Children's use of phonology to infer grammatical class in vocabulary learning

KIMBERLY WRIGHT CASSIDY
Bryn Mawr College, Bryn Mawr, Pennsylvania

and

MICHAEL H. KELLY
University of Pennsylvania, Philadelphia, Pennsylvania

Prior studies of the relationship between phonological information and grammatical category assignment have focused on whether these relationships exist and whether people have learned them. This study investigates whether these relationships affect preschool children's vocabulary acquisition in a laboratory setting. Child participants learned 12 vocabulary words (6 nouns and 6 verbs) under three conditions, in which, (1) the syllable number/grammatical category relationship matched English, (2) the syllable number/grammatical category relationship was opposite to English, or (3) there was no relationship between syllable number and grammatical category. In the initial presentation of the words, children assumed that the novel words matched the pattern found in English. When the syllable number/grammatical category pattern matched that of English, the children learned more of the words. Phonological information also predicted error patterns. These results suggest that any account of vocabulary acquisition should consider the role of phonological information.

In order to learn a word successfully, a person must acquire its semantic, syntactic, and phonological properties. It is well established that semantic and syntactic properties are correlated, so that knowledge of one can be used to infer aspects of the other. Theories of vocabulary acquisition differ about the relative importance of inferring semantics from syntax (e.g., Gleitman, 1990) or syntax from semantics (e.g., Pinker, 1984), but all exploit systematic mappings between the domains. Empirical studies have also shown that children use these mappings to induce lexical properties. For example, in Brown's (1957) classic study of word meaning induction, preschool children inferred that the novel word *seb* was an action when it appeared as a verb in a sentence but was an object when it appeared as a noun.

Whereas extensive research in linguistics and psycholinguistics has examined relationships between semantics and syntax in the lexicon, possible interactions between these domains and phonology have been relatively ignored. We believe this neglect is due to the widespread assumption that word form and function are unrelated. This is not to suggest that there is some sort of inherent connection between the phonology of words that refer to objects or actions and the objects or actions them-

selves. Instead, languages may contain language-specific regularities that result in word forms denoting similar concepts or occupying the same syntactic role being more phonologically similar to each other than to other types of word forms. For many semantic categories, the validity of this arbitrariness assumption has been confirmed statistically, at least in English (Martin, Gagnon, Schwartz, Dell, & Saffran, 1996). However, other research has documented a potentially important violation of arbitrariness. Sentences 1a and 1b provide an example of this arbitrariness violation in English.

Nouns and verbs show a contrast in stress. (1a)

Nouns and verbs contrast in stress. (1b)

Both sentences contain the word *contrast*, which appears as a noun in Sentence 1a but as a verb in Sentence 1b. Note that this grammatical difference is associated with a phonological difference. The noun version of *contrast* has stress on the first syllable, whereas the verb version has stress on the second syllable. The stress difference between disyllabic nouns and verbs permeates the English lexicon (Kelly & Bock, 1988; Sherman, 1975). Further examples can be heard in the pronunciations of the noun-verb homographs *record*, *contest*, *rebel*, and *permit*.

The English noun-verb stress difference is not an isolated case of a phonological cue to grammatical class. Many other predictive patterns have been found in English and other languages (for a review, see Kelly, 1992). These cues are admittedly probabilistic, but, when combined, they could provide quite reliable means for assigning words to grammatical classes. For instance, Shi, Morgan,

The authors thank the children, parents, and teachers of the Phebe Anna Thorne School, Chesterbrook Academy, and The Piaget School for their participation in this project. Correspondence should be addressed to K. W. Cassidy, Department of Psychology/West House, 101 North Merion Avenue, Bryn Mawr, PA 19010 (e-mail: kcassidy@brynmaur.edu).

and Allopenna (1998) found that self-organizing networks could distinguish between functional (e.g., pronouns and prepositions) and lexical (e.g., nouns and verbs) categories in both Mandarin and Turkish with 82%–93% accuracy using a variety of acoustic and phonological features.

Given the existence of phonological cues to grammatical class, the next question concerns whether speakers learn them and use them to solve problems in language acquisition and processing. Many studies have documented that correlations between word phonology and grammatical class are learned. For instance, English speakers are more likely to pronounce pseudowords like *ponveen* with second syllable stress when they act as verbs in sentences rather than when they serve as nouns (Kelly & Bock, 1988; Smith & Baker, 1976). Real words are classified into their grammatical category more quickly and accurately when they have typical phonological properties for their category than when they have atypical properties (Bates, Devescovi, Pizzamiglio, D'Amico, & Hernandez, 1995; Davis & Kelly, 1997; Sereno, 1994). Though children's knowledge has not been investigated as much as adults' knowledge, there is some evidence that children learn phonological cues to grammatical class by at most 4 or 5 years of age (e.g., Cassidy & Kelly, 1991; Levy, 1983; MacWhinney, 1978). Indeed, more recently, Shi, Werker, and Morgan (1999) reported that newborn infants could use perceptual cues to discriminate between functional and lexical words.

In this article, we make an initial attempt to go beyond the issue of information sensitivity to information exploitation. In particular, we present evidence that children not only learn phonological cues to grammatical class, but also apply this knowledge when they learn new words. We will focus on a phonological distinction between nouns and verbs in English—namely, the tendency for nouns to contain more syllables than verbs. We chose this phonological cue because it appears robustly not only in the adult English lexicon, but also in parental speech to young children (Cassidy & Kelly, 1991). Furthermore, evidence indicates that children do learn this cue. Cassidy and Kelly found that 4-year-old children preferred to map novel words to actions (a prototypical verb referent) more often when they contained one syllable rather than three syllables. In contrast, the preference for objects (the typical noun referent) was greater for three-syllable words. In the experiment reported here, we tested whether this knowledge affected children's success in learning the meanings of new words.

Children between 3 and 6 years of age were introduced to a puppet who spoke an unfamiliar language. The child's task was to learn the meanings of some words in the puppet's language. On each trial, the child watched a videotape showing a common object performing or experiencing an action, such as a toy car driving in a circle or a door closing. During the videotape, the puppet would repeatedly shout a pseudoword, such as *varb*, and the child was asked to guess whether the pseudoword referred to the object or action in the videotape. The children were told whether their responses were correct or not,

after which the next trial began. The same set of videotapes and pseudowords were then presented through two more blocks of trials, with the child's accuracy serving as the dependent variable.

The pseudowords either contained one or three syllables, and we systematically varied the relationship between syllable number and word meaning across three conditions. In the *consistent* condition, monosyllabic words were assigned action referents and trisyllabic words object referents just as English tends to do. In the *inconsistent* condition, syllable number predicted word reference, but in the opposite direction from English. Finally, in the *independent* condition, syllable number was unrelated to word reference. If any correlation between word form and function assists learning, performance in the consistent and inconsistent conditions should be better than that in the independent condition. On the other hand, if children apply their knowledge of English principles to the task, they should perform best in the consistent condition. In this case, the inconsistent condition may be most difficult since the child's application of English knowledge would repeatedly lead to error.

The children's performance on the first block of trials provided a replication of Cassidy and Kelly (1991) and thus a measure of children's sensitivity to the relationship between syllable number and grammatical category. The feedback given on all three blocks of trials and the two extra blocks of trials provides important information about how children might use phonological information to learn the meanings of new words.

METHODS

Subjects

Forty-six native English-speaking children participated in the experiment. The ages of the children ranged from 3 years, 1 month to 6 years, 0 months, with a mean of 4 years, 7 months. The children attended one of three preschools in suburban Philadelphia. Eighty percent of the children were Caucasian, 10% were Asian American, and 10% were African American.

Materials

Pilot research indicated that children would tolerate about 12 trials before becoming fidgety. Hence, 12 scenes were videotaped, with each scene consisting of a common object performing or experiencing a common action, such as a car driving or a cup falling. The standard names for both objects and actions were monosyllabic English words. Words were controlled for frequency so that, in half the scenes, the word denoting an object was more frequent and in the other half, the word denoting the action was more frequent. See the Appendix for lists of scenes, along with the words denoting the object and action.

Six monosyllabic and six trisyllabic pseudowords were used for the learning trials. One word was assigned randomly to each scene, with the stipulation that no more than two words with the same number of syllables could appear in consecutive scenes. Hence, the children had to learn the meaning of one pseudoword per scene. This word denoted either the object or the action, depending on condition. See the Appendix for a list of the pseudowords. Although the scenes and pseudowords were the same as those used by Cassidy and Kelly (1991), the word-to-scene pairings were new, thus avoiding any specific word/scene pairing effects.

The study consisted of three learning conditions with the order of scene presentation and word-to-scene pairing constant across con-

ditions. The only factor that differed across conditions was whether the pseudowords denoted the object or action in their associated scene. In the consistent condition, all trisyllabic pseudowords denoted the object in the scene and all monosyllabic words denoted the action. In the inconsistent condition, all trisyllabic pseudowords denoted actions and all monosyllabic words denoted objects. Thus, in both of these conditions, number of syllables was a perfect cue to word meaning, but the mapping in the consistent condition matched the English pattern, whereas the mapping in the inconsistent condition reversed that pattern. In the independent condition, half of the trisyllabic words denoted the object and half denoted the action. The same was true for the monosyllabic words. So, in this condition, syllable number did not predict word meaning. For half of the subjects in the independent condition, the pairings were switched so that the words paired with objects for the first half of the subjects were paired with actions for the second half, and vice versa.

Procedure

Each child was randomly assigned to one condition. The ages and genders of the children did not differ significantly across conditions. Thus, any significant differences between conditions were not due to age differences. There were 16 children in the consistent condition and 15 in each of the other two conditions.

At the beginning of the experiment, each child was introduced to a puppet named Blip who was visiting from another planet and who spoke a foreign language. The children were told that they were going to learn some words from Blip's language. The experimenter then administered the following procedure for each trial, using *varb* as the pseudoword example.

Experimenter: When Blip sees the scene that I am about to show you he says "varb, varb, varb." Can you say "varb?" (The child is asked to say the word to ensure that it has been heard correctly.) Good. Let's see if you can figure out ("remember" was used on later repetitions of the scene) what Blip means when he says, "varb."

(The video tape was then turned on. During the scene Blip says the word three more times.)

Experimenter: When Blip says, "varb," do you think he means (action) or (object)?

The English names of the action and object were given to the child, and each of these names contained one syllable. Hence, the child's responses could not be based on differences in syllable number between the English words. The order of the action and object alternatives was counterbalanced across scenes within subjects and across subjects within scenes.

After each trial, the child was given feedback about the response. Correct responses were acknowledged, and the word/scene pairing was reinforced by saying, "That's right, *varb* means *close*." Errors were corrected by identifying the proper match. All feedback was scripted so that each child received the same level of feedback and reinforcement. At preset points throughout the procedure, all children were reminded that they were to try to learn Blip's language and to try hard to remember and learn as many of the words as they could. Each child was tested by one of two female experimenters. Both experimenters were blind to the experimental hypothesis.

Each child received three blocks of 12 learning trials. Pilot testing revealed that the task became tedious for preschool children unless there were regular interruptions. Hence, every 3 trials, children were rewarded with an opportunity to remove a piece from an attractive four-piece puzzle. A picture was slowly revealed underneath the puzzle as they removed each piece. This picture matched one of three distinctive suitcases that Blip brought from his planet. At the end of each block of trials, when all puzzle pieces had been removed, the child could match the picture to the suitcase and open the suitcase to see what was inside. The suitcases contained unusual and fun objects from Blip's planet that the child could play with for a brief time. This activity maintained the children's interest for the procedure.

RESULTS

Each child received three scores corresponding to the number of word meanings correctly inferred in each learning block. Each pseudoword item received nine scores corresponding to the proportion of children who identified its meaning correctly in each of the cells formed by crossing the factors of learning condition and learning block. These values were submitted to two analyses of variance (ANOVAs) with subjects and items as random variables. Learning condition was a between-subjects and within-items factor. Learning block varied within subjects and items.

Table 1 shows the children's mean number of correct word meaning assignments as a function of learning block and condition. Accuracy improved across the three blocks, indicating that the children were indeed learning the word meanings [subjects, $F(2,86) = 30.89$; items, $F(2,22) = 25.70$; both $ps < .001$]. Learning condition was also significant across subjects [$F(2,43) = 4.45$, $p < .02$] and items [$F(2,22) = 8.24$, $p < .005$]. Performance in the consistent condition was significantly better than that in the inconsistent [subjects, $t(29) = 3.01$, $p < .01$; items, $t(11) = 4.25$, $p < .01$] and independent [subjects, $t(29) = 2.13$, $p < .05$; items, $t(11) = 2.41$, $p < .05$] conditions. The slightly better performance in the independent condition, as compared with the inconsistent condition, was not significant.

The interaction between learning block and learning condition was not significant [subjects, $F(4,86) = 2.01$; items, $F(4,44) = 1.97$; both $ps > .10$]. Hence, the advantage of the consistent condition over the others did not change systematically over learning experience. Indeed, Table 1 shows that the advantage appeared even in the first block of trials [subjects, $F(2, 43) = 4.72$; items, $F(2,22) = 4.89$; both $ps < .05$], in which one would expect that the child's responses would be essentially guesses. This pattern indicates that children used their knowledge about the noun-verb syllable difference to guide their initial inferences about the meanings of the novel words.

To provide further evidence that children entered the experiment with the expectation that object terms are longer than action terms, we computed the number of times each child guessed noun as the meaning of the term for one- and three-syllable words in the first block of trials. A 3×2 (condition $\times$ syllable number) ANOVA revealed that the children were more likely to guess noun for three-syllable words ($M = 3.37$, $SD = 1.66$) than for

Table 1
Children's Mean Number
of Correct Word Meaning Assignments

Condition	Trial					
	Block 1		Block 2		Block 3	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Consistent	6.63	2.03	8.06	1.91	9.19	1.42
Independent	6.07	2.05	7.27	1.44	8.13	1.41
Inconsistent	4.67	1.35	7.60	2.4	7.00	2.62

one-syllable words ($M = 2.48$, $SD = 1.21$) [subjects, $F(1,43) = 12.08$; items, $F(1,10) = 19.78$; both $ps < .001$]. Since there was no syllable number $\times$ condition interaction, the children, regardless of condition, showed this bias.

Given the initial advantage of children in the consistent condition, one might question whether the form-function relations in that case actually facilitated learning beyond the baseline effect. To address this question, we measured each child's improvement between trial Block 1 and trial Block 3. This improvement was determined by calculating the proportion of words missed in Block 1 that were correct in Block 3. The use of a proportion controls for the fact that the children in the consistent condition had the least opportunity to learn the words, since their performance began at a higher level. A one-way ANOVA showed a main effect of condition on the proportion of words learned [subjects, $F(2,43) = 3.24$; items, $F(2,22) = 3.54$; both $ps < .05$]. In the third block, the children in the consistent condition responded correctly to 70% of the items on which they erred in the first block. In contrast, children in the inconsistent and independent conditions learned only 58% and 52%, respectively. Comparisons of the proportion of improvement in Block 3 for those items missed in Block 1 showed that children in the consistent condition showed greater learning success than did the children in both the independent condition [subjects, $t(29) = 2.87$; items, $t(11) = 2.19$; both $ps \leq .05$] and those in the inconsistent condition, although the latter difference was statistically marginal in the subjects analysis [subjects, $t(29) = 1.72$, $p < .1$; items, $t(11) = 2.15$, $p = .05$]. Children in the independent condition were not different from children in the inconsistent condition.

If children are using phonological information when learning pseudowords, certain patterns of errors would be predicted. Specifically, in the independent condition, children should make more errors on monosyllabic nouns and trisyllabic verbs since these words violate the relations between syllable number and grammatical class found in English. The proportion of all errors made on monosyllabic nouns plus trisyllabic verbs was calculated for each subject in the independent condition for the last block of trials. The mean proportion of all errors made for monosyllabic nouns and trisyllabic verbs ($M = 0.70$, $SD = 0.23$) was significantly different from .50 [subjects, $t(14) = 3.29$; items, $t(11) = 2.63$; $ps < .05$], which would be expected if errors were made randomly across all word types.

While the age range of the subjects was broad, a great majority of the subjects were 4-year-olds. The small number of 3- and 5-year-old subjects made it impossible to examine age effects. Future work might investigate developmental differences in the ability to exploit phonological information in vocabulary acquisition.

DISCUSSION

Prior studies of the relationship between phonological information and grammatical categories have focused on

whether or not these relationships exist and whether or not people, particularly adults, have learned them (for a review, see Kelly, 1992). Few studies have investigated whether or not these cues are exploited in language acquisition or processing. Although the data from Cassidy and Kelly (1991) suggest that preschool children have knowledge of the relationship between number of syllables and grammatical category, the data from the present study suggest that preschool children not only are sensitive to the relationship between grammatical category and number of syllables, but that they use this relationship when learning new words.

Children's knowledge of the grammatical category/syllable number relationship seems to affect their language learning in three ways. First, most of the children in this study who were presented with a set of novel words to learn, appear to have assumed that these words would conform to the syllable number/grammatical category pattern found in English. On the initial presentation of the words, the children were more likely to guess that trisyllabic words denoted objects and monosyllabic words denoted actions. Second, the children learned novel word meanings more successfully when the experimental correlation between number of syllables and word meaning matched that found in their native language. This was true even after correcting for the fact that children's assumptions about form-meaning relationships translated into greater accuracy in the consistent condition in the very first block of trials, before any feedback about accuracy had been provided. Not only did the phonological cue/grammatical category relationship determine which children learned the best, it also influenced which items were difficult to learn. For example, in the independent condition, children had more difficulty learning the items that violated the noun/verb syllable pattern found in English.

Third, children in this study appeared to be applying their knowledge of the noun/verb syllable number difference from their native language to the language learning task, rather than exploiting the relationship that was present in the data set. Children's performance in the inconsistent condition was not different from their performance in the independent condition. At least for the brief learning situation presented here, the children were not able to exploit a form-function relationship that violated the pattern of their native language.

If the children's performance on this laboratory learning task reflects their behavior in actual vocabulary acquisition, certain patterns should be present in natural language learning data. For example, children should be less likely to have words in their vocabulary that violate the noun/verb syllable number difference. Although in part this might reflect what is in their input (Cassidy & Kelly, 1991), we predict that the noun/verb syllable number difference should be significantly more pronounced in children's language.

This study examined just one of the many phonological cues that children might exploit in learning language. For example, English nouns and verbs also differ according to stress, duration, and vowel type. In normal vocabulary acquisition, children may exploit multiple cues in making grammatical category assignments.

It might be argued that the only reason phonological information helped children in this study was that the learning situation was impoverished. There was no syntactic information and little semantic context. However, children do hear single words in isolation. Cassidy and Kelly (1991) examined a large corpus of mother-to-child speech collected from naturalistic observation. In this corpus, 15% of all mothers' utterances consisted of a single word, and these words were from many different grammatical categories. We are also not claiming that phonological information is the only source of information that children use in actually making grammatical category assignments. Just as people use multiple cues to perceive a three-dimensional environment, the grammatical class of a word could be determined by simultaneously using various types of phonological, semantic, and distributional information. Our main point is that phonological information has been widely neglected in studies of grammatical category assignment. This is true despite the fact that what little research has been conducted has demonstrated that phonological cues to grammatical category are plentiful and are known by adult native and nonnative speakers and by children (Davis & Kelly, 1997; Kelly, 1992). Our study goes one important step further in showing that children can use these probabilistic cues in acquiring new words. In particular, our results suggest that phonological features that are indicative of grammatical class may increase the speed and accuracy of children's learning the meaning of novel words. More generally, our data indicate that any account of vocabulary acquisition needs to consider the role of phonological information, as well as the more extensively studied domains of syntax and semantics.

REFERENCES

- BATES, E., DEVESCOVI, A., PIZZAMIGLIO, L., D'AMICO, S., & HERNANDEZ, A. (1995). Gender and lexical access in Italian. *Perception & Psychophysics*, **57**, 847-862.
- BROWN, R. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal & Social Psychology*, **55**, 1-5.
- CASSIDY, K. W., & KELLY, M. H. (1991). Phonological information for grammatical category assignments. *Journal of Memory & Language*, **30**, 348-369.
- DAVIS, S. M., & KELLY, M. H. (1997). Knowledge of the English noun-verb stress difference by native and nonnative speakers. *Journal of Memory & Language*, **36**, 445-460.
- GLEITMAN, L. (1990). Structural sources of verb learning. *Language Acquisition*, **1**, 1-54.
- KELLY, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review*, **99**, 349-364.
- KELLY, M. H., & BOCK, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 389-403.
- LEVY, Y. (1983). It's frogs all the way down. *Cognition*, **12**, 75-93.
- MACWHINNEY, B. (1978). The acquisition of morphonology. *Monographs of the Society for Research in Child Development*, **43** (1-2, Serial No. 174).
- MARTIN, N., GAGNON, D. A., SCHWARTZ, M. F., DELL, G. S., & SAFFRAN, E. M. (1996). Phonological facilitation of semantic errors in normal and aphasic speakers. *Language & Cognitive Processes*, **11**, 257-282.
- PINKER, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- SERENO, J. (1994). Phonosyntactics. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound symbolism* (pp. 263-275). Cambridge: Cambridge University Press.
- SHERMAN, D. (1975). Noun-verb stress alternation: An example of lexical diffusion of sound change in English. *Linguistics*, **159**, 43-71.
- SHI, R., MORGAN, J. L., & ALLOPENNA, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language*, **25**, 169-201.
- SHI, R., WERKER, J., & MORGAN, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, **72**, B11-B21.
- SMITH, P. T., & BAKER, R. G. (1976). The influence of English spelling patterns on pronunciation. *Journal of Verbal Learning & Verbal Behavior*, **15**, 267-285.

APPENDIX

Scenes Shown and the Word Paired with These Scenes

-
- Scene 1: A bear puppet runs across the screen. (gebinarf)
- Scene 2: A cup falls from the top of a box. (yolt)
- Scene 3: A remote control car drives in circles. (bligorite)
- Scene 4: A box slides across a table top. (delfinade)
- Scene 5: A kite flies in the air. (varb)
- Scene 6: A hand waves back and forth continuously. (sumtredlin)
- Scene 7: A pig puppet claps its hooves. (janx)
- Scene 8: A sheep puppet bows. (vinractin)
- Scene 9: A door closes. (pell)
- Scene 10: A cow puppet turns around and around. (zorp)
- Scene 11: A ball rolls across a table. (jadbursen)
- Scene 12: A toy frog jumps across the table top. (litch)
-

(Manuscript received March 29, 2000;
revision accepted for publication October 4, 2000)