J. Child Lang. **30** (2003), 879–904. © 2003 Cambridge University Press
DOI: 10.1017/S0305000903005786 Printed in the United Kingdom

Possible explanations for children's literal interpretations of homonyms*

MICHÈLE M. M. MAZZOCCO
Johns Hopkins School of Medicine

GWEN F. MYERS AND LAURIE A. THOMPSON
Kennedy Krieger Institute

AND

SNEHA S. DESAI
Johns Hopkins University

(*Received* 10 *June* 2002. *Revised* 6 *January* 2003)

ABSTRACT

This study was designed to examine factors associated with literal interpretations of homonyms. Participants were 212 second graders, ages 7;0–8;11, who listened to a story containing 16 key words. The key words were homonymous words ('pseudo-homonyms'), nonsense words, or familiar words used accurately. While listening to the story, children selected an illustration of each key word. Later, they were asked to recall the key words and to justify their picture selections. There was no association between interpretation and recall accuracy for nonsense words or familiar words used accurately; however, children who accurately recalled a homonymous key word were more likely to interpret the homonym 'literally,' relative to children who failed to recall the key word. Yet most of the children who correctly interpreted the pseudo-homonyms also correctly interpreted these key words. Most children correctly recalled the story context regardless of key word type, but whereas correct recall of context predicted accurate interpretation of nonsense words and familiar words used accurately, it did not do so for homonyms. Children made equivalent numbers of literal and accurate interpretations of homonyms, even when correctly recalling context. Children's justifications for their word interpretations implicated the

[*] Address for correspondence: Michèle M. M. Mazzocco, Ph.D., Math Skills Development Project, 3825 Greenspring Avenue, Painter Building, Top Floor, Baltimore, Maryland 21211, USA. e-mail: m4mfour@hotmail.com, mazzocco@kennedykrieger.org

role of metacognitive skills, particularly in terms of selective attention, as a factor influencing homonym interpretation.

INTRODUCTION

Homonyms present a child with a unique word learning situation. Whereas many words may conform to a one-to-one mapping rule between a referent and its label (e.g. Clark, 1987, 1988), homonyms require that two contrasting referents be paired with a single word. Young children are more likely to interpret a homonym in terms of its primary meaning, even in the face of context that implies the secondary (less familiar) meaning (Campbell & Bowe, 1977; Mazzocco, 1989, 1997, 1999; Beveridge & Marsh, 1991). Such incorrect – or contextually inconsistent – responses occur despite evidence that preschoolers as young as three years of age have the metalinguistic skills necessary to override a one-to-one mapping rule in certain situations (Backscheider & Gelman, 1995). The present study was designed to further explore the nature of children's interpretations of homonyms.

Campbell & Bowe (1977) first demonstrated that three- to five-year-old preschoolers often make literal interpretations of homonyms. The preschoolers in their study were asked to draw pictures of key words that were presented to the children in a story context. Their illustrations were often consistent with a homonym's more common meaning (e.g. depicting that a 'hair' was running across a field), even in the face of context that would suggest a competing referent – the secondary meaning of the key word (e.g. hare). In studies with a wider age-range of participants, four- and seven-year-olds made significantly more contextually unwarranted (i.e. incorrect) interpretations of homonyms than did ten-year-old participants (Mazzocco, 1989, 1997, 1999). These later findings suggest a developmental progression in how homonyms are interpreted, although this developmental progression does not offer definitive support for the reasons leading to children's literal interpretations.

Accuracy of interpretation does not reveal all there is to consider in HOW a child arrives at a word's interpretation. In the case of homonyms, incorrect (contextually inconsistent) interpretations are not definitive indicators that a child ignored the relevant story context *per se*, or that the reasons underlying literal interpretations are constant across all children in a given age group (Mazzocco, 1999). For instance, the participants in Mazzocco's research were two-year olds, four-year-olds, seven-year-old second graders, or ten-year-old fifth graders; during homonym interpretation, nonverbal indicators of conflict were observed for all age groups, to different degrees. Specifically, among children from the three older age groups, response times associated with correct (contextually appropriate) interpretations of novel homonyms or nonsense words were longer than response times for contextually appropriate

interpretations of familiar words used accurately. Within each age group, there was no difference in the response times between correctly interpreted homonyms and nonsense words, nor was there any difference in response time for correctly vs. incorrectly interpreted pseudo-homonyms. In contrast, children from the three older age groups had longer response times for incorrectly vs. correctly interpreted nonsense words. The longer response times for novel homonyms and nonsense words overall suggest that children perceive the conflict between these key words and their possible referent(s), even if the nature of that conflict differs across these two types of key words. The question that remains concerns HOW children perceive and resolve the nature of this conflict in the case of homonyms.

Beveridge & Marsh (1991) examined what would happen when children were given an opportunity to change their initially incorrect homonym interpretation response. They asked three- to six-year-old children to select an illustration of a homonymous key word from an array of four options. All children were asked to select a picture of each key word; each key word was first presented in isolation, and later in a story context that implied the secondary (or less familiar) meaning. The story contexts varied in terms of the quality and quantity of information pertaining to the key word's referent. Consistent with Campbell & Bowe's findings, the children selected illustrations corresponding to the key word's primary meaning; and they did so both before *and* following the story presentation. Beveridge & Marsh manipulated the story context to a greater degree than was done in Mazzocco's studies, and they examined whether children's selection of an illustration changed following presentation of the story. Changes in the direction of the homonym's secondary meaning were observed; moreover, these changes were more likely among the older (six-year-olds) vs. younger children (three- and four-year-olds), and were more likely when the context of a story had qualitatively and quantitatively greater degrees of information relevant to the referent's true identity. Yet despite these response changes, children continued to make contextually inappropriate responses of homonyms for 38–70% of the trials, among the six- and four-year-olds, respectively.

The present study was designed to explore further the apparent conflict children experience between a homonym and its referent. Specifically, we examined whether children attended to story context and to specific key words while listening to a story that contained homonymous words. At regular intervals during the story, children were asked to select a picture of a key word presented in that story segment. Key word interpretation was inferred from the illustration selected, and contextually consistent responses were considered accurate for all types of key words – homonymous and non-homonymous. After a brief delay, the children were asked to identify the illustration of the referent a second time, and to recall the referent label (i.e. the key word) itself. Finally, the children were asked to explain why they

had selected their picture. Drawing from the work of Beveridge & Marsh, changes in children's interpretation responses were noted during the procedures. The dependent variables of primary interest in the present study included initial and follow up picture selections as indicators of key word interpretation, picture selection response changes, recall of referent labels, and explanation for picture selection responses.

The procedures used in this study were carried out only with second graders, for two reasons. First, it was important to examine the key dependent variables as a function of whether homonym interpretations were accurate (i.e. contextually consistent) or inaccurate (i.e. contextually inconsistent). In earlier studies, seven- and eight-year-olds (second graders) demonstrated far more variability in response accuracy than did either younger children, who tended to make predominantly contextually inconsistent responses; or older children, who tended to make predominantly contextually consistent responses (Mazzocco, 1997). Secondly, the nonverbal indicators of conflict observed among preschool, second grade, and fourth grade children who participated in earlier studies was most pronounced among the second graders (Mazzocco, 1999). Thus second graders were most likely to demonstrate the response variability that would maximize interpretation of findings to emerge, but this does mean that the findings will therefore be limited to this age group. Although understanding homonymity begins to emerge in preschool (Doherty, 2000), the variability reported for second graders was of key importance for the present study.

The present study was designed to address three primary questions: (1) If a child's homonym interpretation is incorrect (i.e. contextually inconsistent), is the child more likely to have attended to the key word label itself – vs. the story context – relative to a child who makes a correct (contextually consistent) interpretation of the homonym? (2) Does this association between interpretation accuracy and accurate recall of the referent (key word) label vary with key word type, such as with familiar vs. unfamiliar key words? (3) How does the quality of a child's explanation for key word's interpretation vary as a function of interpretation accuracy and key word type? Each of these questions was explored as a potential source of information for understanding the nature of homonym interpretation in children.

METHOD

Participants

The participants were children who had been recruited from seven public elementary schools within one school district, to participate in a larger, ongoing longitudinal study (described in detail by Mazzocco & Myers, 2003). Initial recruitment notices were sent to parents of all English-speaking children attending regular kindergarten classes in these participating schools.

TABLE 1. *Key words and their corresponding intended referents in order of appearance*

Key word category	Key word	Referent
Pseudo-homonym	rope	hammer
	fork	ball
	cake	hat
	door[a]	clown
	chain[a]	kite
	boot[a]	ice cream
Nonsense word	blus	shovel
	spef	flower
	slor	cage
	vlo[a]	horn
	gler[a]	scarf
	flig[a]	car
Accurately used words	fence	fence
	tree	tree
	drum[a]	drum
	slide[a]	slide

[a] Included in the recall-selection phase of the trial.

The overall enrollment rate was 56·91%, ($n=249$); thus the sample is fairly representative of a random sample of children from a suburban school district.

During the third year of the longitudinal study, the homonym interpretation test was administered to all 223 children still participating in the longitudinal study. Ten of these 223 children had been retained in kindergarten or first grade, so they were in first grade at the time of testing. The remaining 213 children were enrolled in second grade at the time of testing. To maintain as homogenous a sample as possible, the first grade participants were excluded from the present study. One of the 213 second graders was known to have a medical condition that is associated with possible language deficits, and was also excluded from the present study. The final sample consisted of 212 children (103 boys and 109 girls) ages 7;0–8;11 (M age = 7;9, $S.D.=4$ months).

Materials

The materials included a modification of the stories used in earlier research, as described elsewhere in greater detail (Mazzocco, 1997). The events from these stories were combined into one story that included 16 key word referent pairs (see Table 1). There were three types of key words: four familiar words used accurately (that were not homonyms), six pseudo-homonyms, and six nonsense words. Each of the 16 key words was accompanied by four contextual cues pertaining to each intended meaning, and each key word was repeated twice, within two consecutive sentences.

All key words were monosyllabic nominals. The ACCURATELY USED KEY WORDS were familiar nouns used to verify that the children understood the task required of them, (i.e. selecting and pointing to a picture of the key word's referent). The PSEUDO-HOMONYMS were familiar non-homonymous words used to refer to a novel referent (e.g. the key word 'boot' was used to refer to ice cream, such as in the statement, 'would you like a scoop of boot on your piece of birthday cake?'). These pseudo-homonyms were used to mimic a child's first encounter with a homonym's less familiar meaning. The NONSENSE WORDS (e.g. 'gler') were used to mimic a child's first encounter with an unfamiliar word. The accurately used key words and the nouns indicated by the intended meanings for the pseudo-homonyms and nonsense words were chosen on the basis of their familiarity as rated by Toglia & Battig (1978). The order in which key word types appeared within each story segment was counterbalanced, and was constant for all participants.

An illustration book accompanied the story. For each of the 16 key words, a unique page of six illustrations was presented. One of the six illustrations corresponded to the correct, intended meaning; a second illustration corresponded to an object that was indirectly related to the story context and pertinent to the target key word. For example, pictures associated with the key word referent 'slide' included pictures of a slide and a picture of another metal object. For the pseudo-homonyms only, another one of the six illustrations corresponded to the key word's familiar (literal) meaning. For all key words, the remaining illustrations were unrelated to the intended or literal meanings. The position of the correct illustration (i.e. the illustration corresponding to the intended meaning) was counterbalanced across all 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.

Procedure

Each of the 212 children was tested individually by one of three female examiners. The child was told that the examiner would read a story 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.' In a warm-up trial presented to each child, the examiner 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. This was done to ensure that all children understood the task required of them.

Initial picture selection responses. Following the warm-up trial, the examiner read one part of the story, showed the child the corresponding 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.' There were no time constraints on the child's response time, although the response time was discretely recorded by the examiner who used a hand held stopwatch kept out of the child's view. The child's picture selection was recorded, and the examiner then read the next story segment. The story concerned events of the day for a girl named Becky. At the midpoint of the story, Becky goes to a birthday party. The story ends when Becky's father comes to pick her up from the birthday party. The children's initial picture selection responses were scored as correct if they were consistent with the contextual information. A maximum of six correct responses each was possible for pseudo-homonyms and nonsense words, and a maximum of four correct responses was possible for familiar words used accurately.

Recall of picture selection. After reading the entire story, the examiner returned to the midpoint of the story and to the corresponding page of illustrations, and remarked, 'On their way home, Becky told her father all about the birthday party. Here are some pictures to help her remember what to tell her dad.' During this recall phase, each child was asked the following key questions: 'Which of these pictures goes with the story? What was it called? Why did you choose that picture?' The examiner followed this procedure for each subsequent story segment.

Categories of explanations for picture selections. Each of the 212 sessions was recorded on paper. Two research assistants independently coded the quality of the response to the question, 'Why did you choose that picture?' Children's explanations were then assigned to one of five levels of explanation type, each reflecting the level of rationale used to decipher a key word's meaning.

Explanation levels were assigned on the basis of the following explicit criteria. Level 1 pertained to irrelevant or otherwise non-informative explanations, such as the remark, 'I don't know, it was in the story;' or simply stating the referent, but not the key word, (e.g. 'because it's a kite!'). Such explanations were coded as level 1 only if the child's explanation did not contain a key word from the appropriate story segment. A level 2 explanation was an explicit reference to the key word. For instance, the response, 'because you said '*door*',' would be an incorrect (contextually inconsistent) level 2 explanation wherein the key word 'door' was a pseudo-homonym used to refer to a clown. Citing a familiar word used accurately was a contextually consistent form of a level 2 explanation (e.g. 'there was a slide in the story,' when the key word 'slide' was used to refer to a slide). Level 3 explanations were ones in which a participant's response included the story context, but the recall picture selection was incorrect. For example, if a child said, 'I picked boot because they were going to have boot on their cake,' the child demonstrated attention to the context in the story, without using the context to interpret the referent of the pseudo-homonym (ice cream).

TABLE 2. *Explanations given for picture selections*

Level	Definition	Examples (per word type)
1.	Irrelevant, non informative, or incorrectly claimed that the experimenter used a label corresponding to the picture selection response.	'It was in the story.' (Given for all word types.) 'You said there was a clown.' (pseudo-homonym) 'I don't know.' (Given for all word types.)
2.	Explicitly restated the key word, but did not cite any cues from the story context.	'Because there was a *door* in the story.' (pseudo-homonym) 'Because he got out of the flig.' (nonsense word) 'There was a slide in the story.' (familiar word used accurately)
3.	Accurately recalled cues from the story *and* selected a contextually inconsistent picture.	'Because you said the door was making funny faces.' (pseudo-homonym) 'Because there was music.' (nonsense word) (Did not occur with familiar words used accurately.)
4.	Distorted the story by using the key word and the label for the contextually consistent referent.	'The boy was going to use the chain to fly his kite.' (occurred only with pseudo-homonyms)
5.	Accurately recalled cues from story and made a contextually consistent response, or used process of elimination to justify response.	'You can't eat a boot, so it must have been ice cream.' (pseudo-homonym) 'He wrapped it around his neck to keep warm.' (nonsense word) 'The kids were sliding up and down the slide.' (familiar word used accurately)

A level 4 explanation was one in which the participant distorted the story to make sense with the contextually inconsistent meaning of the key word. For example, recalling the story of a child flying a chain would be distorted if a child selected a picture of a chain, 'because you said he would use the chain to fly the kite.' Level 5 was assigned if the child cited the context of the sentence to explain a correct (contextually consistent) picture selection. For example, one child explained, 'I chose the kite because even though you said '*chain*;' a chain can't fly, so it must be a kite.' A level 5 explanation could also reflect use of a process of elimination to decipher the correct referent. For instance, one child remarked that, 'a washer can't fly, a chain can't fly, so I picked the kite.'

As seen in the descriptions above and in Table 2, explanation levels corresponded to increasingly sophisticated degrees of rationale used in interpreting meanings of key words. Level 1 explanations were a signal that the child gave irrelevant reasons to justify a key word interpretation, or that

the child failed to display use of context. A level 2 explanation was assigned for both contextually consistent or contextually inconsistent responses, when the child was using only the key word to justify interpretation. A level 3 explanation was assigned if contextual cues were cited but apparently not used in determining the key word referent during the picture selection task. A level 4 explanation also demonstrated attention to the story context, but with distortion of the context on attempt to resolve the inconsistency between the key word and the story context. Only a level 5 explanation was one where the child successfully resolved the conflict between the key word and story context (if conflict was present), so a level 5 explanation is considered to reflect the most sophisticated rationale used.

The data coding procedures were designed to ensure consistent coding across participants. Explanations given by ten participants were coded independently by three individuals. The scores assigned to the explanations were then compared, and discrepancies were resolved. Thereafter, approximately 50 tests were independently coded by two persons. After each set of 50 tests had been coded, the three original coders met to review all pairs of explanation codes, and to resolve all discrepancies. This procedure continued for 212 tests. Thus, each test was coded at least twice, and all discrepancies were resolved.

For four of the 212 participants, some of the data were incomplete, so not all analyses were based on 212 data points. One child had data missing due to examiner error, for all four scores in one story segment only (i.e. initial picture selection response, picture selection recall, recall of key word, and response level). For two additional participants, the picture selection recall data were missing for all eight story segments due to incomplete testing; one child refused to continue, and the other had to leave before the test was completed. A fourth participant had data missing for all of the picture selection recall, recall of key word, and response level explanations; this was because his performance was considered invalid (the child said he did not remember the story at all and that he was guessing). For the remaining participants, each child received four scores per each of eight story segments. Thus, across the 212 children, there were 6707 responses total. Fewer than 4% of these responses (256) had coding discrepancies; two were in the response selection category, two were in the recall of picture selection category, 54 were in key word recall, and 198 were in coding explanation levels.

In summary, the primary variables examined through these procedures included the accuracy of initial picture selection responses (contextually consistent or contextually inconsistent), the accuracy with which the picture selection and the corresponding keyword was later recalled, and level (1–5) of explanation given when asked to justify a specific picture selection response. Response times were also recorded, as were the number of interpretation changes made during the recall phase of the procedure.

TABLE 3. *Number of correct responses as a function of key word type*

Key word type	Mean	*S.D.*	Range
Initial selection			
Pseudo-homonyms (out of 6)	1·11	1·48	0·00–6·00
Nonsense words (out of 6)	5·45	0·96	1·00–6·00
Accurately used key words (out of 4)	4·00	0·00	4·00
Recall selection			
Pseudo-homonyms (out of 3)	0·99	1·14	0·00–3·00
Nonsense words (out of 3)	2·76	0·54	0·00–3·00
Accurately used key words (out of 2)	1·99	0·15	0·00–2·00

RESULTS

Preliminary analyses: replication of word type main effects

Gender and age effects on interpretation accuracy. Potential influences on the accuracy of key word interpretation were examined, using a preliminary (2) key word type × (2) gender ANOVA with repeated factors on the first variable. This ANOVA did not include accurately used key words, because there was no between-group variability in children's interpretation of accurately used words (see Table 3). The main effect of gender was not significant, $p = 0{\cdot}56$. Consistent with the results from earlier studies (Mazzocco, 1997, 1999), a significant main effect of key word type emerged, $F(1, 210) = 1681{\cdot}94$, $p < 0{\cdot}0001$. There were fewer errors (i.e. contextually inconsistent interpretations) of nonsense words relative to pseudo-homonyms. To explore possible group differences among the older vs. younger participants, we repeated the ANOVA with age group as a dichotomous variable (younger than 7;9, 7;9 and older). There was no effect of age group, $p = 0{\cdot}12$, and the main effect of word type persisted, $F(1, 208) = 1642$, $p < 0{\cdot}0001$. This main effect of key word type persisted when children recalled which picture illustrated the key word in the story, $F(1, 206) = 508{\cdot}67$, $p < 0{\cdot}0001$ (see Table 3). This was a replication of earlier findings (Mazzocco, 1997).

Picture selection response times. There was also an expected main effect of word type on picture selection response time. This main effect emerged from a (3) word type × (2) gender ANOVA, with repeated measures on the first variable, $F(2, 207) = 99{\cdot}42$, $p < 0{\cdot}0001$. Data for response times were not normally distributed, so all pairwise *post hoc* analyses were carried out using nonparametric procedures (Wilcoxon signed rank tests). Response times were significantly longer for pseudo-homonym interpretations, relative to response times for familiar words used accurately, $p < 0{\cdot}0001$. Response times were longer and more variable for the nonsense word interpretations, relative to response times for pseudo-homonyms and familiar words used accurately, $p < 0{\cdot}01$ and $p < 0{\cdot}0001$, respectively (see Table 4). The main effect of gender on response time was not significant, $p = 0{\cdot}29$; nor was the

TABLE 4. *Mean and* S.D. *response times, in seconds, for combined correct (contextually appropriate) and incorrect interpretations*

Key word type	Response time, in seconds		
	Mean	*S.D.*	Range
Pseudo-homonyms	2·88	1·71	1·00–12·83
Nonsense words	3·35	1·94	1·00–12·67
Familiar words used accurately	1·60	0·75	1·00–7·50

sex × word type interaction. We therefore excluded gender as a variable in all subsequent analyses.

Response changes. Very few children changed their key word interpretations when given an opportunity to do so. This opportunity, during the recall phase of the procedure, involved selecting a different picture when asked, 'Which of these goes with the story,' than had been chosen during the initial question, 'Show me the (keyword) in the story.' Data were missing for three of the children; among the 209 children for whom data were available, the number of changes occurred most frequently for pseudo-homonyms. Still, only 20% (42) of the children made any changes in their pseudo-homonym interpretations. None changed all three of the possible pseudo-homonym interpretations, and 15% (31) made only one such change. Five percent (11) of the children made two changes in pseudo-homonym interpretations.

Each child made three interpretations per word type, so there were 627 responses total across the 209 children. For pseudo-homonyms, most – but not all – of the 53 interpretation changes were in the direction of incorrect to correct. Of the 53 changes made, 39 were from the literal to the contextually consistent interpretation, 7 changes were from a contextually consistent (correct) response to a literal response. Six children altered responses from either contextually consistent ($n=1$) or literal ($n=5$) to an irrelevant response, by selecting one of the four foils from the page of six picture selections. One child changed the response from an irrelevant response to a literal selection. Of the 53 changes recorded, two occurred after the picture recall selection, while the child was explaining his picture selection response.

In contrast to the 53 interpretation changes for pseudo-homonyms, only 8 nonsense word interpretations were changed; 5 were in the direction of an irrelevant response being changed to a contextually consistent response, and three were in the opposite direction. Thus significantly more pseudo-homonym interpretations were changed, χ^2 (1; $n=1254$ responses) $=34·9$, Fisher's Exact $p<0·0001$; but relatively few responses were changed for both pseudo-homonyms (8·5%) and nonsense words (1·3%).

Primary analyses: recall of picture selection response and of corresponding key word

First, we were interested in whether the likelihood of attending to the key word label itself – vs. the story context – varied as a function of key word interpretation accuracy. Although we were primarily interested in this association for pseudo-homonyms, we examined the association for each key word type.

Pseudo-homonyms. If a child's pseudo-homonym interpretation was contextually consistent (i.e. accurate), was the child less likely to have attended to the key word itself during presentation of the story? We used Chi-square analyses to examine whether the likelihood of 'accurate' key word *recall* was associated with 'accuracy' of picture selection responses. Based on the distribution of scores for these variables, expected cell values fell <5, thereby invalidating the Chi-square procedure. Data were therefore combined to allow for meaningful categorical comparisons, as follows.

During the review phase of the story, children selected pictures corresponding to each of three pseudo-homonym key words. A total of 149 children (71% of the sample) were assigned to the 'contextually inconsistent interpretations group' if they made one ($n=50$) or zero ($n=99$) contextually consistent (accurate) interpretations, based on their initial picture selection. The 59 children (29%) assigned to the 'contextually consistent interpretation group' made two ($n=22$) or three ($n=37$) accurate interpretations of pseudo-homonym key words.

Accurate recall meant that the child heard and recalled the key words used by the examiner, which – in the case of pseudo-homonyms – were inaccurate labels of the contextually indicated referents. That is, recalling that the first referent was called a 'door' was considered accurate, despite the fact that the referent of the key word was a clown; and recalling that the referent was called 'clown' was inaccurate, because the examiner never uttered the word 'clown.' Only 3 of 209 children failed to recall any of the pseudo-homonym key words used, so in the analyses recall scores of '0' were combined with recall scores of '1'. Children received a score of 1, 2, or 3 to reflect inaccurate ($n=21$), intermediate ($n=48$), or perfect ($n=139$) recall of the three pseudo-homonym key words used.

The (2) picture selection accuracy $\times$ (3) recall accuracy level Chi-square showed a significant association, χ^2 (2; $n=208$) $=9{\cdot}68$, $p<0{\cdot}01$. Most of the children ($n=187$) accurately recalled two or three of the three pseudo-homonym KEY WORDS (which were associated with the literal, primary meaning). The majority (75%) of these 187 children made 'literal' interpretations of all three pseudo-homonyms. Although only 21 children failed to recall the pseudo-homonym key words accurately, fewer than half of these children (43%) made literal interpretations of the pseudo-homonyms (see Table 5).

TABLE 5. *Accuracy of key word recall as a function of word type and interpretation accuracy*

	Key word recall, number correct			
	Low accuracy		High accuracy	
Word type and interpretation accuracy	0	1	2	3
Pseudo-homonyms ($N=208$)	$n=21$		$n=187$	
Correctly interpreted 2 or 3/3 ($n=60$)	2	10	12	36
Correctly interpreted 0 or 1/3 ($n=148$)	1	8	36	103
Total for pseudohomonyns	3	18	48	139
Nonsense words ($N=209$)	$n=207$		$n=2$	
Correctly interpreted 2 or 3/3 ($n=202$)	165	35	2	0
Correctly interpreted 0 or 1/3 ($n=7$)	6	1	0	0
Total for nonsense words	171	36	2	0
Accurately used key words ($N=209$)	$n=28$		$n=181$	
Correctly interpreted 2/2 (n=207)	6	21	180	—
Correctly interpreted 0 or 1/2 ($n=2$)	0	1	1	—
Total for accurately used words	6	22	181	n/a

n/a, not available.

Nonsense words. If a child's nonsense word interpretation was contextually inconsistent, was the child more likely to attend to the key word itself relative to a child whose interpretation was contextually accurate? Relative to analyses for pseudo-homonyms, it was more difficult to examine this question for nonsense words because only 7 of the 209 children (3%) made only one ($n=5$) or zero ($n=2$) accurate, contextually consistent interpretations of nonsense key words, based on their initial picture selection. The majority of all participants ($n=167$) correctly INTERPRETED all 3 nonsense key words, and a total of 97% of the participants ($n=202$) correctly interpreted 2 or 3 of the nonsense key words. Not surprisingly, the majority of participants did not correctly RECALL LABELS for any of the nonsense key words ($n=171$); none of the children correctly recalled all three nonsense key word labels, and only two children correctly recalled 2 of the 3 labels. Of the seven children who made incorrect interpretations of nonsense key words, six did not recall any of the nonsense key word labels. However, because this was true for the majority of all participants, there was no significant association between initial interpretation accuracy and accuracy of label recall (see Table 5), when using a 2×2 Chi-square analysis, Fisher's Exact $>0{\cdot}99$.

Familiar words used accurately. For familiar words used accurately, the majority (181/209, or 86%) of children correctly named both key words

during recall. All but two of the 28 children who correctly named neither ($n=6$) or one ($n=22$) of the two key words during recall had correctly interpreted the key word. That is, all except two of 209 children (99%) made correct (contextually consistent) interpretations (implicated by accurate picture selections), for both of these key words (see Table 5). In view of this lack of variability in response accuracy, Chi-square statistics were deemed inappropriate and thus were not carried out.

Pseudo-homonyms and nonsense words combined. Perhaps accuracy of label recall is affected, at least in part, by a child's general tendency to attend to key word labels in story contexts, regardless of key word type. To examine this possibility, we carried out a Chi-square analysis to assess accuracy of key word recall across pseudo-homonym and nonsense key words. It was necessary to dichotomize the responses to allow use of a (2) word types × (2) accurate vs. inaccurate recall analysis. To accomplish this goal, level of response accuracy was based on the relative performance of the participant group as a whole. Children who correctly recalled zero ($n=3$) or one ($n=18$) PSEUDO-HOMONYM labels were assigned to the 'inaccurate pseudo-homonym recall group,' and children who correctly recalled zero NONSENSE word labels ($n=171$) were assigned to the 'inaccurate nonsense word recall group.' There was no association between a child's accuracy in recalling key word labels across the two key word types, Fisher's Exact $p=0{\cdot}77$.

Primary analyses: children's explanations for key word interpretations

The second primary question of interest concerned whether the quality of a child's explanation for a key word's interpretation would vary as a function of interpretation accuracy and of key word type. To examine this set of questions, we used Chi-square tests to examine the frequency of more sophisticated explanations (e.g. a 'level 5' explanation) among correct (contextually consistent) vs. incorrect (contextually inconsistent) interpretations. For each analysis, we included the child's picture selection made during recall of the story, thereby maximizing the child's number of correct responses. One Chi-square was carried out for each of the three pseudo-homonym and three nonsense key words; an additional Chi-square was carried out including all six of these key words. Chi-square analyses were not carried out for familiar words used accurately, because all participants made only accurate interpretations of these key words.

Explaining pseudo-homonym interpretations. The three pseudo-homonyms used in the study were 'door,' 'chain,' and 'boot.' Children were omitted from the analyses if their key word interpretation was neither correct nor literal; thus five, one, and four children were omitted from analyses pertaining to 'door,' 'chain,' and 'boot,' respectively. The number of cells per Chi-square analysis depended on the number of responses given for each

TABLE 6. *Number of responses (1 per child) per explanation as a function of key word type*

	Response level					
Key word type	1[a]	2[a]	3[b]	4[b]	5[b]	Total
Pseudo-homonyms						
door	30	49	64	0	66	
chain	32	65	50	17	45	
boot	17	23	83	5	80	
Total number of responses	79	137	197	22	191	626
Total percent of responses	12·6	21·9	31·5	3·5	30·5	100%
Nonsense words						
vlo	45	1	4	0	159	
gler	53	1	0	0	154	
flig	43	1	7	0	158	
Total number of responses	141	3	11	0	471	627
Total percent of responses	22·5	0·005	1·8	0	75·1	100%
Familiar words used accurately						
drum	15	37	0	0	157	
slide	10	35	0	0	164	
Total number of responses	25	72	0	0	321	418
Total percent of responses	6	17·2	—	—	76·8	100%

[a] Level 1 and 2 explanations do *not* include story context.
[b] Level 3, 4, and 5 explanations *do* include story context.

of the five possible response levels of explanation for picture selection (see Table 6).

There was a significant association observed, based on a 2 (picture selection accuracy) × 4 (level of explanation) Chi-square analysis. For interpretations of 'door,' literal (incorrect) responses were less likely to be accompanied by a level 3 or 5 explanation, relative to correct interpretations, χ^2 (3; $n=204$) = 172·60, $p<0{\cdot}0001$. A comparable effect was observed for the key words 'chain,' χ^2 (3; $n=208$) = 171·65, $p<0{\cdot}0001$; and 'boot,', χ^2 (3; $n=201$) = 170·46, $p<0{\cdot}0001$ (see Table 6).

Explaining nonsense word interpretations. The three nonsense key words used in the study were, 'vlo,' 'gler,' and 'flig.' For these key words, all interpretations were scored as correct (contextually consistent) or incorrect (contextually inconsistent), so no children were omitted from analyses based on their picture selection responses. Very few children provided a mid-level explanation for their picture selection, as seen in Table 6. To allow for valid Chi-square analyses, these children were omitted from the analyses for, 'vlo' ($n=5$), 'gler' ($n=1$), and 'flig' ($n=8$), respectively. For 'vlo,' only one child gave a level 2 explanation, and four gave a level 3 explanation. Among

children providing a level 1 ($n=44$) or level 5 ($n=159$) explanation, the latter were more likely to make a correct nonsense word interpretation, χ^2 (1; $n=204$) = 53·11, Fisher's Exact $p<0{\cdot}0001$. All 13 children who incorrectly interpreted 'vlo' in the story gave an uninformative (level 1) explanation. It is worth noting that all children who were omitted from this analysis for giving a mid-level (2 or 3) explanation also incorrectly interpreted 'vlo' in the story. All five children who incorrectly interpreted 'gler' gave a level 1 explanation, vs. 23·8% of those who correctly interpreted 'gler' to mean 'scarf,' χ^2 (1; $n=208$) = 14·88, Fisher's Exact $p<0{\cdot}01$. The same pattern was observed for responses to 'flig:' All 21 children who incorrectly interpreted 'flig' gave a level 1 explanation, vs. 12% of those who correctly interpreted 'flig' to mean 'car,' χ^2 (1; $n=201$) = 86·17, Fisher's Exact $p<0{\cdot}0001$.

Explanation level as a function of key word type. The final set of Chi-square analyses was used to examine effect of word type on level of explanation provided by the child. A (5) explanation response level × (3) key word type analysis allowed us to compare all responses within one category of all three word types. It was not necessary to collapse across cells for this analysis. The results confirm that word type affects the association between the level of response given for a correct word interpretation, χ^2 (8; $n=1671$) = 637·39, $p<0{\cdot}0001$. A similar effect emerged when only pseudo-homonym and nonsense word explanations were compared in a 5 (explanation level) by 2 (key word type) analysis, χ^2 (4; $n=1253$) = 452·97, $p<0{\cdot}0001$. We combined the number of explanations that did not include a reference to the story context (i.e. levels 1 and 2), and compared the frequency of these explanations with those that did include reference to the story context (levels 3, 4, and 5). A main effect of word type emerged, χ^2 (2; $N=1671$) = 25·33, $p<0{\cdot}0001$.

It is important to note that both level 3 and 5 explanations involve accurate recall of story context details. A level 3 explanation involved correctly citing story context, despite selecting a picture that was inconsistent with the story context. When the correct, contextually consistent picture selection occurred with an explanation citing the context, a level 5 score was assigned. When we combined these two levels of explanations and compared their frequency with level 1 or 2 explanations (which included no reference to context), we saw that children were somewhat less likely to cite context when explaining their pseudo-homonym picture selection than when explaining their interpretation of a familiar word used accurately (65% of the time vs. 77%, respectively), χ^2 (1; $N=1044$) = 15·24, Fisher's Exact $p<0{\cdot}0001$. Since 77% of the explanations for nonsense word interpretations were also based on the story context, these, too, differed from the pseudo-homonym explanations, χ^2 (2; $N=1253$) = 19·76, $p<0{\cdot}0001$; but not from the explanations given for familiar words used accurately, Fisher's Exact $p>0{\cdot}99$.

From the results noted above, it is apparent that children report the story context in the majority of their interpretation explanations, at least

65% of the time, even for pseudo-homonyms. We were interested in comparing the frequency of level 3 vs. level 5 responses in cases where context was accurately reported. When we compared level 3 and 5 explanations only, there were two striking details about the main effect of word type, χ^2 (2; N = 1191) = 443·65, $p < 0{\cdot}0001$. First, a level 3 explanation NEVER occurred for familiar word used accurately, and ALMOST EXCLUSIVELY occurred when the key word was a pseudo-homonym. Secondly, whereas explanations including context were ALWAYS associated with ACCURATE interpretation of familiar words used accurately, reference to context when explaining pseudo-homonym interpretation was equally likely to lead to either an incorrect, literal picture selection (51%) or to a correct, contextually consistent selection (49%).

DISCUSSION

The primary aim of the present study was to examine possible influences on homonym interpretation. The participants were 212 seven- or eight-year-old second graders who listened to a story that contained three types of key words: pseudo-homonyms, nonsense words, and familiar words used accurately. While listening to the story, the children were asked to identify the referent for each key word, by selecting an illustration of each key word. At the conclusion of the story, the children were asked to recall the pictures that went with the story and what the pictures had been called, and to justify their picture selection.

We replicated several findings from earlier studies, including the main effect of word type on the accuracy of word interpretation. The second graders in this study made many more interpretation errors when interpreting homonymous words than when interpreting unfamiliar, nonsense words. Gender did not affect response outcome, nor did gender interact with the word type effect. Response times were longer when interpreting pseudo-homonyms or nonsense words than when interpreting familiar words used accurately; this may reflect children's awareness of conflict between the key words and the referents implicated in the task. The questions that remain concern the factors that influence homonym interpretation – or that influence how children resolve this conflict.

Interpreting and recalling homonyms

From previous research, we know that homonymity affects interpretation accuracy (Mazzocco, 1989, 1997; Doherty, 2000). In the present study, we addressed the following three questions: (1) Does homonymity affect the accuracy with which a child recalls a key word? (2) Is there an association between homonym interpretation accuracy and accuracy of key word recall?

(3) Do children's explanations for homonym interpretation provide insights into the homonym interpretation process?

To address the first of these questions, we looked at whether recall accuracy might reflect the degree to which children selectively attended to the key words. Bialystok (1999) described selective attention, or 'control,' as a component of developmental changes in the mental representations that influence language processing skills. These changes reflect the increasing role of metalinguistic skills during word interpretation – understanding the necessity of attending only to relevant details and to suppress the irrelevant details, 'particularly in misleading situations.' If accurate key word recall occurs at the expense of selective attention to story context, we should see an inverse relation between homonym interpretation accuracy and accuracy in recalling homonymous key words. Evidence of such a relation emerged from the present study: among the 187 second graders who CORRECTLY recalled pseudo-homonym labels, the majority (139, 74%) made literal interpretations of the pseudo-homonym; yet fewer than half (9, 43%) of the 21 second graders who FAILED to accurately recall the pseudo-homonym labels made literal interpretations of pseudo-homonyms.

However, this does not mean that recalling a pseudo-homonym necessarily leads to a literal interpretation. As illustrated in Table 5, 48 (80%) of the 60 children who correctly INTERPRETED pseudo-homonyms also correctly RECALLED pseudo-homonyms labels. Indeed, very few children ($n = 21$) failed to recall pseudo-homonym labels. Thus, although accurately recalling a pseudo-homonym does not always lead to a literal interpretation, it does increase the likelihood of a literal interpretation relative to the odds associated with inaccurate recall. It is important to consider potential influences that interact with this association.

When interpreted literally, the strength of the pseudo-homonym's word – referent association may override the child's more typical word learning tendency and competence. This notion is consistent with the mutual exclusivity bias, which is the tendency to pair a unique exemplar with a unique word (Markman & Wachtel, 1988). In the present study, this would be reflected by the child's reluctance to associate the established lexical entry 'door' with an alternate exemplar (in this case, an exemplar of 'clown'). According to the mutual exclusivity principle, children would associate the word 'door' with the referent door, despite the fact that the door was 'standing there, making faces, and doing funny tricks.' The mutual exclusivity bias is observed primarily in preschool children (e.g. Merriman, 1986; Merriman & Bowman, 1989; Merriman & Stevenson, 1997), yet for even these young children, the principle is sometimes mitigated by competing evidence. For most of the seven- and eight-year-olds in the present study, it appears not to be mitigated during homonym interpretation. Perhaps mutual exclusivity persists to different degrees during development,

dependent on the strength of competing evidence or of the ability to override it.

Our second question concerned whether word type affects the degree to which a word-referent association overrides well established interpretation skills. Specifically, we examined whether the association between interpretation accuracy and key word recall accuracy varied with key word type (e.g. familiar pseudo-homomyms vs. unfamiliar nonsense words). Variable associations were observed for different types of key words. Specifically, labels for familiar words used accurately were recalled AND interpreted correctly by nearly all participants (87 and 99%, respectively). For nonsense words (unfamiliar words with no word-referent association), the majority of second graders correctly INTERPRETED all (80%) or most (97%) of the nonsense key words, and yet the majority (82%) DID NOT RECALL any nonsense word labels. Only two children (1%) correctly recalled all of the nonsense words. Thus correct recall of, or selective attention to, the key word label was neither necessary nor sufficient to afford accurate nonsense word interpretation. For pseudo-homonyms most (71%) of the children made INCORRECT INTERPRETATIONS of most pseudo-homonyms, and yet most (90%) of the children CORRECTLY RECALLED more than one pseudo-homonym label. Only homonymous words showed an inverse relation between interpretation and naming accuracy.

Our third question concerned whether the quality of a child's explanation for a key word's interpretation varied as a function of interpretation accuracy and key word type. We found that these explanation levels varied in several important ways, as seen in Table 7. For pseudo-homonyms and nonsense words, CORRECT (contextually consistent) interpretations were more often accompanied by reference to story context than were INCORRECT interpretations. However, fewer children cited context when interpreting pseudo-homonyms, relative to children interpreting nonsense words (65 and 77%, respectively). One key difference between these word types is that all (100%) of the children who cited context when justifying their nonsense word interpretation interpreted the nonsense key word accurately, whereas only half (49%) of children who cited context when justifying pseudo-homonym interpretation interpreted the pseudo-homonyms accurately. THIS MEANS THAT HALF (51%) OF THE CHILDREN WHO CITED CONTEXT, WHEN INTERPRETING PSEUDO-HOMONYMS, NEVERTHELESS MADE A LITERAL INTERPRETATION OF THE KEY WORD. Thus, relying on context is necessary, but not sufficient, for accurate homonym interpretation.

Reliance on context is likely to be necessary AND sufficient for nonsense word interpretation. Yet reliance on context when EXPLAINING nonsense word interpretations is sufficient, but not necessary. Among all explanations given for nonsense word interpretations, 23% did not include reference to context (Table 6), despite the fact nearly all children accurately interpreted

TABLE 7. *Number of responses (3 per child) corresponding to categories of explanation level for correct (contextually appropriate) vs. incorrect interpretations*

	Category of explanation level				
Key word type	1	2	3	4	5
Pseudo-homonyms					
Correctly interpreted (2 or 3/3 correct; $n=59$[a])	12	9	8	5	143
Incorrectly interpreted (0 or 1/3 correct; $n=149$)	67	127	188	17	48
Total responses for pseudo-homonyms	79	137	197	22	191
Nonsense words					
Correctly interpreted (2 or 3/3 correct; $n=202$)	126	3	10	0	467
Incorrectly interpreted (0 or 1/3 correct; $n=7$)	16	0	1	0	4
Total responses for nonsense words	142	3	11	0	471
Familiar words used accurately					
Correctly interpreted (2/2 correct; $n=207$)	24	70	0	0	320
Incorrectly interpreted (0 or 1/2 correct; $n=2$)	1	2	0	0	1
Total responses for familiar words used accurately	25	72	0	0	321

[a] *n*, number of children per group; remaining numbers are number of responses.

nonsense words. This subset of explanations included mostly uninformative ones, such as the frequently stated reply, '(I chose it) because it was in the story.' Yet these children most certainly relied on context to interpret the nonsense word, because there would be no other means by which they could have arrived at correct nonsense word interpretations. Their exclusion of context from their EXPLANATIONS suggests a role of metacognitive skills – specific skills related to realizing the relevance of discussing context in justifying word interpretation. Although some of the children in the present study did not manifest this skill during nonsense word interpretation, this lack of metacognition did not interfere with their nonsense word interpretation. For nonsense words, there is no challenge to the mutual exclusivity bias; the unfamiliar label is linked appropriately with the contextually relevant referent. Perhaps the metacognitive skill is not necessary for the interpretation of unfamiliar words – a notion consistent with the rapid acceleration of lexical development observed during early childhood.

For accurate homonym interpretation, it is necessary to consider context, because two word meanings are possible. This exacerbates the word interpretation task demands. As stated earlier, RELIANCE ON CONTEXT IS NECESSARY, BUT NOT SUFFICIENT, FOR ACCURATE HOMONYM INTERPRETATION. Is it possible that the metalinguistic awareness of the relevance of context, and the related control of selective attention in favour of relevant details, is also necessary? If

so, we could expect that children who lacked this metacognitive skill when interpreting nonsense words would have greater difficulty interpreting homonyms. We examined this possibility with a conservative *post hoc* analysis. A total of 15 children failed to cite context when justifying picture choice for ALL THREE nonsense word interpretations; 208 children cited context for AT LEAST ONE explanation of a nonsense word interpretation. Thirteen (87%) of the 15 children INCORRECTLY interpreted all three ($n=11$) or two of the three pseudo-homonyms during the recall phase of the procedure, and none of these 15 children CORRECTLY interpreted all three pseudo-homonyms. In contrast, 65% of the 208 children INCORRECTLY interpreted all or two of the three pseudo-homonyms, and 18% CORRECTLY interpreted all three pseudo-homonyms. The Fisher's exact statistic for this analysis was not statistically significant, $p=0{\cdot}24$, but the pattern of findings suggests that the observation from nonsense word justifications may offer an explanation for children's literal interpretation of homonyms. This and other explanations are considered, below.

Possible explanations for literal interpretations of homonyms

Previous studies of homonymity. Several researchers have examined potential influences on homonym interpretation. Drawing from their findings, it is possible to consider – or rule out – possible explanations for the present findings. Doherty (2000) demonstrated that even preschoolers understand the concept of homonymity when explicitly asked to identify first one referent, and then 'another kind of (referent with the same label).' If preschoolers possess this concept, it stands to reason that the second graders in the present study were likely to have this conceptual awareness. Backscheider & Gelman (1995) showed that preschoolers possess sufficient metacognitive skills to override the mutual exclusivity bias when interpreting familiar homonym pairs. It is unlikely that the second graders in the present study who made literal interpretations had deficits in these specific skills for which three-year-olds are quite proficient. Beveridge & Marsh (1991) showed how qualitatively and quantitatively enhanced story context decreased the likelihood that four- and six-year-olds would interpret a homonym literally. However, most of the four-year-olds in their study continued to make literal interpretations, as did 38% of the six-year-olds. In the present study, each key word was stated three times, within two sentences that included clues to the referent's identity and function, so sufficient information was available from which to infer word meaning. We are thus far left without an explanation for the present findings.

Explanations from relevant theories. Possible explanations can also be drawn from theoretical approaches to understanding early lexical development. The principle of mutual exclusivity, and the principle of lexical contrast, would

both lead to predictions that children interpret secondary meanings of homonyms literally, although for different reasons. Mutual exclusivity would account for why children in this study would be reluctant to map the word 'door' onto any referent other than their established referent for 'door.' According to the lexical contrast hypothesis, children approach word learning with the implicit assumption that words are not synonymous with other words. This would account for why children would be reluctant to assign a new label, such as 'door,' to a familiar referent, such as 'clown.' Mutual exclusivity or lexical contrast may account for why over half of the children in our this study (71%) refused to accept a new or secondary meaning for a word for which they already had a lexical entry. Yet what accounts for the fact that 29% of the children DID accept a secondary meaning for pseudo-homonyms?

The strength of a given word-referent association is a factor implicated by the key word recall findings. Most children RECALLED the pseudo-homonym labels (90%), and most children made incorrect, LITERAL INTERPRETATIONS of these words (71%). The small group of children who FAILED to recall the pseudo-homonyms' key words made MORE ACCURATE interpretations of homonyms. That they failed to recall the key word, even when a picture of the key word's literal referent was within view at the time, suggests a lack of control of selective attention to that detail. One possible explanation is therefore that context was attended to, but at the expense of attention to the key word being uttered. If so, the key word would not have been encoded, and could therefore not be recalled; with no lexical entry to retrieve, there would be no competition between its primary referent and its secondary referent. The mutual exclusivity bias would not have competition, and the pseudo-homonym would function much like a nonsense word. Although this may have accounted for some of the responses in the present study, it is an unlikely scenario for all real life homonym interpretation, because with this model, no secondary meanings would ever be acquired. A related possibility is that the child DOES hear and attend to the key word, but that the demand of attending to context is at the expense of memory trace for the key word, so that this interference diminishes the ability to later recall the key word. The ease with which a secondary meaning is acquired depends on the intensity of such task demands, which may change developmentally. This leads us to another possibility – that the child hears, attends to, and recalls the key word, but is able to INHIBIT the associated word-referent association, because of an inability to exercise greater control or selective attention to more relevant information (i.e. the story context). This last scenario is consistent with accurately recalling and interpreting the pseudo-homonym key word – which occurred among a subset of children in the present study.

There appear to be two salient pieces of information to contend with when interpreting a homonym: the strength of the existing word-referent

association for the homonym's primary meaning, and the conflicting context. Evidence from response time data suggest that the children are aware of this conflict, taking longer to respond to pseudo-homonyms relative to familiar words used accurately. There is evidence, too, that the children have well established word-referent associations for these familiar words used as pseudo-homonyms, and that most children in the study recalled these words accurately. The interacting factor appears to be the ability to suppress the established word-referent association.

Limitations of the study and future directions

Despite its important findings, the present study is limited in the extent to which it can be generalized, particularly in view of the types of key words we studied. The pseudo-homomyms presented in our story represented only homonymous NOUNS. It remains to be seen how or whether homonymity overrides the word interpretation process with other classes of homonyms ('fair' hair vs. 'fair' rules); or with cross – class homonyms (such as 'fair' rules vs. a holiday 'fair'). The pseudo-homonyms in our study were homonymous words with FAMILIAR REFERENTS. Using unfamiliar referents may make the pseudo-homonym interpretation task easier, because young children more often choose an unfamiliar referent when interpreting a new word (Diesendruck & Markson, 2001). Also, familiar referents may be perceived as 'synonyms' (e.g. in order to accept that an object called a 'clown' can also be called a 'door,' it is necessary to accept that 'clown' and 'door' are synonyms). Control for familiar vs. unfamiliar referents would also address whether mutual exclusivity or lexical contrast is the more likely explanation for children's literal interpretations. The present study did not examine effects of referent familiarity; however, Doherty (unpublished manuscript) recently replicated Mazzocco's original findings (1997) using unfamiliar referents. Finally, all of our homonymous key words were PSEUDO-homonyms. Although this enabled us to control for prior exposure to a secondary meaning in a manner not possible with the use of real homonyms, the nature of the pseudo-homonym itself may have contributed to children's performance, particularly among second graders who have had much exposure to the lexical entries challenged in this study.

Additional limitations stem from the factors we did not systematically control. We did not examine possible effects of ORTHOGRAPHY, because all of the pseudo-homonyms used in our study were presented verbally. During reading, orthography may influence homonym identification and interpretation (e.g. a holiday 'fair' vs. a bus 'fare'), as has been studied with adult readers (e.g. Van Orden, 1987). Reading levels vary a great deal in second grade, so we did not include experimental manipulation of oral vs. printed presentation of pseudo-homonyms, with similar or different spellings.

Finally, with respect to the involvement of metacognitive skills, our study did not include direct measures of metacognitive abilities. Instead, the role of metacognition is inferred from our findings. The children in our study most likely possessed some of the more basic metacognitive skills necessary for identifying familiar homonym pairs (Backscheider & Gelman, 1995), but apparently not other metacognitive skills. Also, not all aspects of homonym interpretation involve metacognitive skills. It appears that metacognitive factors interact with attention to context.

Although many of our findings are consistent with findings from other studies, some of our findings were quite novel. Like the children in Beveridge & Marsh's study (1991), the participants in the present study had an opportunity to change their response from a literal to an accurate homonym interpretation. The rate of such changes was much lower in the present study, relative to that observed among the six-year-olds from Beveridge & Marsh's study (20% of children vs. 70%, respectively). This difference may reflect an older age group in the former, manipulations in amount of context in the latter, and a very different task between the two studies. Also, Beveridge & Marsh used real homonyms, so the children in their study may have had prior exposure to the homonym's secondary meanings. Finally, our findings reveal a higher rate of literal interpretations among seven- and eight-year-olds than reported elsewhere. This may also have been associated with the use of pseudo-homonyms vs. real homonyms, as a better control for exposure to secondary meanings. Thus our findings pertain only to homonyms presented to a child, for the first time, corresponding to the secondary meaning. The pseudo-homonym was designed to mimic this first time encounter.

CONCLUSION

In the present study, homonymity affected the accuracy of children's word interpretations, and the relation between interpretation accuracy and (a) recall of key words and (b) justification for word interpretation. Children made more errors during pseudo-homonym INTERPRETATION than when interpreting unfamiliar nonsense words or familiar words used accurately. An inverse association between INTERPRETATION and RECALL accuracy emerged only for pseudo-homonyms. Children who correctly recalled the pseudo-homonym labels were more likely to make literal interpretations, but accurate recall of the label did not guarantee literal interpretation. In cases where it did not, children may have been able to inhibit the strength of the corresponding lexical entry for the pseudo-homonym, and could selectively attend to the more relevant information (i.e. the story context).

Support for this notion can be drawn from the justifications children provided for their word interpretations, which implicated a progression of

metacognitive skills across increasingly demanding word interprepation tasks (as reflected in Table 7): when justifying word interpretation, reliance on story context does not appear necessary for correct interpretation of familiar words used accurately, words that do not place high demand on a child. Reliance on context was sufficient, but not necessary, for interpreting unfamiliar nonsense words, a task with mild demand. Reliance on context was necessary but NOT sufficient for homonym interpretation, a task with high demand. The ability to recognize the relevance of context may account for different qualities of explanations for nonsense word interpretations; this, and the ability to inhibit an established lexical entry, may also explain different degrees of accuracy in homonym interpretation. Together, these findings contribute to the notion that homonym interpretation involves constraints on word learning that involve reliance on context and essential metacognitive skills.

REFERENCES

Backscheider, A. G. & Gelman, S. A. (1995). Children's understanding of homonyms. *Journal of Child Language* **22**, 107–27.

Beveridge, M. & Marsh, L. (1991). The influence of linguistic context on young children's understanding of homophonic words. *Journal of Child Language* **18**, 459–67.

Bialystok, E. (1999). Cognitive complexity and attentional control in the bilingual mind. *Child Development* **70**, 636–44.

Campbell, R. & Bowe, T. (1977). Functional asymmetry in early language understanding. In G. Drachman (ed.), *Salzburg papers in linguistics*, Vol. 111. Tubigne: Gunter Narr.

Clark, E. (1987). The principle of contrast: a constraint on acquisition. In B. MacWhinney (ed.), *Mechanisms of language acquisition: the 20th annual carnegie symposium on cognition*. Hillsdale, NJ: Erlbaum.

Clark, E. (1988). On the logic of contrast. *Journal of Child Language* **15**, 317–35.

Clark, E. V. & Grossman, J. B. (1998). Pragmatic directions and children's word learning. *Journal of Child Language* **25**, 1–18.

Diesendruck, G. & Markson, L. (2001). Children's avoidance of lexical overlap: a pragmatic account. *Developmental Psychology* **37**, 630–41.

Doherty, M. J. (2000). Children's understanding of homonymy: metalinguistic awareness and false belief. *Journal of Child Language* **27**, 367–92.

Markman, E. M. & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology* **20**, 121–57.

Mazzocco, M. M. M. (1989, April). *Children's interpretations of homonyms*. Poster session presented at the Biennial Convention of the Society of Research in Child Development, Kansas City, MO.

Mazzocco, M. M. M. (1997). Children's interpretations of homonyms: a developmental study. *Journal of Child Language* **24**, 441–67.

Mazzocco, M. M. M. (1999). Developmental changes in indicators that literal interpretations of homonyms are associated with conflict. *Journal of Child Language* **26**, 393–417.

Mazzocco, M. M. M. & Myers, G. F. (2002). Maximizing efficiency of enrollment for school-based educational research. *Journal of Applied Social Psychology* **32**, 1577–1587.

Merriman, W. E. (1986). Some reasons for the occurrence and eventual correction of children's naming errors. *Child Development* **57**, 942–52.

Merriman, W. E. & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning. *Monograph of the Society for Research in Child Development* **54**, 1–132.

Merriman, W. E. & Stevenson, C. M. (1997). Restricting a familiar name in response to learning a new one: evidence for the mutual exclusivity bias in young two year olds. *Child Development* **68**, 211–28.
Toglia, M. P. & Battig, W. F. (1978). *Handbook of semantic word norms*. New York: Erlbaum.
van Orden, G. C. (1987). A ROWS is a ROSE: spelling, sound, and reading. *Memory and Cognition* **5**, 181–98.