How the Melody Facilitates the Message and Vice Versa in Infant Learning and Memory

Erik D. Thiessen\textsuperscript{a} and Jenny R. Saffran\textsuperscript{b,c}

\textsuperscript{a}Carnegie Mellon University, Pittsburgh, Pennsylvania, USA
\textsuperscript{b}University of Wisconsin–Madison, Madison, Wisconsin, USA
\textsuperscript{c}Waisman Center, University of Wisconsin–Madison, Madison, Wisconsin, USA

Infants are often presented with input in which there are multiple related regularities, as is the case in musical input with both melodic and lyrical structure. Adult learners often learn more easily from complex input containing multiple correlated regularities than from simplified input. Do infants also capitalize on complexity, or instead do they benefit from simplified input? In this series of experiments, infants were presented with music in which melodic and lyrical structure predicted each other, or in which only one type of regularity was presented in isolation (melodies alone, or lyrics presented with no melody). Infants learned lyrics more easily when they were paired with a melody than when they were presented alone; similarly, they learned melodies more easily when they were paired with lyrics than when they were presented alone. There are several potential mechanisms that could explain how infants’ learning is facilitated by complex input, suggesting important implications for learning in infants’ natural environments.

Key words: infant learning; infant memory; lyrics; melody; redundant cues

William James famously described the infant’s world as a “blooming, buzzing confusion” (p. 462).\textsuperscript{1} This quotation encapsulates the common intuition that the complexity of the environment is likely to confuse or overwhelm infants. Consequently, it is often assumed that simplified input is most optimal for infant learning. For example, infant-directed speech has been argued to promote infants’ learning because it is less complex than adult-directed speech.\textsuperscript{2} Similarly, Newport’s “less is more” hypothesis\textsuperscript{3} suggests that infants learn language more successfully than adults because their cognitive limitations force them to process a more limited portion of the input.\textsuperscript{4}

However, in some situations, infants actually learn more successfully from complex input, in which there are multiple correlated possible relations to be learned, than they do from simplified input. Complex input can provide infants with multiple redundant cues, and these cues can facilitate learning, especially when the regularities to be learned are related to each other. For example, redundant information helps infants make connections between auditory and visual stimuli more easily, an ability that is important for identifying the referents of words.\textsuperscript{5} Older children learning novel nouns also take advantage of redundant cues.\textsuperscript{6} Similarly, infants’ search for hidden objects is more successful if there are more cues identifying the location of the object.\textsuperscript{7} In each of these situations, infants benefit from richer, more complex input that provides multiple sources of information.

A second reason why complex input may be beneficial to learners is that more complex stimuli are often more interesting or arousing than simple stimuli. Consider prosody (the
rhythm and pitch of spoken language) as an example. While prosody increases the variability and complexity of speech, infants listen longer to speech with the exaggerated prosody characteristic of speech directed to infants than they do to speech with more monotonous prosody. Infants also learn more quickly from infant-directed speech than more monotonous speech. This is likely due to the fact that infant-directed speech maintains infants’ attention, and infants learn more quickly in a state of sustained attention. As such, there are at least two ways in which complex input can facilitate learning: by providing additional redundant cues and by supporting infants’ attention.

Note that in complex input—with correlated regularities available to be learned—there are multiple potential facilitative relations. Consider the relation between melody and lyrics in music. It is well known that melodic structure facilitates adults’ learning of lyrics. Melody provides a redundant cue to the identity of the lyrics; for example, in the song *Mary had a little lamb*, both the word *little* and the notes that accompany it redundantly predict the subsequent word *lamb*. But just as melody may facilitate learning about lyrics, lyrics can facilitate learning about melody. This kind of bidirectional facilitation—where related aspects of a complex stimulus reciprocally facilitate learning of each other—is clearly possible for adult learners. In addition to music, it can be seen in language. Consider the relation between semantics and syntax. Many words in English are inherently ambiguous (e.g., *watch* can refer either to a timepiece or to the act of gazing). The syntactic structure in which a word occurs helps to disambiguate potential meanings of a word. At the same time as syntax facilitates semantic retrieval, adults’ knowledge of semantics influences judgments about syntactic structure. Adults expect that a phrase like *the policeman arrested* is much more likely to take a direct object (such as *the perpetrator*) than a phrase like *the criminal arrested*. To account for findings like this, many models of adult linguistic processing incorporate bidirectional connections between levels of linguistic organization, such as phonemes and words, or between semantics and syntax.

However, it is unclear whether infants would benefit from the possibility of bidirectional facilitation in the same way as adults would. While prior research demonstrates that infants benefit from multiple cues, these experiments typically present infants with multiple cues that point toward a single learning problem (such as sequence learning), rather than allowing infants to solve multiple problems simultaneously. Infants are cognitively limited relative to adults, and these limitations may prevent them from detecting or fully benefiting from bidirectional relations between aspects of complex stimuli. Additionally, infants’ inexperience in particular domains may mean that they fail to use some of the available cues. Thus, our question in this series of experiments is whether infants, like adults, show bidirectional facilitation when learning regularities are simultaneously available in complex input. To explore this question, we assessed whether infants show evidence of reciprocal facilitation between melody and lyrics when learning simple songs. In Experiment 1, we tested whether infants learn the serial order of lyrics more easily when they are accompanied by melody, as adults do. In Experiment 2, we reversed the direction of effect, and asked whether the serial order of lyrics facilitates learning about melody, as it does in adults.

**Experiment 1**

**Method**

**Participants**

Forty infants between 6.5 and 8.0 months old from English-speaking families were assigned to either the Spoken (*M* = 7.2 months) or Sung (*M* = 7.3 months) condition, with half of the infants in each condition. To obtain data from 40 infants, it was necessary to test 54. The other
14 were excluded (8 from the Spoken condition and 6 from the Sung condition) for the following reasons: crying or squirming too vigorously to respond to the lights (7), looking times of less than 3 s to one or both side lights (6), and falling asleep (1).

**Stimuli**

We familiarized infants with two strings of digit names during the familiarization period: 9-7-3-1-5 and 6-2-8-0-4. In the Spoken condition, 9-7-3-1-5 and 6-2-8-0-4 were spoken in an adult-directed register. In the Sung condition, each of the two word strings was sung with a different melody. The sequence 9-7-3-1-5 was sung to the melody B4-G4-C #4-A3-E4, and the sequence 6-2-8-0-4 was sung to the melody A4-D4-C4-Bb3-F4. The first melody can be perceived either as atonal or as modulating from one key to another (it ends in A major, but also contains G, which is not a part of A major). The second melody is more tonal, and can be perceived as either D minor or B-flat major; the noncanonical ordering of the pitches suggests a possible modulation from D minor to B-flat major. The melodies were similar in contour, and contained the same number of syllables. In each sequence, four of the digits were monosyllabic. The bi-syllabic digits in each string were produced with equivalent emphasis on the first and the second syllable, but each syllable was shortened so that the bi-syllabic digits were equivalent in duration to the monosyllabic digits. The speaker was instructed to produce numerical strings at the same tempo, whether speaking or signing. Her productions were subsequently digitally edited to ensure concordance in rate and duration of the items across conditions. Each string lasted approximately 2.5 s, with a pause of 1 s between strings. In both conditions the strings alternated, and infants heard each of the strings 12 times, for a total duration of 1 min 23 s. The rate and amplitude of the word strings were equal across conditions.

All test items were spoken in an adult-directed register similar to the Spoken condition, with a duration of 2.5 s each. Test items were produced by a different female speaker than the strings infants heard during familiarization. Familiar items were numerical sequences identical to those that infants heard in the familiarization phase. Novel items consisted of the same numbers with the 2nd and 4th numbers inverted. Thus, the two Familiar test items were 9-7-3-1-5 and 6-2-8-0-4, while the two Novel test items were 9-1-3-7-5 and 6-0-8-2-4.

Previous research indicates that infants have more difficulty distinguishing Novel from Familiar items when the test items differ from the familiarization stimuli on some dimension. For example, when learning labels for novel objects, infants are more likely to successfully discriminate between Familiar and Novel labels if the words are produced in isolation during both training and testing. If the words are embedded in a sentence at training, but produced in isolation during test trials, infants are less likely to succeed. By presenting all of the infants with spoken test items (which are more similar to the materials in the Spoken familiarization condition), we ensured that demonstrating learning would be more difficult for infants in the Sung condition than in the Spoken condition. This allows a more conservative test of the hypothesis that melodies help infants learn the order of words.

**Procedure**

Infants were tested using the Headturn Preference Procedure in a sound-attenuated room while seated on a parent’s lap. An experimenter outside the room coded the infant’s looking behavior on a video monitor. To eliminate bias, the parent inside the room listened to masking music over headphones.

During the familiarization period, the infant heard one set of numerical strings (either spoken or sung). During this phase, the strings played continuously while the lights were lit and extinguished contingent on the infant’s looking behavior. After familiarization, 12 test trials were presented. Six of the trials were Familiar
trials, and six were Novel trials. Test items were presented in random order. A test trial began with a central flashing light to draw the infant’s gaze forward. When the experimenter signaled the computer that the infant had fixated on the center light, one of the side lights began to flash, and the center light was extinguished. When the infant turned in the direction of the flashing side light, repetitions of one of the test items were presented from the speaker beneath the light. When the infant looked away for more than 2 s, the test item stopped playing, and the center light began to flash.

Results and Discussion

We averaged listening times to the set of Familiar (6.1 s, SE = 0.5) and the set of Novel (6.2 s, SE = 0.4) test items for infants in the Spoken condition (Fig. 1). Similarly, we averaged listening time to Familiar (6.6 s, SE = 0.6) and Novel (8.1 s, SE = 0.5) test items for infants in the Sung condition. To assess whether there were significant differences between these two conditions, we performed a 2 × 2 (Condition by Item) ANOVA. There was no main effect of condition: F(1, 28) = 1.14, n.s. Even though the test items were spoken (and thus more similar to the Spoken condition), this did not drive overall longer looking time for infants in the Sung condition. This may suggest that infants’ responses at test were more driven by the sequential information in the test items than their surface features. This suggestion is further supported by a main effect of test item F(1, 28) = 6.34, P < 0.05, indicating that infants differentiated between test items that followed or violated the sequence they had been exposed to during familiarization. This effect was largely due to infants’ preference for Novel items in the Sung condition; infants in the Spoken condition showed little difference in their looking time to Novel or Familiar test items. Because of this difference across conditions, there was a significant interaction between condition and test item: F(1, 38) = 4.46, P < 0.05. This interaction is consistent with our prediction that infants would learn more successfully in the Sung condition than in the Spoken condition.

To further explore this interaction, we performed planned comparisons on infants’ looking times to Familiar and Novel test items in the two conditions. A t-test (all t-tests were two-tailed) indicated that the difference in looking times between test items in the Spoken condition was not significant: t(19) = 0.21, n.s. Infants’ failure to discriminate between test items that follow or violate the word order presented during familiarization suggests that exposure to spoken lyrics, without melody, was not sufficient for learning these lyrics given the brevity of familiarization. By contrast, infants in the Sung condition looked significantly longer during Novel test items than during Familiar test items: t(19) = 3.31, P < 0.01. Unlike infants in the Spoken condition, infants who heard each of the two-digit strings paired with a unique melody distinguished between Familiar and Novel trials, despite the fact that the test items were spoken and not sung. These results indicate that the presence of melody facilitated infants’ learning about the order of the digit strings. Only infants who heard these lyrics paired with melodies learned them well enough to detect when the digit order was violated.

Experiment 2

The results of Experiment 1 indicate that the presence of melody facilitates infants’ learning
of lyrical order, consistent with prior demonstrations of facilitation between melody and lyrics for adult learners. Note that this indicates that infants’ learning is facilitated by more complex input; while lyrics alone are a simpler input than lyrics plus melody, the presence of melody facilitates learning of lyrics. Our next question is whether the reverse relation is true: does the presence of lyrics facilitate melodic learning? To address this question, we reversed the design of Experiment 1, and asked whether the presence of word sequences facilitates learning about melodies. Infants in the No Lyrics condition heard two melodies sung with a single syllable, thus introducing no word-order information. This condition parallels the Spoken condition of Experiment 1. Infants in the Lyrics condition heard each melody sung with a unique digit string—akin to the Sung condition in Experiment 1. If word-order information facilitates learning about melodies, then infants in the Lyrics condition should learn more successfully than infants in the No Lyrics condition.

Method

Participants

Thirty infants between 6.5 and 8.0 months old were randomly assigned to either the No Lyrics (M = 7.5 months) or Lyrics (M = 7.2 months) condition, half to each condition. To obtain data from 30 infants, it was necessary to test 40. The other 10 were excluded (5 from each condition) for the following reasons: crying or squirming too vigorously to respond to the lights (4), parent stopping the experiment (3), and looking times of less than 3 s to one or both side lights (3).

Stimuli

We familiarized infants with two five-note melodies during the familiarization period: G4-E4-C4-D4-A4 and C5-G4-E4-F4-C4 (melodies were sung in the octave above middle C). Each melody lasted approximately 2.6 s, and there was a pause of 1 s between melodies. The melodies alternated in both conditions, and infants heard each of the melodies 12 times, for a total duration of 1 min 29 s. The rate and amplitude of the melodies was equal across conditions.

Unlike the melodies used in Experiment 1, these melodies were tonal in the key of C major. While the results of Experiment 1 demonstrate that infants can learn from atonal melodies, we chose to use tonal melodies in Experiment 2 in an effort to reduce the rate of infant drop-out. In the No Lyrics condition, the two melodies G4-E4-C4-D4-A4 and C5-G4-E4-F4-C4 were both sung using the repeated syllable doo. In the Lyrics condition, the two melodies were associated with two different five-digit strings: 9-7-3-1-5 (G4-E4-C4-D4-A4) and 6-2-8-0-4 (C5-G4-E4-F4-C4).

The test items were sung by a different adult female than the one who sang the familiarization melodies. All test items were sung using the syllable doo, with a duration of 2.6 s. As in Experiment 1, the simple familiarization materials (here, the No Lyrics condition) were most similar to the test materials, providing a conservative test of the hypothesis that complex input facilitates learning. Familiar items were note sequences identical to those that infants heard in the familiarization phase. In Novel test items, the 2nd and 4th notes were inverted. Thus, the two Familiar test items were G4-E4-C4-D4-A4 and C5-G4-E4-F4-C4, while the two Novel test items were G4-D4-C4-E4-A4 and C5-F4-E4-G4-C4. The test items were designed to create a difficult test contrast; in particular, the note changes used to produce the Novel test items did not alter the original melodic contours.

Procedure

The procedure was identical to that of Experiment 1.

Results and Discussion

We averaged listening times to the set of Familiar (7.5 s, SE = 0.9) and the set of Novel (6.3 s, SE = 0.7) test items for infants in the No
Infants’ looking time to Familiar and Novel test items in Experiment 2. Error bars represent standard error.

Figure 2. Infants’ looking time to Familiar and Novel test items in Experiment 2. Error bars represent standard error.

Lyrics condition (Fig. 2). Similarly, we averaged listening time to Familiar (5.3 s, SE = 0.7) and Novel (6.1 s, SE = 0.7) test items for infants in the Lyrics condition, and then performed a 2 × 2 (Condition by Item) ANOVA. There was no main effect of condition or item. However, there was a significant interaction: $F(1, 28) = 12.68, P < 0.01$. As in Experiment 1, we explored this interaction using planned comparisons of looking to Familiar and Novel test items in each condition. Infants in the No Lyrics condition looked significantly longer to familiar test trials: $t(14) = 2.53, P < 0.05$. Infants in the Lyrics condition looked significantly longer to novel test trials: $t(14) = 2.57, P < 0.05$. These results contrast with those of Experiment 1, where infants (in the Sung condition) showed a familiarity preference, and infants in the Spoken condition showed no preference.

These results indicate that infants in both the No Lyrics and the Lyrics condition learned the melodies—they were both able to detect the change in the sequence of notes between the Familiar and Novel test trials. However, the results also suggest that infants in the Lyrics condition learned more successfully than infants in the No Lyrics condition. Note that infants in the No Lyrics condition showed a familiarity preference, listening longer to the test items that match their exposure. Infants in the Lyrics condition showed a novelty preference, listening longer to test trials that mismatched the tone sequence to which they were exposed. A wide array of experimental data indicates that the difference between a novelty preference and a familiarity preference is an important one. In particular, when a learning task is difficult, learning is often incomplete during the familiarization phase. This leads infants to prefer the familiar test items, which they have yet to completely process. $^{10,22,24–26}$ When a learning task is easy, infants are more likely to display a novelty preference. The fact that infants showed a familiarity preference in the No Lyrics condition, and a novelty preference in the Lyrics condition, indicates that infants found the tonal sequence easier to learn in the Lyrics condition—a particularly striking result given that the test materials were sung on *doo* and thus more similar overall to the No Lyrics exposure condition than the Lyrics condition. This pattern of results is consistent with the hypothesis that the presence of lyrics facilitates learning about melodies, in much the same way as the presence of melody facilitates learning about lyrics.

Discussion

These experiments demonstrate that complexity can sometimes benefit infant learners. Infants learn the serial order of words most easily not when words are presented alone, but when those words are presented as lyrics paired with a consistent melody. Similarly, the presence of a consistent lyrical order facilitates learning about melody. Even in infancy, the simplest input does not necessarily lead to the best learning outcomes. Across a variety of learning situations, complexity can actually be facilitative for infant learners. For example, infants identify categories more easily when they are exposed to variable exemplars of a category rather than to a single repeated exemplar. $^{27,28}$ Similarly, infants identify nonadjacent regularities (of the form X-A-Y, where X predicts Y) if the elements intervening between the nonadjacent items are highly variable. $^{29}$ But these results do more than demonstrate that complexity can be facilitative, even for
the youngest learners. Additionally, these results demonstrate that different aspects of complex input can be mutually facilitative. Melody facilitates learning about lyrics—but lyrics also facilitate learning about melody. This finding extends previous research demonstrating that infants can benefit from unidirectional facilitation, especially in language: using one aspect of the linguistic input to facilitate learning about other aspects.\(^{30,31}\) If infants can also take advantage of bidirectional facilitation, complex input may accelerate learning in a dynamic, nonlinear manner.

There are several domains in which infants may benefit from bidirectional facilitation. In language, there are correlated regularities (such as word meaning and word order) that infants could learn simultaneously.\(^{32}\) Similarly, in musical input, infants are exposed to both lyrical and melodic regularities simultaneously presented in the input. Of course, the melodies and lyrics used in these experiments are vastly simplified compared to the complex input that infants experience in their natural environment. Similarly, while redundancy between aspects of different parts of complex input exists in natural systems, it is rarely as clear as the completely deterministic relation between melody and serial word order that infants experienced in these experiments. Although these experiments demonstrate that complexity is not always a hindrance to learning, there are clearly many situations in which learning from simpler input is more successful than learning from more complex input.\(^{33,34}\) One possibility is that natural systems, such as music and language, present infants with a near-optimal level of complexity: enough to facilitate, but not enough to overwhelm. Natural musical and linguistic systems may have been shaped, by successive generations of learners, into a form that is well calibrated to infant learners.\(^3,35\) Only by better understanding the abilities and limitations of infant learners will it be possible to assess this hypothesis.

There may be multiple mechanisms enabling the bidirectional facilitation infants demonstrated in these experiments. One possibility is that infants learn more easily when melody and lyrics are both present because they find the combined input more interesting than either kind of stimulus alone, and this increased interest facilitates learning.\(^{11}\) If this is the sole mechanism at work, infants might learn equally well from numerical sequences spoken in an infant-directed register as from sung numerical sequences, as both infant-directed speech\(^{10}\) and singing may promote attention equally. Another possible mechanism of facilitation is that infants learn more quickly when there is a second source of information that provides redundant cues to identify structure in the input.\(^{36}\) The current results do not distinguish between these two accounts. However, it is unlikely that these mechanisms are ever completely dissociated in the natural environment. Both mechanisms may well combine in a complementary manner, both in the current experiments and in infants’ learning from natural input. While the underlying mechanism or mechanisms remain ambiguous, both achieve the same end: infants learn more efficiently when both melodic and word-order information are presented simultaneously than when either is presented alone.

The current results, along with others indicating that infants are able to take advantage of multiple converging cues,\(^{37}\) suggest that infants are capable of attending to multiple related regularities simultaneously. This may enable infants to learn about complex stimuli, such as music or language, in ways that would be impossible if infants attended solely to any single source of information at a time. When the input is organized in such a way that learning yields bidirectional facilitation between different aspects of the input, then learning may be accelerated far beyond what would otherwise be possible. Research that focuses on infants’ ability to learn about any single type of information may underestimate infants’ learning abilities. If this is the case, then complexity in the input may not be an obstacle for infant learners to overcome. Instead, complexity in the input may
support learning by providing infants with an opportunity to use what they have learned to facilitate subsequent learning.

Conflicts of Interest

The authors declare no conflicts of interest.

References


