Developmental changes in indicators that literal interpretations of homonyms are associated with conflict

MICHELE M. M. MAZZOCCO
Kennedy Krieger Institute,
Johns Hopkins University School of Medicine

(Received 11 February 1998. Revised 7 October 1998)

ABSTRACT

Processes by which children interpret homonyms were examined. Participants were 16 two- and three-year-olds, 32 four-year-olds, 32 seven-year-olds and 32 ten-year-olds. Each child individually was asked to interpret keywords from stories read aloud by an examiner. Keywords were homonyms, nonsense words, or unambiguous words. For the three older groups, response times (RTs) to unambiguous words were significantly shorter than RTs to homonyms and nonsense words, when interpretations for all three-word types were consistent with the story contexts. Seven-year-olds had longer RTs for homonyms versus nonsense words. RTs did not vary among responses to homonyms. Four- and seven-year-olds had longer RTs for inappropriate, versus appropriate, nonsense-word interpretations. The number of cues recalled from a story about a keyword did not differ across appropriately versus inappropriately interpreted homonyms, but did differ across appropriately versus inappropriately interpreted nonsense words. These findings have implications for understanding how children arrive at literal interpretations of homonyms.

INTRODUCTION

When a young child hears a new word, the word’s unfamiliarity is a signal that a new word meaning must be derived (Carey, 1978; Carey & Bartlett, 1978). Children successfully use linguistic context to derive a word’s
meaning across a wide range of linguistic and situational settings. In the case of homonym interpretation, the familiarity of the word itself may interfere with a child’s well-mastered use of linguistic context. Indeed, there is a limit to the effect that linguistic context has on increasing the likelihood of correctly selecting the less familiar meaning of a homonym (e.g. hare vs. hair; Beveridge & Marsh, 1991). Young children often persist with inferring a familiar but incorrect referent (Beveridge & Marsh, 1991; Campbell & Bowe, 1977; Mazzocco, 1997) in the face of incompatible linguistic context. This study was designed to examine whether time to respond in an interpretation task varies as a function of whether the word meaning in question is ambiguous, and specifically whether response times were longer for (1) homonyms vs. non-homonyms, and (2) familiar vs. contextually appropriate interpretations of homonyms. Also examined was whether the amount of story information recalled varied as a function of appropriateness of homonym interpretation.

Children’s difficulty with homonym interpretation has been illustrated in several studies. Campbell & Bowe (1977) found that, when asked to illustrate a story, three- to five-year-olds were more likely to depict a familiar vs. appropriate interpretation of a homonym embedded in the story context even when the story context did not support the familiar interpretation. For instance, some of the four-year-olds in their study drew a hair (vs. a hare) running across a field. Mazzocco (1997) showed that children who successfully interpreted novel, nonsense words on the basis of story context were nonetheless unsuccessful when asked to interpret ‘pseudo-homonyms’—familiar words used to refer to a new referent. The story contexts used in Mazzocco’s study, across nonsense and pseudo-homonym conditions, were identical; however, four-year-olds and many seven-year-olds interpreted the pseudo-homonyms literally, by indicating responses that were consistent with the pseudo-homonyms’ familiar meanings. In contrast, ten-year-olds’ and adults’ responses were consistent with the story contexts. Even when asked to enact the stories with paper dolls, the younger children made literal interpretations of pseudo-homonyms. Similarly, Beveridge & Marsh (1991) demonstrated that although four-year-olds do benefit when the linguistic context supports a homonym’s secondary meaning, 70% of the four-year-olds in their study persisted with the primary meaning when presented with a simple story, and 38% of the six-year-old children persisted with primary meanings even when presented with elaborate stories.

Backscheider & Gelman (1995) showed that three-year-olds have both the metalinguistic awareness and cognitive flexibility necessary for accurate homonym interpretation. In their study, three-year-olds succeeded when asked to ‘find another kind of (item)’ when first shown another item with the same name. The three-year-olds in their study demonstrated knowledge that homonyms could indicate referents across different categories rather than
referents within a category (such as that a flying bat and a baseball bat are not members of the same object category). Thus a one-to-one mapping rule does not appear to govern preschoolers’ naming or interpretation responses overall. Despite the clear effect of homonymy on word interpretation demonstrated across these studies, the processes that underlie homonym interpretation are less clear. One question that remains is, why do preschoolers persist with familiar interpretations of homonyms despite having necessary prerequisite skills to do otherwise?

Some insight into this question may be revealed through observation of behaviours that occur during interpretation tasks, including response times and recall of story information. Examining the accuracy of interpretation per se does not provide sufficient information for understanding the nature of homonym interpretation. For instance, although a child’s correct word interpretation may imply that the child derived meaning from the contextual information of the story, an incorrect response could indicate either that (a) the child was unaware of the conflict between the familiar word’s well established meaning (e.g. hair) and the intended referent of the word (e.g. hare), or that (b) the child was aware of this conflict between meanings without being able to resolve it. Instead, a measure of response time may provide evidence for these important scenarios. A longer response time may reflect that the child was considering alternative choices or simply experiencing indecision. For a response time assessment, it is important to have a baseline measure of response times for easy, or more automatic, word interpretation responses. In the study, this was achieved by measuring children’s response times when interpreting accurately used familiar words.

In addition to response times, the variables examined in the present study included the quantity of linguistic context recalled when retelling stories that contain homonymous words, and the quality of the recall with respect to accuracy or distortion of story information. The use of these variables has been useful for understanding children’s interpretation of other linguistically ambiguous forms, such as metaphors. Developmental changes reported for metaphor interpretation include an increase in the frequency of responses consistent with the metaphors’ meanings from six to nine years of age, and from nine to fourteen years of age (Winner, Rosentiel & Gardner, 1976; Winner, Engel & Gardner, 1980). Young children’s ‘magical’ interpretations appear to suggest unquestioning belief in what adults (or others in general) report. Older children’s ‘metonymic’ interpretations reflect a distortion of the story context in such a way that the metaphor makes sense. Perhaps children who resort to a more familiar interpretation of a homonym do so in a similar fashion; it was hoped that whether this is true might be revealed by children’s verbal recall of the stories presented to them that contained homonyms. An example of recall reflecting a ‘magical’ interpretation would be simply that a hair was running across a field. An example of a ‘metonymic’
response in which the story and the homonym’s more familiar meaning become compatible is that the hair on a rabbit was blowing in the wind as the rabbit ran across the field. In the present study, children’s recall of the stories presented to them was examined as a means by which to examine whether either of these sorts of responses was observed when the more familiar interpretation was given to an ambiguous word.

Similarly, response time has been successfully used as an indicator of different levels of awareness (Patterson, Cosgrove & O’Brien, 1980). Patterson et al. asked four-, six-, eight-, and ten-year-olds to select a stimulus from an array of four choices that included a plain and a striped triangle, and a plain and striped circle. They found that the children’s response time varied with the degree to which the selection instruction was informative. The instruction to select the ‘striped triangle’ was informative, whereas reference to ‘the triangle’ was partially informative, and reference to ‘a shape’ was uninformative. In addition to recording which item each child selected as the speaker’s referent, Patterson et al. recorded the children’s response times, eye contact with the experimenter, body movement, and hand movement. Their results indicated that each behaviour was related to message type for at least some age group(s), and that response time varied significantly for all four age groups. The informative messages led to the shortest response times, and the uninformative messages led to the longest response times. Of critical
importance is the fact that the four-year-olds exhibited this effect of message type on reaction time regardless of the accuracy of their referent selection responses. The implication to be drawn from these results is that children may indicate an awareness of the differences between the ambiguous and non-ambiguous words through slower response times for the former, even when they opt for the more familiar interpretation. To examine this possibility, response time measures were employed in the present study, and were examined across word type and interpretation accuracy.

In summary, the primary objectives of the present study were to further explore the factors associated with young children’s interpretations of homonyms. Of specific interest was whether children’s retelling of stories varied as a function of the familiarity of the key word present in the story (homonym, nonsense word, or familiar word used accurately); and whether response times differed as a function of whether the word interpretation was contextually appropriate vs. contextually inappropriate. These objectives were met through analyses of video taped sessions of the children who participated in a previous study (Mazzocco, 1997), across whom the aforementioned developmental sequence in homonym interpretation was demonstrated.

METHODOLOGY

Subjects
Sixteen two- and three-year-olds (age 2;9 to 3;3, mean 3;0), 35 four-year-olds (age 4;0 to 4;11, mean 4;6), 32 seven-year-olds (age 7;0 to 8;11, mean 7;7) from grade 2, and 32 ten-year-olds (age 10;1 to 11;8, mean 10;8) from grade 5 participated in the study. Three of the 35 four-year-olds were omitted from the analyses due to failure with the audio visual recording of their session. One additional toddler (age 3;0) was not tested because he was unable to follow the instructions for the task. The two-, three- and four-year-olds attended one of two laboratory preschools at a large university. The seven- and ten-year-olds were recruited from several public elementary schools, and from one parochial school. All of the children were from the same metropolitan region in the United States, and each child was tested on the University Campus where the laboratory preschools were located. All spoke American English as a primary language. There were no regional accents noted that would have significantly influenced the children’s or examiner’s performance during any single session.

Materials
The materials, described in greater detail elsewhere (Mazzocco, 1997), included six stories coordinated with one book of illustrations. Each of the six stories included three monosyllabic key words, one each of three types:
pseudo-homonyms, nonsense words, and familiar words (that were not ambiguous) used accurately. Thus there were 18 key words, six falling in each of three types (see Table 1). The contextual information accompanying each key word included two sentences. This information referred to a function of the object denoted by the key word and physical features of that object. There were four contextual cues per intended meaning, and each key word was repeated two times.

All key words were nominals. The unambiguous accurately used words were familiar words used to provide a check on whether the children understood the task. The pseudo-homonyms were familiar nouns used to simulate a child’s first encounter with a homonym’s less familiar meaning. These pseudo-homonyms were familiar non-homonyms used to refer to a novel referent (e.g. the word ‘chain’ was used to refer to a kite, such as in the statement, ‘Ann opened a package she got and said, ‘Look, Aunt Beth sent me a cloth chain and some string so that I can fly the chain on a windy day’
). The nonsense words (e.g. ‘flig’) were used to simulate a child’s first encounter with a new unfamiliar word. The unambiguous, accurately used words, and the nouns indicated by the intended meanings for the nonsense words and pseudo-homonyms, were chosen on the basis of their familiarity as rated by adults (Toglia & Battig, 1978).

Composition of illustration book. For each of the six stories, there were three pages of illustrations, one corresponding to each of the three types of key word. For instance, in the birthday party story that appears in the appendix, there were three pages of pictures. These three pages corresponded to the story segments concerning playing on a slide, seeing a clown, and having ice cream on birthday cake. Six colour illustrations were presented on each of the illustration pages. Each illustration was a static depiction of an object, such as an illustration of a door, wheelbarrow, or a saucepan. There were no persons or associated items that appeared in these illustrations, and no action was depicted. One of the six illustrations corresponded to the intended meaning (e.g. ice cream); and a second illustration corresponded to an object that was indirectly related to the context (e.g. a saucepan, which is found in the kitchen). For the pseudo-homonyms, one of the six illustrations corresponded to the key word’s familiar meaning (e.g. a door). The remaining illustrations were unrelated to the intended meanings.

The position of the illustration corresponding to the intended meaning was counterbalanced across the six trials for each key word category. In the pseudo-homonym condition, the position of the illustration that corresponded to the familiar meaning of the key word was also counterbalanced. The same 18 key words and intended meanings were included in each of these story sets, and 18 pages of illustrations were used.

Four sets of the six stories were constructed by varying the respective positions of the nonsense word and pseudo-homonym within each story's
format, as seen in the appendix. The order in which key word types appeared within each story was counterbalanced within subjects. Six possible orderings of the three key word types were used, one for each subject.

A second book of pictures was modified so that each picture was cut out from its background and placed in a plastic photograph sleeve. The two picture books were identical in format. However, it was possible to easily remove the cut out pictures only from the latter.

**Procedure**

Each of the 112 children was tested individually by the same female experimenter. The child was told that the examiner would read stories from a book that did not contain any pictures, and that the child’s task was to ‘…figure out which one picture (from a page of the illustration book) is the picture that goes with the story.’ The transition from one story to the next was marked by the experimenter’s comments (e.g. ‘Now I have another story. This one is about Ann,’) so that contextual information from a preceding story was not confused with a subsequent story.

A warm-up trial was presented to each child, during which the experimenter read a portion of ‘Little Red Riding Hood’, and asked the child to select a picture of the flower that Little Red Riding Hood had picked for her grandmother. Following the warm-up trial, the experimenter read one part of the first story, showed the child one page of six possible illustrations, and asked the child to ‘look at all of the pictures on this page, then show me the picture of the (key word) in the story’. The child’s picture selection was recorded, and the examiner read the next story segment.

For some of the children, the instructions were modified in that they were asked to choose ‘picture toys’ and to act out what happened in the story with paper dolls. All 16 of the two- and three-year-olds were in the ‘select picture’ condition. For the three older age groups, 16 children were in the select picture condition, and 16 children were in the act out condition. Two- and three-year-olds were excluded from the act out condition because the additional task demands and time requirements would have been too difficult for these children to follow. In the previous study, the effect of response mode was shown not to significantly affect interpretation accuracy (Mazzocco, 1997).

**Recall.** After one story (composed of three segments) was completed, the experimenter turned to the first page of illustrations/picture toys and asked the child, ‘Now I want you to tell me which one of these pictures/picture toys was in the story, and tell me everything that you can remember about that part of the story’. This recall procedure was repeated for each of the three story segments. It was included in the procedure for all of the children except the two- and three-year-olds. The above sequence of reading each
story segment, having the child select a picture, and having the child recall all three story segments was repeated for each of the six stories. The children’s interpretation responses were scored as correct if they were consistent with the contextual information. A maximum of six correct responses was possible per word type, per child.

Coding
Each of the 112 sessions was video- and audio-taped by a research assistant who was viewing the session through a one-way-mirror. The following behaviours were coded from the videotapes of the sessions: (1) response time, (2) recall score and (3) insertion score (see below). Each of these behaviours was coded by one of four coders. To establish reliability, the behaviours from 32 of the 112 sessions were coded independently by two different coders. The intercoder reliability obtained was at 92% for response time (measured as the same if the two scores were similar within one second), 88% for the number of cues recalled (out of four), and 99% for the number of insertions. The reliability figures were calculated as scores in agreement as a percentage of scores in and out of agreement.

Response times were measured by a hand-held stopwatch, for each picture selection response. Response time was defined as the amount of time that the page of illustrations was in the child’s view before the child made a selection response (by pointing to a picture/picture toy). In the event that two coders reported response times that differed by more than one second, and in the event that recall or insertions scores differed, the session in question was viewed again, and consensus reached.

Recall scores were defined as the number of cues, out of four, that a child reported during recall of each story segment, per key word. Insertions were defined as the number of pieces of information reported by a child during the recall phase of the session that had not been included in the story as it was read by the experimenter. If a child paraphrased a portion of the story as had been read by the Experimenter or used a synonym, no insertion score was assigned. For example, consider the following story segment: ‘...Brad and his sister were playing when they heard their father drive up the driveway in his boot. Their father opened the door of the boot and got out.’ If a child recalled that ‘Brad’s father came home’, this was considered a paraphrase and no insertion score was assigned. If, however, the child reported that ‘Brad saw the boot come out of the car’, the phrase ‘boot come out of the car’ would be considered an insertion because it is information that was not presented in the story, nor could it be inferred from the story, as read by the experimenter.
RESULTS

Interpretations were judged to be ‘correct’ only when consistent with the story context. Although a child’s ‘correct’ interpretation of a word may suggest that the child used contextual information to derive the word’s meaning, an ‘incorrect’ interpretation does not necessarily suggest that the child completely ignored, or did not hear, the context surrounding the word. How a child arrives at a word interpretation counter to the context is the focus of the analyses described below. Comparisons are drawn between behaviours corresponding to correctly and incorrectly interpreted words within and across word types, within each age group. Three categories of behaviour are examined in these analyses: (1) response times to interpretation tasks, and (2) the amount and (3) accuracy of information reported by the children during their recall of the stories.

Response time to interpretation task

Children’s response times for choosing a picture or picture toy for each of 18 word interpretation tasks (six for each of three word types) were examined. Of interest in this set of analyses was whether children’s response times differed as a function of the type of word being interpreted and/or as a function of the contextual appropriateness of the response (whether the child’s interpretation was consistent with the context). There was no significant effect of response mode (select picture or act out) on response time, so responses across these conditions were collapsed.

In order to eliminate statistical outliers from the data, the 18 individual response times per child were evaluated in terms of whether they deviated from both (1) the group mean for a given child’s age group, and (2) the mean of the individual child in question, by more than three standard deviations. The means and standard deviations on which these evaluations were based were derived only from responses from the same age group, word type, and accuracy of interpretation as the response time in question. That is, for each age group, separate mean response times were calculated for word interpretations that were consistent or inconsistent with the story context, for each of three word types. Thus, within each age group, there were response means for six categories: pseudo-homonym interpretations that were (1) consistent or (2) inconsistent with the story context; nonsense word interpretations that were (3) consistent or (4) inconsistent with story contexts, and (5) correct or (6) incorrect interpretations of words with familiar usage. A child’s response time was replaced if it exceeded the corresponding group mean by three or more standard deviations; and also exceeded, by three or more standard deviations, his or her own mean score for the word type and response accuracy (correct or incorrect) in question. Only 26 (1.3%) of the 2016 individual response times were changed. In each case, the outlier was
replaced with the individual child’s next longest response time for a word of the same type and accuracy of interpretation as that which was replaced. These outliers occurred in each of the four age groups, among all word types, and for both correct and incorrect responses. Thus it was not the case that any particular type of response or type of word accounted for the outliers. In fact, six of the 26 outliers occurred with correct interpretations to accurately used words, which suggests that the outliers occurred because of occasional inattentiveness, and not because of the nature of any individual word interpretation task.

Response time as a function of word type and accuracy of interpretation. The response time data were examined as a function of both word type and contextual appropriateness of word interpretation. These analyses addressed three specific questions: (1) whether response times to correctly interpreted words vary as a function of the type of word being interpreted; (2) whether children’s response times to nonsense word interpretation tasks differ as a function of the contextual appropriateness of the interpretation; and (3) whether children’s response times to pseudo-homonym interpretation tasks differ as a function of the contextual appropriateness of interpretation.

In order to compare response times for correct and incorrect responses, it was necessary to eliminate from the analyses those children who made only correct, or only incorrect, responses. Similarly, when comparing response times for correct (or incorrect) responses among different word types, it was necessary to exclude from the analyses those children who did not make any correct (or any incorrect) interpretations for one or more word types. If accuracy of interpretation and word type were to be included in a single analysis, the number of subjects who could be included (those who responded both incorrectly and correctly to some of each of the three-word types) would have been very small. Therefore, it was necessary to conduct separate analyses to investigate the effects of word type and appropriateness of interpretation on response time. The following comparisons, as described below, were included in this second set of analyses: (1) four-year-olds’ and seven-year-olds’ response times for contextually appropriate responses to nonsense words, pseudo-homonyms, and accurately used words; (2) four-year-olds’ and seven-year-olds’ contextually appropriate and inappropriate responses to nonsense words (ten-year-olds were eliminated because only 2 of the 32 ten-year-olds made any contextually incorrect interpretations of nonsense words); and (3) four-year-olds’, seven-year-olds’, and ten-year-olds’ contextually appropriate and inappropriate responses to pseudo-homonyms.

Response times for correct interpretations. To examine the response times for correct interpretations made of all three-word types, two orthogonal contrasts were carried out. Examined were the possible differences in response times for (1) pseudo-homonyms and nonsense words combined vs. accurately used
TABLE 2. Mean and standard deviation response times, in seconds, for correct (contextually appropriate) interpretations

<table>
<thead>
<tr>
<th>Word type</th>
<th>Age group</th>
<th>Two- and three-year-olds</th>
<th>Four-year-olds</th>
<th>Seven-year-olds</th>
<th>Ten-year-olds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pseudo-homonyms</td>
<td></td>
<td>3.6 (1.4)</td>
<td>4.8 (2.3)</td>
<td>5.4 (2.5)</td>
<td>3.7 (2.1)</td>
</tr>
<tr>
<td></td>
<td>n = 13</td>
<td>n = 16</td>
<td>n = 24</td>
<td>n = 29</td>
<td></td>
</tr>
<tr>
<td>Nonsense words</td>
<td></td>
<td>4.7 (2.2)</td>
<td>4.8 (3.6)</td>
<td>4.1 (3.6)</td>
<td>2.8 (1.8)</td>
</tr>
<tr>
<td></td>
<td>n = 16</td>
<td>n = 31</td>
<td>n = 32</td>
<td>n = 31</td>
<td></td>
</tr>
<tr>
<td>Accurately used</td>
<td></td>
<td>4.8 (4.4)</td>
<td>2.8 (1.4)</td>
<td>2.2 (0.8)</td>
<td>1.9 (0.8)</td>
</tr>
<tr>
<td>words</td>
<td></td>
<td>n = 16</td>
<td>n = 31</td>
<td>n = 32</td>
<td>n = 31</td>
</tr>
</tbody>
</table>

TABLE 3. Nonsense words: mean response times (standard deviations), in seconds

<table>
<thead>
<tr>
<th>Appropriateness of interpretation</th>
<th>Age Group*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Three-year-olds (^a)</td>
</tr>
<tr>
<td>Inconsistent with story context</td>
<td>7.8 (4.6)</td>
</tr>
<tr>
<td>Consistent with story context</td>
<td>4.8 (2.2)</td>
</tr>
</tbody>
</table>

\(^a\) No significant differences between response times for correct versus incorrect responses.
\(^b\) difference between response times for correct versus incorrect responses, \(p < 0.05\).
\(^c\) differences between response times for correct versus incorrect responses, \(p < 0.001\).
\(^d\) differences between seven-year-olds and the two remaining groups, on response times for incorrect responses, \(p < 0.001\).

* Data reported are only from children who made both correct and incorrect responses to nonsense words.

words, and (2) pseudo-homonyms vs. nonsense words (see Table 2). Only children who made at least one correct interpretation of each of the three-word types were included in these analyses. Thus 13 of the 16 two- and three-year-olds, 16 of the 32 four-year-olds, 24 of the 32 seven-year-olds, and 29 of the 32 ten-year-olds were included.

Response times to contextually correct interpretations of pseudo-homonyms and nonsense words combined were significantly longer than response times to correctly interpreted accurately-used words, \(F(1, 78) = 19.6, p < 0.0001\). There was a significant age X word type interaction, \(F(3, 78) = 4.85, p < 0.01\). Tests of simple effects within this contrast revealed that the effect
of word type was present for the four-year-olds, seven-year-olds, and ten-year-olds; $F$'s (1, 78) = 14.01, $p < 0.001$; 19.56, $p < 0.0001$; and 7.52, $p < 0.01$, respectively. This effect was not significant for the two- and three-year-olds, $F(1, 78) = 1.11$. Thus for the three older age groups, response times to accurately used words were significantly shorter than were response times to pseudo-homonyms and nonsense words, when these words were interpreted correctly (consistent with story context).

Response times to correctly interpreted pseudo-homonyms were contrasted with response times to ‘correctly’ interpreted nonsense words. The overall response times for these two word types did not differ significantly, $F(1, 78) = 0.57$. However, there was an interaction between word type and age, $F(3, 78) = 3.36$, $p = 0.029$. Tests of simple effects for this contrast revealed that response times to these two word types were significantly different for the seven-year-olds only, $F(1, 78) = 6.49$, $p = 0.013$, whose mean response times were 5.4 seconds and 4.1 seconds for pseudo-homonyms and nonsense words, respectively. There were no significant differences for response times to pseudo-homonyms versus nonsense words for the three remaining age groups. Thus only seven-year-olds seemed to respond with slightly different speed to these two types of words.

Response times for nonsense words. Response times for contextually correct vs. incorrect interpretations of nonsense words were examined. Fifteen of the 16 two- and three-year-olds, 23 of the 32 four-year-olds, and seven of the 32 seven-year-olds made both types of interpretations of nonsense words, so these children were included in this analysis. Only two of the ten-year-olds made any incorrect responses to nonsense words, so this age group was omitted. A (3) age × (2) (accuracy of response, correct vs. incorrect) ANOVA showed significant main effects of both factors: Age, $F(2, 42) = 4.39$, $p < 0.05$, and accuracy, $F(2, 42) = 28.93$, $p < 0.0001$. The age by accuracy interaction was significant, $F(1, 42) = 5.39$, $p < 0.01$. (See Table 3.) Tests of simple effects showed that the effect of accuracy was significant for the four-year-olds, $F(1, 42) = 5.53$, $p = 0.024$; and for the seven-year-olds, $F(1, 42) = 22.82$, $p < 0.0001$; but not for the two- and three-year-olds, $F(1, 42) = 2.22$. Furthermore, the effect of age was significant for the incorrect responses, $F(2, 42) = 5.63$, $p < 0.01$, but not for the correct responses, $F(2, 42) = 0.01$. Thus it takes children of all three of these age groups longer to make an inappropriate, rather than an appropriate, interpretation of a nonsense word, and the differences are significant for four-year-olds and seven-year-olds. Post-hoc comparisons (Newman–Keuls) showed that, for the inappropriately interpreted nonsense words, seven-year-olds’ average response time was significantly longer (17.91 seconds) than were the average response times of four-year-olds (8.43) and two- and three-year-olds (7.84), $p < 0.01$. The response times for the latter two age groups did not differ significantly.

Response times for pseudo-homonyms. In the second set of analyses, response
times to correct and incorrect interpretations of pseudo-homonyms were examined. Thirteen of the 16 two- and three-year-olds, 16 of the 32 four-year-olds, 24 of the 32 seven-year-olds, and 23 of the 32 ten-year-olds made at least one incorrect and one correct pseudo-homonym interpretation, and were included in this analysis. Response times to correct and incorrect interpretations of pseudo-homonyms were examined in a (4) age × (2) accuracy of response ANOVA, with repeated measures on the second factor. There were no significant main effects or interactions involving these factors. Thus children from all four age groups took an equivalent amount of time (3–5 seconds) to respond to pseudo-homonyms regardless of the accuracy of their interpretations. (See Table 4.)

**TABLE 4.** Pseudo-homonyms: mean response times (standard deviations), in seconds

<table>
<thead>
<tr>
<th>Appropriateness of interpretation</th>
<th>Age group*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Three-year-olds</td>
</tr>
<tr>
<td>Inconsistent with story context</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>(28)</td>
</tr>
<tr>
<td>Consistent with story context</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>(14)</td>
</tr>
</tbody>
</table>

* No significant differences between response times for correct vs. incorrect responses.
* Data reported are only from children who made both correct and incorrect responses to pseudo-homonyms.

**Recall of stories**

Only the children in the three older age groups who were asked to recall the stories read to them were included in these analyses. 

_Cues recalled_. It is possible that children’s interpretations reflect whether the children were not attentive to the stories, rather than that they failed to use contextual information to derive word meanings. It was important to investigate whether the amount of information recalled for a story segment corresponded to the accuracy of interpretation.

_Recall scores as a function of accuracy of interpretation_. A (3) age × (2) response mode × (2) accuracy ANOVA, with repeated measures on the third factor, was conducted to compare the number of cues recalled after correct interpretations with the number of cues recalled after incorrect interpretations. All three-word type categories were collapsed. All of the children in each age group made at least one correct and one incorrect response; however, recall data were missing for four of the 32 ten-year-olds due to a loss of sound in one of the videotapes. This means that 32 four-year-olds, 32...
seven-year-olds, and 28 ten-year-olds were included in this and the following analyses.

Main effects were found for age $F(2, 80) = 20.46, p < .0001$, with older children generally recalling more than younger children, and accuracy, $F(1, 80) = 9.34, p < .001$, with more information recalled when word interpretations were correct. (See Table 5.) There was no effect of response condition, $F(1, 80) = 1.25$; nor were there any significant interactions.

The relation between the amount of information recalled by children and the accuracy of their word interpretation was investigated for pseudo-homonyms and nonsense words individually. Accurately used words were not included in this set of analyses because of the low incidence of incorrect, or contextually inappropriate, responses for this word type.

Recall scores for story segments containing nonsense words. A $(2) \text{ age} \times (2) \text{ accuracy}$ ANOVA, with repeated measures on the second factor, was carried out for responses to nonsense words. Only children who made both contextually correct and incorrect responses to nonsense words were included. This meant that 23 of the 32 four-year-olds and seven of the seven-year-olds were included. (Only two of the ten-year-olds made any incorrect interpretations of nonsense words, so this age group was not included.) The results showed a main effect of accuracy, $F(1, 28) = 116.13, p < .0001$, with more information recalled when interpretations were correct, or consistent with the story context. There was no main effect of age, $F(1, 28) = .07$. There was, however, an age by accuracy interaction, $F(1, 28) = 21.93, p < .0001$, of magnitude. Tests of simple effects showed that the effect of word type was significant for both the four-year-olds and the seven-year-olds, $F$’s $(1, 28) = 41.05$ and $76.96, p’s < .0001$, respectively. As seen in Table 5, when seven-year-olds made contextually incorrect interpretations of nonsense words, they also received a score of zero on their recall for the corresponding story segment. This suggests that their incorrect interpretation could have resulted from a momentary lack of attention to the stories.

Recall scores for story segments containing pseudo-homonyms. The next question to be addressed was whether findings comparable to those described for nonsense words, above, would be found for recall for story segments containing pseudo-homonyms. Only those children who made both correct and incorrect interpretations of pseudo-homonyms were included in this analysis. Sixteen of the 32 four-year-olds, 25 of the 32 seven-year-olds, and 20 of the 28 ten-year-olds for whom recall data were available were included in a $(3) \text{ age} \times (2) \text{ accuracy}$ ANOVA, with repeated measures on the second factor. The findings from this ANOVA were quite different from those obtained for the analysis of recall following nonsense words. There was a main effect of Age, $F(1, 58) = 17.85, p < .0001$, but no effect of accuracy, $F(1, 58) = 3.91$. The age by accuracy interaction was also not significant, $F(2, 58) = .35$. 

406
TABLE 5. Mean number of cues recalled, out of 4, and standard deviations, per word type

<table>
<thead>
<tr>
<th>Appropriateness of interpretation</th>
<th>Four-year-olds</th>
<th>Seven-year-olds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pseudo-homonyms*</td>
<td>n = 16</td>
<td>n = 25</td>
</tr>
<tr>
<td>Contextually consistent</td>
<td>0.88 (0.60)</td>
<td>1.58 (0.53)</td>
</tr>
<tr>
<td>Contextually inconsistent</td>
<td>0.78 (0.48)</td>
<td>1.24 (0.67)</td>
</tr>
<tr>
<td>Nonsense words*</td>
<td>n = 23</td>
<td>n = 7</td>
</tr>
<tr>
<td>Contextually consistent</td>
<td>1.14 (0.59)</td>
<td>1.67 (0.21)</td>
</tr>
<tr>
<td>Contextually inconsistent</td>
<td>0.46 (0.44)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>Accurately used words</td>
<td>n = 32</td>
<td>n = 32</td>
</tr>
<tr>
<td>Correct</td>
<td>0.80</td>
<td>1.49</td>
</tr>
<tr>
<td>Incorrect</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

* No significant difference in number of cues recalled for correctly vs. incorrectly interpreted key words.

b More cues recalled for correctly vs. incorrectly interpreted key words, p < 0.0001.

* Data reported only from children who made both correct and incorrect responses.

TABLE 6. Number of children making insertions per age group and word type*

<table>
<thead>
<tr>
<th>Group</th>
<th>Number of insertions made across 6 trials</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 1 2 3 4 5 6 7 8</td>
</tr>
<tr>
<td>Four-year-olds (n = 32)</td>
<td></td>
</tr>
<tr>
<td>Pseudo-homonyms</td>
<td>8 15 4 1 2 0 1 1 0</td>
</tr>
<tr>
<td>Nonsense words</td>
<td>15 13 0 1 0 0 0 0 0</td>
</tr>
<tr>
<td>Accurately used</td>
<td>26 5 1 0 0 0 0 0 0</td>
</tr>
<tr>
<td>(No insertions overall: 3)</td>
<td></td>
</tr>
<tr>
<td>Seven-year-olds (n = 32)</td>
<td></td>
</tr>
<tr>
<td>Pseudo-homonyms</td>
<td>10 12 7 2 0 1 0 0 0</td>
</tr>
<tr>
<td>Nonsense words</td>
<td>24 2 2 0 0 0 0 0 0</td>
</tr>
<tr>
<td>Accurately used</td>
<td>27 5 1 1 0 0 0 0 0</td>
</tr>
<tr>
<td>(No insertions overall: 8)</td>
<td></td>
</tr>
<tr>
<td>Ten-year-olds (n = 28)</td>
<td></td>
</tr>
<tr>
<td>Pseudo-homonyms</td>
<td>21 7 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>Nonsense words</td>
<td>25 3 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>Accurately used</td>
<td>28 0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>(No insertions overall: 17)</td>
<td></td>
</tr>
</tbody>
</table>

* Insertions are reported for correctly and incorrectly interpreted key words, combined.

Insertions

Perhaps it is the case that how information is recalled (rather than only how much information is recalled) is affected by the type of word used in a story setting. In the case of uncertainty presented by a pseudo-homonym, is a
child’s recall of the story information distorted in any systematic way? To examine this possibility, the number of insertions made by children was examined as a function of the word type category and the accuracy of interpretations corresponding to the story segment for which insertions were made. When interpreting these data, it is important to consider that not all children made any insertions during recall (see Table 6).

In order to include both word type and accuracy of interpretation in a single analysis, the resulting number of subjects who made both correct and incorrect responses to all three-word types would have been very small. Therefore, it was necessary to conduct separate analyses of word type and accuracy on insertion scores.

Insertion scores corresponding to correct interpretations. A (3) age × (2) word type ANOVA, with repeated measures on the second factor, was carried out with only those children who made correct responses to both pseudo-homonyms and nonsense words. This meant that only 16 of the 32 four-year-olds, 25 of the seven-year-olds, and 24 of the 32 ten-year-olds were included. There was a main effect of age, F(2, 61) = 9.47, p < .001, but no effect of word type, F(2, 61) = .02. There was no significant interaction, F(2, 61) = .36. The results show that younger children made more insertions to story segments corresponding to responses, regardless of word type (see Table 6).

Insertion scores corresponding to incorrect interpretations. A (3) age × (2) word type ANOVA, with repeated measures on the second factor, was carried out. Only children who made contextually incorrect responses to pseudo-homonyms and nonsense words were included in this analysis. This meant that 23 of the 32 four-year-olds and seven of the 32 seven-year-olds were included. (Only two of the ten-year-olds made any incorrect interpretations of nonsense words, and they were not included.) There was no effect of age on insertion scores, F(1, 28) = .18, although there was an effect of word type, F(1, 28) = 7.27, p < .05. Both four-year-olds and seven-year-olds (in this partial sample) made more insertions to story segments corresponding to incorrect interpretations of pseudo-homonyms relative to story segments corresponding to inappropriate interpretations of nonsense words. The age x word type interaction was not significant, F(1, 28) = 3.12.

Correlations between recall scores and insertion scores
Of interest was whether the 23 four-year-olds and the 22 seven-year-olds who made insertions did so in a way that was systematically related to the amount of information they recalled about a story segment. Correlations were not examined among the ten-year-olds because the number of ten-year-olds who made insertions when retelling the stories was insufficient to allow for statistical analyses.

Correlations with recall scores were obtained among four-year-olds who made (a) contextually inappropriate and (b) contextually appropriate non-
sense word interpretations; and (c) inappropriate and (d) appropriate pseudo-homonym interpretations. The sample sizes for these four analyses were 31, 22, 31 and 16, respectively. Correlations with recall scores were obtained also for the seven-year-olds who made (a) contextually appropriate nonsense-word interpretations; and (b) inappropriate and (c) appropriate pseudo-homonym interpretations. The sample sizes for these three analyses were 32, 32 and 24, respectively. The number of seven-year-olds who made contextually inappropriate nonsense-word interpretations was insufficient to allow for statistical analyses. Of these seven correlations examined between recall scores and number of insertions, one reached statistical significance. Among four-year-olds who made incorrect interpretations of pseudo-homonyms, lower recall scores were associated with higher insertion scores, Pearson’s $r = -0.420$, $p < 0.01$. The remaining correlations coefficients ranged from $-0.30$ to $0.108$.

**Discussion**

The participants in this study included two- and three-year-olds, four-year-olds, seven-year-olds, and ten-year-olds. In a previous study, it was shown that the four-year-olds and seven-year-olds made significantly more contextually unwarranted interpretations of homonyms than did older participants (Mazzocco, 1997). However, such an interpretation of a homonym in context is not a sufficient indicator that context was ignored. Nor is it an indicator of similar processes across developmental levels. For this reason, the present study was designed to examine the nature of children’s literal interpretations of homonyms in a story context. The findings are limited to words that are nominals, and to words used to indicate familiar referents. The results are also primarily limited to words that are not phonologically similar to their referents. The pseudo-homonyms and nonsense words used in the stories were not manipulated for phonological similarity with their corresponding referents, as all but one of the word-referent pairs differed markedly in sound from each other. As seen in Table 1, the exception to this pattern is the word ‘cake’ used to refer to a ‘cage’. This pseudo-homonym was more likely to be interpreted correctly than were the remaining five pseudo-homonym and referent word pairs, as reported earlier (Mazzocco, 1997). Thus phonological similarity may play an important role in homonym interpretation, and in this study this role was not examined.

*Response times within and across groups*

The finding that response times to word interpretation tasks diminish with age is not surprising. However, the response time data are most informative when considered within, rather than across, age groups; and within age
groups across levels of response accuracy and word type. The response times from the children in the three older age groups did vary as a function of word type. For all three age groups, accurately used words were responded to more quickly than were the remaining two word types. Similarly, response times associated with contextually appropriate interpretations of both pseudo-homonyms and nonsense words were longer than response times for correct (contextually appropriate) interpretations of accurately used words, among the children from all four age groups. For all age groups, there was no difference in the response times for correctly interpreted pseudo-homonyms and nonsense-word interpretations consistent with the story context.

One possible explanation for these findings is that characteristics common to pseudo-homonyms and nonsense words account for the comparably longer response times. One such characteristic pertains to the principle of contrast (Clark 1987, 1988; Au & Markman, 1987; Markman & Wachtel, 1988), which refers to the conflict that occurs when a referent that is already associated with an established label is referred to by a new label. This is also compatible with assumption that words do not have common exemplars, one component of the mutual exclusivity bias (e.g. Merriman & Stevenson, 1997). This type of conflict with the referent labels was present for the pseudo-homonym and nonsense words in the study, but it does not occur for familiar, accurately used words. Perhaps the conflict with referent labels accounts for the longer response times for both pseudo-homonyms and nonsense words, relative to the response time for accurately used words.

An alternative explanation for the longer response times for both pseudo-homonyms and nonsense words is that a conflict – but not necessarily the same conflict – occurred for each of these two word types. Whereas the conflict with referent labels can occur for both word types, a conflict between meanings is possible only for the pseudo-homonyms. The lack of significant differences in response times to these two types of word suggests that, when children make literal interpretations of pseudo-homonyms, they do so without being aware of a conflict unique to homonyms. However, the possibility of different conflicts leading to comparably longer response times cannot be ruled out on the basis of the present data. A study based on homonyms and nonsense words with unfamiliar referents would contribute additional information regarding these alternatives.

The fact that none of the above effects on response times was significant within the youngest age group suggests that the two- and three-year-olds relied on the story context regardless of the type of word they were interpreting. However, another possibility is that response time is not a sufficiently sensitive measure to detect differences in interpretation strategies for children of this age. This may be true because the children in this age group were often distracted (e.g. by sounds outside the testing room, by other pictures in the illustration book, by the examiner’s earrings, etc.) and
did not respond to requests to select a picture as promptly as did the older children. As seen in Table 4, mean response times for contextually correct responses to pseudo-homonyms and nonsense words were comparable across these two word types, and were similar across groups; but only the two- and three-year-olds had similar response times across all three-word types. Measures other than response time may be needed to explore the processes that underlie two- and three-year-olds’ homonym interpretation processes.

Response times to pseudo-homonyms did not differ as a function of interpretation accuracy, for any of the four age groups, among children who made both accurate and inaccurate responses across the six stories. Thus it took children of all four ages just as long to make literal interpretations of pseudo-homonyms as it did for them to appropriately interpret these words, as seen in Table 4. This finding is in direct contrast with the finding that response times for nonsense words varied as a function of interpretation accuracy, for four-year-olds and seven-year-olds.

If a longer response time reflects a struggle with deriving word meaning, the differences in response time for contextually correct versus incorrect nonsense-word interpretation suggest that children who make errors are aware of the conflict with referent. However, the lack of response time differences for literal versus accurate pseudo-homonym interpretation suggests that children who make literal interpretations do not experience a conflict unique to the processes that lead to a literal interpretation. If incorrect interpretations had required more time, one could argue that the children were considering the conflict between meanings. However, it is also possible that the children were considering the conflict or indecision in both scenarios, but were uncertain about how to respond and thus sometimes made literal interpretations. What is clear is that the pattern of response times differed markedly across pseudo-homonyms and nonsense words, for four-year-olds and seven-year-olds. What is unclear is why this is so.

Only the seven-year-olds had significantly longer response times on correct pseudo-homonym versus nonsense-word interpretations. This is an interesting finding in light of the fact that accurate homonym interpretation appears to be emerging, but not mastered, in this age group. That is, most four-year-olds made predominantly literal interpretations of pseudo-homonyms, most ten-year-olds made predominately accurate interpretations of pseudo-homonyms, and performance was mixed among the seven-year-olds (Mazzocco, 1997). The longer response times seen only in seven-year-olds, when correctly interpreting pseudo-homonyms versus nonsense words, may reflect this emerging awareness and its associated indecision. Consistent with this notion is the observation that their response times to literally interpreted pseudo-homonyms were longer than seven-year-olds’ response times to accurately used words; whereas these two response times were the same among four-year-olds (as seen in Tables 2 and 3).
With respect to response times to nonsense words, seven-year-olds had much longer response times than the four-year-olds and two- and three-year-olds when their interpretations were incorrect. However, this effect was seen only among the seven seven-year-olds who made both inaccurate and accurate interpretations of nonsense words across trials. Perhaps response time is a particularly sensitive marker of interpretation uncertainty in this age group. An alternative explanation is drawn from Table 5, in which it is noted that seven-year-olds who inappropriately interpreted nonsense words recalled no details about the corresponding story segment. This suggests that the long response times reflected confusion that resulted from momentary inattention to the examiner’s reading of the story. Thus, it is important to consider response time data and recall data together when evaluating interpretation processes.

Recall of story details
When asked to recall a story, older children tended to recall more information overall than younger children, yet within each age group there was no significant difference as a function of interpretation accuracy when the word being interpreted was a pseudo-homonym. For example, when children made literal interpretations of pseudo-homonyms, they recalled as much of the contextual information as they did when making accurate, contextually-based interpretations. This important finding of no difference in amount of information recalled leads us to hypothesize that the low number of correct interpretations of pseudo-homonyms was not due to children’s lack of attention to the stories. This is in contrast to the finding that more cues were recalled when nonsense words were correctly interpreted, relative to when nonsense words were incorrectly interpreted, among both four-year-olds and seven-year-olds.

The pattern exhibited by the recall data is similar to the pattern exhibited by the response time data, as follows: there are differences between responses associated with contextually correct and incorrect interpretations of nonsense words, but not between correct and incorrect interpretations of pseudo-homonyms. When a child perceives that a new word meaning is intended, it appears that an unfamiliar nonsense word promotes the creation of a new lexical entry, whereas homonymy interferes with the establishment of a new word meaning unless the child is aware of the conflict between meanings. The fact that both response times and recall scores for story segments containing pseudo-homonyms are the same regardless of interpretation accuracy suggests that children who incorrectly interpret homonymous words are not always able to resolve this conflict. This effect of homonymy on response time and recall is apparent for each of the three age groups for whom response time and recall were investigated. These findings
suggest that the strategies used by children to interpret words are affected by homonymy.

**Children's insertion of information during story recall**

Although the insertion data provide important information, it is important to consider the data in light of the fact that not all of the children made insertions, and that not all of the insertions corresponded to the meaning conveyed by the child’s response or the story context (see Table 6). If the data on story insertions reveal any pattern of systematic story distortion, such a pattern would represent only some of the children from each age group, rather than the overall age group.

The insertion data reveal that younger children make more insertions to story segments corresponding to correct responses, regardless of word type. More insertions were made when recalling story segments corresponding to incorrectly interpreted pseudo-homonyms versus incorrectly interpreted nonsense words; this effect was seen among the four-year-olds and the seven-year-olds. Correlations were calculated to examine whether four- and seven-year-olds who made insertions did so in a manner related to the amount of information they recalled about a story segment. Among the seven correlations examined, the only significant correlation to emerge was that seen among four-year-olds who made incorrect interpretations of pseudo-homonyms, among whom lower recall scores were associated with higher insertion scores, \( r = -0.420 \). This was not observed among the seven-year-olds who made incorrect interpretations of pseudo-homonyms.

One possible explanation for this negative correlation is that when young children hear a word for which they have an established meaning, information corresponding to the established meaning is considered in addition to – or instead of – the contextual information concerning the intended novel meaning. This may lead to a distortion of the information presented, and this possibility is exemplified by three children who, after hearing the story segment wherein Brad’s father ‘…drove up the driveway in his ‘boot’,… opened the door of the ‘boot’ and got out,’ stated that (child 1:) ‘…he put the boot on’, or that (child 2:) ‘the father got out (of the car) with his boot,’ and finally (child 3:) that the father ‘drove over (the boot)’.

A response of this type is similar to the metonymic responses reported by Winner et al. (1976) in their investigation of how children interpret metaphors, as was reported earlier. Of the 28 insertions made by four-year-olds to literally

---

[1] In the United States, which is where the study was carried out, the word ‘boot’ does not correspond to automobiles. The rear storage compartment is commonly referred to as a ‘trunk’ of an automobile. This fact helps clarify the extent to which the story about driving up in the ‘boot’ presented contextual inconsistency for the child participants in this study.
interpreted pseudo-homonyms, 14 were distortions of this type; and of the 27 insertions made by seven-year-olds to literally interpreted pseudo-homonyms, 16 were distortions of this type.

However, it is important to keep in mind that the analysis of the number of cues recalled for pseudo-homonyms, as reported earlier, revealed that accuracy of interpretation did not affect the number of cues recalled. Moreover, not all of the four-year-olds’ insertions corresponded to the literal meaning of the key words.

Another explanation for the negative correlation between the recall and insertion scores is that children who have forgotten information from the stories contribute their own information (i.e. make insertions) simply as a result of being asked to report story information. This is exemplified by the fact that some of the four-year-olds’ responses were unrelated to both the key word’s literal and intended meanings. For instance, a child who heard a story wherein ‘Ann (had received) a cloth ‘chain’ and some string so that she could fly the ‘chain’ on a windy day,’ reported that Ann ‘…dropped it (the chain)’. The act of dropping a chain was not described in the story, so it is considered an insertion. However, dropping the chain is not relevant to the word’s literal or intended meaning, and the illustration corresponding to the literal interpretation of a chain depicted a heavy, thick metal chain.

The fact that only one of the seven correlations examined reached significance suggests that there is no strong relation between the amount of information that a child recalls about a story segment and the number of insertions that a child makes when retelling a story segment. Nevertheless, the quality of insertions corresponding to literally interpreted pseudo-homonyms reveals important misinterpretations of not only words but the context in which the word is presented.

Explanations for accurate interpretations of pseudo-homonyms

For the children who made contextually-based (i.e. accurate) interpretations of pseudo-homonyms, can we conclude that the conflict between meanings was apparent, and that the context was favoured over the familiar meaning? This is one possible explanation for why some four-year-olds made contextually appropriate, accurate interpretations of pseudo-homonyms. Another explanation for the occurrence of contextually-based interpretations is that some children simply believed that the speaker erred (e.g. that the speaker meant to say ‘clown’ but said ‘door’ instead). In fact, one child corrected the experimenter who had just read to the child, ‘…she saw that her bird was not in the metal cake where it lived.’ This preschooler, who made a contextually-based interpretation of the key word (cake) by choosing the picture of the cage, told the experimenter, ‘It’s supposed to be ‘cage’; you didn’t write it very well.’ Although this instance may have occurred because of the aforementioned phonological similarity between the pseudo-
homonym and its referent, it nevertheless illustrates one ‘route’ to a contextually-based word interpretation decision.

CONCLUSION

The results from this study illustrate the importance of studying accuracy of word interpretation in the context of indicators of what underlies the word interpretation process. The results also illustrate that the degree to which nonverbal indicators are informative may vary across age groups; in this study, response time appeared less informative among two- and three-year-olds than it was among the remaining three age groups. This finding is consistent with earlier work on nonverbal indicators of awareness (Patterson et al., 1980).

The amount of information recalled about the stories indicates that literal, inaccurate interpretations of homonyms made by children were not associated with attention to story details, unlike the case with nonsense words for which no details were recalled by seven-year-olds who made inaccurate interpretations. The types of insertions made when recalling the stories indicate that a significant number of four-year-olds and seven-year-olds distorted the stories they heard to conform to the familiar meaning associated with the pseudo-homonym. The children altered the stories rather than change their evidently well-established association between the pseudo-homonym and its familiar referent. This finding is consistent with the literature on children’s ‘metonymic’ interpretation of metaphor (Winner et al., 1976, 1980).

The response time data for four-year-olds and seven-year-olds suggests that children in these age groups experience a conflict, or indecision of some sort, when interpreting homonymous words and unfamiliar words. However, the specific nature of this conflict may differ across these two age groups. The seven-year-olds showed more evidence of an emerging awareness of the conflicts posed by homonyms, and had longer response times potentially evident of indecision. Four-year-olds, on the other hand, showed evidence of indecision when interpreting nonsense words, but less so when interpreting pseudo-homonyms.

It appears that literal interpretations of homonyms occur among some four-year-olds and seven-year-olds because the interpretation of the context in which the homonym appears is misdirected and distorted by an established word meaning. Moreover, children who ‘choose’ literal interpretations of homonyms do so despite evidence that they experience conflict or indecision, and the evidence for this conflict is more marked among seven-year-olds than among four-year-olds. This developmental trend is not unlike those reported for higher order cognitive problem-solving skills, and may reflect an interaction between linguistic and problem-solving skills when faced with interpreting lexically ambiguous information. At the very least, the results
from this study illustrate the complexities and variations that can lead to a literal interpretation of homonyms.

REFERENCES


APPENDIX

Story 2 as it appeared across the four story sets with key words in italics.

*Set A*

When James got to Joey’s birthday party, he saw children playing on the metal *slide* outside in Joey’s yard. James climbed up the ladder of the *slide* and went down, very fast.

Then James saw that a *door* was standing there making funny faces and doing tricks. James laughed, because the *door* looked so funny.

When it was time for dessert, Joey’s mother took some *gler* out from the freezer and said, ‘Do you want a scoop of *gler* on your piece of birthday cake?’
**Set B**

When James got to Joey’s birthday party, he saw children playing on the metal *slide* outside in Joey’s yard. James climbed up the ladder of the *slide* and went down, very fast.

Then James saw that a *gler* was standing there making funny faces and doing tricks. James laughed, because the *gler* looked so funny.

When it was time for dessert, Joey’s mother took some *door* out from the freezer and said, ‘Do you want a scoop of *door* on your piece of birthday cake?’

**Set C**

When James got to Joey’s birthday party, he saw children playing on the metal *slide* outside in Joey’s yard. James climbed up the ladder of the *slide* and went down, very fast.

When it was time for dessert, Joey’s mother took some *gler* out from the freezer and said, ‘Do you want a scoop of *gler* on your piece of birthday cake?’

Then James saw that a *door* was standing there making funny faces and doing tricks. James laughed, because the *door* looked so funny.

**Set D**

When James got to Joey’s birthday party, he saw children playing on the metal *slide* outside in Joey’s yard. James climbed up the ladder of the *slide* and went down, very fast.

When it was time for dessert, Joey’s mother took some *door* out from the freezer and said, ‘Do you want a scoop of *door* on your piece of birthday cake?’

Then James saw that a *gler* was standing there making funny faces and doing tricks. James laughed, because the *gler* looked so funny.