Cognition, 16 (1984) 285–303

Integration of melody and text in memory for songs*

MARY LOUISE SERAFINE
ROBERT G. CROWDER
Yale University

BRUNO H. REPP
Haskins Laboratories

Abstract

Two experiments examined whether the memory representation for songs consists of independent or integrated components (melody and text). Subjects heard a serial presentation of excerpts from largely unfamiliar folksongs, followed by a recognition test. The test required subjects to recognize songs, melodies, or texts and consisted of five types of items: (a) exact songs heard in the presentation; (b) new songs; (c) old tunes with new words; (d) new tunes with old words; and (e) old tunes with old words of a different *song from the same presentation ('mismatch songs'). Experiment 1 supported the integration hypothesis: Subjects' recognition of components was higher in exact songs (a) than in songs with familiar but mismatched components (e). Melody recognition, in particular, was near chance unless the original words were present. Experiment 2 showed that this integration of melody and text occurred also across different performance renditions of a song and that it could not be eliminated by voluntary attention to the melody.*

Introduction

Song is a universal artform that consists of two seemingly separate components, melody and text. In practice a song may derive from a pre-composed melody to which words are added, or from a pre-existing text later set to music. In fact a song may be the work of two artists, a composer and a poet or librettist. Yet the relationship between melody and text raises interesting

*This research was supported by NSF Grant 8005838 to Robert G. Crowder and by NICHD Grant HDO1994 to Haskins Laboratories. Reprint requests should be sent to Mary Louise Serafine who is now at the Department of Psychology, Vassar College, Poughkeepsie, New York 12601, U.S.A.

0010-0277/84/$6.20

questions in the domains of both aesthetics and cognitive psychology.

One of the aesthetic issues is how the artform should be defined: whether it is simply a pairing of independent components or an integral whole that transcends its parts. This issue has implications for the analysis of songs from a music-theoretic viewpoint—for example, whether the components can be considered or analyzed separately.

A parallel issue can be raised from a cognitive viewpoint: To what degree are melody and text independent or integrated in perception and memory? While there is substantial literature both on linguistic memory and on musical memory (Deutsch, 1969; Dowling, 1973, 1978; Dowling and Fujitani, 1971), research thus far does not indicate how a hybrid form such as songs might be represented in memory. Indeed, research on hemispheric differentiation, especially that which suggests left-hemisphere dominance for language and right-hemisphere dominance for music (e.g., Best *et al.*, 1982; Kimura, 1967), leaves entirely open how melody and text in songs might be processed.

Our interest in this issue was generated by informal observations suggesting that, in memory for songs, melody and text form an integrated unit, such that people find it difficult to separate the two components. For example, if asked to recite the words of their national anthem, many people would have to sing the song, or at least rehearse it subvocally, in order to generate the words. Also, people may not immediately recognize that two different songs have the same melody if their texts are different. 'Twinkle, Twinkle, Little Star' and 'Baa, Baa, Black Sheep' are a case in point, where identical melodies are part of what are considered entirely different songs. Yeston (1975) provides the example of the well-known theme of the Mozart C major piano sonata (K. 545), which (with slight changes in rhythm) is rarely recognized as the melody of 'Hey There, You With the Stars in Your Eyes.' Finally, the first author has found instances of profound melody/text integration in informal experiments with a young child. In these experiments a 2-year-old, who could repeatedly and accurately perform a large body of songs, was nevertheless incapable of singing the melodies on the syllable 'la' without the words. Instead, she simply spoke the syllable in rhythm. Similarly, she was either unwilling or unable to repeat the words without the melody.

These examples argue for some form of integration of melody and text in memory for songs, although it is also true that adults, at least, can voluntarily separate a melody from its text and *vice versa* in singing and recognition. Thus in theory the memory representation for songs might consist of: (1) independent components, (2) integrated components, or (3) a non-decomposable whole (an extreme form of integration). If melody and text were stored as independent components, we would expect that memory for songs could be predicted by the independent probabilities of memory for melody and

memory for text. On the other hand, if the components were integrated, we would expect that memory for one component facilitates memory for the other. Finally, if songs were stored as non-decomposable wholes, we would expect that melodies cannot be recognized as familiar when their words are different, and *vice versa*. This last hypothesis is clearly false in many situations: Words are easy to recognize in new contexts, and most people can probably recognize a tune when the words are different if the tune is pointed out to them. (Musicians and experienced listeners can often do so in any case.) Nevertheless, it is worth investigating the degree to which a holistic representation may characterize novel songs, when attention is not explicitly drawn to tune-similarity.

Note that the issue of integration can be distinguished, at least conceptually, from two related issues: *compatibility* and *association*. Melodies and texts are often compatible rhythmically in that higher pitches, longer durations, and musical-metric stresses tend to occur on accented syllables. Similarly, melodies and texts may be compatible 'semantically' in that the tempo and musical mood seem to fit the meaning of the words. However, it is possible that a cognitive form of integration occurs irrespective of the compatibility of components. Indeed, whether compatibility is necessary or sufficient for integration is an empirical question not under consideration in the present experiments.

Also, integration of melody and text in song can be distinguished from association as mere knowledge of co-occurrence. That melody and text co-occur is undeniable. Yet association may occur without integration. Indeed, it is possible to imagine other co-occurring events (e.g., speech and background music) that do not give rise to integration.

The purpose of these experiments, then, was to investigate the degree to which melody and text are independent, integrated, or nonseparable ('holistic') in memory for songs.

Experiment 1

Subjects heard 24 consecutive excerpts from folksongs, followed by a 20-item recognition test. The test items were of two types: (1) excerpts that had been heard in the presentation ('old songs') and (2) excerpts that had not been heard in the presentation ('new songs'). Further, new songs were of four types: (a) new tune with new words; (b) old tune with new words; (c) new tune with old words; and (d) old tune with old words that had been sung to a different tune in the original presentation ('tune and words mismatched'). In the remainder of the paper, the terms 'tune' and 'words', as used with

subjects, are interchangeable with 'melody' and 'text.'

The main prediction was that, if subjects integrate melody and text in memory, they should recognize previously heard melodies or texts more accurately when they are paired with their original companion (text or melody) than when they are paired with a different companion. On the other hand, if melody and text are stored as independent components, then subjects should recognize previously heard melodies or texts equally well whether paired with the same or with a different companion. Finally, if songs are stored as holistic units, then subjects should not be able to recognize melodies (or texts) at all, except when they are paired with their original companion.

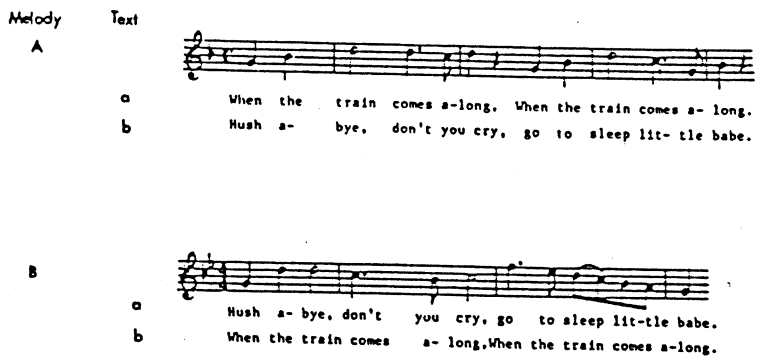
Method

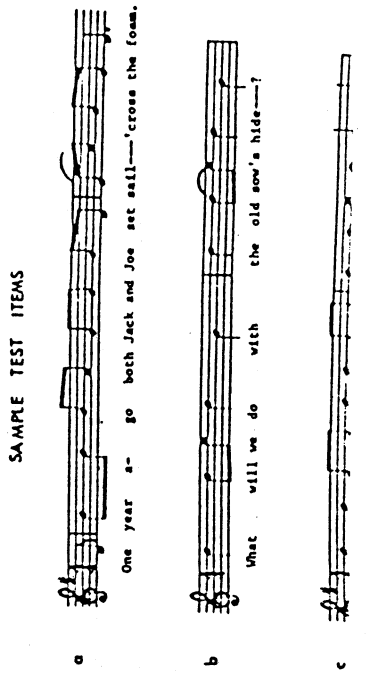
Materials

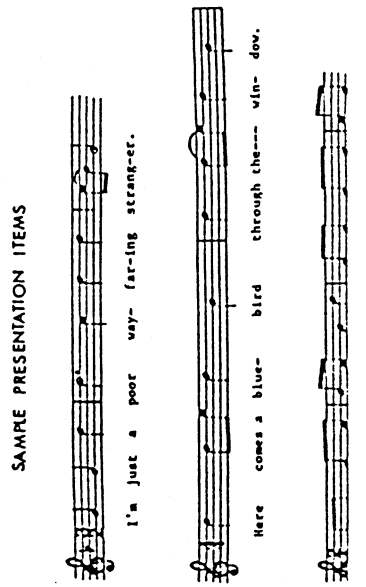
Songs that we considered unfamiliar to the average listener were drawn from a collection of indigenous American folksongs compiled by Erdei (1974). Twenty pairs of excerpts with interchangeable melodies and texts were chosen, each excerpt consisting of the opening two to four measures of a song (see list in appendix). Thus each pair of excerpts yielded four different songs, a total of 80. Figure 1 shows a sample pair of interchangeable melodies and texts. Examples of the five test-item types are shown in Fig. 2.

In some cases minor alterations were made to the original melody or text to ensure a rhythmic fit with its companion (see appendix). For example, 'across' from one original text was changed to 'cross' in our experiments (Fig.

Figure 1. *Sample pair of songs with interchangeable texts. (Aa and Bb denote original songs; Ab and Ba denote derivatives.)*







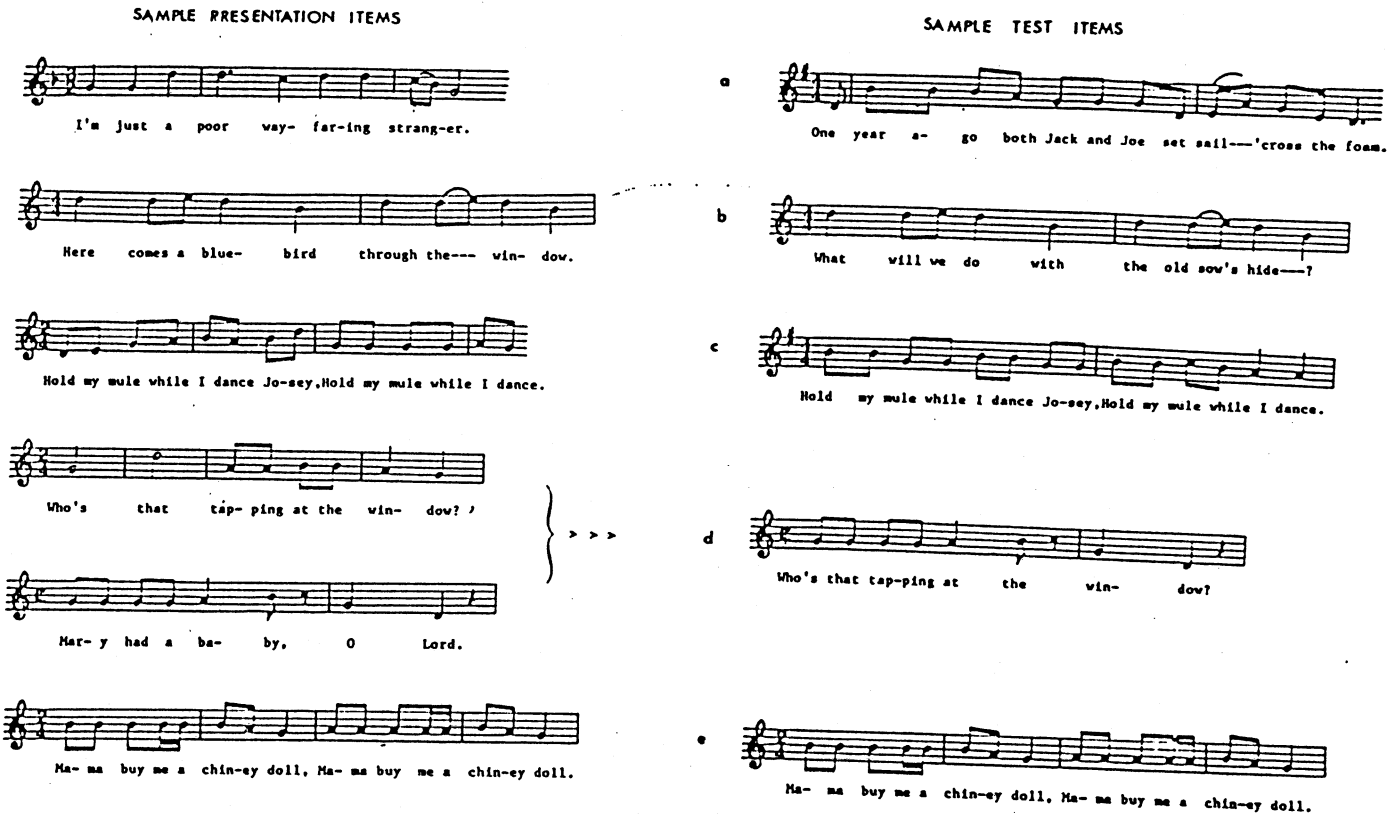


Figure 2. *Sample presentation and test items. (a: new tune, new words; b: old tune, new words; c: new tune, old words; d: old tune, old words—mismatched; e: old song.)*

2, test item a). However, in all cases the texts and melodies were identical across parallel presentation and test versions of a song.

The excerpts were recorded on tape, sung by a tenor with vocal training. They were sung as notated, except transposed down a fifth (or twelfth) to the tenor range, and at a tempo of one beat per second (MM = 60). The excerpts varied in key and mode, but were notated with G as the tonic in each case. The tapes were recorded with a 5-second interval of silence between presentation items and a 10-second response interval after each test item.

Design

From the bank of 80 song excerpts, five parallel sets of presentation and test sequences were constructed. Each set was administered to a different group of subjects.

In the presentation sequences (24 items) half the excerpts were tunes with their original words (type Aa or Bb in Fig. 1), and half were tunes with words borrowed from their companion song (type Ab or Ba in Fig. 1).

In the test sequences (20 items) each of the five test item types (a through e in Fig. 2) occurred four times. Moreover, across the five subject groups presentation items were paired with each of the five possible test item types, following a Latin square design. For example, Table 1 shows the generation of possible presentation and test items from two of the song pairs.

Table 1. *Presentation and test items from sample song pairs*

Subject group	I	II	III	IV	V
SONG PAIR*	*Pres./Test*	*Pres./Test*	*Pres./Test*	*Pres./Test*	*Pres./Test*
	c	*a*	*b*	*d*	*e*
$[Aa \text{ and } Bb]_X$	Aa Ba	Aa Bb	Aa Ab	Aa Ab Bb	Aa Aa
	b	*c*	*e*	*a*	*d*
$[Aa \text{ and } Bb]_Y$	Ab Aa	Ab Bb	Ab Ab	Ab Ba	Ab Aa Ba

a, New tune, new words.
b, Old tune, new words.
c, New tune, old words.
d, Old tune, old words, mismatched.
e, Old song.
* Two examples (X, Y) from 20 song pairs.

As shown in Table 1, a mismatch test item (type d: old words paired with old tune of a different song) required *two* presentation excerpts. Whenever two such items were required in the presentation, they immediately followed each other on the tape. Thus each presentation sequence required four *pairs* of songs for the mismatch test items, plus 16 songs for the other test item types (total of 24 songs).

In all presentation and test sequences, the excerpts were generated successively from Song Pairs 1 through 20, in the order listed in the Appendix. Thus, the interval between each presentation item and its corresponding test item was roughly constant.

Procedure

Subjects were tested in small groups in a quiet room, except for one large group that was tested in a classroom. Presentation and test tapes were played back over loudspeakers. Subjects were instructed to listen carefully to a presentation of 24 excerpts from simple folksongs and told that their memory for them would be tested later. After the presentation they were asked whether any of the excerpts were familiar, and if so, to estimate their number on the answer sheet. Following that, the answer sheet was explained and the test sequence was presented. About 5 minutes elapsed between the presentation and test sequences.

For each test item subjects were asked to indicate on the answer sheet whether they had "heard that exact excerpt before," and if not, whether they had heard either the tune or the words. In advance of the test, subjects were given an explanation of the term 'tune' (melody) and a description of the five types of items they could expect on the test (types a through e). Thus, they were prepared for the test of recognizing tune, words, or exact song, but they did not have knowledge of this requirement prior to the presentation.

Subjects

Subjects were 32 undergraduate students with varying degrees of musical training. The first 16 subjects, who were tested together in a classroom environment, necessarily were all assigned to one particular presentation/test condition. Sixteen additional subjects were divided equally among the remaining four conditions.

Results and discussion

Subjects' post-presentation estimates of the number of songs that seemed familiar averaged 1.4 (out of 24 presentation items, 12 of which were original folksongs). This result confirmed the relative unfamiliarity of the materials.

For the discussion of recognition scores, we adopt this terminology: If subjects indicated that they recognized an exact test excerpt as one that had been heard in the presentation, this is called an 'old song' response. Similarly, if they indicated recognition of just the melody or the text, this is called an 'old tune' or 'old words' response, respectively.

Recognition of old songs

Table 2 lists the mean proportion of 'old song' responses made to the five types of test items. Subjects correctly recognized old songs 85% of the time, a surprisingly high recognition rate, given that the presentation excerpts had been heard only once. Incorrect responses were lowest (0.07 and 0.06) whenever new words were heard, and highest (0.39) for 'mismatched' tune and words, where both components had been heard originally in *different* presentation songs.

It might be argued that this high false alarm rate for mismatched old tune and words (0.39) indicates some measure of independent storage for song components. To some degree, subjects erroneously thought they recognized the mismatch songs, apparently because the components were familiar, though never paired in the original presentation. Note, however, that this effect was largely due to a high false-alarm rate for all items containing old

Table 2. *Mean proportion of 'old song' responses (Exp. 1)*

New songs				
		Words		
		New	Old	Mean
Tune	New	0.07	0.25	0.16
	Old	0.06	0.39	0.22
	Mean	0.06	0.32	
Old songs				0.85

words, which probably reflects the fact (discussed below) that words were much easier to remember than tunes. Nevertheless, this false alarm rate is far below the hit rate for original old songs, which indicates that subjects were more likely to retain the association of a presented melody with its presented text than to retain the components independently.

This is necessary but not sufficient evidence that subjects integrated melody and text. To address the issue of integration, we must examine melody and text recognition separately, and determine whether these components were more accurately recognized in old songs than in any type of new song.

Recognition of components

Table 3 shows the mean proportion of responses to the questions regarding recognition of old songs, tunes, and words. In this table 'old song' responses are included both in 'old tune' and in 'old words' responses, for a response of 'old song' indicates that subjects recognized both the tune and the words.

The main hypothesis was that, if melody and text are integrated in memory, old tunes should be recognized more accurately in old songs than in mismatch songs (or songs with new words). Similarly, old words should be recognized more accurately in old songs than in mismatch songs (or songs with a new tune).

Consider first the 'old tune' responses. Old tunes were recognized more accurately in old songs (0.92) than in mismatch songs (0.63) or songs with new words (0.53). The advantage for old songs over mismatch songs was highly significant across subjects, $t(32) = 5.27$, $p < 0.001$, and across test items, $F(1,18) = 16.20$, $p < 0.001$. The advantage was equally large for old

Table 3. *Mean proportions of 'old tune' and 'old words' responses (Exp. 1)*

New songs		'Old tune' responses				'Old words' responses		
		Words				Words		
		New	Old	Mean		New	Old	Mean
Tune	New	0.44	0.40	0.42	New	0.10	0.78	0.44
	Old	0.53	0.63	0.58	Old	0.13	0.85	0.49
	Mean	0.48	0.52		Mean	0.12	0.82	
Old songs				0.92				0.92

songs presented in their original folksong version and for old songs constructed by recombining the melodies and texts of different original folksongs, $F(1,18) = 0.05$.

Consider now the 'old words' responses. Words were recognized more accurately in old songs (0.92) than in mismatch songs (0.85) or songs with new tune (0.78). Because of ceiling effects, the advantage for old songs over mismatch songs fell just short of significance across subjects, $t(32) = 2.00$, $p < 0.06$, but it was significant across items, $F(1,18) = 6.35$, $p < 0.02$. Once again, it did not matter whether or not the old song was a real folksong, $F(1,18) = 0.18$.

These results suggest that melody and text are integrated in memory to a considerable degree. One component is recognized better in the context of the other, original component, than in some new context. The advantage for original contexts (old songs) holds even over new contexts in which the components are just as familiar (mismatch songs). The deciding factor seems to be not whether the components are *familiar*, but rather whether they had been *paired* in the initial perception. Thus melody and text appear not to be stored independently; the components are stored in some integrated fashion.

Responses to new songs

The data from Experiment 1 allow for a further clarification of the integration effect. In the remaining discussion we consider 'old tune' and 'old words' responses to new songs only. Two issues are of interest here. First, were tunes and words recognized better than chance in these new contexts? A strong form of the integration hypothesis—that is, a 'holistic' conception—would predict that tunes and words cannot be recognized at all outside of their original contexts (old songs). Second, aside from the integration effect described above, there might also be a 'contamination' effect from companion components at the recall (not storage) stage. That is, were tune and word judgments (whether correct or incorrect) influenced by the familiarity of the other component?

To examine the above issues, separate 2 × 2 ANOVAs were performed on 'old tune' and 'old words' responses, both across subjects and across items, with the factors of tune (old *versus* new) and words (old *versus* new) whose combination represents the four types of 'new song' test items. With regard to tunes, Table 3 shows a mean hit rate of 0.58 and a mean false alarm rate of 0.42, which represents rather poor performance. The difference between hits and false alarms was significant across subjects, $F(1,30) = 9.61$, $p < 0.01$, but not across items, $F(1,18) = 2.99$, $p < 0.11$. Thus, tune recognition in new songs was near chance. The recognition score for words was much

higher: a mean hit rate of 0.82 *versus* a mean false alarm rate of only 0.12. This difference was highly significant, of course.

Thus the strong form of the integration argument—a 'holistic' conception—does not hold up to test here. Certainly, texts were recognized better than chance in new contexts, and there seemed to be some minimal memory for tunes as well, indicating some degree of independent storage of components. As discussed earlier, components are *more* accurately recognized in original contexts (old songs), but they may also be recognized to some degree in new contexts.

The second issue concerns the influence of one component's familiarity on judgments of the other component—a 'contamination' effect. With respect to tunes, Table 3 reveals that subjects responded 'old tune' more frequently when the words were old (mean of 0.52) than when the words were new (mean of 0.48). This small effect was significant across subjects, $F(1,30) = 5.91$, $p < 0.05$, but not across items, $F(1,18) = 1.35$. With respect to words, subjects responded 'old words' somewhat more frequently when the tune was old (mean of 0.49) than when it was new (mean of 0.44). This effect was also significant across subjects, $F(1,30) = 5.52$, $p < 0.05$, but not across items, $F(1,18) = 0.05$.[1]

In summary, Experiment 1 yielded the following results. The main finding was that recognition of one component (melody or text) was facilitated by the simultaneous presence of the other, original component (in old songs). This effect argues for an integrated representation of melody and text in memory for songs. However, there is evidence that tunes, and especially words, can be recognized to some degree when paired with new components. While this does not contradict the integration hypothesis, it does indicate some measure of separation of components and argues against the stronger 'holistic' conception of melody/text relations.

In addition, we found that recognition memory for old songs was excellent, even after a single presentation. Our casual observation was that this excellent performance was accompanied by rather low confidence. Many subjects felt they were just guessing.

Experiment 2

The integration effect in Experiment 1 leaves open at least three questions. First, perhaps the effect was induced by the requirements of a song (rather

[1] We noted that both of these 'contamination' effects were exhibited only by one group of subjects (the large group assigned to Condition I) but not by the other groups. We have no explanation for this difference. With respect to all other effects of interest, the subject groups gave equivalent results.

than melody or text) recognition task. The testing procedure in Experiment 1 required primarily 'old song' recognition, and only conditionally tune and word recognition. Thus tune and word recognition scores were based in large part on correct 'old song' responses. It remains to be determined whether subjects would recognize tunes or words more accurately in old than in new [illegible] were asked to judge only these components. In other words, if old song responses were not permitted, would there still be an advantage for tune or word recognition in old songs?

A second issue concerns the extent to which the integration effect is sensitive to subjects' strategies at the presentation stage. In the first experiment, subjects listened to the presentation songs with the knowledge that their memory for *songs* would be tested. Perhaps this instruction engendered a global, integrated memory for melody and text at the presentation stage. What remains to be determined is whether this integration is optional or obligatory. In other words, would the integration effect still hold if subjects were given the instruction to listen analytically? For example, if subjects were told at the presentation stage that their memory for *tunes* would be tested, would they be able to ignore the words?

A third question concerns the generality of the integration effect. In the first experiment, the presentation and test tapes were recorded by the same performer. Thus vocal inflection, timbre, and other variables in the performance of melodies and texts would be similar across presentation and test songs. It remains to be determined whether the integration effect is sufficiently abstract to hold across different performance renditions of a song. In other words, would the integration effect hold even for a recognition test in which the items are sung by a different performer? Moreover, a possible danger to avoid is that old song recognition might be an artifact of the acoustical identity of old songs across the presentation and test tapes. Any physical identity, even an accidental or musically irrelevant one, could have contributed to the old song recognitions in Experiment 1. If the integration effect were found to hold across different performers, it would prove to be abstract as well as unattributable to the acoustical identity of old songs.

Experiment 2 was designed to address these issues. Specifically, Experiment 2 sought to determine (1) whether the integration effect would hold in a melody-only rather than song recognition task; (2) whether it would hold even in the face of instructions to listen analytically—that is, to tunes only—at the presentation stage; and (3) whether it would hold across different performers (and performance renditions) of the presentation and test songs.

Method

Materials

The materials were the same as in Experiment 1, except that the five sets of presentation and test sequences were recorded a second time, this time by a female vocalist in the alto range, a perfect fifth higher than the male tenor recordings. While the same general guidelines were followed as to tempo and other notated musical factors, no attempt was made to imitate the tenor's performance renditions.

Design

The recordings by male and female vocalists allowed for four combinations of male and female presentation and test sequences (M/M; M/F; F/M; F/F). These four conditions were further subdivided into two instruction conditions. The resulting eight conditions were applied across the five sets of presentation/test sequences. This resulted in 40 conditions.

Procedure

The procedure was similar to that of Experiment 1, with the following differences: Half the subjects received the same instructions as in Experiment 1. The other half received 'analytic' instructions: "Listen carefully to these songs and your memory for the tune or melody only—that is, just the musical portion—will be tested later. You can ignore the words because you will not be tested on these." At the test stage, all subjects made a written response to the question, "Did you hear this exact melody before?" for each item.

Subjects

Subjects were 48 undergraduate students of varying musical backgrounds. Each of 40 conditions contained one subject. Eight additional subjects were assigned to the first set of the presentation/test tapes, distributed across the eight conditions of performance rendition and instruction.

Results and discussion

Subjects in Experiment 2 found the folksong materials as unfamiliar as had subjects in Experiment 1. After the presentation of 24 songs, subjects reported a mean of 1.2 familiar songs.

Recognition of tunes

[illegible] recognition in mismatched new songs and in old songs for two conditions of performance rendition (same voice *versus* different voice) and two conditions of instruction (general *versus* analytic). These data were analyzed in a three-way ANOVA across subjects. For reasons having to do with the design of the experiment, the performance and instruction factors were not included in the ANOVA across items.

The results confirmed the integration effect found in Experiment 1. That is, even in this tune recognition task, subjects recognized tunes more accurately in old songs (mean of 0.84 across all conditions) than in mismatch songs (mean of 0.64), $F(1,40) = 17.19$, $p < 0.001$, across subjects and $F(1,18) = 13.42$, $p < 0.002$, across items. Moreover, the integration effect was maintained even for the analytic condition, where subjects were told to pay attention only to tunes at the presentation stage. There was no significant main effect for instructions or any interaction in this analysis.

Further, the integration effect held to a considerable degree even across different performance renditions. Although Table 4 suggests that the advantage for old songs was reduced in the different-performer condition, this interaction of test item type and performance rendition was not significant, $F(1,40) = 3.86$, $p < 0.10$. Finally, as in Experiment 1, whether or not an old song was a real folksong made no difference, $F(1,18) = 1.31$.

We thus conclude that the integration effect is robust. Melody and text appear to be integrated in memory, even in the face of attempts to focus on or separate the melody at the presentation stage, and even when the performer is different at the recognition stage.

Responses to new songs

It remains to be determined how the effects of instruction and performance rendition influenced (1) the accuracy of tune recognition in new songs, and (2) the 'contamination' effect of words on 'old tune' responses. The relevant

Table 4. *Mean proportion of 'old tune' responses (Exp. 2)*

Performance:	Same		Different	
Instructions:	General	Analytic	General	Analytic
New songs (mismatch)	0.60	0.69	0.69	0.60
Old songs	0.94	0.94	0.73	0.77

data are shown in Table 5, separately for the two instruction conditions but averaged over performance conditions, which showed no effect here.

The data for new songs were analyzed in an ANOVA across subjects on 'old tune' responses with the factors of tune (old *versus* new), words (old *versus* new), instruction condition (general *versus* analytic), and performance rendition (same *versus* different). In the ANOVA across items, only the first two factors were included.

As Table 5 shows, tune recognition in new songs was poor, and worse than in Experiment 1. The main effect for tunes was not significant in either analysis. (That is, the mean hit rate for tunes was not significantly higher than the mean false alarm rate.) Although it may appear that subjects had some success in recognizing tunes when the words were old (compare hits with false alarms in 'old words' columns), in fact the tune × words interaction was not significant. Thus subjects did not recognize tunes better than chance in new song contexts. Moreover, tune recognition was equally poor regardless of instructions or performance renditions.

However, there was a highly significant main effect for words, with subjects giving many more 'old tune' responses when the words were old (mean of 0.59) than when the words were new (mean of 0.28), $F(1,40) = 50.01$, $p < 0.0001$, across subjects, and $F(1,18) = 37.58$, $p < 0.0001$, across items. In addition, this effect interacted with instructions, $F(1,40) = 4.15$, $p < 0.05$, in that it was less pronounced in the analytic instruction condition.

In summary, Experiment 2 showed that the integration effect for memory of original melody and text is both obligatory and abstract. Analytic instructions did not reduce the integration effect; subjects were unable to ignore the words in storing melodies at the presentation stage. Moreover, the integration effect is generalizable across different performance renditions in the presentation and test stages.

That tunes were recognized so poorly in new contexts would seem to argue

Table 5. *Mean proportion of 'old tune' responses: new songs (Exp. 2)*

		General Instructions				Analytic Instructions		
		Words				Words		
		New	Old	Mean		New	Old	Mean
Tune	New	0.20	0.57	0.38	New	0.36	0.50	0.43
	Old	0.23	0.64	0.44	Old	0.35	0.64	0.50
	Mean	0.21	0.61		Mean	0.35	0.57	

for an even stronger form of the integration hypothesis—a 'holistic' conception of melody/text relations in memory. Even instructions to listen analytically did not improve tune recognition. While it seems possible that there is an asymmetry in memory integration, such that tunes are more dependent on the words than *vice versa*, our findings may simply reflect the fact that the tunes were much harder to remember than the words. This may be an artifact of the folksong genre, since the melodies were in many ways similar (small range; G tonic; homogeneous rhythm; mostly step-wise melodic motion), but the texts were very different from each other. Moreover, texts could be recognized by a single salient word (e.g., 'Babylon' or 'turkey'), but the tunes had no such advantage. Ultimately, the question of which component is more memorable boils down to the nature of the materials. We might imagine a reversal of the memory advantage for words if we had selected texts that were very similar to each other, and melodies that were widely discrepant. But in the case of the folksongs used in our experiments, a natural asymmetry exists in the salience of texts and melodies.

General discussion

We conclude from these experiments that melody and text are integrated in memory to a considerable degree. We found that the familiarity of old tunes and words (when mismatched) was an insufficient predictor of the superior recognition for original old songs. Moreover, we found no evidence that subjects can voluntarily reduce the degree of integration of melody and text. Indeed, what was surprising was not only the size of the integration effect, but that subjects seemed to be unaware of it. Thus melody and text appear not to be stored as independent components. On the other hand, a stronger or 'holistic' form of integration appears to be untenable, at least as far as the text is concerned. Our results leave open the possibility that, under certain conditions, the melody may be completely integrated with the text (but not *vice versa*).

In addition to this integration, there appears to be a reciprocal 'contamination' in familiarity judgments of melodies and texts. This effect may be voluntarily reduced, though not entirely removed. The effect itself may be an artifact of the selected materials, and it may depend on other factors that are not clear at present; we have no explanation for the difference between Experiments 1 and 2 in the magnitude of the influence of word familiarity on tune judgments.

One question that is left unresolved by the present experiments is the degree to which tune recognition in old songs may have been due to subtle

changes imposed on a melody by the specific texts employed. Two possibilities are a semantic effect and a prosodic effect. For example, specific semantic connotations may become associated with a melody when it is heard in connection with a text about animals, cobblers, lullabies, dancing, and so forth. These connotations may facilitate tune recognition in old songs or hinder recognition when the text is different. To take an extreme example (Fig. 2, item b), it may be difficult to recognize a melody originally heard in connection with a bluebird coming through a window, when that melody is later heard in connection with an "old sow's hide."

An alternative hypothesis is that different texts impose prosodic or submelodic variations on melodies. A change in text results in a drastic change in the segmental structure of the words, which may have modified to some extent what was nominally the same melody. For example, different patterns of consonants, vowels, stresses, and voicing may influence the onset and decay characteristics of tones and the precise degree of stress given to them. Thus similarity of submelodic structure may have facilitated tune recognition, even across different performance renditions, although it can hardly account for the whole old song advantage.

We note here a natural asymmetry in the relation between (audible) melody and text: While a tune can exist perfectly well without any words (when played on a musical instrument, for example), words always have some kind of 'tune,' if only the nonmusical one provided by the prosody of spoken language. In the context of a song, the musical tune in large measure takes over the function of prosody and thus becomes an aspect of the suprasegmental properties of the words. Viewed in this way, it is quite conceivable that memory for tunes is more dependent on memory for words than *vice versa*; certainly, outside the realm of music the prosody of speech is remembered, if at all, only as an aspect of the words by which it is carried. We hope to investigate this interesting parallel between speech and music in future experiments.

Appendix

Pairs of folksong excerpts with interchangeable texts

All folksongs from Erdei (1974).

Number/title	*Number/title*
1 9: Hunt the slipper	92: Cape Cod girls
2 12: Let us chase the squirrel*	73: Christ was born*
3 15: Who's that tapping at the window?	82: Mary had a baby
4 16: How many miles to Babylon?*,†	120: Nuts in May
5 21: Poor little kitty puss*	80: Turn the glasses over
6 22: Down in the meadow	68: The old woman and the pig
7 27: Hush little baby	13: Bye, bye baby
8 32: Bluebird	55: The old sow
9 38: Ida Red*,†	39: Mama, buy me a chiney doll
10 52: Dear companion	88: Wayfaring stranger
11 67: I lost the farmer's dairy key	128: Watch that lady
12 69: Old turkey buzzard	72: My good old man
13 78: Hold my mule	102: Needle's eye
14 99: When the train comes along	132: Hushabye*,†
15 103: Housekeeping	147: My old hen*
16 148: I'm goin' home on a cloud	138: The raggle taggle gypsies
17 110: Give my love to Nell*	137: Blow, boys, blow
18 122: Cripple Creek	129: The little dappled cow
19 142: Goodbye girls, I'm going to Boston	144: Cradle hymn
20 2: The boatman	86: The Derby ram

* Minor alteration was made in text.
† Minor alteration was made in melody.

References

Best, C.T., Hoffman, H. and Glanville, B.B. (1982) Development of infant ear asymmetries for speech and music. *Percep. Psychophys.*, *31*, 75–85.

Deutsch, D. (1969) Music recognition. *Psychol. Rev.*, *76*, 300–307.

Dowling, W.J. (1973) Rhythmic groups and subjective chunks in memory for melodies. *Percep. Psychophys.*, *4*, 37–40.

Dowling, W.J. (1978) Scale and contour: Two components of a theory of memory for melodies. *Psychol. Rev.*, *85*, 341–354.