Infants can use distributional cues to form syntactic categories*

LOUANN GERKEN, RACHEL WILSON

University of Arizona

AND

WILLIAM LEWIS

California State University, Fresno

ABSTRACT

Nearly all theories of language development emphasize the importance of distributional cues for segregating words and phrases into syntactic categories like noun, feminine or verb phrase. However, questions concerning whether such cues can be used to the exclusion of referential cues have been debated. Using the headturn preference procedure, American children aged 1;5 were briefly familiarized with a partial Russian gender paradigm, with a subset of the paradigm members withheld. During test, infants listened on alternate trials to previously withheld grammatical items and ungrammatical items with incorrect gender markings on previously heard stems. Across three experiments, infants discriminated new grammatical from ungrammatical items, but like adults in previous studies, were only able to do so when a subset of familiarization items was double marked for gender category. The results suggest that learners can use distributional cues to category structure, to the exclusion of referential cues, from relatively early in the language learning process.

INTRODUCTION

Our human ability to comprehend and produce an infinite number of sentences depends critically on the ability to treat semantically varying words as belonging to discrete syntactic categories, such as noun, noun phrase, verb, etc. How infants and young children come to employ the linguistic category structure of their language community has been a

---

topic of heated debate. We can identify two main positions in this debate: SEMANTIC BOOTSTRAPPING and DISTRIBUTIONALLY BASED CATEGORY FORMATION. On the semantic bootstrapping hypothesis (Grimshaw, 1981; Pinker, 1984), language learners are born expecting to find a set of syntactic categories in their language input. They are also born knowing how to link actual words that they encounter with these categories. Focusing for the moment on the categories noun and verb, the innate linking knowledge is in the form of an expectation that words referring to objects are nouns and that words referring to actions are verbs.[1] How, on this view, do learners determine the lexical category of words other than object and action labels (e.g. 'The situation justified the measures')? Once they have a large enough set of objects and actions categorized, they notice distributional information in the sentences containing these words. Distributional information can be the location of the word in the sentence (e.g. English nouns are often sentence final), phonological properties (e.g. English two syllable verbs are often stressed on the second syllable), or marker elements (e.g. English nouns are frequently preceded by 'the' or 'a,' and verbs are often ended in '-ed', '-s', or '-ing'). Thus, the process of assigning words to categories occurs by first attending to referential properties and then distributional ones.

If distributional properties are sufficient for assigning words to categories, why are they not used from the beginning? The semantic bootstrapping hypothesis presents two arguments for beginning with reference, one empirical and one logical. The empirical reason for beginning with reference is that very young children often fail to produce grammatical morphemes, which are the most frequent marker of syntactic categories. However, since the time that semantic bootstrapping was proposed, there has emerged ample evidence that young children's failure to produce grammatical morphemes does not reflect a lack of knowledge of these elements. Children show poorer comprehension of even simple sentences when these elements are missing or incorrect (Shipley, Smith & Gleitman, 1969; Petretic & Tweney, 1977; Gerken & McIntosh, 1993). Further, even infants show sensitivity to various properties of grammatical morphemes (Shady, Gerken & Jusczyk, 1995; Shady, 1997; Shafer, Shucard, Shucard & Gerken, 1998; Hoehle & Weissenborn, 1999; Shi, Werker & Morgan, 1999). In addition, a variety of studies suggest that gender categories in particular are acquired based on the presence of grammatical morpheme markers, not on reference (e.g. Karmiloff-Smith, 1979; Levy, 1983; Mills, 1985). In sum, these studies suggest that we can dismiss the empirical reason for beginning with reference.

---

[1] Note that the expectation is unidirectional, such that learners are not saddled with the false belief that nouns are words referring to objects.

The logical reason for beginning categorization with reference is that the ubiquitous nature of major linguistic categories like noun and verb in descriptions of language patterns has led to assumptions that they are part of the vocabulary of innate knowledge, or SUBSTANTIVE UNIVERSALS (Chomsky, 1965). Consider the following four proposed innate syntactic parameters from Baker (2001):

   1a.   Only one verb can be contained within each verb phrase or more than one verb can be contained in a single verb phrase.
   1b.   The subject of a clause is merged with the verb phrase or it is merged with the auxiliary phrase.
   1c.   Adjectives are treated as a kind of verb or adjectives are treated as a kind of noun.
   1d.   If agreement with a noun phrase X is not required, use the agreement to show that the noun phrase X is animate and/or definite in its reference.

Within a parameter setting framework, a language learner uses some piece of data from her input to decide between the choices offered by innate parameters like these. Importantly for our discussion, these four parameters refer to the linguistic entities verb, verb phrase, subject, auxiliary phrase, adjective, noun, and noun phrase. Positing innate knowledge in such a form requires that the child have *a priori* knowledge of these entities and a way of linking them to particular words and phrases encountered in her language. The distribution of sentence position, phonological cues, and marker elements may give learners the ability to segregate nouns into different groups than verbs or adjectives. However, these distributional cues on their own cannot provide the necessary link to the innate noun, verb or adjective categories referred to in the above parameters. It is this linking problem that Semantic Bootstrapping attempts to solve.

It is important to keep in mind that the linking problem is only a problem so long as we assume innate categories and rules that refer to them. Proponents of Distributionally Based Category Formation question the assumption of innate substantive universals (e.g. Maratsos & Chalkley, 1980; Braine, 1992). Rather, they argue that words can be segregated into unlabeled classes based on distributional information. They further suggest that syntactic roles, like sentential subject, can be discerned when distributional information is combined with referential information, a claim that eliminates the linking problem. A full account of how syntax acquisition might work without assuming innate syntactic categories has yet to be proposed. Indeed, the logical possibility of a workable account of this sort is questioned by many nativists.

One currently testable difference between the two approaches concerns their relative reliance on reference in the early stages of category formation.

Under Semantic Bootstrapping, early segregation of words into categories and linking these categories to innate substantive universals occurs simultaneously, via reference. In contrast, under Distributionally Based Category Formation, learners may segregate their input into distributionally distinct sets without relying on reference. The distributionally based sets may later be linked to sentence structure (e.g. subject) via reference. The notion that learners can extract purely formal properties of their language before linking them to meaning has been proposed in at least two recent discussions of language development (Gómez & Gerken, 2000; Naigles, 2002). The goal of the research reported here was to determine whether infants are able to show evidence of syntactic category formation based on distributional properties alone. Although a positive finding would not rule out Semantic Bootstrapping, it would allow us to consider alternative accounts in which segregating words into sets and linking these sets to sentence structure are potentially independent.

Before we describe our approach, let us consider what we might mean by distributionally based category formation. Consider the fact that most English verbs can end in the present tense in '-ing.' A learner sensitive to such distributional information might form a category of words ending in '-ing.' However, such a category might incorrectly include the words 'spring' and 'sing' (see Gleitman & Wanner, 1982 for an excellent discussion of these issues). The problem would be much worse if the distributional information attended was the inflection '-s,' because it not only marks present tense on verbs, but plural and possessive on nouns. The problem of overgeneralization indicates that noting single category-marker relations is not a viable approach to distributionally based category formation.

A different view of distributionally based category formation can be found in the seminal work of Maratsos & Chalkley (1980), who propose that syntactic categories are nothing more than PARADIGMS of distributional cues. Thus, most English verbs can end in either '-ing' or '-s' (running, runs, considering, considers), and it is exceedingly unlikely that a non-verb will participate in this paradigm. However, 2a–f, below, demonstrate that simply noting distributional paradigms in which category markers appear in complementary distribution is also not sufficient. A learner who noted 2a–2d might hear 2e and come to the incorrect conclusion that 2f is a possible phrase of English.

2a.   the brush
2b.   can brush
2c.   the paint
2d.   can paint
2e.   the cat
2f.   *can cat

Interestingly, human learners may be protected from making a category overgeneralization error such as the one in 2f, because they are unlikely to discern any morphosyntactic paradigm in which there is only a single cue to category in each utterance. Early studies (Braine, 1966; Smith, 1966) presented adults with four types of elements (e.g. M, N, P and Q) that could occur only in MN and PQ combinations. Participants learned that M and P came first in phrases, but they did not learn the co-occurrence restriction that M could only occur with N and P with Q. Because co-occurrence restrictions are at the heart of syntactic category formation, the early studies cast doubt on adults' ability to learn categories via distributional cues alone. Braine (1987) dubbed the failure to learn co-occurrence restrictions in a paradigm learning task the 'MN/PQ problem.' In order to overcome this problem, Braine suggested that single distributional cues, such as suffixing morphemes, could be supplemented with either referential cues, additional morphophonological cues, or statistical family resemblance relations not governed by grammar (Zubin & Kopcke, 1981). For example, the fact that only 2a, 2c, and 2e can occur with both a plural marker and determiner might be sufficient information to allow learners to solve the MN/PQ problem without making the overgeneralization in 2f.

In one study testing the multiple cues hypothesis, Braine (1987) presented participants with an MN/PQ type language with MN and PQ phrases each comprising two auditory nonsense words. Each phrase was presented with a picture. Half of the N words was accompanied by pictures of women and half of the Q words by pictures of men. The other half of the pictures illustrated inanimate objects with no apparent semantic category. Additionally, the M and P words corresponded to numerosity in the pictures. Thus, there was an M and a P word for 'one,' for 'two,' etc. Some of the possible MN and PQ pairings were withheld. Participants made grammaticality judgments of the phrases and were correctly able to generalize to the withheld cases. Generalization was even correct when the phrase corresponded to a picture of an inanimate object, suggesting that participants had formed categories of M, N, P and Q words and did not need reference to access the categories, once they had been formed. Braine speculated that the mechanism learners used to solve the MN/PQ problem is to first note that N and Q words behave as categories, based on referential or other information. Learners then note that the referentially based categories co-occur with morphophonological markers (in this case, number words differentially marked for gender), and ultimately use the markers themselves as the basis of categorization.

A variety of researchers followed up on Braine's (1987) study, attempting to show category formation without referential cues. Until recently, the results have, at best, been equivocal (Brooks, Braine, Catalano, Brody &

Sudhalter, 1993; Frigo & McDonald, 1998; Kempe & Brooks, 2001).[2] Gerken, Wilson, Gómez & Nurmsoo (2002) suggested that the failures of earlier studies using purely distributional cues to syntactic categories were due to the fact that participants in these studies were presented with a referential field for each familiarization phrase that was irrelevant to category formation. Gerken *et al.* argued that participants might have focused more fully on learning the referent for words in the familiarization phrases than on inducing the structure of the language. This explanation is supported by recent sets of studies (Gerken *et al.*, 2002; Mintz, 2002; Wilson, 2002) in which adults were familiarized with paradigms containing two fully or partially correlated cues to categories. The paradigms examined two different types of artificial languages as well as a part of the Russian noun gender system. As in the previous studies, some members of the paradigm were withheld during familiarization and presented, along with ungrammatical items, at test. Participants across several studies using different types of materials discriminated new grammatical from ungrammatical test items, suggesting that they had learned the paradigm structure. However, they were only able to do so when at least a subset of the familiarization stimuli contained two category cues. They failed to discriminate the same test items when familiarization stimuli contained only one distributional cue to categories (Gerken *et al.*, 2002; Wilson, 2002). The adult learning data complement a variety of studies of distributional analysis over corpora by computer suggest that such information results in accurate word sorting (Brent, 1992; Cartwright & Brent, 1997; Mintz, 2002).

Although the multiply-cued paradigm approach to syntactic category formation solves both the MN/PQ problem and the problem of over-generalization, it puts a heavy computational burden on the learner, who must remember the various morphosyntactic contexts in which a particular word or stem has occurred. Recent research in which infants are exposed to artificial languages and tested on their ability to generalize suggests that they are indeed computationally more capable than detractors of distributional approaches might believe (Saffran, Aslin & Newport, 1996; Gómez & Gerken, 1999; Marcus, Vijayan, Rao & Vishton, 1999; Gómez, 2002). The infant studies, coupled with recent studies showing adult participants' ability to use multiply-cued paradigms, allows us to now ask whether infants are able to show sensitivity to syntactic categories familiarized with input containing partially overlapping cues to category structure.

---

[2] By equivocal, we mean that participants' performance on test items with only a single marking to category structure is not reported to be above chance.

EXPERIMENT 1

In all three experiments presented here, the syntactic category paradigm we used was Russian gender. We chose this domain for three reasons. First, Maratsos & Chalkley used the Russian gender marking system as an example of a rich and complex morphophonological paradigm that real children learn. It therefore seemed fitting to ask whether infants could learn a part of such a paradigm in a brief laboratory exposure. Second, we wanted to familiarize infants with a category that did not have a counterpart in English. Finally, we know that American adults can learn a Russian gender paradigm system if they are familiarized with stimuli containing partially correlated cues to syntactic categories, giving us a point of comparison for the infants' behaviour.

In Experiment 1, one group of infants was familiarized with the Russian gender paradigm prior to test. It was predicted that these infants would discriminate grammatical from ungrammatical test trials. A second group of infants were not familiarized, but simply given the test. This group was included as a control to ensure that the grammatical and ungrammatical items were not discriminable on acoustic grounds alone.

METHOD

Participants

Two groups, each with 14 infants, contributed data for Experiment 1. The Familiarization group (8 female) ranged in age from 1;4.17 to 1;5.7, with a mean of 1;4.28. An additional seven infants were tested in the Familiarization group but were not included due to failure to remain attentive during the two-minute familiarization (n=1) or failure to look toward the blinking light for at least two seconds on at least nine of the 12 test trials (n=6). The no-familiarization Control group (7 female) ranged in age from 1;4.23 to 1;5.24, with a mean of 1;5.5. An additional 13 infants were tested in the Control group but were not included due to failure to look toward the blinking light for at least two seconds on at least nine of the 12 test trials. Infants were excluded from either group for the following reasons: exposure to a language other than English for more than five hours per week, history of ear infections, under 38 weeks gestation, birth weight under 5·5 pounds, family history of speech, language or hearing disorders.

Stimuli

A Russian gender paradigm was created in which six masculine and six feminine lexical stems appeared with two different case endings each, for a total of 24 words (see Table 1). Words were recorded by a fluent speaker of Russian (RW). The feminine case endings were *-oj* and *-u*, and the

TABLE 1. *Stimuli used in Experiment 1*. Bolded words were withheld during familiarization and comprised the grammatical test items. Ungrammatical words were *ruchkyem, vannya, zhitelyu, tramvayoj*

		Feminine words			
polkoj	rubashkoj	ruchkoj	**vannoj**	knigoj	korovoj
polku	rubashku	**ruchku**	vannu	knigu	korovu
		Masculine words			
uchitel'ya	stroitel'ya	zhitel'ya	**tramvaya**	korn'ya	pisar'ya
uchitel'yem	stroitel'yem	**zhitel'yem**	tramvayem	korn'yem	pisar'em

masculine endings were -*ya* and -*yem*. Three of the feminine words ended in the consonant /k/, and three of the masculine words ended in the syllable /tɛl'/.[3] The suffix /k/ without a preceding vowel in Russian is a productive diminutive inflection for feminine words; /k/ is added with an epenthetic vowel (e.g. /ok/) on masculine words. Even words that are not productively diminutivized in modern Russian can only appear with a final /k/ and no preceding vowel if they are feminine, perhaps because they are frozen forms of older diminutivized forms. For example, the modern root *ruchk-* meaning 'pen' is derived from 'little hand.' Thus, /k/ without a preceding vowel can be thought of as a derivational morpheme for feminine nouns, which is either currently productive or which resulted in the past in a now-frozen derived form. The syllable /tɛl'/ is a derivational inflection meaning roughly 'a person who does x' (Townsend, 1968). We refer to these six words as 'double marked,' because both the stem ending and case inflection indicate their gender. Two words of each gender were withheld from familiarization and served as the grammatical test items. One withheld word of each gender was double marked and one was marked only with the case inflection.

The familiarization stimuli, which were only heard by infants in the Familiarization group, consisted of 20 words in four different random orders, for a total of 80 words. Words ranged in duration from approximately 750 to 1000 msec. and were separated by a 500 msec. pause. Thus, the entire familiarization set lasted approximately two min. During test, in which all infants participated, grammatical trials comprised four tokens of each of the four withheld words with 375 msec. pauses between words, for approximately 20 sec. total duration. Two different versions of grammatical trials were presented in each third of testing, with versions exhibiting different random orders of the 16 words. Ungrammatical test trials had the same

[3] Throughout the text and tables, an apostrophe after a Russian consonant indicates that it is palatalized.

format, but were created by putting incorrect case endings on the four withheld words.

Each infant was tested individually while seated on the caregiver's lap in a sound-proof booth using the head-turn preference procedure (see Kemler Nelson, Jusczyk, Mandel, Myers, Turk & Gerken, 1995). The booth contained an amber light directly in front of the infant and two red lights above speakers to the infant's right and left. An observer outside the test booth watched the infant on a video monitor with the sound off and recorded the infant's looking behaviour using a button box connected to an Apple Power Macintosh computer. The experimental control programme initiated trials and scored head-turn responses. The caregiver listened to masking stimuli over headphones.

During familiarization, stimuli were presented simultaneously from the two speakers. The center light began to flash, and when the observer indicated that the infant looked at the light, one of the two side lights began to flash and continued to flash throughout the familiarization. There was no relationship between lights and sound during familiarization. During the test, each trial began with the light blinking at center. When the observer indicated that the infant was looking at the light, one of the two side lights began to flash. When the infant looked toward the flashing side light, auditory stimuli began to play from the corresponding speaker until the infant looked away for two seconds or until the stimulus ended. The computer programme tracked looking times, the amount of time looking away from the source of sound (terminating trials after 2 sec.), and controlled the randomization and presentation of stimuli. The dependent measure was amount of time an infant oriented toward the light on each trial type.

RESULTS AND DISCUSSION

Trials under two seconds in duration were excluded from the analysis, and as noted above, all infants included in the study contributed at least nine useable trials.[4] It is typical in experiments with infants for looking times to decrease over test trials. In order to take this source of variance into account, test trial block (1, 2, 3) was included as a factor in the data analysis. Each infant's looking time for the two versions of each trial type were averaged for each block. Thus, each infant contributed six looking times

---

[4] We used a 2 sec. trial criterion, because that duration is standard in the field, and it takes into account a 375 msec. silence at the beginning of each trial, 1000 msec. Russian words, and time for the observer at the start of each trial to determine if infants were moving but still looking at the light or whether they were looking away.

(2 grammaticality × 3 block). Three infants (2 Familiarization, 1 Control) were missing a data point for one of the six blocks, and these were filled in with the mean for that variable. A (2) group (Familiarization vs. Control) × (2) grammaticality (grammatical vs. ungrammatical) × (3) block analysis of variance performed on the looking time data revealed no significant effects. Although there was no interaction involving group in the ANOVA, a chi-square test on the number of positive vs. negative differences in listening times for grammatical minus ungrammatical trials was significant ($\chi^2 = 5 \cdot 14$, n$=28$, $p=0 \cdot 02$). Therefore, separate ANOVAs were performed on the Familiarization and Control groups. The ANOVA on the Familiarization group revealed a significant main effect of grammaticality, such that ungrammatical trials (10·16 sec.) garnered longer looking times than grammatical trials (8·91 sec., $F(1, 13)=4 \cdot 51$, $p=0 \cdot 05$).[5] Neither the effect of block nor the grammaticality × block interaction was significant. In contrast, the ANOVA on the Control group revealed no significant main effects or interactions (mean grammatical$=8 \cdot 05$ sec., mean ungrammatical$=8 \cdot 46$ sec.).

The results of Experiment 1 show that, by 1;5, children are able to distinguish grammatical from ungrammatical Russian words, none of which they have heard before, after having been exposed to a Russian gender paradigm for about two minutes.[6] In contrast, infants who were not exposed to the paradigm were unable to discriminate grammatical from ungrammatical items, suggesting that the effect in the Familiarization group was indeed due to familiarization with the paradigm.

However, before we can consider the possibility than infants are able to form proto-categories comprising feminine vs. masculine Russian words, we need to consider two potential artifacts, either of which could have caused the effects observed. First, a subset of the test items, as well as the familiarization items, were double marked. Therefore, infants might have responded based on familiar phonological patterns. For example, infants might have noted the phonological pattern *tel'yem* during familiarization. Therefore, even though they had never heard *zhitel'yem* before, the familiarity of the final two syllables could have caused them to discriminate this grammatical test item from the ungrammatical counterpart *zhitel'yu*. Because double marked and single marked grammatical test items occurred

---

[5] Note that the same analysis of variance performed on the data with participants with missing data omitted also showed a statistically significant main effect of grammaticality ($F(1, 11)=6 \cdot 15$, $p=0 \cdot 02$).

[6] We tested a group of one-year-olds with the same familiarization and test stimuli used in Experiment 1. They showed no evidence of learning, suggesting that whatever abilities are required to discern morphophonological paradigms, they are not available to learners of all ages. Gómez & Lakusta (2004) have been successful at showing in one-year-olds an apparent precursor of the sort of category formation reported here.

TABLE 2. *Stimuli used in Experiment 2 and in the double marking condition of Experiment 3.* Bolded words were withheld during familiarization and comprised the grammatical test items. Ungrammatical words were *vannya, korovyem, medevedoj, pisaru*

		Feminine words			
polkoj	rubashkoj	ruchkoj	**vannoj**	knigoj	korovoj
polku	rubashku	ruchku	vannu	knigu	**korovu**
		Masculine words			
uchitel'ya	stroitel'ya	zhitel'ya	**medved'ya**	korn'ya	pisar'ya
uchitel'yem	stroitel'yem	zhitel'yem	medved'yem	korn'yem	**pisar'yem**

in each block of grammatical trials, the discrimination effect could have been carried by the double marked cases alone.

The second and related potential artifact concerns Russian palatalization. Five of masculine items and none of the feminine items ended in palatalized consonants. Palatalization affects the pronunciation of the following case ending, making it sound to the English ear like it begins with a glide. The only masculine item not ending in a palatalized consonant already ended in a glide (*tramvay-*). Therefore, when the feminine case endings were added to masculine words to make ungrammatical test items, the resulting inflections might have sounded to infants like they began with glides (*zhitelyu, tramvayoj*). In contrast, the same feminine endings did not begin with glides when they appeared on feminine words during training. The difference in the apparent phonological shape of the ending may have cued infants to differences between grammatical and ungrammatical items. Both of these potential artifacts were controlled in Experiment 2.

EXPERIMENT 2

Experiment 2 was designed to eliminate the potential artifacts discussed with respect to Experiment 1. As in Experiment 1, infants were familiarized with masculine and feminine stems that were both single and double marked for gender. However, to eliminate the possibility that infants in Experiment 1 responded based on familiar phonological sequences comprising the derivational inflection plus case inflection (e.g. *tel'ya*), test items were only marked with the case inflection (see Table 2). Further, to eliminate a potential confound involving palatalization, masculine cased endings appearing on feminine stems in the ungrammatical test items began with glides (i.e. *vannya* and *korovyem*), and feminine case endings appearing on masculine words were not palatalized (i.e. *medvedoj* and *pisaru*). Note that the masculine word *tramvay* was replaced with *medved* so that the feminine case ending *-oj* could be added to a word not ending in a glide.

METHOD

Participants

Eighteen infants (10 females) ranging in age from 1;4.16 to 1;5.27, with a mean of 1;4.29 participated in the experiment. An additional seven infants were tested but not included due to failure to look toward the blinking light during any test trials (n=2) or for at least two seconds on at least nine of the 12 test trials (n=5). Exclusionary criteria were the same as in Experiment 1.

Stimuli

The 20 familiarization stimuli were identical to those used in Experiment 1, except that, as noted above, *medved'yem* replaced *tramvayem*. The grammatical and ungrammatical test items are given in Table 2. Both familiarization and test stimuli were presented in the manner described in Experiment 1.

Procedure

The procedure was identical to that of Experiment 1.

RESULTS AND DISCUSSION

As in Experiment 1, each infant's looking times for the two versions of each trial type were averaged for each block, resulting in six looking times per infant. Two infants were missing a data point for one of the six blocks, and these were filled in with the mean for the variable. A 2 grammaticality × 3 block analysis of variance performed on the looking time data revealed a marginal grammaticality × block interaction $(F(2, 34)=2\cdot68, p=0\cdot08;$ Figure 1). The interaction was followed up with pairwise comparisons. They revealed a significant effect of grammaticality in the third block of trials, such that, as in Experiment 1, infants produced longer looking times for ungrammatical trials (6·93 sec.) than grammatical trials (4·95 sec.). As shown in Fig. 1, looking times for grammatical and ungrammatical trials were nearly equal, until in the third block, looking times for grammatical items decreased. Neither the effect of grammaticality or block approached significance.

   Experiment 2 removed two potential artifacts found in Experiment 1 and nevertheless replicated the effect that infants discriminated grammatical from ungrammatical, at least for the third block of trials. The combined data from the first two experiments suggest that infants in the second year of life can discern enough information about a partial gender paradigm to generalize to new grammatical words of the paradigm. Two questions
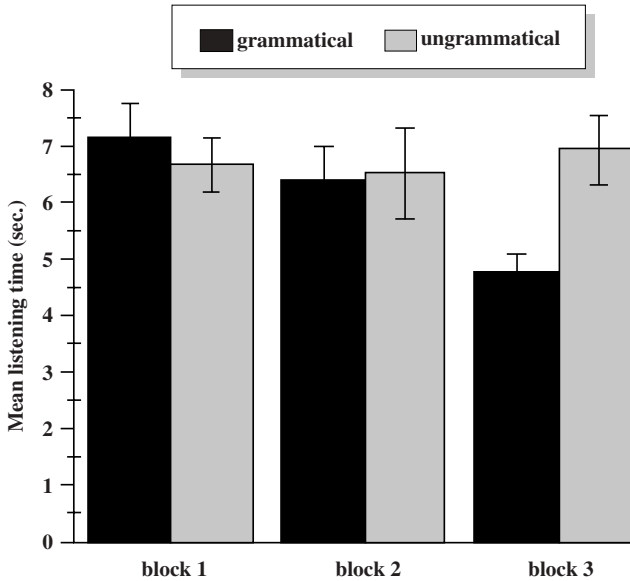
Fig. 1. Mean listening times and SE for grammatical vs. ungrammatical trials by block in Experiment 2.

remain. First, infants' ability to discriminate grammatical from ungrammatical test trials was relatively weak in Experiment 2, appearing only during the last block of testing. This pattern of results is puzzling, because the same physical stimuli occurred in all three blocks. In any case, it is important to know whether the effect can be replicated in another group of infants. Second, do infants, like adults in previous studies, require a subset of the familiarization stimuli to be double marked for category in order to learn the category structure? These questions were addressed in Experiment 3.

EXPERIMENT 3

To address the two questions raised by Experiment 2, Experiment 3 tested two groups of infants. The first group had an identical familiarization and test experience as the infants in Experiment 2 and thereby potentially allowed for a replication of that study. The other group had the same test experience, but was familiarized with a set of stimuli with no double marking of gender category. Thus, the only marking came from the two feminine and two masculine case inflections. We predicted that only infants in the double marking condition would discriminate grammatical from

ungrammatical test items. If the predicted result is borne out, we will have evidence that familiarization affects infants' performance during test, because both the single and double marking groups are tested on the same items.

METHOD

Participants

In the Double Marking condition, 18 infants (8 female) ranging in age from 1;4.10 to 1;5.20, with a mean of 1;4.22, contributed analysable data. An additional 12 infants were tested but not included due to failure to remain attentive during the two minutes familiarization (n=1) or failure to look toward the blinking light for at least two seconds on at least nine of the 12 test trials (n=11). In the Single Marking condition, 16 infants (6 female) ranging in age from 1;4.3 to 1;5.13, with a mean of 1;4.24 contributed analysable data. An additional 13 infants were tested but not included due to experimenter error (n=1), failure to remain attentive during the two min. familiarization (n=2), or failure to look toward the blinking light for at least two seconds on at least nine of the 12 test trials (n=10).[7] Exclusionary criteria were the same as in Experiment 1.

Stimuli

For infants in the Double Marking condition, familiarization and test stimuli were identical to those in Experiment 2 (see Table 2). For infants in the Single Marking condition, two of the three feminine nouns ending in /k/ and the two of the three masculine nouns ending in /tɛl'/ were replaced with feminine and masculine nouns, respectively, that did not have a phonological cue to gender category (see Table 3). These changes eliminated one cue to gender and left only the case marking. The test items were identical in both single marked and double marked conditions.

Procedure

The procedure was identical to that of Experiment 1.

---

[7] The total number of infants tested in the single and double marking conditions were the same (30 each). We decided that we would not continue testing infants in the single marking condition until we had 18 infants who met inclusion criteria, because even if both additional infants had longer looking times for ungrammatical test items, the grammaticality effect could not reach statistical significance. Further, the chi-square on infants from the two familiarization conditions showing the majority vs. minority pattern would still show a significant difference (6:12 vs. 14:4).

TABLE 3. *Stimuli used in the singly marked condition of Experiment 3*. Bolded words were withheld during familiarization and comprised the grammatical test items. Ungrammatical items were identical to those in the double marking condition: *vannya, korovyem, medevedoj, pisaru*

		Feminine words			
lapoj	malinoj	ruchkoj	**vannoj**	knigoj	korovoj
lapu	malinu	ruchku	vannu	knigu	**korovu**
		Masculine words			
tramvaya	iul'ya	zhitel'ya	**medved'ya**	korn'ya	pisar'ya
tramvayem	iul'yem	zhitel'yem	medved'yem	korn'yem	**pisar'yem**

RESULTS AND DISCUSSION

As in Experiments 1–2, each infant's looking times for the two versions of each trial type were averaged for each block, resulting in six looking times per infant (2 grammaticality × 3 block). Three infants from the Single Marking group and one from the Double Marking were missing a data point for one of the six blocks, and these were filled in with the mean for that variable. A (2) group (single marking vs. double marking) × (2) grammaticality × (3) block analysis of variance performed on the looking time data revealed a significant effect of block ($F(2, 64) = 3.45$, $p < 0.05$; Figure 2). The main effect of grammaticality was not reliable. There was a trend toward an interaction between group and grammaticality ($F(1, 32) = 2.64$, $p = 0.11$). As in Experiment 1, we also probed this interaction with a non-parametric test. A chi-square test comparing the number of positive vs. negative grammatical minus ungrammatical differences was highly significant ($\chi^2 = 9.47$, n = 34, $p = 0.002$). Returning to the results of the analysis of variance, the main effect of group did not approach significance, nor did any of the other interactions.

Because our examination of the group by grammaticality interaction suggested that the two groups showed different patterns of discrimination during test, and because we predicted that only infants in the Double Marking group would demonstrate an effect of grammaticality, separate (2) grammaticality × (3) block analyses of variance were carried out for each group. For infants who were familiarized with double marked stimuli, there was a significant effect of grammaticality ($F(1, 17) = 6.60$, $p < 0.02$). Neither the effect of block nor the interaction approached significance. For infants who were familiarized with stimuli with only a single cue to gender category, looking times for the same test items experienced by the Double Marking group revealed no effect of grammaticality ($F < 1$). Infants in the Single Marking group did demonstrate a marginal effect of block ($F(1, 15) = 2.68$, $p < 0.09$). The interaction was not significant.
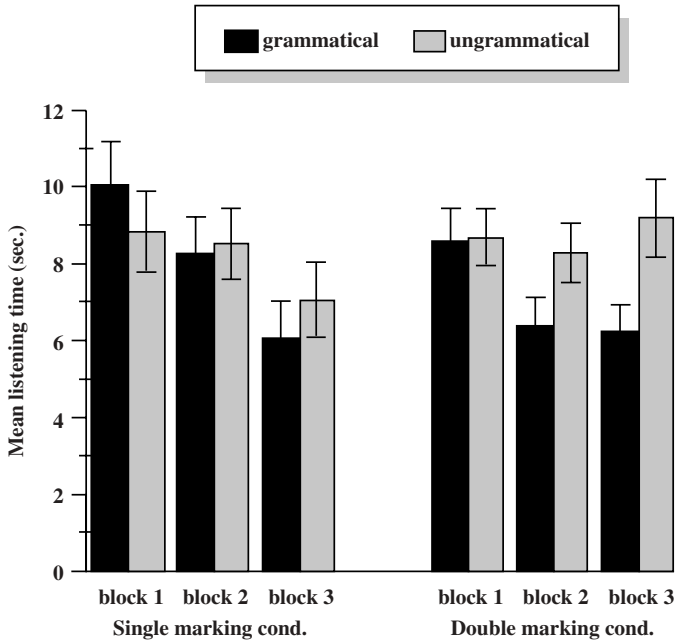
Fig. 2. Mean listening times and SE for grammatical vs. ungrammatical trials by block for Single Marking and Double Marking conditions of Experiment 3.

The results from the Double Marking condition of Experiment 3 strongly replicate the effect found only in the third block of Experiment 2. Although the effect of grammaticality was significant across blocks in the Double Marking condition of Experiment 3 (no grammaticality × block interaction), an examination of Fig. 2 reveals that the grammaticality effect grew stronger over blocks. Perhaps infants were able to consolidate the learning that began during familiarization over the three blocks of test trials. In contrast, infants in the Single Marking condition failed to discriminate the same test stimuli discriminated by the infants in the double marking group. Although the direction of the effect changed over blocks (longer listening times on grammatical trials in block 1 and on ungrammatical trials in block 3), the grammaticality effect did not approach significance in any block. The lack of even a hint of a grammaticality effect in the Single Marking condition demonstrates that infants, like adults in previous studies, are more likely to discern the category structure of a paradigm when more than one cue to the structure is available on a subset of the familiarization items.

TABLE 4. *Likelihood in written Russian corpora of the double gender markings used in Experiments 1–3*

Feminine noun markings			
Derivational marker given case marker		Case marker given derivational marker	
# -oj	7489	# k	37061
# -k given -oj	1508	# -oj given k	3359
likelihood of k given -oj	20%	likelihood of -oj given k	9%
# -u	5188	# k	37061
# k given -u	871	# u given k	1499
likelihood of k given -u	17%	likelihood of -u given k	4%
Masculine noun markings			
Derivational marker given case marker		Case marker given derivational marker	
# -ya	35314	# tel	2509
# -tel' given -ya	195	# -ya given tel'	127
likelihood of tel' given -ya	0·5%	likelihood of -ya given tel'	5%
# -em	2818	# tel	2509
# tel' given -em	60	# em given tel	265
likelihood of tel' given -ya	2%	likelihood of -em given tel'	10%

CONCLUSION

The results of Experiments 1–3 clearly show that infants in the second year of life are able to discern the gender category structure embodied in a morphophonological paradigm, provided that at least part of the paradigm has correlated cues to category. Although the familiarization stimuli used may have provided a near ideal learning situation, the cues were ones that actually exist in Russian. Therefore, it seems possible that infants could make use of such cues in a real language learning situation. To examine the cue value of the morphological markers we used in actual Russian, we used a programme to find instances of the inflections –*tel'* and -*k* and the case markers -*oj*, -*u*, -*ya*, and -*yem* on Russian language web sites. Two kinds of searches were conducted. First, given the total number of instances of the case-markings, a search was conducted to determine the proportion of those instances that also included the phonological marker. This search allowed us to calculate, for example, the probability of -*tel'* given -*ya*. The second search considered the question in reverse. That is, given the total number of instances the phonological markers, what is the proportion of those occurrences that are followed by the case markers? For example, what is probability of -*ya* given -*tel'*?'.

For these counts, the programme searched approximately 13540 Russian-language web sites and included approximately 1600000 lexical items. Each search is different because the World Wide Web is constantly in flux, rendering a constantly changing corpus. Some searches are larger than others because the initial seed site happened to yield web sites that

contained more valid links to other sites. Despite the differences in total number of pages searched, the two searches should be comparable since it is the percentages that are compared and not the raw counts. The results of the search are shown in Table 4. The highest instance of one type of marking, given another, was 20% (likelihood of -*k* given -*oj*), and the lowest was 0·5% (likelihood of -*tel'* given -*ya*). When considering the lowest percentage, it is important to note that these counts do not tell us how frequently Russian nouns are double marked for gender, because other phonological correlates of gender exist but were not included in the analysis. Both figures are lower than the 50% double marking used during familiarization in our studies. However, it appears that double marking is a relatively frequent component of Russian, especially when one takes into account both masculine and feminine diminutive endings (e.g. Kempe & Brooks, 2001; Kempe, Brooks, Mironova & Fedorova, 2003). A question for future research is whether the degree of double marking found in Russian corpora is sufficient for the sort of rapid learning we observed in Experiments 1–3.

Let us end with a discussion of what the current results might mean for a theory of language acquisition. It is clear that learners do not need referential cues to form proto-categories from paradigms. Rather, consistent with recent adult research, it appears that partially overlapping cues of a variety of kinds, including referential, phonological, morphological and positional, can be used in combination (also see Morgan, Meier & Newport, 1987, 1989). However, proponents of semantic bootstrapping have not argued that learners cannot discern categories from distributional information alone. Rather, they argue that referential information is required to solve the problem of linking distributionally based categories with innate syntactic categories like subject and noun (Grimshaw, 1981; Pinker, 1984). The general interpretation of Semantic Bootstrapping, however, is that referential information is used prior to distributional information. The current data argue against this particular version of Semantic Bootstrapping. Although researchers working in less strongly nativist traditions have argued that syntactic categories could be learned from the input, a full account of how syntactic categories could be derived from input has yet to be proposed. The current data suggest that, if such an account is forthcoming, it will be strengthened by the fact that young learners appear to be able to discern syntactic category structure from a natural language paradigm without a need for reference.

REFERENCES

Baker, M. C. (2001). *The atoms of language*. New York, NY: Basic Books.
Braine, M. D. S. (1966). Learning the positions of words relative to a marker element. *Journal of Experimental Psychology* **72**(4), 532–40.

Braine, M. D. S. (1987). What is learned in acquiring word classes – a step toward an acquisition theory. In B. MacWhinney (ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum, 1987.

Braine, M. D. S. (1992). What sort of innate structure is needed to 'bootstrap' into syntax? *Cognition* **45**(1), 77–100.

Brent, M. R. (1992). Automatic acquisition of subcategorization frames from unrestricted English. *Dissertation Abstracts International* **53**(2-B), 929.

Brooks, P. J., Braine, M. D., Catalano, L., Brody, R. & Sudhalter, V. (1993). Acquisition of gender-like noun subclasses in an artificial language: the contribution of phonological markers to learning. *Journal of Memory and Language* **32**, 76–95.

Cartwright, T. A. & Brent, M. R. (1997). Syntactic categorization in early language acquisition: formalizing the role of distributional analysis. *Cognition* **63**(2), 121–70.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Frigo, L. & McDonald, J. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language* **39**, 218–45.

Gerken, L. A. & McIntosh, B. J. (1993). Interplay of function morphemes and prosody in early language. *Developmental Psychology* **29**(3), 448–57.

Gerken, L. A., Wilson, R., Gómez, R. L. & Nurmsoo, E. (2002). *Linguistic category induction without reference: the importance of correlated cues*. Unpublished manuscript, University of Arizona.

Gleitman, L. & Wanner, E. (1982). The state of the state of the art. In L. Gleitman (ed.), *Language acquisition: the state of the art*. Cambridge, England: CUP.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science* **13**(5), 431–6.

Gómez, R. L. & Gerken, L. A. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition* **70**(2), 109–35.

Gómez, R. L. & Gerken, L. A. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences* **4**, 178–86.

Gómez, R. L. & Lakusta, L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science* **7**, 567–80.

Grimshaw, J. (1981). Form, function, and the language acquisition device. In J. J. McCarthy (ed.), *The logical problem of language acquisition*. Cambridge, MA: MIT Press.

Hoehle, B. & Weissenborn, J. (1999). Discovering grammar: prosodic and morpho-syntactic aspects of rule formation in first language acquistion. In A. D. Friederici & R. Menzel (eds), *Learning: rule extraction and representation*. Berlin, Germany Walter de Gruyter & Co, 1999.

Karmiloff-Smith, A. (1979). Micro- and macrodevelopmental changes in language acquisition and other representational systems. *Cognitive Science* **3**, 91–117.

Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A. & Gerken, L. A. (1995). The head-turn preference Procedure for testing auditory perception. *Infant Behavior and Development* **18**, 111–6.

Kempe, V. & Brooks, P. J. (2001). The role of diminutives in the acquisition of Russian gender: can elements of child-directed speech aid in learning morphology? *Language Learning* **51**(2), 221–56.

Kempe, V., Brooks, P. J., Mironova, N. & Fedorova, O. (2003). Diminutivization supports gender acquisition in Russian children. *Journal of Child Language* **30**, 471–85.

Levy, Y. (1983). It's frogs all the way down. *Cognition* **15**, 75–93.

Maratsos, M. & Chalkley, M. (1980). The internal language of children's syntax. In K. Nelson (ed.), *Children's language* (Vol. 2). New York: Gardener Press.

Marcus, G. F., Vijayan, S., Rao, S. B. & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science* **283**(5398), 77–80.

Mills, A. E. (1985). The acquisition of German. In D. I. Slobin (ed.), *The crosslinguistic study of language acquisition*. Hillsdale, NJ: Erlbaum.

Mintz, T. H. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition* **30**(5), 678–86.

Morgan, J. L., Meier, R. & Newport, E. (1987). Structural packaging in the input to language learning. *Cognitive Psychology* **22**, 498–550.

Morgan, J. L., Meier, R. P. & Newport, E. L. (1989). Facilitating the acquisition of syntax with cross-sentential cues to phrase structure. *Journal of Memory & Language* **28**, 360–74.

Naigles, L. (2002). Form is easy, meaning is hard: resolving a paradox in early child language. *Cognition* **86**, 157–99.

Petretic, P. A. & Tweney, R. D. (1977). Does comprehension precede production? The development of children's responses to telegraphic sentences of varying grammatical adequacy. *Journal of Child Language* **4**(2), 201–9.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* **274**(5294), 1926–8.

Shady, M. E. (1997). Infants' sensitivity to function morphemes. *Dissertation Abstracts International: Section B: The Sciences & Engineering* **58**(1-B), 0441.

Shady, M. E., Gerken, L. A. & Jusczyk, P. W. (1995). Some evidence of sensitivity to prosody and word order in ten-month-olds. In D. MacLaughlin & S. McEwan (eds), *Proceedings of the 19th Boston University Conference on Language Development* (Vol. 2). Somerville, MA: Cascadilla Press.

Shafer, V. L., Shucard, D. W., Shucard, J. L. & Gerken, L. A. (1998). An electro-physiological study of infants' sensitivity to the sound patterns of English speech. *Journal of Speech, Language & Hearing Research* **41**(4), 874–86.

Shi, R., Werker, J. & Morgan, J. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition* **72**, B11–21.

Shipley, E., Smith, C. & Gleitman, L. R. (1969). A study in the acquisition of language. *Language* **45**, 322–42.

Smith, K. H. (1966). Grammatical intrusions in the recall of structured letter pairs: mediated transfer or position learning? *Journal of Experimental Psychology* **72**, 580–8.

Townsend, C. E. (1968). *Russian word formation*. New York: McGraw Hill.

Wilson, R. (2002). *Syntactic category learning in a second language*. Unpublished PhD dissertation, University of Arizona, Tucson.

Zubin, D. A. & Kopcke, K. M. (1981). Gender: a less than arbitrary grammatical category. In R. A. Hendrick, C. S. Masek & M. F. Miller (eds), *Papers from the seventh regional meeting of the Chicago Linguistics Society*. Chicago: University of Chicago.