.g., Collins & Loftus, 1975), these priming effects can be explained by activation travelling from the presented item up to a higher-level node (e.g., following the word *canary* up to its higher-level classification of "animal," in order to decide that it is a concrete object). McRae and Boisvert's results highlight the spread of activation across concepts on the basis of semantic similarity, and they suggest a graded characteristic to this spread: Closer relationships allow for more rapid spread of activation than do farther (but still semantically related) relationships.

Unlike with words, there are no norms to distinguish which tune pairings within a category have strong associative connections and which do not. However, either associative priming or pure semantic priming would reflect conceptual organization rather than musical organization (specifically, organization via pitch intervals), and thus priming via either mechanism would reveal new information about the organization of familiar songs in semantic memory. From findings with category-coordinate relationships in the word-priming literature, such as those of McRae and Boisvert (1998), we expected faster processing for target songs following prime songs from the same category than for targets following prime songs from a different category when deciding whether a particular melody or set of lyrics was from a real song.

Accordingly, our goal in the present series of experiments was to examine across-song priming effects with familiar tunes and lyrics of varying levels of conceptual similarity, in relation to each other and to their conceptual categories. A basic priming effect across songs has not been shown before and would be of theoretical interest, further confirming the

semantically based organization of songs in memory and extending the role of priming mechanisms to a novel domain. We began by exploring tune-to-tune priming in the first experiment, using the categorical relationships established in Halpern's (1984a) study as the basis for the tune–category associations (e.g., "The Star Spangled Banner" as an example of the Patriotic category, or "Row Row Row Your Boat" as an example of the Children's category). We expected to find semantic priming across tunes within a superordinate category, in parallel with the dominance of conceptual organization exhibited when people are asked to consciously sort the songs.

Experiment 1: tune-to-tune priming

In this experiment, we contrasted RTs for prime–target target pairs that had close relationships on the basis of Halpern's (1984a) cluster analysis (e.g., "Jingle Bells" and "Rudolf the Red-Nosed Reindeer"; that is, related primes, as both are Christmas tunes) versus those that did not have close relationships (e.g., "Jingle Bells" and "My Darling Clementine"; that is, unrelated primes, as the latter is a folk song). These conditions are comparable to those used in other priming studies focusing on the dimension of semantic similarity. The task was equivalent to a lexical decision task—that is, a real-song judgment ("Is the song a real tune?"). False tunes were created by reorganizing the notes and rhythms of actual songs used in the study, thus creating new unfamiliar songs that were well matched to the real tunes in musical qualities. The expected priming effect would consist of faster RTs for related than for unrelated pairs.

Method

Participants The participants were 20 undergraduate students from Bucknell University (17 females, 3 males). The reported age range was from 18 to 20 years. Although the majority of participants did indicate some sort of musical experience in their background, they were not selected on the basis of musical experience, and on the whole, the sample was not especially musically skilled. We recruited only participants who were raised for the majority of their lives in the United States, so that all participants would be highly familiar with the tunes being used.

Materials The stimuli were the melodies of highly familiar tunes from a number of categories (see Appendix A). The songs were selected from the 59 songs included in the musical sort portion of Halpern's (1984a) first experiment (see Fig. 2 from that study). To ensure that those songs were still familiar to our population, familiarity ratings based on song titles were made by a separate group of 24 college-aged participants using a 7-point scale (from 1= *Never heard of this song* to 7= *So familiar I can "hear" it in my head*). The participants rated all 59 of Halpern's (1984a) songs, and only songs with an average score of 5 or greater were included in the present experiment.

For each song, two audio files were created in standard piano timbre using the Songworks software, one for the real version of the tune and one for the false version. The songs used are listed by category in Appendix A. The real melodies ranged from 3.5 to 8.3 s (8 to 20 notes) in length, corresponding to the first phrase of the song. False tune targets were derived from real tunes by permuting the pitch and rhythmic values (i. e., reorganizing the notes) so as to create plausible melodies that were nonetheless unfamiliar but that conserved the musical characteristics and length of the originals. None of the false tunes began in a way that paralleled a real tune, beyond the first interval, to our knowledge. Some false tunes had been used in prior experiments (Bartlett, Halpern, & Dowling, 1995); others were newly created for this experiment. All of the tunes were rated by a separate group of student listeners for musicality and nonidentifiability. The MIDI files created from this software were compatible with the SuperLab software used for stimulus presentation and response recording.

Two stimulus sets were created, in order to counterbalance the directionality of the prime–target pairings for the real-tune targets. Thus, for every A–B tune pairing found in the first stimulus set, there was a corresponding B–A tune pairing in the second stimulus set. Participants were assigned to one of the two sets in a block-randomized fashion. Each stimulus set contained 58 real-tune target trials (29 related-prime and 29 unrelated-prime trials) and 60 false-tune target trials. Thus, the relatedness proportion for this study was .5.[1]

Primes were always real tunes. For related-prime trials, the prime tune and target tune were from the same semantic category, as determined by Halpern's (1984a) cluster analysis (e.g., "Star Spangled Banner"–"God Bless America"). For the unrelated-prime trials, the prime and target tunes reflected different semantic categories (e.g., "Old MacDonald"–"God Bless America"). Among false-tune targets, half were permutations of tunes that were related to the prime, and half were permutations of tunes that were unrelated to the prime. However, apart from controlling for general musical characteristics in this way, the relatedness distinction was not meaningful for the false tunes, and we combined all 60 false-tune targets for the analysis.

Within each set, every tune appeared equally often in all four possible conditions, which was done by crossing the two key item variables: real and permuted targets, and

---

[1] We calculated the relatedness proportion on the basis of the proportion of related real tune–real tune pairings as compared to all real tune–real tune pairings. This calculation excluded false-tune targets in the same way that word-priming studies exclude nonword targets when calculating the relatedness proportion (e.g., Lucas, 2000).

related and unrelated primes. There were only a couple of exceptions to this rule, involving tunes from categories that were too small to fully counterbalance in this way without repeating specific tune pairings. Any given pairing of tunes only occurred once per participant.

Procedure The prime stimulus was played in its entirety, followed by a 550-ms interstimulus interval, and then the target tune began. Participants were instructed to press either "yes" or "no" on a button box to indicate whether the second tune was a real or a false tune. They were encouraged to press a button as soon as they had made their judgment, and there was no time limit for their response. After responding, there was a 1,500-ms intertrial interval before the next trial began.

After the priming phase was completed, participants filled out a musical background survey, which included several standard demographic questions as well as questions about musical experience, after which they were debriefed.

Results

One of the participants' data were excluded due to an unusually low overall accuracy rate for the task (69%; >3 SDs below the average across all participants). The remaining 19 participants had a high accuracy rate overall (93%, ranging from 83% to 99%). We analyzed accuracy for the related versus unrelated conditions in order to determine whether a speed–accuracy trade-off could account for any priming effect in RTs. Accuracy was no different for targets following related primes than following unrelated primes, $t < 1$, $d = .06$ (see Table 1). Thus, any priming could not be attributed to a speed–accuracy trade-off. Furthermore, accuracy was quite high overall, suggesting that participants were in fact very familiar with the tunes.

Analyses were conducted for correct responses only. Outlier RTs of 2.5 SDs or more either above or below the

overall RT, determined on a participant-by-participant basis, were excluded (total of 1.9% of the data). Following this, mean RTs for the two prime conditions (related vs. unrelated) for real tunes were compared. Relative to the false-tune RTs ($M = 4,745$ ms, $SD = 979$ ms), participants were much faster at responding to real tunes ($M = 3,037$ ms, $SD = 468$ ms). More importantly, the RTs for real tunes following related primes were significantly faster than RTs for real tunes following unrelated primes, $t(18) = 2.25$, $p < .05$, $d = .52$—that is, a priming advantage of 98 ms for conceptually related primes (see Table 2 for the means).

Discussion

The significant priming observed here reinforces the claim by Halpern (1984a) that semantic organization is a key functional element of memory for familiar music. Faster judgments for a melody after hearing a related melody (as compared to an unrelated one) support the idea that semantic networks allow for spreading activation from one tune to another following thematic ties.

We chose the pairings solely on the basis of semantic similarity metrics. However, it was important to verify that the related tune pairs were not more musically similar to each other than were the unrelated tune pairs. This seemed unlikely, due to the wide variety of tempos, contours, and harmonies found within the genres. However, to confirm this, we submitted the MIDI files to an objective musical similarity analysis using the SIMILE toolbox, devised by Müllensiefen and Frieler (2007). We used their recommended similarity measure, called opti3, which is a weighted combination of measures of *n*-gram (motive), rhythmic, and harmonic similarity between two tunes. This measure can range from 0 (*no similarity*) to 1 (*identity*). The mean opti3 for related pairs was .119 ($SD = .08$), and for unrelated pairs

Table 1 Percent accuracy (standard deviations in parentheses) for real-tune targets following related and unrelated primes in Experiments 1–4

	Related	Unrelated	Priming[a]	Effect Size
Within-Modality Experiments				
Experiment 1: Tune primes, tune targets	94.4 (4.8)	94.2 (3.8)	+0.2	.06[d]
Experiment 3: Category primes, lyric targets				
Real and false lyrics combined[b]	91.9 (5.7)	89.2 (7.5)	+2.7[c]	.22[e]
Across-Modality Experiments				
Experiment 2: Category primes, tune targets	92.2 (8.8)	91.5 (8.3)	+0.7	.08[d]
Experiment 4: Tune primes, lyric targets				
Real lyrics	90.7 (6.4)	88.2 (6.8)	+2.5[c]	.32[e]
False lyrics	91.8 (9.1)	92.1 (7.9)	−0.3	.003[e]

[a] Accuracy for related trials – unrelated trials; positive difference = priming. [b] No interaction between type of prime and type of lyric. [c] $p < .05$. [d] Cohen's *d*. [e] Partial η^2

Table 2 Mean reaction times (in milliseconds; standard deviations in parentheses) for tune and lyric targets following related and unrelated primes in Experiments 1–4

	Related	Unrelated	Priming[a]	Effect Size
Within-Modality Experiments				
Experiment 1: Tune primes, tune targets	3,003 (407)	3,101 (500)	+98[b]	.52[c]
Experiment 3: Category primes, lyric targets				
Real lyrics	2,606 (512)	2,575 (531)	−31	.01[d]
Time-locked false lyrics	840 (286)	930 (341)	+90[b]	.25[d]
Across-Modality Experiments				
Experiment 2: Category primes, Tune targets	2,999 (468)	2,949 (476)	−50	.27[c]
Experiment 4: Tune primes, Lyric targets				
Real lyrics	2,681 (596)	2,664 (553)	−17	.01[d]
Time-locked false lyrics	908 (267)	948 (310)	+40	.06[d]

[a] Unrelated RT – related RT; positive difference = priming.
[b] $p < .05$. [c] Cohen's d. [d] Partial η^2

it was .112 ($SD = .07$).[2] Thus, the tune sets were essentially identical in pairwise musical similarity, and the tune pairs were only modestly musically related on average.

Having established that priming among familiar tunes can result from conceptual relationships, we next addressed one explanation for how this priming is accomplished. One possibility that seemed likely, given the prior findings in the literature on the semantic priming of words (e.g., McRae & Boisvert, 1998), was the implicit activation of the target via a shared superordinate node. For example, hearing the melody for "Jingle Bells" may cause activation to spread to the general node of Christmas tunes, and from there to other nodes strongly related to the larger category (e.g., "Rudolf the Red-Nosed Reindeer"). To test this possibility, we next explored the priming of melodies via the superordinate musical categories to which they belong.

Experiment 2: category-to-tune priming

The primes in this study were thematically associated with the song but did not reflect simple verbal associations (e.g., the word *patriotic* is not found in the lyrics of the song "The Star Spangled Banner"). Following a logic similar to that of Experiment 1, we expected a priming effect reflecting faster RTs for making a real-tune judgment for tunes preceded by related category primes as compared to unrelated category primes.

Method

Participants The participants were 25 undergraduate students from Moravian College (16 females, 9 males). The reported age range was 18 to 26 years. As in Experiment 1, the participants'

musical backgrounds varied, but the sample was not especially musically skilled overall. Again we made an effort to recruit only participants who were raised for the majority of their lives in the United States. Some participants failed to indicate this on their survey, but the majority of the sample (19 participants) did indicate that they were raised in the United States.

Materials The tunes were substantially the same as in Experiment 1, with a few changes based on ratings of the appropriateness of the category/genre labels made by a new group of 16 participants. The songs rated included 34 items from the first experiment, as well as 8 additional items taken from Halpern's (1984a) original list, for a total of 42 items. The category labels were also taken from that study, which characterized major branches of the cluster solution to the sorting data. Participants scored each category–song pairing on a 4-point scale indicating how strongly associated the category label was to the song (from 1 = *Not at all; I would never think of this category for this song* to 4 = *Very strongly; I would immediately think of this category for this song*). Participants were also encouraged to offer suggestions for alternative category labels that might be as good as, or better than, the one listed for the song.

Songs were kept if the mean category association rating was 3 or higher. We also looked at the alternative categories offered by participants and removed items from our list if multiple participants listed one of the other categories being used in the experiment as a viable alternative for that song (e.g., "Jimmy Crack Corn" was dropped from the Folk category due to multiple participants listing Children's as an alternative category). On the basis of these category ratings, it was necessary to drop several items, and even some categories (e.g., the Old Time category from Halpern's, 1984a, study was not considered a strong label for any of its songs). After completing these filtering processes, we were left with 32 songs from six categories (see Appendix A).

---

[2] We thank Daniel Müllensiefen for carrying out this analysis.

For each song, two audio files were used, as in Experiment 1: a real-melody version and a permuted version. The stimulus sets closely paralleled those used in Experiment 1, except that category labels were used as primes instead of melodies. For example, if the prime–target pairing in Experiment 1 had been "Mary Had a Little Lamb"–"Old MacDonald," in Experiment 2 the related pairing would be *Children's* (a verbal label)–"Old MacDonald." Thus, for related-prime trials, the category correctly reflected the tune it was paired with, as in the example above, and for unrelated-prime trials, the prime was from a different category (e.g., *Patriotic*–"Old MacDonald"). Although some songs could belong to more than one category, we used the dominant category, and we avoided potential cross-category associations when making the pairings for the unrelated-prime trials.

Each stimulus set contained 58 real-tune trials (half related-prime and half unrelated-prime) and 60 permuted-tune trials. The majority of tunes appeared equally often in the related- and unrelated-prime conditions for a given participant (with a small number of exceptions, following the same reasoning as described for Experiment 1).

Procedure As in Experiment 1, the participants were instructed to press either "yes" or "no" on a button box to indicate whether the tune was real or false (i.e., permuted). The prime (category label) appeared for 200 ms centered on the screen, followed by a 550-ms interstimulus interval, and then the target tune began. Apart from the prime, the timings were the same as in Experiment 1. After this phase was completed, participants filled out the musical background survey.

Results

Three of the participants' data were excluded due to overall accuracy rates that were more than 2 *SD*s below the average across all participants. The remaining 22 participants (13 female) had an overall accuracy of 90% (ranging from 79% to 98%). Accuracy was compared for the related and unrelated prime conditions, and there was no significant difference, $t < 1$, $d = .08$ (see Table 1). Thus, a speed–accuracy trade-off cannot account for the RT results. Once again, we note the high overall accuracy in this experiment.

As in Experiment 1, analyses were conducted for correct responses only, and outlier RTs were dropped (1.3% of all items) on the basis of the same criteria as before. Following this exclusion, we compared mean RTs for the two prime conditions. The results revealed the same relationship between real- and false-tune RTs, with responses to false tunes ($M = 4,368$ ms, $SD = 1,056$ ms) occurring much more slowly than those to real tunes ($M = 2,976$ ms, $SD =$

463 ms). The key comparison between RTs for targets following related primes versus those for targets following unrelated primes was not significant, $t(21) = 1.25$, $p = .23$, $d = .27$, and was in the nonpredicted direction (see Table 2).

Discussion

In contrast with the priming found between tunes from the same superordinate category, the tunes in this experiment were not primed by their categorical labels. This lack of priming is conspicuous in light of the absolutely parallel structure of the priming paradigm, the target stimuli, and the judgment being made in this experiment as compared to Experiment 1. One potential concern might be the aptness of the labels. However, because they were initially based on categories from a successful cluster solution (high proportion of variance accounted for, as well as reflecting a logical category distribution; Halpern, 1984a) and were cross-checked by independent raters, we think that the labels are accurate representations of the song categories.

A more likely explanation involves modality. Experiment 2 involved cross-modality priming from a verbal label to an auditory stimulus. As described previously, Peretz et al. (2004) found significant priming between the musical and lyrical components within the same song, although cross-modality priming was not always of the same strength as within-modality priming. In the domain of environmental sounds, Friedman, Cycowicz, and Dziobek (2003) found reductions in novelty P3 amplitudes for sound–word conceptual repetitions—for instance, a word (e.g., *pig*) preceded by a matching sound (e.g., an oink) rather than a nonmatching sound (e.g., an airplane roar). However, Stuart and Jones (1995) found repetition priming for environmental stimuli (sound–sound and word–word), but no cross-modal priming (sound–word or word–sound). Therefore, in the next experiment, we explored priming from the verbal category label to the song lyrics, a comparison parallel to the one in Experiment 2, but within a modality.

Experiment 3: category-to-lyric priming

The paradigm in this study was identical to that of Experiment 2, but here the targets were written lyrics and the task was a real-lyric judgment—that is, were the words true lyrics from a well-known song? On the basis of the implication from Experiments 1 and 2 that priming occurs within a modality but not across modalities, we did expect a priming effect: faster RTs for tunes preceded by related rather than unrelated category primes.

Our false lyrics were newly written in such a way that the words were correctly ordered at the start (see Appendix B). We could not use permutations of the real lyrics because to

do so would allow participants to reject false lyrics purely on the basis that the item was nonsensical. In this way, the decision task was not quite analogous to the one used in the word-priming literature, but it did mirror the decision tasks in Experiments 1 and 2, in that rejection of the false melodies in those experiments did not reflect a judgment about non-sensicality, but rather a failure of identification. In addition, whereas many songs begin with the same opening melodic interval (e.g., "London Bridge Is Falling Down" and "Silent Night"; or "The Little Drummer Boy," "Do Re Mi," and "Oh Susanna"), lyrics are unique to particular songs. To avoid extremely fast rejections for the false lyrics on the basis of a single word, we elected to make the false lyrics begin with a few real lyrics. This also allowed the general sense to be conserved between the true and false lyrics.

Method

Participants The participants were 22 students enrolled in Moravian College courses (17 female, 5 male). Their ages ranged from 18 to 22 years. Scores on the musical background questionnaire were similar to those from the sample in Experiment 2, albeit with a higher percentage of participants indicating that they had received private instrument or voice lessons at some point in their lifetime (82%, as compared to 32% in Experiment 2). Apart from this, participants were not especially musically skilled. All participants indicated that they had been born or raised from a young age in the United States.

Materials The categories and tunes were identical to those used in Experiment 2. However, in this experiment, rather than using two different sound files per tune for the targets, we used two different lyric sets per tune: one real and one false. The true lyrics were always the first phrase of the song (corresponding to the notes that were used in the melodies in Experiments 1 and 2). False lyric versions of each tune were created to conform to the following constraints: The first few words were always accurate, for the reasons stated above. False lyrics also kept the meter and syllabic pattern of the original. Semantic category was also roughly conserved, as was the general sense of the original lyric, so that a serious song stayed serious, and the same for more frivolous songs. For example, the true lyrics for "Jingle Bells" were *Dashing through the snow, on a one-horse open sleigh*, and the corresponding false lyrics were *Dashing through the frost, on a pony swift and sure*. See Appendix B for a full list of the true and false lyrics. Category labels were used to create related-prime and unrelated-prime trials in the same manner as for Experiment 2.

Procedure As in Experiment 2, participants were instructed to make true/false judgments by pressing either "yes" or "no" on a button box. This time, however, the task was to indicate whether the lyrics were real or fake. The

category label appeared for 200 ms, followed by a 550-ms interstimulus interval, and then the lyrics began. In order to closely parallel the presentation of melodies in the earlier experiments, the lyrics were presented in a word-by-word fashion, similar to the rapid serial visual presentation technique used in some sentence processing paradigms (e.g., Juola, Ward, & McNamara, 1982). The words were presented one at a time in the center of the screen for 400 ms each, with no interitem interval. Participants were encouraged to press a button as soon as they had made their judgment, allowing themselves enough time to feel fairly confident in their answer. They were given up to 2 s following the final word in the longest set of lyrics to make a response. For shorter lyrics, the final word simply stayed on the screen longer, to create equal total presentation times for all lyrics. Following a 1,500-ms intertrial interval, the next trial began. Participants were provided with six practice trials, using categories and tunes not used in the main experiment, to familiarize them with the word-by-word presentation of the target. All participants indicated feeling comfortable with the task instructions and with their ability to read the lyrics before proceeding to the actual trials. Finally, participants filled out the same musical background survey as in the previous experiments.

Results

Two of the participants' data were excluded, one due to an overall accuracy rate that was more than 2 SDs below the average across all participants, and one due to long RTs, suggesting that the participant waited until the end of each trial before making a response. The remaining 20 participants (16 female) had an overall accuracy of 91% (ranging from 70% to 97%). Accuracy rates were lower for the real-lyric items ($M = 86.7\%$, $SD = 10.3\%$) than for the false-lyric items ($M = 94.3\%$, $SD = 5.2\%$), $F(1, 19) = 10.10$, $p < .01$, $\eta_p^2 = .35$, and there was no interaction between type of lyric and type of prime, $F(1, 19) = 1.01$, $p > .05$, $\eta_p^2 = .05$. Accuracy was significantly, but only slightly, higher for targets following related as compared to unrelated primes, $F(1, 19) = 5.39$, $p < .05$, $\eta_p^2 = .22$ (see Table 1 for the accuracy rates for related and unrelated primes, combined across the real- and false-lyric trials). Thus, any potential priming effect was not contaminated by a speed–accuracy trade-off. Finally, the accuracy remained as high as in the first two experiments, suggesting that participants found the true lyrics and tunes equally familiar.

RTs were discarded for incorrect answers and outlier RTs (2.3% of the correct responses). Notably, the RTs for false lyrics in this experiment were significantly shorter than the

RTs for the real lyrics [$t(19) = 2.26$, $p < .05$, $d = 0.50$] and substantially shorter than the RTs for false tunes in the prior experiments (cf. 2,418 ms from this experiment with 4,745 and 4,368 ms in Experiments 1 and 2, respectively). The short RTs for false-lyric trials suggest that participants were taking a different approach to rejecting lyrics as false than the strategy (or strategies) used during a similar judgment with melodies. The false-lyric trials began with the correct lyrics up to a point and then switched to being false lyrics (e.g., the false lyrics for "Twinkle Twinkle Little Star" were *Twinkle twinkle distant light*; thus, at the third word, someone familiar with the song could immediately reject the lyric as false). In essence, the false lyrics were *correct* lyrics until the point at which the first false word occurred. By contrast, the false tunes were permutations of the real tunes that did not start with the correct notes for that tune, although they did preserve the basic characteristics of the original tune and were, therefore, very tune-like in quality.

Thus, for each false lyric, there was one clear point at which the correct rejection could be made. Some priming work does support the idea that initial sentential context can influence feature activation for later words in the sentence (e.g., Moss & Marslen-Wilson, 1993), which might help explain fast "false" judgments at the moment of the first false (i.e., unexpected) word. In contrast, for the real lyrics, participants had to build up enough confidence to assume that the tune was fully correct and that the lyrics wouldn't become false in the next word or two. The accuracy rates described above were consistent with the idea that false-lyric rejection was completed more easily than real-lyric acceptance. Following the assumption that participants made their decision to reject false lyrics at a distinct point, we time-locked the RTs for the false-lyric trials to reflect the finite nature of their decision. The time-locking was performed on a lyric-by-lyric basis, such that the RT clock started at the first incorrect word in that particular lyric. RTs for items that were correctly rejected as false before the first false word was presented (i.e., guesses) were dropped from the analyses. This amounted to dropping, on average, 1.15 items per person. Mean RTs were taken for each of the prime conditions, and we compared the RTs for related versus unrelated primes for both real and false lyrics, using planned contrasts to directly compare the two prime conditions while controlling for multiple comparisons.[3]

For true lyrics, there was no significant priming effect for targets following related primes versus those following unrelated primes, $F(1, 19) = 0.28$, $p = .60$, $\eta_p^2 = .01$. On the other hand, for false lyrics, there was a significant priming effect: False lyrics following related primes were rejected more quickly than false lyrics following unrelated primes, $F(1, 19) = 6.18$, $p = .02$, $\eta_p^2 = .25$ (see Fig. 1 and Table 2). Thus, the category label did lead to significant priming of false lyrics when RTs were time-locked (and, in fact, the priming was nearly significant when no time-locking procedure was used).[4]

Discussion

Overall, these results support the conclusion that priming of the components of familiar songs occurs within but not across modalities. Tunes primed tunes, for songs within the same category (Experiment 1), but the verbal category label itself did not prime the tunes (Experiment 2). The category label did prime the lyrics from those same songs, however. Although this priming occurred only in false-lyric trials, we argue that this was due to criterion effects during true-lyric trials, in which participants' confidence in a "real" judgment gradually increased with time (i.e., with additional words). Such confidence effects might well create individual differences in the threshold for responding "yes" (i.e., that the lyrics were real), adding variability to the scores in a way that overshadowed the relatively small priming effects. In contrast, false-lyric trials were demonstrably false starting at a discrete point in time, and "no" (or "false") judgments could be made accurately right at that moment. The average response times of ~900 ms following the first false word were not greatly longer than the responses times seen in traditional semantic-priming studies involving lexical decisions for single words (e.g., Neely, 1977). These response times support the inference that participants made their responses essentially immediately upon seeing that word.

To test the hypothesis of successful within- versus unsuccessful cross-modality priming, we conducted one additional experiment to look at the priming from tunes to lyrics. On the one hand, given that tunes prime other tunes within a category and that evidence supports associative links between a tune and its own lyrics (Peretz et al., 2004), it would be reasonable to expect priming from tunes to the lyrics of other songs within the category (e.g., from the tune for "Row Row Row Your Boat" to the lyrics for "Twinkle Twinkle Little Star"). On the other hand, the modality explanation of the results in Experiments 1–3 predicts no priming from tunes to lyrics.

---

[3] Planned contrasts were used because RT differences between true and false items were not meaningful following the time-locking procedure, nor would it be appropriate to draw conclusions about any interaction between item type (true vs. false) and prime in this case.

[4] As noted, a similar planned contrast of the raw RTs (prior to time-locking) for the false lyrics in Experiment 3 revealed a nearly significant priming effect, $F(1, 19)=4.26$, $p=.053$, $\eta_p^2=.18$, mean RT difference=75 ms.

Fig. 1 (a) Mean reaction times for real lyrics and false lyrics following related and unrelated primes in Experiment 3. (b) Mean time-locked reaction times for false lyrics following related and unrelated primes in Experiment 3. Error bars reflect SEMs. *$p \leq .053$
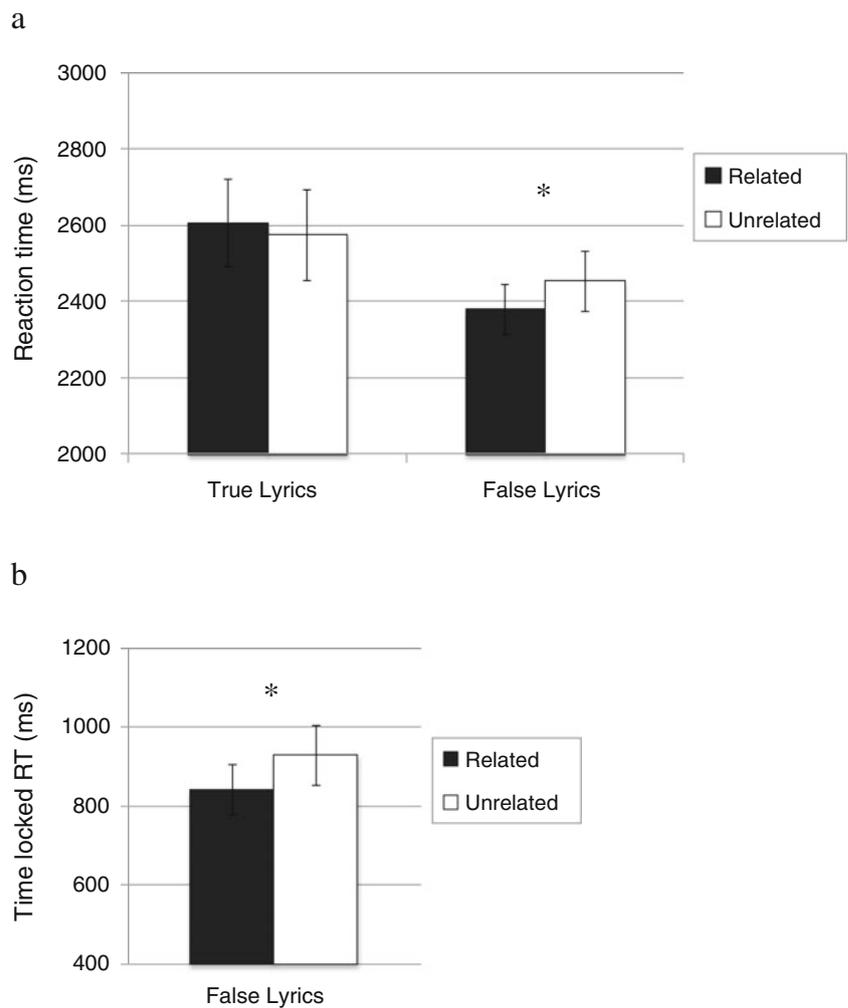
Experiment 4: tune-to-lyric priming

In the final experiment, the same structure of prime–target pairings was used as in Experiment 1, with tunes as primes, but with lyrics replacing tunes as targets. Participants made a real-lyric judgment for targets, as in Experiment 3. The presence or absence of a priming effect would address the modality explanation applied to the results from the earlier experiments.

Method

Participants The participants were 23 students enrolled in Moravian College courses (18 female, 5 male). Their ages ranged from 17 to 24 years. Once again, participants on the whole were not skilled musicians. The scores on the musical background questionnaire were similar to those in Experiment 3, with slightly higher rates of participation in amateur choirs (74%) and bands/orchestras (83%) in the present experiment than in the previous one (45% and 68% for choir and band/orchestra participation, respectively). All

participants indicated that they had been born or raised from a young age in the United States.

Materials The lyric targets were identical to those used in Experiment 3, and the tune primes were nearly identical to those used in Experiment 1. Tunes were paired with lyrics to create related-prime and unrelated-prime trials in the same manner as the tune-to-tune pairings in Experiment 1 (and, again, these were always across-song pairings).

Procedure Participants were instructed to make a true/false lyric judgment by pressing "yes" on the button box if the lyrics were from a real song and "no" if they were not. The tune prime played fully, followed by a 550-ms interstimulus interval, and then the lyrics began, presented word by word in the center of the screen at a rate of 400 ms each (no interval between words). Participants were encouraged to press a button as soon as they had made their judgment and were given the same time frame as in Experiment 3. A 1,500-ms interval separated the trials. Participants went through six practice trials, using categories and tunes not

used in the main experiment, prior to the real trials. After finishing the real trials, participants filled out the musical background survey.

Results

Four of the participants' data were excluded, three due to low overall accuracy rates (all three at chance-level performance) and one due to difficulty following the task instructions. The remaining 19 participants (16 female) had an overall accuracy of 91% (ranging from 81% to 97%). Accuracy rates were similar for the real-lyric ($M = 89.5\%$) and false-lyric ($M = 92.0\%$) items, $F < 1$, $\eta_p^2 = .0$. In addition, accuracy was no different for targets following related primes ($M = 91.3\%$) or unrelated primes ($M = 90.2\%$), $F(1, 18) = 1.41$, $p > .05$, $\eta_p^2 = .07$. There was, however, a significant interaction between type of lyric and type of prime, $F(1, 18) = 6.56$, $p < .05$, $\eta_p^2 = .27$. Analyses of the simple main effects revealed no effect on false-lyric trials of related versus unrelated primes, $F < 1$, $\eta_p^2 = .003$, but a significant effect of prime for the true-lyric trials, $F(1, 18) = 8.34$, $p < .05$, $\eta_p^2 = .32$. The pattern reflected higher accuracy for the true-lyric targets following a related prime relative to those following an unrelated prime (see Table 1 for the accuracy rates for related and unrelated primes, separated for real- and false-lyric trials).

RTs were discarded for incorrect answers and for outlier RTs (again, 2.3% of the correct responses). Mean RTs were compared for each of the prime conditions for both real and false lyrics, including, as was done in Experiment 3, time-locking the RTs for false lyrics (see Fig. 2 and Table 2). Again, we used planned contrasts to specifically compare the two prime conditions for each item type while controlling for multiple comparisons. The RTs for real lyrics following related primes and unrelated primes were not significantly different, $F(1, 18) = 0.20$, $p = .66$, $\eta_p^2 = .01$. The time-locked RTs for false lyrics were also not significantly different for targets following related versus unrelated primes, $F(1, 18) = 1.18$, $p = .29$, $\eta_p^2 = .06$ (and the same was true for the non-time-locked RTs).[5]

Discussion

While priming did occur in accuracy, we speculate that this reflected a decrease in accuracy following unrelated primes as opposed to an increase in accuracy following related primes, on the basis of a comparison of the accuracy rates

---

[5] As noted, a planned contrast of the raw RTs (prior to time-locking) for the false lyrics in Experiment 4 revealed no significant priming, $F(1, 18) = 2.01$, $p = .17$, $\eta_p^2 = .10$, mean RT difference = 53 ms.

for the real-lyric trials with the false-lyric trials, as well as with the tune-target accuracy rates from Experiments 1 and 3. However, even accepting this result as priming in favor of related pairings (i.e., increased accuracy for related tune–lyric pairs), this one positive cross-modal result stands alone. As can be seen in Tables 1 and 2, all other cross-modality comparisons were not significant, and multiple within-modality comparisons were significant. Thus, the absence of priming of RTs from tune to lyric, in combination with the pattern of results found in the first three experiments for both RTs and accuracy, is consistent with stronger within-modality priming and weaker cross-modality priming. Although, within a song, the tune and the lyrics may be associated strongly enough to allow for RT priming (Peretz et al., 2004), across-song priming of RTs appears to be influenced by modality.

General discussion

These results support the claim that memory for familiar songs—and for individual components of songs (melodies and lyrics)—is organized along conceptual lines. The fact that songs within a category may not share much surface-level similarity either in lyrics or melodic qualities, as was the case with the pairings used here (e.g., "Rudolph the Red-Nosed Reindeer" and "The First Noel"), makes these categorical associations even more interesting. Shared category memberships create relationships between songs in a way that can be seen both with conscious, intentional measures (e.g., the sorting tasks used by Halpern, 1984a) and with less intentional measures, such as the priming shown here. We maintain that the benefit for targets following related primes derives from spreading-activation processes, or coactivation of shared features, via shared membership in a musical category. However, the precise nature of how this spreading/shared activation occurs is not obvious.

Our initial interpretation of the tune-to-tune priming was that it occurs via spreading activation through a superordinate conceptual node, one that could presumably be captured via a category label. However, if a superordinate conceptual node is the bridge between tunes, why then would the label fail to prime the tune? One possibility is that the conceptual nodes are multifaceted and, in tune-to-tune priming, these nodes are activated in such a way as to specifically evoke the musical aspects of the node (including, one would assume, the relevant verbal label); however, when the labels are presented by themselves as primes, they do not evoke those musical aspects exclusively. For example, if the tune "Rudolph the Red-Nosed Reindeer" plays, it may cause activation to spread to the concept of Christmas (and the word *Christmas*) in a way that particularly activates features of Christmas related to music (e.g., caroling). On

Fig. 2 (a) Mean reaction times for real lyrics and false lyrics following related and unrelated primes in Experiment 4. (b) Mean time-locked reaction times for false lyrics following related and unrelated primes in Experiment 4. Error bars reflect *SEM*s
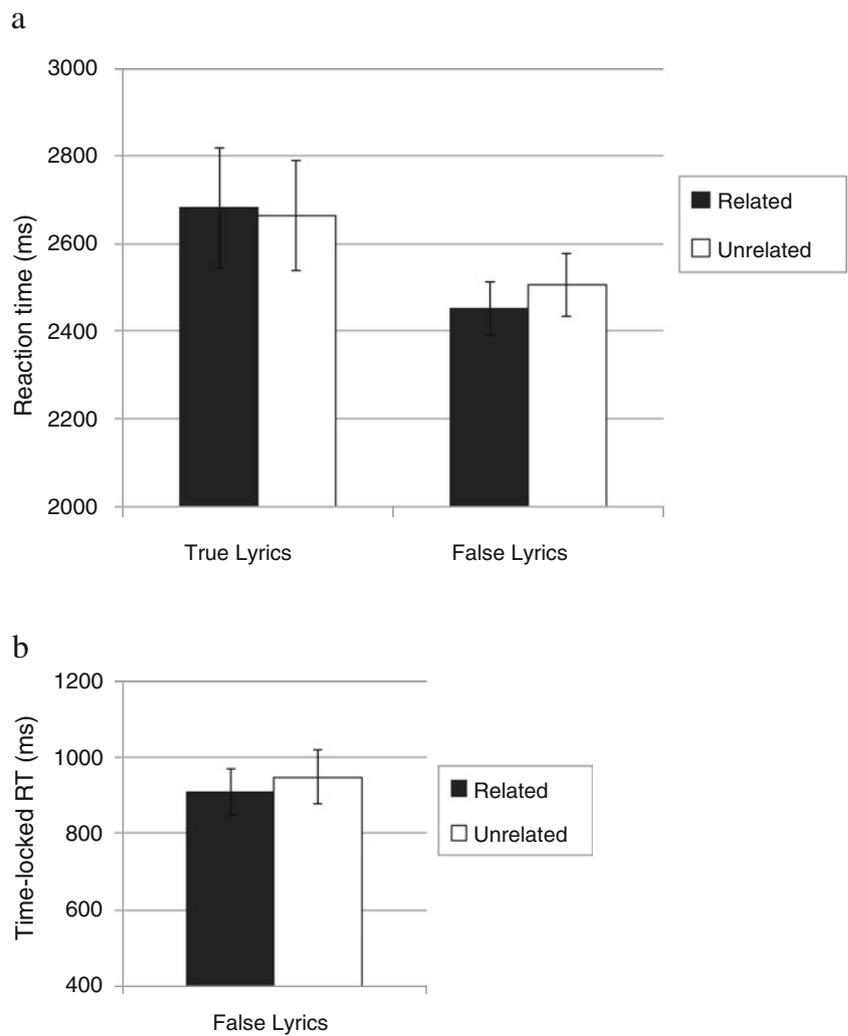
the other hand, when the word *Christmas* is presented, it may result in activation of a variety of relevant concepts (e.g., trees, presents, or vacation), only some of which are musical. Essentially, this account suggests that hearing a tune leads to a biasing of the superordinate concept (Christmas) toward music, allowing activation to spread to other tunes from that concept. Such a bias would not be instantiated when the verbal label is provided first as the prime. Note that a similar explanation applies with a distributed-network account, whereby the biasing determines which features (musical or nonmusical) are activated and can lead to priming, given that those features are shared between prime and target. Presenting the word *Christmas* may result in activation of features that are not shared with the tune because they are not primarily music-related.

This idea could be tested by deliberately biasing a conceptual node toward its musical features. We could first present each category label with some associated tune (but not the tunes used as targets later in the study). If participants first activated the concept of Patriotic in the context

of hearing "God Bless America," we might see more priming thereafter from *Patriotic* to the tune of "The Star Spangled Banner," for example. Alternatively, the musical aspect of the conceptual node could be instantiated more actively by adding a music-related task with the primes themselves. For example, participants could perform a "musical-genre verification" response for each prime (e.g., "Is the presented word a musical genre? Yes or No"), and this might encourage priming from the category label to the tunes.

Category-label-to-lyric priming, as was shown in Experiment 3, could occur via musical or nonmusical features of the superordinate node—that is, the words to a well-known song could be activated via linguistic components, allowing for priming whether or not the conceptual node is biased toward music. Thus, the musical-bias interpretation is compatible with the results from Experiments 1–3; however, it does not explain the lack of priming of RTs in Experiment 4. The effect on RTs for this experiment was in the expected direction, but cross-modal priming is notoriously weak. Prior work suggests weaker cross-modal associations for environmental sounds

than within-modality associations (Stuart & Jones, 1995). Furthermore, both behavioral evidence (Bonnel, Faïta, Peretz, & Besson, 2001) and neurological evidence (Besson, Faïta, Peretz, Bonnel, & Requin, 1998; Hébert & Peretz, 2001) support the independence of the linguistic and melodic components within a song. Thus, priming across these components in different songs, and in different modalities, may be hard to achieve.

A point of interest in explaining the priming results is the role of strategic processes. Melodies and lyrics are notably longer than the stimuli used in the majority of the priming literature. Therefore, we may well expect more strategic processing to occur by the time an excerpt is recognized or identified. Neely and Keefe's (1989) hybrid model of semantic priming, which combines spreading activation with the strategic processes of expectancy and semantic matching, may be important in fully characterizing priming across musical excerpts (whether within a song or across songs). A key timing component in relation to the use of strategic processing is the length of time between presentation of the prime and selection of a response. When category labels were used as primes, the interval between offset of the prime and onset of the target was 550 ms, relatively long for a priming paradigm. Strategic processing is more likely to occur when longer SOAs are used (e.g., Neely, 1977). In addition, when tunes were used as primes, the SOA became longer and variable, as it incorporated the length of the prime tune, resulting in an SOA anywhere from 4,050 to 8,850 ms long. Furthermore, because the targets were also lengthy, relative to the single words used in the majority of priming work, participants had additional time to process both prime (held in mind) and target up to the point at which they gave their response.

Expectancy is the most likely candidate here in terms of strategic processes: Participants may form an idea or guess about what target will follow when the prime appears, and therefore they would experience facilitation if they guess correctly. The relatedness proportion in the present experiments was .5, which is not extraordinarily high but, coupled with long SOAs, could induce a guessing strategy. Presumably, participants could start to coactivate a couple of context-appropriate guesses, such that they were likely to receive priming for a number of different targets. Equally, some interference or inhibition might occur if the guess were incorrect. Whether the source of the priming is facilitation, inhibition, or both, the conclusion regarding the conceptual organization of songs remains the same. Future work might attempt to disentangle the roles of these two sources in the priming of music. We predict that both facilitation and inhibition play a role, given the likelihood that some degree of strategic processing is occurring (see McNamara, 2005, for an in-depth discussion of this issue).

The compound-cue theory of priming (Ratcliff & McKoon, 1988; see McNamara, 2005, for a review) may prove especially helpful in explaining our findings, in light of the fact that no response was made to the primes themselves in these tasks. This theory suggests that only basic perceptual features of the prime are processed in paradigms in which responses are made only to the target, which could explain the modality-specific effects found here. The compound-cue model suggests that priming results from the familiarity of the target and cue together, as opposed to spreading activation from one activated concept to the next. This account does not incorporate a clear explanation for differential priming on the basis of a modality match between the prime and target. However, when category labels are used as primes, low-level perceptual processing may shift focus to the verbal attributes of the category label, and processing of music-related information may be limited, similar to the cue-biasing idea discussed above. As suggested, a task involving the musical nature of the prime might extend processing, in order to allow for cross-modality priming to become evident.

Even accepting some debate as to the manner of the priming, a priming effect is informative about the nature of relationships across these stimuli. Specifically, it is of clear interest that functional relationships occur via conceptual ties, not just melodic ones. These ties do not need to be explicitly based in hierarchical relationships to a category node, but they do seem to reflect categorical relationships, if indirectly. For example, the priming in Experiment 3 may reflect direct associational priming across words (e.g., from *Christmas* to *snow* or to *Santa*), as opposed to a relationship between the lyrics and a categorical, genre-based node. Whether such conceptual priming would replicate for songs that lack a linguistic component, such as classical pieces, remains to be explored. The tune-to-tune priming found in Experiment 1 is not likely to be due to similar lyrics, given the lack of tune-to-lyric priming in Experiment 4; therefore, we suggest that priming across tunes within a conceptual category would occur even for purely instrumental pieces.

The results presented here show clear evidence of thematic organization in familiar verbal songs. People can process tunes on the basis of musical qualities, as well, such as key, rhythm, and so forth (see, e.g., Halpern, 1984b; Schulkind, 2004). Nonetheless, the ease with which people treat familiar songs on the basis of conceptual similarities (Halpern, 1984a) and the priming based on these conceptual similarities, as shown here, present strong evidence that meaning may dominate both the organization of familiar music and the functionality of memory for music.

Appendix A

Table 3 Categories and songs

Children's
　Three Blind Mice
　London Bridge
　Mary Had a Little Lamb
　Old MacDonald
　Pop Goes the Weasel
　Row Row Row Your Boat
　Twinkle Twinkle Little Star

Christmas
　The First Noel
　O Come All Ye Faithful
　Deck the Hall
　Hark the Herald Angels Sing
　Jingle Bells
　Rudolph The Red-Nosed Reindeer
　Santa Claus Is Coming to Town
　Silent Night
　The Little Drummer Boy
　White Christmas
　We Wish You a Merry Christmas

Folk
　My Darling Clementine[**]
　Oh Susanna
　I've Been Working on the Railroad[*]

Movie
　Do Re Mi
　We're Off to See the Wizard
　Singing in the Rain

Old Time[*]
　When the Saints Come Marching In[*]
　You Are My Sunshine[*]

Party
　Happy Birthday
　For He's a Jolly Good Fellow

Patriotic
　America the Beautiful
　My Country 'tis of Thee
　God Bless America
　The Star Spangled Banner
　This Land Is Your Land
　The Yankee Doodle Boy
　Yankee Doodle Went to Town

[*] Used in Experiment 1 only. [**] Used in Experiments 2–4 only

Appendix B

Table 4 True and false lyrics in Experiments 3–4[COMP: This table will likely break across two pages]

Song	True Lyrics	False Lyrics
America the Beautiful	Oh beautiful for spacious skies, for amber waves of grain	Oh beautiful for sun-streaked plain, for yellow fields of wheat
Deck the Hall	Deck the hall with boughs of holly, fa la la . . .	Deck the hall with strands of tinsel, fa la la . . .
Do Re Mi	Do, a deer, a female deer. Re, a drop of golden sun	Do, a deer, a timid deer. Row, a way to steer a boat
For He's a Jolly Good Fellow	For he's a jolly good fellow, for he's a jolly good fellow	For he's a wonderful neighbor, for he's a wonderful neighbor
God Bless America	God bless America, land that I love	God bless America, long may she stand
Happy Birthday	Happy birthday to you, happy birthday to you	Happy future to you, wealth and happiness too
Hark the Herald Angels Sing	Hark the herald angels sing glory to the newborn king	Hark the herald cherubs sing, praises worthy of the king
Jingle Bells	Dashing through the snow, on a one-horse open sleigh	Dashing through the frost, on a pony swift and sure
London Bridge	London Bridge is falling down, falling down, falling down	London Bridge is crumbling round, crumbling round, crumbling round
Mary Had a Little Lamb	Mary had a little lamb, little lamb, little lamb	Mary had a pretty bird, pretty bird, pretty bird
My Country 'tis of Thee	My country 'tis of thee, sweet land of liberty	My country pure and free, kind home of charity
My Darling Clementine	In a cavern, in a canyon, excavating for a mine	In a cavern in a valley, coming in to spend some time
O Come All Ye Faithful	O come all ye faithful, joyful and triumphant	O come all the humble, worthy and deserving
Oh Susanna	Oh I come from Alabama with a banjo on my knee	Oh I come from California with a saddle on my mule
Old MacDonald	Old MacDonald had a farm, e-i-e-i-o	Old MacDonald ate some grits, e-i-e-i-o
Pop Goes the Weasel	All around the mulberry bush, the monkey chased the weasel	All around the evergreen tree, the rabbit ate the flowers
Row Row Row Your Boat	Row row row your boat gently down the stream	Row row row the raft, all the way to town
Rudolph the Red-Nosed Reindeer	Rudolph the red-nosed reindeer had a very shiny nose	Rudolph the big-eyed panda had a friendly goofy grin
Santa Claus Is Coming to Town	You'd better watch out, you better not cry, you better not pout, I'm telling you why	You'd better take care, you better beware, watch for the man with presents to share
Silent Night	Silent night, holy night	Silent night, heavenly bright
Singing in the Rain	I'm singin' in the rain, just singin' in the rain	I'm singin' in the show, now singin' in the show
The First Noel	The first noel, the angels did say	The first noel, the wise men did pray
The Little Drummer Boy	Come, he told me, pa rum pa pum pum	Come, he beckoned, tra la la la la
The Star Spangled Banner	Oh say can you see by the dawn's early light	Oh say do you know what the flag means to us
The Yankee Doodle Boy	I'm a yankee doodle dandy, yankee doodle do or die*	I'm a yankee old frontiersman, a yankee native to the bone
This Land Is Your Land	This land is your land, this land is my land	This land is freedom, this land is glory
Three Blind Mice	Three blind mice, three blind mice, see how they run, see how they run	Three blind ants, three blind ants, see how they climb, see how they climb
Twinkle Twinkle Little Star	Twinkle twinkle little star, how I wonder what you are	Twinkle twinkle distant light, what a pretty sparkling sight
We Wish You a Merry Christmas	We wish you a merry Christmas, we wish you a merry Christmas	We wish you a jolly noel, we wish you a jolly noel
We're Off to See the Wizard	We're off to see the wizard, the wonderful wizard of Oz	We're off to find the bishop, the scary but kindly old man
White Christmas	I'm dreaming of a white Christmas	I'm dreaming of a soft snowfall
Yankee Doodle Went to Town	Yankee Doodle went to London, riding on his pony	Yankee Doodle came to dinner, had a nice hot supper

*The "a" was inadvertently omitted in the second line of the true lyrics. This one item did not affect the patterns of priming for accuracy or RT in Experiments 3 and 4

References

Bartlett, J. C., Halpern, A. R., & Dowling, W. J. (1995). Recognition of familiar and unfamiliar melodies in normal aging and Alzheimer's disease. *Memory & Cognition, 23,* 531–546. doi:10.3758/BF03197255

Bartlett, J. C., & Snelus, P. (1980). Lifespan memory for popular songs. *American Journal of Psychology, 93,* 551–560. doi:10.2307/1422730

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. doi:10.1111/1467-9280.00091

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. doi:10.3758/BF03194534

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82,* 407–428. doi:10.1037/0033-295X.82.6.407

Friedman, D., Cycowicz, Y. M., & Dziobek, I. (2003). Cross-form conceptual relations between sounds and words: Effects on the novelty P3. *Cognitive Brain Research, 18,* 58–64. doi:10.1016/j.cogbrainres.2003.09.002

Halpern, A. R. (1984a). Organization in memory for familiar songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10,* 496–512. doi:10.1037/0278-7393.10.3.496

Halpern, A. R. (1984b). Perception of structure in novel music. *Memory & Cognition, 12,* 163–170. doi:10.3758/BF03198430

Hébert, S., & Peretz, I. (2001). Are text and tune of familiar songs separable by brain damage? *Brain and Cognition, 46,* 169–175. doi:10.1016/S0278-2626(01)80058-0

Juola, J. F., Ward, N. J., & McNamara, T. (1982). Visual search and reading of rapid serial presentations of letter strings, words, and text. *Journal of Experimental Psychology: General, 111,* 208–227. doi:10.1037/0096-3445.111.2.208

Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review, 7,* 618–630. doi:10.3758/BF03212999

McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition.* Hove, U.K.: Psychology Press.

McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 558–572. doi:10.1037/0278-7393.24.3.558

Moss, H. E., & Marslen-Wilson, W. D. (1993). Access to word meanings during spoken language comprehension: Effects of sentential semantic context. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1254–1276. doi:10.1037/0278-7393.19.6.1254

Müllensiefen, D., & Frieler, K. (2007). Modelling experts' notion of melodic similarity. *Musicae Scientiae, 11*(Suppl.), 183–210. doi:10.1177/102986490701100108

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General, 106,* 226–254. doi:10.1037/0096-3445.106.3.226

Neely, J. H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G. W. Humphreys (Eds.), *Basic processes in reading: Visual word recognition* (pp. 264–336). Hillsdale: Erlbaum.

Neely, J. H., & Keefe, D. E. (1989). Semantic context effects on visual word processing: A hybrid prospective-retrospective processing theory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory (Vol. 24)* (pp. 207–248). New York: Academic Press.

Peretz, I., Radeau, M., & Arguin, M. (2004). Two-way interactions between music and language: Evidence from priming recognition of tune and lyrics in familiar songs. *Memory & Cognition, 32,* 142–152. doi:10.3758/BF03195827

Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review, 95,* 385–408. doi:10.1037/0033-295X.95.3.385

Schulkind, M. D. (2004). Serial processing in melody identification and the organization of musical semantic memory. *Perception & Psychophysics, 66,* 1351–1362. doi:10.3758/BF03195003

Stuart, G. P., & Jones, D. M. (1995). Priming the identification of environmental sounds. *Quarterly Journal of Experimental Psychology, 48A,* 741–761. doi:10.1080/14640749508401413