

## On the Nature of Melody-Text Integration in Memory for Songs

MARY LOUISE SERAFINE

*Vassar College*

JANET DAVIDSON AND ROBERT G. CROWDER

*Yale University*

AND

BRUNO H. REPP

*Haskins Laboratories*

In earlier experiments we found that the melodies of songs were better recognized when the words were those that had originally been heard with the melody than when they were different. Similarly, song texts were better recognized when sung with their original melodies. Some possible causes of this "integration effect" were investigated in the present experiments. Experiment 1 ruled out the hypothesis that integration was due to semantic connotations imposed on the melody by the words, since songs with nonsense texts yielded the same effect. Experiments 2 and 3 ruled out the possibility that the earlier results were caused by a decrement in recognition when a previously heard component is tested in an unfamiliar context. The results support the notion of an integrated memory representation for melody and text in songs. © 1986 Academic Press, Inc.

Songs consist of two components, melody and text, which seem to be separable in a number of ways. They can be performed, perceived, and notated separately, and in practice may be composed by different artists. At least intuitively, however, the melody and text of a song seem more tightly related than two arbitrary simultaneous events. The components of a song seem more integrated, for example, than a spoken voice with background music. These observations raise questions about the memory representation for songs and whether it consists of independent (separate) or integrated components.

In a previous study (Serafine, Crowder, & Repp, 1984) we found evidence for what

we termed the *integration effect*—the tendency for a melody to be better recognized when the text was the one with which the melody was originally heard than when the text was different. Similarly, there was a tendency for the text to be better recognized when sung with the original melody than with a different melody. The effect for melody recognition was very robust. It held across performances by different singers and could not be eliminated voluntarily by our subjects when we instructed them to focus on melody only. We concluded that melody and text form an integrated memory representation.

Integrated memory for melody and text may explain some of the experiences that people commonly have in recalling and recognizing song components. For example, if asked to recite the words to their national anthem, many people would have to sing the song or at least rehearse it subvocally in order to generate the words. Also, many people do not recognize even a very fa-

This research was supported by NSF Grant 82-19661 to R. Crowder and by NICHD Grant HD-01994 to Haskins Laboratories. Requests for reprints may be sent to M. L. Serafine, Department of Psychology, Vassar College, Poughkeepsie, NY 12601.

miliar melody if it is sung with different words. Examples are the folksong "Baa, Baa Black Sheep," which has the same melody as "Twinkle, Twinkle Little Star," and the folksong "Merrily We Roll Along," which has the same melody as "Mary Had a Little Lamb." The integration effect may also underlie the informal observation (Gottlieb, 1984) that young children are frequently unable to sing only the melody of a song if asked to replace the words with a repeated syllable such as "la." Their tendency to respond by speaking the syllable, by singing some spontaneous, unrecognizable melody, or by refusing to respond altogether may be evidence that they are unable to access the melody without its text.

Our previous study of melody-text integration employed the following method, which was similar to that used in the present experiments. Subjects heard a serial presentation of excerpts from 24 largely unfamiliar folksongs. The presentation was immediately followed by a recognition test in which two types of items were heard: (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 melody with new words, (b) old melody with new words, (c) new melody with old words, and (d) old melody with old words that had been sung to a different melody in the original presentation ("mismatch songs"). The critical finding was that recognition of a melody (or text) under the old song condition was superior to recognition under the mismatch condition. That is, recognition of a component was better when it was paired with its original component than with a different, even if equally familiar component. The experiments reported here were intended to evaluate two interpretations of the obtained integration effect.

*The semantic hypothesis.* The integration effect could be caused by the semantic connotation that words impose on a melody. In the more usual cases a melody

may be imbued with qualities implied by the text's meaning, even if the melody on its own would not normally convey that meaning. For example, words may make some aspect of the melody particularly salient. In the present folksongs, reference to a cobbler may make a repetitive melodic pattern seem to suggest hammering; reference to a bluebird may make higher pitched or ascending tones seem to imply flying, birdsong, etc. In some admittedly more rare cases, the melody may overtly mimic the meaning of the words, as when a repeated eighth-note figure appears on the words "tapping at the window." More generally the text of a sea chantey, hymn, lullaby, or other stylized song could trigger (even unconscious) recognition of the special tonal and rhythmic conventions that are characteristic of such songs.

Once the melody of a song is taken to be especially related to a particular meaning, its recognition may be inhibited in the context of a different, especially if incongruous, meaning. What has suggested hammering or a birdsong is less recognizable in the context of Cape Cod or an old sow's hide. The semantic hypothesis, then, accepts the reality of melody-text integration and attributes it to the semantic level. (Note that this hypothesis could account only for the integration effect in melody recognition, not for that in text recognition.)

*The decrement hypothesis.* By contrast, a second interpretation denies that the observed integration effect implies an integrated memory representation. Rather, the integration effect could be an artifact of the deleterious, distracting influence that a "wrong" component has on an already familiar component. For example, the memory representation of a melody may be quite independent of its text, and under normal circumstances may be just as easily recognized in one condition as another. However, the mismatch condition, precisely because it contains different words, may distract or confuse subjects and de-

press melody recognition. In such a case the integration effect would be only an experimental artifact. The decrement hypothesis can be tested by comparing recognition of components in old songs and mismatch songs to the recognition of melodies and texts presented alone (hummed or spoken respectively).

Experiment 1 addressed the semantic hypothesis for melody recognition. Experiments 2 and 3 addressed the decrement hypothesis for melody and text recognition, respectively. All three experiments employed the same general procedure: Subjects heard a serial presentation of folksong excerpts, followed immediately by a recognition test for melodies or words in which the items represented different combinations of old and new components. Because all three experiments employed variations of the same musical materials, these are described in some detail before the experiments proper.

#### GENERAL METHOD

Songs that we believed would be unfamiliar to the average listener were drawn from a collection of indigenous American folksongs compiled by Erdei (1974).<sup>1</sup> Twenty pairs of song 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). Interchangeability of words and melodies within a pair was crucial to the construction of plausible recognition foils. Thus, with two exceptions each text within a pair contained the same number of syllables, and each text contained a suitable stress pattern that would fit with either melody. The exceptions were Song Pairs 11 and 17, where one text was shorter by a syllable, and thus one syllable was sung across two tones ("slurred"), as

<sup>1</sup> In our earlier studies subjects had estimated the number of songs that seemed familiar to them after a presentation of 24 excerpts from these songs, and the means of these estimates were 1.4 and 1.2, respectively, in different experiments.

is normally the case in the different verses of a song. (The opening "O-oh" of our national anthem is an example.)

Each pair of excerpts yielded four different songs, a total of 80. Figure 1 shows a sample pair of interchangeable melodies, and Figure 2 shows examples of the five types of test items that can be generated from each pair. These materials allowed for counterbalancing so that every presentation item could be tested against every possible test item type. Thus, natural variations among the folksongs were controlled.

In some cases minor alterations were made to the 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. 2, test item a). However, in all cases the texts and melodies were identical across presentation and test versions of a song.

The excerpts were recorded on tape, sung by a female in the alto range, at a tempo represented by one beat per second. A silent metronome was employed to ensure an accurate beat, but because of normal metric variations in the songs (e.g., "double time") the subjective tempo of the excerpts was not necessarily uniform. All songs were notated with G as the tonic, although they varied in key, mode, and starting tone. The excerpts were sung as notated, except transposed down a fifth or twelfth to the appropriate range. A pitch pipe was used to ensure starting pitch accuracy. The experimental tapes were dubbed from a master tape, with a 5-s interval of silence between presentation items and a 10-s response interval after each test item.

#### EXPERIMENT 1

The semantic hypothesis holds that the integration effect is due to semantic connotations that the words of a song impose on its melody. If this hypothesis were correct, the integration effect should disappear when the semantic meaning of the words is eliminated. In the present experiment sub-

Melody	Text
A	a When the train comes a-long, When the train comes a-long.
	b Hush a- bye, don't you cry, go to sleep lit- tle babe.
B	b Hush a- bye, don't you cry, go to sleep lit- tle babe.
	a When the train comes a- long, When the train comes a-long.

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

SAMPLE PRESENTATION ITEMS

I'm just a poor way- far-ing strang-er.

Here comes a blue- bird through the- win- dow.

Hold my mule while I dance Jo-sey, Hold my mule while I dance.

Who's that táp- ping at the win- dow? }

>>>

Mar- y had a ba- by, O Lord.

Ma- ma buy me a chin-ey doll, Ma- ma buy me a chin-ey doll.

FIG. 2. Sample presentation and test items (a, new melody, new words; b, old melody, new words; c, new melody, old words; d, old melody, old words—mismatched; e, old song).

jects heard a presentation of 24 folksong excerpts in which the words had been translated into nonsense. The presentation was followed immediately by an 18-item recognition test comprising 6 each of the following types of items:

(a) old songs (old melody, old nonsense words) exactly as heard in the presentation;

(b) new songs (new melody, new nonsense words) that had not been heard in the presentation;


(c) mismatch songs (old melody with old nonsense words that had been sung to a *different* melody in the presentation).

The main prediction was that if the semantic hypothesis were correct, melody recognition should not be better in the old song condition than it is in the mismatch condition. On the other hand, if the integration effect is due to factors other than the semantic connotation of words, then the effect should still hold when nonsense words are employed.

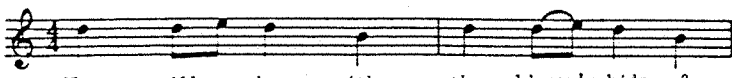
#### Method

*Materials.* Eighteen of the 20 pairs of interchangeable folksong excerpts listed in the Appendix were used to generate presentation and test stimuli (Song Pairs 4 and 10 were omitted, since these each con-

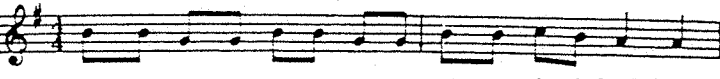
#### SAMPLE TEST ITEMS

a 


One year a- go both Jack and Joe set sail---'cross the foam.

b 


What will we do with the old sow's hide---?

c 

Hold my mule while I dance Jo-sey, Hold my mule while I dance.

d 

Who's that tap-ping at the win- dow?

e 

Ma- ma buy me a chin-ey doll, Ma- ma buy me a chin-ey doll.

FIG. 2—Continued.

tained a song that was more frequently identified as familiar by subjects in our earlier studies). Each of the 36 texts was translated into a nonsense text by applying the following rules:

1. Vowels remain the same.
2. Consonants are interchanged according to the following list, where, if the right-listed consonant appears it is changed into the left-listed consonant and vice versa. Phonetic classes are preserved.

B	G
K (QU, C)	T
L	Y (or F)
M	N
P	D
S (C)	F
H	J
R	W
Z	V
Sh, Th	Ch

3. Whenever necessary, license was taken with the above rule to ensure pronounceability and to eliminate accidental semantic meaning.

The following are examples of translated texts:

Original: Cobbler, cobbler, make my shoe.  
 Nonsense: Tog-glue, tog-glue, nate nie choo.  
 Original: Cape Cod girls they have no combs.  
 Nonsense: Tade top berf shey jaze mo tong.

The excerpts were sung and recorded on tape as described under General Method.

*Design.* Three parallel sets of presentation and test sequences were constructed from the set of 18 pairs of excerpts. Each set was administered to a different group of subjects. In the presentation sequences (24 items), half the excerpts were melodies with nonsense words derived from their original texts (type *Aa* or *Bb* in Fig. 1), and half were melodies with nonsense words derived from the companion, interchangeable text of the pair (type *Ab* or *Ba* in Fig. 1). In the test sequences (18 items), each of

the three types of test items (old, new, and mismatch song) occurred six times. Further, across the three subject groups, each presentation excerpt was tested against each of the three test item types. For Test Tape 1, the three item types were assigned at random to the 18 items available (for example, old, new, and new for the first 3 items). Thereafter Test Tapes 2 and 3 were derived accordingly (for example, mismatch, old, old, and new, mismatch, mismatch, respectively).

The presentation and test excerpts were generated successively from Song Pairs 1 through 20 (omitting 4 and 10), in the order listed in the Appendix. Thus, the interval between each presentation item and its corresponding test item was roughly constant. Note that each of the "mismatch" test items required 2 presentation excerpts, since the old words of one excerpt would be paired with the old melody of another excerpt. When two such presentation excerpts were required, they immediately followed each other on the tape. (If anything this convention would inflate performance in the mismatch condition, working against the hypothesis of an integration effect.) The resulting total of 24 presentation excerpts represents the 12 excerpts necessary for the old and new test items (six each), plus the 12 excerpts necessary for six mismatch items which each requires 2 excerpts.

*Procedure.* Testing was conducted individually in a quiet laboratory in which presentation and test tapes were heard over loudspeakers. Subjects were instructed to listen carefully to a presentation of 24 songs that sound like folksongs, except that the words have been changed to nonsense. They were told that their "memory for the songs would be tested later," but they were given no further information. The test sequence followed immediately. For each item, subjects were asked to indicate on the answer sheet whether they had "heard that exact melody before—that is, just the

musical portion" (yes or no), and to indicate the degree of confidence they felt in their judgment by marking a three-point confidence rating scale (1 = not very confident, 3 = very confident). No advance information was given about what types of items would occur on the test.

*Subjects.* Thirty-seven Yale undergraduates with undetermined levels of musical training were paid to participate. The three subject groups contained 13, 12, and 12 subjects, respectively.

### *Results and Discussion*

Yes/no responses with confidence ratings were translated into a single rating that ranged from 1 to 6, where 1 represents *very confident no* (did not hear melody), and 6 represents *very confident yes* (did hear melody). Mean ratings for the old, new, and mismatch conditions were 4.47, 2.60, and 3.76, respectively. The results of two analyses of variance for the three conditions were significant: With subjects as the sampling variable,  $F(2,72) = 51.94$ ,  $p < .001$ , and with the 18 song pairs as the sampling variable,  $F(2,34) = 38.35$ ,  $p < .001$ . Post hoc analyses (Scheffé procedure) revealed that melody recognition under the old song condition (mean = 4.47) was significantly better than it was under the mismatch condition (mean = 3.76), both across subjects,  $p < .01$ , and across song pairs,  $p < .05$ .

Thus, the integration effect was confirmed with the new materials used here. Melodies were recognized better when they were paired with their original text than when paired with another, even if equally familiar text. Since this effect held when nonsense texts were used, the semantic hypothesis must be ruled out as an explanation for the integration effect. This does not imply, however, that semantic integration of melody and text never occurs. Indeed, especially in those cases where the melody directly symbolizes textual meaning (e.g., repeated eighth notes on

"tapping"), integration on the semantic level seems likely. What Experiment 1 does show, however, is that integration does not *depend* on semantic factors.

### EXPERIMENT 2

Thus far, we have attributed the performance advantage in old songs over mismatch songs to a recognition *superiority* in the former condition. The decrement hypothesis, on the other hand, holds that the seeming advantage in old songs is due to the deleterious, distracting effect that "wrong" words have on melody recognition under the mismatch condition. If this hypothesis were correct, it could account for the performance advantage in old songs without recourse to an integrated memory representation. Perhaps the melody by itself could be recognized well without the original words, but adding new or mismatched words somehow disguises the retained melodic information.

In the present experiment, subjects heard a presentation of 24 consecutive folksong excerpts, followed by a 20-item recognition test. (Normal texts, not nonsense, were used throughout.) The test items were of five types:

- (a) old songs (exactly as heard in the presentation);
- (b) mismatch songs (old melody with old words from a *different* song in the presentation);
- (c) old words with new melody;
- (d) a hummed version of an old melody from the presentation ("old hum"); and
- (e) a hummed version of a new melody that had not been heard in the presentation ("new hum").

The decrement hypothesis predicts that melody recognition when old words are present (as measured by responses to mismatch songs and old words with new melody) will be poorer than melody recognition when no words are present (as mea-

sured by responses to old hum and new hum). In essence, the hummed conditions provide a baseline against which to measure two influences. First, if there is a decrement caused by "wrong" words, then discrimination of old and new melodies should be better when they are hummed than when they are presented with old (but mismatched) words. Second, if melody-text integration has a positive or facilitative effect on melody recognition, then old (intact) songs should have a recognition advantage over old hummed melodies.

#### Method

*Materials.* The materials consisted of the same set of 20 pairs of folksongs with interchangeable texts (not nonsense) that were described previously, except that additional recordings were made by the same female alto of hummed versions of the melodies. In this experiment, two recordings done on separate occasions were made of each stimulus. This allowed for different performances to be used across presentation and test items, thus eliminating the possibility that the physical identity of old song and old hum test items (including even accidental sounds) could contribute to superior melody recognition on those items.

*Design.* Five parallel sets of presentation and test sequences were constructed using (in the order listed) the 20 pairs of folksong excerpts in the Appendix. Each set was administered to a different group of subjects. In the presentation sequences (24 items), half the excerpts were melodies with their original texts (type *Aa* or *Bb* in Fig. 1) and half were melodies with texts borrowed from their companion song (type *Ab* or *Ba* in Fig. 1). In the test sequences (20 items), each of the five types of items (old song, mismatch, old words with new melody, old hum, new hum) occurred four times. Across the five subject groups each presentation item was tested against each of the five possible test item types, which were assigned by following a Latin square de-

sign. Each of the mismatch test items required two presentations, which immediately followed one another on the tape.

*Procedure.* The procedure was the same as that used in Experiment 1. Subjects were told to listen carefully to a presentation of 24 excerpts from simple folksongs and that their "memory would be tested later." They were not told that only melody recognition would be tested. Prior to the test they were told that items on the test would be either hummed melodies or melodies with words, but in all cases they were to disregard the words and indicate whether they had "heard this exact melody before—that is, just the musical portion." Subjects indicated "yes" or "no" on the answer sheet and gave a confidence rating.

*Subjects.* Forty Yale undergraduates with undetermined levels of musical training were paid to participate in the study. They were divided equally among the five presentation/test sequences.

#### Results and Discussion

As in the first experiment, subjects' responses were translated into ratings ranging from 1 to 6 where 1 represents *very confident no* (did not hear melody) and 6 represents *very confident yes* (did hear melody). Means for the five conditions—old songs, mismatch songs, old words with new melody, old hum, and new hum—were 4.71, 3.73, 3.21, 3.99, and 3.11, respectively. The results of analyses of variance on these means were significant both across subjects,  $F(4,156) = 26.76$ ,  $p < .001$ , and across song pairs,  $F(4,76) = 17.37$ ,  $p < .001$ .

*Confirmation of the integration effect.* Post hoc analyses (Scheffé procedure) revealed that melody recognition under the old song condition (mean = 4.71) was superior to that in the mismatch condition (mean = 3.73), both across subjects,  $p < .01$ , and across song pairs,  $p < .01$ . This confirms the integration effect found in the previous experiment.

*Disconfirmation of the decrement hy-*



*pothesis*. For this analysis subjects' melody recognition performance was measured by difference scores with a theoretical range of  $-5$  to  $+5$ , where incorrect recognitions were subtracted from correct recognitions (hits minus false alarms). The mean difference score when *old words* were present (rating for mismatch minus rating for old words/new melody) was  $.52$ . The mean difference score when *no words* were present (rating for old hum minus rating for new hum) was  $.88$ . The difference between these means narrowly missed the conventional level of significance,  $t(39) = 1.89$ ,  $.06 < p < .07$  (with subjects as the sampling variable), indicating that melody recognition was not significantly lower when old words were present than when no words were present. This result fails to support the decrement hypothesis, which holds that poorer recognition in the mismatch than in the old song condition (the integration effect) could be due to the fact that wrong words depress melody recognition performance. On the other hand, because the difference was close to statistical significance, we should leave this hypothesis tentatively open, the more so because melody recognition in both conditions was near chance.

The alternative hypothesis, however, that original old songs have a positive, facilitative effect on melody recognition was supported by the following results. The mean difference score when *original old words* were present (rating for old song minus rating for old words/new melody) was  $1.49$ , which is significantly greater than the mean difference score when *no words* were present ( $.88$  as above),  $t(39) = -2.61$ ,  $p < .02$  (with subjects as the sampling variable). Thus melodies were better recognized in the presence of their original old words than on their own, without words.

*Criterion effects.* To assess criterion effects, we analyzed the tendency to respond "yes, I heard the melody," whether correct or incorrect, when *old words* were

present and in the hummed conditions. The overall rating when old words were present (mean of mismatch and old words/new melody) was  $3.47$ , which is not significantly lower than the overall rating of  $3.55$  in the hummed conditions (mean of old hum and new hum). The Scheffé procedure yielded no significant difference across subjects or across song pairs. Thus, by itself, the presence of old words did not increase subjects' tendency to respond "yes, I heard this melody" when they heard a particular song.

#### Summary

The decrement hypothesis was not supported in the present experiment and the positive, facilitative effect of original old words on melody recognition was confirmed. Even leaving open the possibility that a larger experiment would show a significant performance decrement when familiar-but-wrong words are present (relative to hummed conditions), we can conclude that the advantage of original old songs over mismatch songs does not *depend* on such a decrement in the latter condition.

#### EXPERIMENT 3

The purpose of Experiment 3 was to test the decrement hypothesis for text recognition rather than melody recognition. In order to conduct a rigorous test of this hypothesis and because our earlier studies had shown that recognition for our folk-song texts was near ceiling, nonsense texts were used in the presentation and test sequences. Following a 24-item presentation of folksongs with nonsense texts, subjects heard a 20-item test comprising the following types of test items: (a) old songs; (b) mismatch songs; (c) old melody with new words; (d) a spoken rendition of an old nonsense text ("old words"); and (e) a spoken rendition of a new nonsense text ("new words").

The decrement hypothesis holds that text recognition is poorer in the mismatch

than in the old song condition not because melody and text are integrated, but rather because the presence of a wrong melody in the mismatch condition depresses text recognition. Thus the decrement hypothesis predicts that text recognition will be poorer when an old melody is present (as measured by responses to the mismatch songs and old melody with new words) than it is when no melody is present (as measured by responses to old words and new words).

#### Method

*Materials.* We used the same set of 20 folksong pairs described previously, except that songs were sung with nonsense texts derived in the manner of Experiment 1. As much as possible, spoken texts used the rhythm of the first melody of each pair, so that spoken test items did not deviate rhythmically from the original presentation. Because of the difficulty of duplicating exact pronunciations of nonsense words, we did not record duplicate performances of all the stimuli. Thus, in the case of "old songs" and "old words" conditions, identical performances were used in the presentation and test.

*Design.* The design was exactly analogous to that of Experiment 2.

*Procedure.* The procedure was identical to that of Experiment 2, except that subjects were asked, "Did you hear this exact text before—that is, just the words?"

*Subjects.* Twenty Yale undergraduates with undetermined levels of musical training were paid for participating in the study. Subjects were equally divided among the five presentation/test sequences.

#### Results and Discussion

Responses were translated into text recognition ratings ranging from 1 to 6, as in the previous experiments. Means for the five conditions—old songs, mismatch songs, old melody with new words, old words, and new words—were 4.66, 3.73, 3.90, 2.90, and 3.23, respectively. The re-

sults of two analyses of variance were significant across subjects,  $F(4,76) = 16.18$ ,  $p < .001$ , and across song pairs,  $F(4,76) = 11.14$ ,  $p < .001$ .

*Confirmation of the integration effect.* The results were analogous to Experiment 2. Text recognition in the old song condition (mean = 4.66) was superior to that in the mismatch condition (mean = 3.73). The Scheffé procedure was significant across subjects,  $p < .01$ , and across song pairs,  $p < .05$ . This result confirms the integration effect: A nonsense text is easier to recognize when paired with its original melody than with a different, even if equally familiar melody.

*Disconfirmation of the decrement hypothesis.* Subjects' text recognition can be measured by difference scores (hits minus false alarms). The mean difference score when an *old melody* is present (mismatch minus old melody/new words) is  $-.18$ , which is not lower than  $-.33$ , the mean score when *no melody* is present (old words minus new words). This result fails to confirm the decrement hypothesis because the presence of a wrong melody does not depress text recognition below what it is when no melody is present. However, text recognition was so poor that old words—whether paired with a melody or not—were not rated as more familiar than new words.

On the other hand, the hypothesis that the original old melody has a positive, facilitative effect on text recognition was supported by the following results. The mean difference score when the original old melody was present (rating for old song minus rating for old melody/new words) was  $.76$ . This is significantly higher than the mean difference score when no melody was present, in the spoken condition ( $-.33$  as above),  $t(19) = -4.52$ ,  $p < .001$  (with subjects as the sampling variable). Thus, nonsense texts were better recognized in the presence of their original old melody than on their own, in spoken form.

*Criterion effects.* A look at the overall

means suggests that familiarity ratings were subject to a criterion effect. Subjects were more likely to respond "yes, I heard that text" when an old melody was present (mean of mismatch and old melody/new words = 3.81) than when just the spoken text was present (mean of old words and new words = 3.06). The difference between these means is significant (Scheffé procedure across subjects,  $p < .01$ , and across song pairs,  $p < .01$ ). Thus, the presence of a familiar melody makes the text *seem* more familiar, whether or not it was heard in the original presentation. This effect must be distinguished from the integration effect, which is the facilitative effect that the *original* melody, as opposed to a new one, has on recognition of a text that *has* been heard before.

#### GENERAL DISCUSSION

Integration of melody and text in memory for songs is an experimental result, not an explanation, and a full account of it remains to be articulated. In the present experiments we have clarified it in two ways. First, Experiment 1 showed that the ordinary semantics of language are not required for integration. However much the lyrics for a well-known song seem to "fit" the music, the robust effects we obtained across all of the experiments in this and the previous article must be caused by something else. This is not to say that perhaps in ways more subtle than those evidenced here, the emotional tone of a melody could not affect subjects' interpretation of a text and hence their memory representation. But the integration effect, at least with the present materials, does not depend on such factors.

Second, Experiments 2 and 3 showed that integration of components in song recognition is a genuine advantage of hearing the song exactly as it was before, not confusion or interference produced by a novel setting. This conclusion must be tempered by the results obtained in Experiment 2, where the decrement hypothesis was not

strongly disconfirmed. Nevertheless, the advantage for "exact" old songs cannot be wholly or even primarily an artifact of interference, because positive facilitation occurred apart from this nonsignificant decrement.

By hearing the song "exactly as it was before," however, we mean the song as an abstraction rather than as an acoustic event. In Experiment 2 of the present paper and in Experiment 2 of Serafine, Crowder, & Repp (1984), different recorded performances of the songs were used in presentation and testing. This is important in ruling out what could be called an "acoustic" hypothesis—people otherwise might recognize old songs well by seizing on some performance artifact such as a note out of tune, a vocal glitch, or even an extraneous background sound.

Clearly, melody-text integration depends neither on the acoustic identity of a reheard song nor on semantic interaction between the components. Rather, we suggest that integration in memory may result from other, more subtle effects that melody and text have on each other. These may be thought of, broadly, as prosodic effects in that they concern the nonsemantic sound pattern of either melody or text. For example, a text's consonant pattern, vocal timbres, and accents may affect the attack and decay patterns, stresses, or other aspects of tones in a melody. Consider consonant patterns. Changing "Tea for two" to "Me for you" entails changing the sound pattern from one of sudden onsets and short durations to one of gradual onsets and more prolonged durations. Such changes, even if they were to occur on melody tones that were nominally identical, would in fact change the musical quality of the tones in question. What this means is that a melody is physically different depending on the words to which it is sung. In a similar way, melody can exert an effect on the words. Patterns of pitch, loudness, stress, and articulation (e.g., staccato and legato) in a melody may affect

pronunciation of individual words as well as prosody of the entire text.

If such effects were substantial, it should not be surprising that melodies are better recognized with their original words; they are in a sense "more" the same melodies than with different words or a hummed version. Likewise, a text is "more" the same words when sung to the same melody than when not.

If this reasoning is correct, then some transformation such as that used to generate nonsense words in Experiment I could be informative. If the mismatch conditions were constructed so that the degree of change in melody or text is minimized (by comparison to the old song) then the integration effect should be much reduced. In the example above we noted the consequences of changing "Tea for two" to "Me for you." If we changed "Gee zor goo" to "Bee vor boo" there should be much less change and correspondingly less integration.<sup>2</sup>

#### *On Association*

We began this program of experiments out of curiosity about an unexplored point in music cognition concerning songs. Almost at once, however, we found ourselves up against fundamental issues in the ancient concept of association. We readily conceded that melody and text could become connected in the sense that presentation of one would lead to retrieval of the other. We never tested for this simple connectionism, but have no doubt our materials could be presented as paired associates and would yield, eventually, associations by this definition. Melody and text could theoretically be associated, in this sense, and yet still be represented independently; that is, each could retain its integrity as a single component and yet be attached to the other.

<sup>2</sup> However, such a manipulation would also increase the tendency to confuse old and new texts, which may be an insurmountable methodological problem.

Our approach has insisted, at least in principle, on a different and a considerably stronger result. We require, instead, that the individual components be to some extent unrecognizable on their own, as opposed to when paired with their original companion. Thus, in this paper, we were at pains to show that the melody on its own, when hummed, was not recognized as well as when restored to its original wording: in fact recognition was close to chance. If the melody could have been recognized independently of the words, then people would have been able to do as well in the hummed condition as they did in the old song condition. This distinction between independent units attached to each other and units that undergo transformation by virtue of having been combined corresponds to the distinction between "mental compounding" and "mental chemistry" in the psychologies of James and of John Stuart Mill, respectively (see Boring, 1957, chap. 12).

In contemporary work on human learning and memory, our research is most closely related to Tulving's on *encoding specificity* (Tulving & Thomson, 1973). He, too, capitalizes on the result that when a word occurs in a particular learning context, that context can be a better aid to retrieval than the target word itself. For example, Thomson and Tulving (1970) presented the word *glue* as a potential learning aid next to the target word CHAIR. Later, people were better able to recall CHAIR, given the cue *glue*, than they were able to remember CHAIR when it was presented alone for recognition. The context apparently had changed the representation of the target (encoding specificity), just as we claim the text and melody change each other when presented together in a song. Of course, the type of change involved is quite different in songs. While Tulving's results reflect mental changes, melody and text (perhaps in addition) have physical effects on each other. What remains for future research is whether and how such changes affect the memory representation for songs.

## APPENDIX

*Pairs of Folksong Excerpts with Interchangeable Texts*

	Number/title	Number/title
1.	9: Hunt the Slipper	92: Cape Cod Girls
2.	12: Let Us Chase the Squirrel <sup>a</sup>	73: Christ Was Born <sup>a</sup>
3.	15: Who's That Tapping at the Window?	82: Mary Had a Baby
4.	16: How Many Miles to Babylon? <sup>a,b</sup>	120: Nuts in May
5.	21: Poor Little Kitty Puss <sup>a</sup>	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 <sup>a,b</sup>	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 <sup>a,b</sup>
15.	103: Housekeeping	147: My Old Hen <sup>a</sup>
16.	148: I'm Goin' Home on a Cloud	138: The Raggle Taggle Gypsies
17.	110: Give My Love to Nell <sup>a</sup>	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

Note. All folksongs from Erdei (1974).

<sup>a</sup> Minor alteration was made in text.

<sup>b</sup> Minor alteration was made in melody.

## REFERENCES

- BORING, E. G. (1957). *A history of experimental psychology* (2nd ed.). New York: Appleton-Century-Crofts.
- ERDEI, P. (1974). *American folksongs to read, sing, and play*. New York: Boosey and Hawkes.
- GOTTLIEB, L. (1984). *Text dependence in the young child's performance of songs*. Unpublished undergraduate research project. Vassar College, Poughkeepsie, NY.
- SERAFINE, M. L., CROWDER, R. G., & REPP, B. H. (1984). Integration of melody and text in memory for songs. *Cognition*, 16, 285-303.
- THOMSON, D. M., & TULVING, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology*, 86, 255-262.
- TULVING, E., & THOMSON, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352-373.

(Received June 11, 1985)

(Revision received November 1, 1985)

