

Contextual Constraints on Noun Distributions to Some English Verbs by Children and Adults¹

ROBERT J. JARVELLA²

Case Western Reserve University, Cleveland, Ohio 44106

AND

JOAN SINNOTT

University of Michigan, Ann Arbor, Michigan 48104

The cloze procedure was used to obtain noun distributions for subjects and objects of 36 common English transitive verbs in active and passive sentence frames. For third and sixth grade and college students, the subject nouns selected were more human and animate than object nouns, linguistic selectional rules correctly predicted relative animacy, humanness, and informational uncertainty for each function, and subject-object interacted on all these properties with active-passive. Distributions obtained from adults contained more inanimate nouns (where permissible), more informational uncertainty, and had smaller intersections with each other than complementary distributions obtained from children.

A general conclusion of recent reviews of work in developmental psycholinguistics (e.g., McCarthy, 1954; McNeill, 1970) is that the grammar of English is largely acquired by children before age five. Vocabulary size, however, continues to grow at a substantial rate throughout the usual school years (Dolch, 1927), and some lexical growth is probably maintained throughout life (Riegel, 1965). As the child's lexicon expands, it also seems to become further differentiated and generalized (Riegel, Riegel, Quarterman & Smith, 1968; Anglin, 1970). However, it is reasonable to hypothesize that throughout the later process

of lexical expansion, representations for individual words in the child's vocabulary will be largely retained. The aim of the present research was to determine to what extent certain representations apparent in the adult lexicon of English are already present in the lexical systems of children 9-12 yr old. In this case, the representations concerned were some contextual features of English verbs inferred from nouns which were selected as their subjects and objects.

There are at least three sources of contextual constraint likely to influence the selection of noun subjects and objects by children as well as adults. Two of these are general grammatical restrictions and some related response biases. The grammatical constraints relevant here involve three aspects of Chomsky's (1965) description of English. First, subjects and objects are taken to have special status in the deep structure of sentences, but which is not necessarily maintained in their surface structure. Thus, the active and passive sentences *John loves girls* and *Girls are loved by John* have the same logical subject (*John*) and logical object (*girls*), but different superficial

¹ This paper was supported by Grant DAHC 15 to the Rockefeller University from the Advanced Research Projects Agency, and was prepared while the senior author was a Visitor at the Institute for Advanced Study, Princeton, NJ. The authors are particularly grateful to David B. Pisoni for his generous help in planning the experiment, and to Beth G. Greene and Ronald S. Tikofsky for collecting the data. George A. Miller is thanked for his helpful suggestions and criticisms.

² Requests for reprints should be sent to Robert J. Jarvella, Department of Psychology, Case Western Reserve University, Cleveland, OH 44106

subjects (*John* and *girls*, respectively). Second, nouns may be categorized on the basis of abstract features such as animacy and concreteness. Thus, *boy* is [+Human] (unlike *dog*), [+Animate] (unlike *chair*), and [+Concrete] (unlike *sincerity*). And third, English verbs have selectional rules containing these features which limit the scope of nouns which may be their logical subjects and objects. In the case of the large majority of verbs, subject nouns must be animate; and in many cases, object nouns only animate or inanimate (Hall, 1965). The verb *read*, for example, requires an animate subject but inanimate object, as [+Animate] Aux —, and — Det [-Animate] in Chomsky's notation; the string *The book read the boy* is therefore ungrammatical because both the verb's subject and object selectional rules are broken. Any mature speaker of English can be expected to honor its selectional rules in a close procedure task; in practice, these rules may also restrict the diversity of nouns selected by a group of speakers as subjects or objects for particular verbs.

Previous psychological research does indicate the operation of several related biases in English speakers' selection of nouns to serve as subjects and objects. It is believed that almost all subject nouns occurring in the earliest sentences of young children are animate, and that object nouns in these sentences are mainly inanimate (Brown, Cazden & Bellugi, 1969; Bloom, 1970). In a cloze procedure task (Taylor, 1953, 1956), Clark (1965) found that subject nouns filled into active and passive sentence frames by high school students also tended to be animate and object nouns to be inanimate. In further experiments, in which adult *Ss* rated how well sentences with subject nouns of different subcategories made sense, and in which *Ss* replaced subject nouns to improve sense, Clark and Begun (1971) found that human nouns were most preferred as subjects, and inanimate nouns among the least preferred. Two additional biases relevant here were also discovered by Clark

(1965). Nouns selected as logical subjects and objects were both found to be more animate when they occurred as surface structure subjects, and both were less diverse (as measured by informational uncertainty) in this role. The present study provides a test for developmental gradients along both subject-object and active-passive dimensions. In general, lexical growth would seem to imply greater diversity and greater uniqueness among adult than child response distributions. However, as in previous research, subjects should be found more animate (and more human) overall than objects, and subject-object and active-passive should interact on both noun subcategorization and response diversity measures.

The third likely source of contextual constraint on noun selection here is semantic information which is more specific to individual verbs. Most completions by adults of sentence frames giving a single content word (Rosenberg & Koen, 1968; Rosenberg, Jarvella & Pisoni, 1969) do seem to reflect either the linguistic meaning of the words given, or familiarity with their referents (Rosenberg & Jarvella, 1970). The cloze procedure also seems from these studies to be an effective method for eliciting much of the same contextual semantic information about specific lexical items from different *Ss*. In the present study, it is plausible that shared semantic information would be reflected by substantial intersections among certain pairs of response distributions obtained. For example, proportionally more responses should be shared by two subject noun distributions obtained for a verb than by a subject and an object noun distribution. Similarly, sizable intersections between complementary noun distributions for particular verbs by children and adults would suggest the presence of much of the same semantic information in the lexical system at these different developmental levels.

METHOD

Subjects

Forty-five third grade, 45 sixth grade, and 45 college students participated in the experiment, held near the

end of the school year. The children were pupils at a public elementary school in New York City, the adults were students taking a course on language acquisition at the University of Michigan. Approximately half of the *Ss* at each educational level were female.

Materials

The stimuli were active and passive sentence frames containing simple past tense forms of 36 common English transitive verbs, together with unfilled slots (20 typewritten spaces long) for both subject and object nouns. For all verbs, sentence frames looked like the following sample ones for the verb *thank*: *The ——— thanked the ———*, and *The ——— was thanked by the ———*. The 36 verbs employed are shown partitioned in Table 1 on the basis of apparent grammatical constraints on the animacy of nouns which may serve as their subjects and objects in simple sentences, although these selectional rules may sometimes be violated in metaphorical, nonstandard, or syntactically more complex usage. For example, life may be attributed to things generally inanimate. By the *Es*' judgment, those verbs in the upper half of Table 1 grammatically would usually require animate (and mainly human) nouns as their subjects, while verbs in the lower half of the table would permit at least some inanimate subjects. Subject noun responses to these two groups of verbs will be called +Animate-Subjects and ±Animate-Subjects, respectively. Similarly, those verbs in the left column of Table 1 seem to require animate nouns as their objects, while those in the center column permit inanimate (or animate) objects, and verbs in the right column usually require inanimate objects. Object noun responses to these groups of

verbs correspondingly will be called +Animate-Objects, ±Animate-Objects, and -Animate-Objects.

For each bloc of six verbs defined jointly by animacy rules for subjects and objects in Table 1, the upper three verbs shown have an average Lorge magazine count frequency of about 2000, as opposed to about 200 each for the lower three verbs. Perhaps because all verbs were quite common, however, they were not noticeably differentiated by frequency on any response measures used. Thorndike-Lorge frequency will not be further discussed as a variable.

Active and passive frames were used for all 36 verbs given in Table 1. The sentence frames were printed six to a page for the child *Ss* and 18 to a page for the adult *Ss*, and assembled into booklets. Each *S* completed one booklet containing a randomized order of all 36 verbs, with active and passive frames alternating every sentence. The experimental frames were preceded by two practice items, and active and passive frames for each verb were counterbalanced within each age group.

Procedure

The child *Ss* were read instructions aloud by their teachers. These were printed as part of adult *Ss*' booklets. The instructions were to make up some sentences by filling in the blank spaces in some "incomplete" sentences. The only restrictions made were that only one word be printed in each blank space, and that the words chosen make sense with the other words in the sentence. The teachers showed their pupils how to do the task by writing the two practice items on the blackboard, and demonstrating how they might be completed, these examples were explained on the front page of adult *Ss*' booklets. Finally, all *Ss* were told to open their booklets and begin filling in the blank spaces. They were given ample time to complete the task.

TABLE 1

VERBS USED AND THEIR SELECTIONAL RESTRICTIONS ON SUBJECT AND OBJECT NOUNS

Function	Animacy	Object		
	Rule	+	±	-
Subject	+	thank	want	read
		command	love	learn
		marry	kiss	design
		hire	attack	sing
		arrest	kick	rent
		punish	steal	repair
	±	fool	pull	make
		surprise	lift	open
		excite	push	break
		entertain	strike	bend
	scare	scratch	fasten	
	confuse	transport	plow	

RESULTS

At each educational level, a distribution of between 21 and 24 responses was obtained for each noun slot for all 36 verbs. Four different response measures were derived from these distributions. Two of these were the proportion of human nouns and the proportion of (nonhuman) animal nouns in each distribution. The third measure was an estimate of the distribution's diversity of nouns; this was its relative entropy (*R*), an informational uncertainty statistic:

$$R = - \frac{\sum_{i=1}^k p_i \log_2 p_i}{\log_2 N}$$

or the ratio of actual information to maximum possible information (Attneave, 1959). The range of R is from 0, when all noun responses in a distribution are the same, to 1, when they are all different. The fourth set of measurements obtained were proportions of response agreement between pairs of noun distributions for the same verb; these were calculated by dividing the number of exact matches between distributions by the smaller N , or maximum number of matches possible. In Tables 2 and 3, the first three measures described are shown averaged for subjects and

TABLE 2

LOGICAL SUBJECTS: MEAN PROPORTION HUMAN AND ANIMAL NOUNS AND RELATIVE ENTROPY FOR RESPONSE DISTRIBUTIONS ($N = 18$)

Measure	Animacy		Educational level		
	rule	Sentence	3	6	12+
Human	+	Active	.943	.930	.918
	+	Passive	.889	.936	.932
	±	Active	.708	.626	.534
	±	Passive	.561	.508	.472
Animal	+	Active	.055	.063	.072
	+	Passive	.067	.058	.063
	±	Active	.154	.140	.100
	±	Passive	.172	.150	.096
Relative entropy	+	Active	.553	.603	.686
	+	Passive	.633	.664	.723
	±	Active	.607	.682	.747
	±	Passive	.712	.764	.783

objects independently following the partition of verbs given in Table 1. In Table 4, the response agreement figures are shown averaged over all verbs.³ All statistical tests made were performed by verbs, and only those with usual significant probability levels are reported in the following. Unless otherwise noted, p levels shown were obtained using Friedman nonparametric analyses of variance (Bradley, 1968).

³ For the response distribution proportions for each of the 36 verbs, order document No. 01682 from ASIS National Auxiliary Publications Service, c/o CCM Information Sciences, Inc., 22 West 34th Street, New York, NY 10001, remitting \$1.00 for microfiche or \$3 00 for photocopies.

TABLE 3

LOGICAL OBJECTS: MEAN PROPORTION HUMAN AND ANIMAL NOUNS AND RELATIVE ENTROPY FOR RESPONSE DISTRIBUTIONS ($N = 12$)

Measure	Animacy		Educational level		
	rule	Sentence	3	6	12+
Human	+	Active	.833	.858	.909
	+	Passive	.874	.889	.936
	±	Active	.341	.353	.297
	±	Passive	.522	.519	.396
	-	Active	.043	.030	.003
	-	Passive	.128	.058	.004
Animal	+	Active	.107	.116	.069
	+	Passive	.087	.090	.042
	±	Active	.140	.139	.051
	±	Passive	.138	.112	.086
	-	Active	0	.011	0
	-	Passive	.019	.016	0
Relative entropy	+	Active	.650	.727	.792
	+	Passive	.611	.647	.732
	±	Active	.774	.770	.814
	±	Passive	.681	.684	.763
	-	Active	.617	.616	.657
	-	Passive	.624	.617	.612

Virtually all nouns chosen were concrete, about 80% of subject nouns and 50% of object nouns were animate, and 85-90% of these animate nouns were human. On noun sub-categorization measures, as predicted by selectional rules, more +Animate-Subjects were human and animate (human plus animal noun proportions) than ±Animate-Subjects; more +Animate-Objects were human and animate than ±Animate-Objects, and more

TABLE 4

MEAN PROPORTIONS FOR INTERSECTIONS BETWEEN PAIRS OF NOUN DISTRIBUTIONS ($N = 36$)

Type of intersection	Educational level		
	3	6	12+
Subject (Active \cap Passive)	.506	.473	.391
Object (Active \cap Passive)	.496	.499	.410
Preverb (Active \cap Passive)	.317	.213	.062
Postverb (Active \cap Passive)	.196	.148	.054
Active (Subject \cap Object)	.216	.143	.053
Passive (Subject \cap Object)	.213	.158	.052

of these were human and animate than -Animate-Objects ($p < .001$ in all cases by Wilcoxon rank sum tests). Subjects were found to be more human and animate than objects ($p < .025$), and subject-object interacted with active-passive ($p < .001$), so that, for both functions, there were slightly more human and animate nouns in the preverb than postverb position in sentences. Animacy, however, was found to decrease with increased age ($p < .001$). This decrease occurred for both \pm Animate-Subjects and \pm Animate-Objects ($p < .01$ for each).

On response diversity, as measured by R , +Animate-Subjects were, as predicted, less diverse than \pm Animate-Subjects, and +Animate-Objects and -Animate-Objects were together less diverse than \pm Animate-Objects ($p < .001$ and $p < .025$, respectively, by Wilcoxon rank sum tests). Subject-object was again found to interact with active-passive ($p < .001$), so that both functions were less diverse in the preverb than postverb position in sentences. R also increased with age ($p < .001$). This increase was significant with $p < .05$ for all classes of subjects and objects separately except -Animate-Objects. As anticipated, the entropy of distributions was thus somewhat greater for adults than for children.

Substantial proportions of agreement between active and passive frame distributions for subjects and objects separately were obtained for both children and adults; these were almost invariably greater than intersections between subjects and objects ($p < .001$). However, for all six sets of proportions shown in Table 4, significantly less agreement ($p < .001$) existed among adult than child responses; for all sets of proportions but the two within functions, less agreement ($p < .05$ by sign tests) existed among sixth than third grade responses. Although greater differentiation and uniqueness of individual distributions are thus supported with increased age, often the same responses were elicited from children as from adults. The average proportion of agreement between sixth grade and adult

responses was about .35 for each of the four noun distributions in sentence frames, while between sixth and third grade responses, it was about .50 in each case (all differences significant with $p < .001$ by sign tests).

DISCUSSION

The present findings may be divided for discussion into (1) effects which were relatively independent of age, and which suggest quite mature language use by the children tested, and (2) developmentally related effects which may serve to qualify this maturity in performance.

First, many of the results obtained suggest that children of the ages tested (9-12 years) are not appreciably less sensitive than adults to relatively abstract constraints on subjects and objects in English. Briefly, all predictions based directly on, or ultimately derived from Chomsky's (1965) description of selectional rules, were strongly confirmed for both children and adults. Almost all nouns selected contained animacy features suggested appropriate by the partition of verbs given in Table 1. Subjects without obvious selectional constraints were also distinguishably more animate than unrestricted objects. However, as predicted, restricted distributions for both functions contained less informational uncertainty than unrestricted distributions. Thus, linguistic selectional rules on potential subjects and objects of verbs appear to have measurable consequences for the diversity of even relatively small samples of cloze responses.

Second, all groups of Ss tended to use somewhat smaller and more animate sets of nouns for both logical functions when they were superficial subjects than when they were not. These function-by-voice interactions very likely reflect a known bias for speakers of English to superimpose features of the logical subject function on preverb nouns, both in speech production (Clark, 1965; Goldman-Eisler & Cohen, 1970) and perception (Fraser, Bellugi & Brown, 1963; Bever,

1970). However, in the present study, the interaction obtained for animacy was limited to less than half the percentage difference previously obtained by Clark (1965) for similar sentence frames not containing verbs; this reduction is probably due both to selectional rules and added constraints on left-to-right selection of words. Finally, from comparisons of agreement within and between noun functions, it is obvious that children as well as adults largely distinguished subjects from objects for particular verbs. Since children's responses matched those of adults for corresponding distributions about one third of the time, it can be concluded moreover that they also showed sensitivity to other, more specific, semantic information about these verbs.

On the other hand, developmental gradients are indicated in some respect for most of the variables measured. With increased age, response diversity within distributions increased, and agreement between distributions decreased. Both findings would be expected if, along with the addition of vocabulary, lexical growth included some elaboration or refinement of the contextual semantic structure of particular words after their original acquisition. This developmental hypothesis seems to be specifically supported by an additional finding: With increased age, average agreement between noun functions decreased at a much faster rate than it decreased within functions. While quite substantial subject-with-object intersections were obtained among child responses for most verbs, the corresponding intersections for adults were usually quite negligible. This disparity suggests that lexical representations for subjects and objects of particular verbs become further differentiated from each other well into adolescence.

Child and adult performance also diverged in two (perhaps interrelated) ways not so clearly implied by progressive differentiation of the lexicon. First, there was an approximately 50% reduction with increased age in the number of animal name nouns used as

both subjects and objects; while children supplied some animal names as subjects or objects (or both) for almost all verbs, adults used animal names almost exclusively with verbs referring to actions customarily involving animals as well as people (*scratch, kick, pull, and attack*). The relatively prominent place of animals in the daily lives of children (e.g., as their pets or fictional heroes) may partly explain this developmental difference. Second, children used about 50% fewer inanimate nouns than adults as logical subjects. It appears from recent linguistic work on the semantics of verb-noun relations (e.g., Lakoff, 1968; Fillmore, 1968a, 1968b; Chafe, 1970) that inanimate subjects do not (or can not) typically perform agentive semantic functions in sentences. For example, the subject of the sentence *The knife sliced the salami* is the instrument used by someone to do the cutting. The proliferation of inanimate noun subjects with increased age may consequently reflect a trend away from making subjects semantic actors. This implies that, in addition to being more lexically diverse, adult responses in sentence frames also held more diverse semantic roles with respect to the actions described by verbs.

REFERENCES

- ANGLIN, J. M. *The growth of word meaning*. Cambridge: M.I.T. Press, 1970.
- ATTNEAVE, F. *Applications of information theory to psychology*. New York: Holt, 1959.
- BEVER, T. G. The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Wiley, 1970.
- BLOOM, L. *Language development*. Cambridge: M.I.T. Press, 1970.
- BRADLEY, J. V. *Distribution-free statistical tests*. Englewood Cliffs: Prentice-Hall, 1968.
- BROWN, R., CAZDEN, C., & BELLUGI, U. The child's grammar from I to III. In J. P. Hill (Ed.), *Minnesota symposium on child psychology*. Minneapolis: University of Minnesota Press, 1969.
- CHAFE, W. L. *Meaning and the structure of language*. Chicago: University of Chicago Press, 1970.
- CHOMSKY, N. *Aspects of the theory of syntax*. Cambridge: M.I.T. Press, 1965.

- CLARK, H. H. Some structural properties of simple active and passive sentences. *Journal of Verbal Learning and Verbal Behavior*, 1965, **4**, 365-370.
- CLARK, H. H., & BEGUN, J. S. The semantics of sentence subjects. *Language and Speech*, 1971, **14**, 34-46.
- DOLCH, W. E. Grade vocabularies. *Journal of Educational Research*, 1927, **16**, 16-26.
- FILLMORE, C. J. Lexical entries for verbs. *Foundations of Language*, 1968, **4**, 373-393. (a)
- FILLMORE, C. J. The case for case. In E. Bach & R. T. Harms (Eds.), *Universals in linguistic theory*. New York: Holt, 1968. (b)
- FRASER, C., BELLUGI, U., & BROWN, R. Control of grammar in imitation, comprehension, and production. *Journal of Verbal Learning and Verbal Behavior*, 1963, **2**, 121-135.
- GOLDMAN-EISLER, F., & COHEN, M. Is N, P, and PN difficulty a valid criterion for transformational operations? *Journal of Verbal Learning and Verbal Behavior*, 1970, **9**, 161-166.
- HALL, B. C. Subject and object in modern English. Unpublished Doctoral dissertation, Massachusetts Institute of Technology, 1965.
- LAKOFF, G. Instrumental adverbs and the concept of deep structure. *Foundations of Language*, 1968, **4**, 4-29.
- MCCARTHY, D. Language development in children. In L. Carmichael (Ed.), *Manual of child psychology*. (2nd ed.) New York: Wiley, 1954.
- MCNEILL, D. The development of language. In P. H. Mussen (Ed.), *Carmichael's manual of child psychology*, I (3rd ed.) New York: Wiley, 1970
- RIEGEL, K. F. Speed of verbal performance as a function of age and set: A review of issues and data. In A. T. Welford & T. E. Birren (Eds.), *Behavior, aging, and the nervous system*. Springfield: Thomas, 1965.
- RIEGEL, K. F., RIEGEL, R. M., QUARTERMAN, C. J., & SMITH, H. E. Developmental differences in word meaning and semantic structure. *Human Development*, 1968, **11**, 92-106
- ROSENBERG, S., & JARVELLA, R. J. Semantic integration as a variable in sentence perception, memory, and production. In G. B. Flores d'Arcais & W. J. M. Levelt (Eds.), *Advances in psycholinguistics*. The Netherlands: North Holland Publ. Co., 1970
- ROSENBERG, S., JARVELLA, R. J., & PISONI, D. B. Transformational effects upon the strength and content of semantic constraints in sentences. In S. Rosenberg & A. P. Van Tessaar (Eds.), *Studies in language and language behavior*, Progress Report VIII. Ann Arbor: University of Michigan, Center for Research on Language and Language Behavior, 1969.
- ROSENBERG, S., & KOEN, M. J. Norms of sequential associative dependencies in active declarative sentences. In J. C. Catford (Ed.), *Studies in language and language behavior*, Supplement to Progress Report VI. Ann Arbor: University of Michigan, Center for Research on Language and Language Behavior, 1968.
- TAYLOR, W. L. Cloze procedure. A new tool for measuring readability. *Journalism Quarterly*, 1953, **30**, 415-433.
- TAYLOR, W. L. Recent developments in the use of cloze procedure. *Journalism Quarterly*, 1956, **33**, 42-48.

(Received July 7, 1971)