Model learning in syntactic development: Intransitive verbs [*]

Anat Ninio

The Hebrew University of Jerusalem

Journal of Bilingualism, 1999, in press.

Address for correspondence: Anat Ninio, Department of Psychology, The Hebrew University, Jerusalem 91905, Israel. E-mail address: msninio@mscc.huji.ac.il. Address for correspondence during 1 October 1998 - 30 September 1999: Anat Ninio, Department of Psychology, Macquarie University, Sydney, New South Wales 2109, Australia. E-mail address: aninio@bunyip.bhs.mq.edu.au.

Short title: **MODEL LEARNING**

ABSTRACT

Previous findings suggest that in natural language, frequency may be inherently correlated with generality, prototypicality or simplicity. The most frequently used instances in each domain may be the "best models" of that domain, so that the effect of input frequency is to have children learn the best exemplars of a domain first. To test this hypothesis, the first intransitive verbs to participate in multiword combinations were investigated, based on longitudinal observations of 20 children acquiring Hebrew. The results were compared to the frequency of intransitive verbs in multiword utterances in a speech corpus based on recordings of 48 mothers. The first verbs that begin the acquisition of intransitive word-combinations are the most frequent intransitive verbs in the input. These verbs also tend to be relatively non-specific, general or generic intransitive verbs with light semantics. The same verbs that children first combine with subjects or locative adjuncts, and which are the most frequent in the speech of mothers, are typical grammaticalized markers of intransitivity or directional-movement in many languages. These results join many previous findings in showing that the statistical structure of natural languages makes possible, at the onset of acquisition, a kind of lexical learning which is at the same time a type of category- or rule-formation.

Model learning in syntactic development: Intransitive verbs

INTRODUCTION

Model-learning

The goal of this paper is to argue for a novel conceptualization of the learning process in language acquisition. According to this proposal, the learning procedure is neither purely lexical learning nor purely inductive rule-formation (or concept-formation) but rather, a hybrid of the two. The principle is that in any linguistic domain, field or paradigm -- whether of lexemes within a given semantic field, of bound morphemes within a paradigm, or of words belonging to the same form-class, in a given syntactic construction -- the first instances of the domain learned will be the items which are the best representations of the domain as a whole. These will be the most general, prototypical, central, members of the domain, the least complex and the least specific. These "best instances" of the linguistic domain are -- in computational terms -- its MODELS; and the hypothesized learning process is, therefore, one of model-learning at the initial stages of acquiring a novel kind of linguistic knowledge.

A model is a good schematic approximation to the kind of "object" we are dealing with. It does not, of course, represent all the intricacies of all the different instances of the "object", and in general, a model is less elaborate than the modeled objects, but it does capture the fundamental characteristics of the latter. Most importantly, a model is the same kind of object as the things it represents, so a model apartment is still an apartment, not a brochure or an advertizing video clip. It has the same logical status as the things it models; it is not, in linguistics or elsewhere, a definition of the class, nor is it any kind of superordinate category designation.

The learning of these "models" or "best instances" is almost equivalent to learning the rule

or the concept on the abstract; the acquisition of all further exemplars within the same domain is a question more of specification of existing knowledge than of independent learning. Or, in another formulation, the first items provide a great deal of facilitation for the learning of other instances; to learn the best items is to learn the "principle" of the domain. The rule or concept is not INDUCED from concrete examples individually learned, but rather, it is INHERITED by later-acquired items from the most prototypical, model exemplars which were learned initially.

Although we shall focus on the working of this learning principle in the domain of the acquisition of syntax, the principle is perhaps best illustrated by an example from vocabulary acquisition. It is a well-established research finding that the first size adjectives acquired by children are *big* and *little*. Other terms from the same semantic field like *tall/short, high/low, long/short,* and *wide/narrow* are acquired later (e.g., Bartlett, 1976; Clark, 1972; Eilers, Oller & Ellington, 1974). One explanation for the order of acquisition used by, e.g. Ingram (1989, p.418) is that of semantic complexity: he cites Bierwisch (1967) as claiming that *big* and *little* are the least complex size terms because they are global and not limited to a specific dimension; by comparison, *tall/short,* for example, are more complex because they are restricted to the dimension of height. Although this explanation is compatible with the present proposal, it seems to miss an important feature of *big* and *little*: they can be and indeed are used as pro-words for size differences along all the specific dimensions listed above. They are not just more GLOBAL words than the others, they are more GENERAL words. This point, and the developmental phenomenon that goes with it, is captured beautifully by Campbell and Wales:

> "In Edinburgh we have obtained some data on the development of a sub-system of English adjectives, the adjectives used to describe variations of size (i.e. *big, tall, fat,* etc., and

> their antonyms). The adjective *big* stands in a peculiar semantic relation to the others. Although the relation between each of *tall, fat, long* etc. and *big* is not strictly one of hyponymy it has similar properties, since the range of application of these terms is smaller than that of *big* and the range of application of *big* overlaps with each of their ranges. Now we have observed an interesting progression in the application of these adjectives with a number of children (too small, unfortunately, on which to base hard and fast conclusions). Initially, *big* (or its antonym *wee*) is used with reference to almost all differences of size. As the other more specialized adjectives are learned, however, *big* may fall out of use or may be restricted to cases of complex differences in size (e.g. to cases where the objects being compared both vary along two or more dimensions). (1970, p.259).

The present approach to the first-acquired model items of a domain is extremely close in spirit to Campbell and Wales' analysis. Although the first-acquired models or general items are not superordinate terms labelling the class as a whole, and the later-acquired, specific items are not strictly speaking their hyponyms, nevertheless, there is an overlap in the "range of application" (whatever that may mean in the different linguistic domains) of the model items with that of each of the more specific ones. This overlap is what makes the hypothesized transference of knowledge, generalization, facilitation, or inheritance of properties from the initial items to all later items possible.

The fact that such facilitation indeed takes place was demonstrated in the syntactic domain. In a recent study (Ninio 1996, 1999), it was found that children acquiring Hebrew and English begin verb-object (VO) and subject-verb-object (SVO) constructions with one or two verbs such as *want, make, do, get, give* or *take*, that remain for a month or two the only

verbs combining in the relevant pattern. These are all rather general, "all-purpose", or semantically light transitive verbs, and, as it will be shown presently, they are indeed the models for verbs combining in transitive verb-object or SVO combinations. Only later are the constructions applied also to other -- non-model -- transitive verbs like *read, hit, lock, say* and so on. When the temporal parameters of the spread of the VO and the SVO patterns over the first 20 different verbs was estimated, both graphs showed a rising exponential or geometrical function, starting very slowly and accelerating gradually. The graphs have the characteristic shape of typical learning curves: The time it takes to apply the new rule to yet another verb is much longer at the beginning of acquisition of that rule; the more verbs the children have already learned to produce in the relevant pattern, the shorter it becomes. Moreover, most of the massive facilitation was provided by the first and second verbs in each type of word-combination, and even the third verb already had a very small additional facilitating effect. This implies that most of the general or abstract knowledge about the VO and the SVO positional patterns is acquired in the context of the first two combining verbs in each pattern.

The results of this study suggested that children learn new combinatorial rules first for a few verbs in a piece-meal way, but immediately begin transferring some more general and abstract principle to other verbs so that applying the same combinatory principle to new verbs becomes progressively easier. This process is equivalent to the gradual consolidation of an abstract grammatical relation such as the verb-object relation, as well as to the consolidation of a similarity-class of verbs to which the relevant principle applies, namely, a lexical form-class which is relative to, and specific to, the syntactic rule applying to its members. These results strongly suggest the surprising conclusion that breaking into a new syntactic combination means solving the conceptual problems associated with that pattern once and for

all. This puts a heavy burden on the first two concrete instances of the pattern -- in this case the first two combining verbs -- as they should provide strongly characteristic, typical, and representative instances of the relevant combinatorial pattern. The proposed process of language acquisition claims that the first two concrete instances will indeed be models of the domain, and in this case, model transitive verbs.

The first two transitive verbs in combination are model transitive verbs

There is independent evidence that the first-acquired transitive verbs are indeed model transitive verbs, namely, crosslinguistic grammaticalization evidence. In many different languages and language-families of the world, the transitivity construct has been grammaticalized. That is, the language utilizes some closed-class or bound-morpheme verbs to express a transitive action-object relation in the required context. These are auxiliary verbs, or bound morphemes derived from verb-stems, serving a transitivity-establishing, defining or enhancing function. They surface as typical verb-transitivizers, bound-verb verbalizers on nouns, extension-markers and so forth, namely, verbs serving to establish the objective grammatical relation, or to define or re-define the argument-structure of other verbs as transitive. In many languages these verbs undergo reanalysis as grammatical morphemes, with the "semantic bleaching" typical of grammaticalization, and thus come to represent "pure transitivity" rather than their own specific semantics.

The relevant verb stems in different languages are remarkable similar. Over and over, the verbs *want, make/do, have, take, give, get, bring, hold, put, see, consume*, and so on, are used as transitivity-defining morphemes (Blake, 1994; Foley & Olson, 1985; Givon, 1991; Lord, 1982; Traugott & Heine, 1991, p.7-8; for a detailed exposition see Ninio, 1999). The grammaticalization evidence suggests that these verbs express a fundamental, unalienable, core

notion of transitivity, one that can be made use of when such a semantic-syntactic component is needed in a clause that otherwise does not possess it. These are, as we have seen, precisely the verbs with which young children begin the production of transitive VO and SVO combinations. It follows that the syntactic construction of transitive verb-object relations begins, in children acquiring English and Hebrew, with verbs which are literally model transitive verbs, in the sense in which a truly good boy-scout is said to be a "model boy-scout".

We may further examine the phenomenon of grammaticalized transitivity markers to understand what makes a linguistic item a good model of its domain. In many languages, the transitivity-inducing verbs have undergone full grammaticalization into adpositions, casemarkers, particles or bound verbal affixes (e.g. Blake, 1994; Givon, 1975; Lord, 1982). In other languages, they are intermediate between full verbs and auxiliaries (e.g. Hindi "vector verbs", Hopper & Traugott, 1993, p.109). The grammaticalization phenomenon suggests that the concept of transitivity takes up a larger proportion of the meaning of these verbs than in other transitive verbs, and that they have fewer unique semantic features -- besides the shared meaning of a transitive act on some object -- than other verbs. Grammaticalized lexemes have to undergo "semantic bleaching" in order to be applicable to a wide range of different contexts, and the more specific the original semantics of a verb, the less likely that it would be a good candidate for grammaticalization. This means that the verbs we are concerned with tend to have extremely general meanings; most of their semantics consists of some schematic notion of transitivity, with the addition of a minimal specific element. This is what makes them into, relatively speaking, the most GENERIC transitive verbs available in languages: Among all other verbs, they represent in the purest way the core notion of syntactic transitivity. This

generic feature is of course only relative, because these verbs do have specific semantics; moreover, it is well documented that they tend to keep some of their non-generic semantics even under grammaticalization, a phenomenon known as "persistence" in the literature (e.g. Hopper & Traugott, 1993). In general, these model verbs may be simpler or more general than other verbs but they are -- unless grammaticalized to the limit, into bound affixes -- still verb-lexemes just like any other verbs. This is true in particular in languages like Hebrew and English where these model transitive verbs are only semi-grammaticalized, where all of them remain free lexemes. However, even in these languages there are grammaticalized uses for these verbs, such as *do* serving as an auxiliary and pro-verb, and *take*, *get*, *make* and so forth serving as pivots of idiomatic verb-object constructions (e.g. Quirk, Greenbaum, Leech, & Svartvik, 1972, p.817). To summarize, language-internal as well as crosslinguistic grammaticalization evidence shows that the transitive verbs with which children begin VO and SVO constructions are model transitive verbs.

Other cases of model-first learning

There have been many other reports in the literature of developmental ordering that seems to represent the same phenomenon of "models first" within a linguistic domain. Researchers have found that development proceeds from the less complex items to the more complex (semantically, cognitively); from prototypical instances to more marginal ones; from the more general, unmarked, nonspecific forms of expression like proforms (pronouns, deictic locatives, pro-verbs and the like) to the use of specific terms (such as common nouns and verbs) in the same communicative circumstances. Some examples are English grammatical morphemes (Brown, 1973); color terms (De Villiers, 1980, p. 6; Heider, 1971; Johnson, 1977); syntactic sentence types like simple, active, affirmative, declarative sentences vs. questions, passives and

negations (Brown & Hanlon, 1970); wh-questions (Forner, 1979); and types of expressions for specific communicative intents such as requests (cf. Barrett, 1981; Bloom, 1973; Clark, 1978; Greenfield & Smith, 1976; Griffiths, 1985; Halliday, 1975; Ninio, 1993; Weisenberger, 1976).

The generality of these findings lends support to the hypothesis formulated in this Introduction to the effect that model-learning is a quite general phenomenon that regularly occurs in different linguistic domains. Before we test this hypothesis in the domain of intransitive verbs in multiword combinations, it is worth asking the self-evident question, why would children learn model instances first, before less model-like, more specific instances of the same domains?

The characteristics of models

Let us start with the case of transitive verbs and the VO and SVO syntactic combinations. We have seen that the first verbs found in these combinations are verbs like *want, get, give, take, do, make* and so on. There are two ways to characterize the set of verbs initiating the VO and SVO transitive patterns. On the one hand, these verbs are among the most general transitive verbs in the language, with very little specific semantics. In English and Hebrew, they are the verb stems closest to generic transitive verbs. On the other hand, these verbs appear to be extremely frequent transitive verbs, perhaps the most frequent in parental speech which serves as linguistic input to the acquisition process.

At first glance, the two features seem like competing alternative explanations as to why these verbs are the earliest ones acquired as combining transitive verbs. A cognitively-oriented approach (like prototypicality theory, Rosch, 1975) would say that semantically light verbs, whose meaning is almost like that of "logical predicates" for transitivity (Lehmann,

1991, p.231) would be the most salient verbs in the input to combine with direct objects. Perhaps the notion of simplicity would be evoked as well: semantically light verbs have little in the way of unique semantics, so they are simpler, semantically, than other verbs with the same syntactic features but with "heavier", more specific semantics.

A learning-theory oriented approach would say that input frequency, a major learning-theoretical factor, is sufficient in itself to explain the order of acquisition, regardless of the distinctive standing of these transitive verbs among all other transitive verbs on a dimension of complexity, generality or any other feature of content.

However, if the statistical structure of natural language is such that as a rule, the most frequent instances in a domain are also the most general, central, prototypical and model-like, there is no logical means of separating the two kinds of effects on the order of development. Rather, there is a "collusion" of simplicity, generality, prototypicality, and frequency, to increase the likelihood that the first instances acquired from each linguistic domain will in fact be its best models.

There are very good chances that such a frequency-prototypicality correlation indeed exists in natural languages. First, this correlation has been explicitly pointed out for some near-synonyms of the term "model"; for example, it has been claimed that prototypes (Rosch, 1975) or "nuclear" items (Dixon, 1982, p.73, p.122) are the most frequent instances of a category. It should be noted that the field is not very comfortable with this correlation; Dixon, for example, says:

> "A nuclear verb normally has a greater frequency of occurrence than its non-nuclear hyponyms. This is not however a criterion of nuclearity; it is rather an as yet improperly understood empirical consequence of it." (p.122)

Second, the fact that the phenomenon of model instances is connected to grammaticalization shows that model exemplars are by necessity very frequent lexemes. In the first place, grammaticalized uses increase the frequency of the relevant lexemes in speech. As it was pointed out above, grammaticalization of model lexemes is not confined to "exotic" languages: for example, English and Hebrew use model transitive verbs such as *do*, *take*, *get*, *make* as auxiliaries, pro-verbs, and pivots of idiomatic verb-object constructions. In the second place, frequency is a condition for grammaticalization. As Traugott and Heine (1991, p.9) say:

> "Given that a form A is a candidate for grammaticalization both because of its semantic content and its salience, a further condition has to apply for grammaticalization to take place: The form has to be used frequently. The more grammaticalized a form, the more frequent it is (Bybee at al., Givon on serial verbs, Hook). The seeds of grammaticalization are therefore in a correlated set of phenomena: Semantic suitability, salience, and frequency. Only the third actually leads to grammaticalization and hence to fixing, freezing, idiomatization, etc."

Thus, in the grammaticalization literature the correlation of model content and frequency is a well-motivated phenomenon that has an explicit role in the theory. We may conclude that in natural languages, the prototypical or model-like instances in any domain tend to be the most frequently used instances of that domain.

Confounding of the effect of simplicity and input frequency in developmental ordering

The statistical structure of natural languages may explain the fact that the effect of input simplicity or prototypicality has been apparently confounded with the effect of input frequency in many developmental studies besides that of order of acquisition of transitive verbs in VO

and SVO combinations (Ninio, 1996, 1999). A few examples follow.

1. Size-adjectives:

Blewitt (1982) pointed out that although it is true that the relative semantic complexity of size-adjectives such as *big/little* predicts the order of their acquisition, "...relative frequency also could account for the observed order of acquisition..." (p.169). Blewitt based this observation on a published data-base of adult-adult speech, where the relative frequency with which the different size-terms were used matched their relative complexity. Other investigators used the frequencies in the more directly related child-addressed speech by parents and found the same confounding.

2. Sentence-types by derivational complexity:

Brown and Hanlon (1970) checked the frequency in the parental input of simplex structures (simple, active, affirmative, declarative) and of more complex structures such as yes/no questions, negations, tag-questions, and so on. They point out that although frequency is not a cognitive variable (p.37), "clearly, frequency and derivational complexity are closely related" (p.39).

3. English grammatical morphemes:

Forner (1979) and Moerk (1980) reanalyzed Brown's (1973) data on the input frequency of these morphemes and they found that the order of acquisition can be predicted not only by their linguistic complexity but equally well by their input frequency (see also Wolff, 1988, p.201).

4. Wh-questions:

Forner (1979) found that both input frequency and cognitive complexity can predict the order of emergence of different wh-questions, typically *what* > *where* > *who* > *how* > *why* >

when. Her data cover both German and Serbo-Croatian.

5. Prototypical core of word-classes:

Berman (1988, p. 49; see also De Villiers, 1980; Slobin, 1987) pointed out that acquisition of the major word-classes starts with a prototypical core and the less central stems are learned only later. For instance, count nouns initiate the acquisition of the noun-class; action verbs, that of the verb class, and verb related adjectives, that of the adjective class. In all these cases, however, the more prototypical and earlier acquired kind of words are also more frequent in the input.

In summary, the literature has repeatedly pointed out the confounding of content-factors like simplicity and prototypicality and of input frequency on the order of acquisition of linguistic knowledge belonging to the various domains of language: syntax, vocabulary, and morphology. Perhaps confounding is not the correct term. Given the statistical structure of language in which frequency and model-like features are correlated, it is expected that both, together, influence the order of acquisition. Apparently, the very frequent, general and prototypical instances that serve as good models of a domain are the optimal instances to learn initially, when starting to learn the domain for the first time.

<u>Acquisition and model intransitive verbs</u>

In the present study, this hypothesis will be tested in yet another domain -- that of intransitive verbs in multiword combinations. We predict that the most frequently modelled combining intransitive verbs in the linguistic input provided by mothers will be the same verbs as the intransitive verbs first produced in word-combinations by children, and that these verbs will be the translational equivalents of frequently-grammaticalized verb-stems.

Intransitive verbs are usually considered to fall into two major semantic-syntactic

categories. The first subtype is variously called active, dynamic, voluntary, "unergative" or agentive intransitive verbs; the second subtype are eventive, static, involuntary, inactive, patientive, "unaccusative" intransitive verbs. Because the two subtypes are so divergent semantically (and perhaps also syntactically, in most languages), it is expected that each subtype would receive at least one "model" of its own in acquisition.

Crosslinguistic evidence suggests that the intransitive verb stems most likely to undergo grammaticalization (and therefore to be the closest to "model" intransitive verbs for the two subtypes) are, on the one hand, the active stems *come* and *go*, and, on the other hand, the static stem *fall*. The equivalents of one or the other of these verbs typically serve as grammaticalized intransitive auxiliaries e.g. in Chinook (Silverstein, 1972); generalized prepositions, e.g. in Kwakiutl (Boas, 1947); locative extension-markers, e.g. in Mandarin (Li & Thompson, 1981); or bases for compound verbs, e.g. in Ungarinyin (a non-Pama Nyungan language of North Australia (Dixon, 1980). Foley and Olson (1985) explicitly point out the privileged position of *come* and *go* for grammaticalization: "Of all verbs, the most favored verbs for serializing constructions are the basic active intransitive motion verbs, *come* and *go*." (p.41). *Fall* is also special. Eventive, static, involuntary-motion verbs tend not to be grammaticalized, nor to serve as the basis to idiomatic compounds. The exception is *fall*, which stands out as the only verb of its kind in several different contexts of grammaticalization. For example, this is the only involuntary-motion, intransitive verb that serves as a copular verb in English, as in "*he fell sick*" (Quirk, Greenbaum, Leech, & Svartvik, 1972, p.821); and it is the only inactive verb -- among many transitive verbs and some active intransitives -- that serves as an obligatorily compounded verb-base in the North Kimberley group of languages, generating compound verbs similar to *fall in love* and the like (Dixon,

1980, p.427). Hebrew, too, has many compound verbs on the basis of *nafal* 'fall', compared, for example, to the very few or nonexistent ones on the basis of *hexlik* 'slip' or *hitpocec* 'burst'.

Besides these three, *come, go,* and *fall,* there are other intransitive verb-stems that tend to undergo grammaticalization, and in particular other active movement-verbs such as *move, stay* (or *sit), stand, lie* (or *sleep)* (see for example Dixon, 1980, p.427; Foley and Olson, 1985; Givon, 1991; Bybee, Perkins and Pagliuca, 1994, p.128). As it is difficult to quantify the crosslinguistic tendency to be grammaticalized, and as there are no other, more precise indicators of the status of "model" intransitive verbs, we shall opt for the cautious hypothesis that among the active, dynamic, "unergative" or agentive intransitive verbs, *come* and *go* are the closest to models (compared, for instance, to *ascend, descend, enter, exit, jump, run, fly, swim* and so on); and among the static, non-active, involuntary, event-like, patientive, "unaccusative" intransitive verbs, *fall* is the closest to a model. Others of this group -- *burst, broke, spilled, rolled* and *spoiled* -- are less model-like, less generic, more specific. We thus predict that the most frequently modelled combining intransitive verbs in the linguistic input will be the same verbs as the intransitive verbs first produced in word-combinations by children, and that, more specifically, *come, go* and *fall* will be more frequent in the input and earlier acquired than the less model-like stems from the same sub-class of intransitive verbs.

As to the other, less frequently, grammaticalized verbs -- *move, stay* or *sit, stand, lie* or *sleep* -- we expect them still to be relatively frequently used in the input, and to be relatively often among the earliest verbs learned in combination, but less often than the apparent favorites *come, go* and *fall*.

A preliminary study indicated that these verbs may indeed be acquired very early. In a cross-sectional sample of eighty 18-month-old girls video-recorded for 20 minutes interacting

with their parents, the intransitive verb-subject combinations produced by the largest number of children were of the roots *go* and *fall* (Ninio, 1996). These data are suggestive but not conclusive, as there is a great deal of individual difference in the rate of development around this age, and in the absence of longitudinal data it is unknown whether the verbs used by this sample at 18 months were the earliest acquired in combination for each child, or later acquisitions.

In the present paper, we will present detailed data from a longitudinal corpus taken from twenty children acquiring Hebrew as their first language. In addition, the speech of 48 Hebrew-speaking mothers talking to young children will be analyzed. A few words on Hebrew are in order. Modern Hebrew is a nominative/accusative language with a basic SVO word-order which is, however, much less rigid than that of English. Verbs are crossmarked for the person, gender and number of the subject in most tenses, and for gender and number in the present tense. Hebrew is a partially subject pro-drop language, basically preferring not to express the subject by free pronouns whenever it is clearly crossmarked on the verb. In addition, it is much more acceptable than in English to drop lexical object-phrases whenever the contextual conditions permit. For example, it is quite normative in colloquial speech to say *kax!* ' take!' , when handing over something to the listener. Third, the direct object noun, if definite, is preceded by the preposition *et* (ACC). Beyond these differences and a few others less relevant to the present topic, Hebrew is typologically fairly similar to English as far as its syntax is concerned.

Morphologically however, there is a important feature particular to Hebrew and other Semitic languages. All Hebrew verbs consist of a ' root' made up of a skeleton of consonants (mostly three but sometimes two or up to five), which is inserted into one of seven prefix-and-

vowel patterns or templates (traditionally called a binyan, plural binyanim). The consonant-root and the vowel-pattern are types of bound morphemes, and they cannot occur independently. There is a difference among these binyanim in their relative morphophonemic complexity: in its form, pa'al (usually deriving active intransitive stems) is the simplest binyan, with each of the others including some extra consonantal affix in the vowel template in addition to the basic three-consonantal root.

This system of verb-stem derivation is relatively regular; still, there are irregular gaps and inconsistencies in the semantics associated with a particular root in a particular binyan. On the whole, however, the verb-roots are neutral as to valency and voice, which are introduced by the choice of the binyan. Each binyan has some prototypical general meaning, defining a verb-class type, with the binyanim pa'al (or kal) generating active intransitive stems; pi'el active transitives, and hif'il active causatives. Each of these has a matching passive, nif'al, pu'al, and huf'al respectively. Lastly, there is a reflexive (and reciprocal) binyan, hitpa'el Intransitive verb stems are derived in all 7 binyanim, with the causative hif'il the least likely to derive intransitive verbs. Most verbs cannot actually appear in every one of the binyanim, or if they do, they may change their basic semantics. Each of the binyanim have their distinctive person, number and tense inflections; some have prefixes, some suffixes and some both. On the (relatively late) acquisition of the binyanim system, see Berman (1985).

METHOD

Subjects and language recording

Maternal speech

Maternal speech samples were taken from a videotaped observational study (Ninio, 1984). Two samples of Hebrew-speaking mother-infant dyads were observed. The first was a cross-

sectional sample of 24 dyads who were observed only once. All mothers had post high-school education, most full college education. Eight of the infants were approximately 10 months old, 8 were 18 months old, and 8 were 26 months old at the time of the observation. Half of each subsample of infants were males, half females. The second, longitudinal, sample consisted of 24 dyads of mothers and their infants who were observed once every two months for a year, 6 times in all. Three observations were missed because of illness or family travel. The 24 infants constituted three sub-samples of 8 infants each. Each sub-sample consisted of 4 males and 4 females, two each of a middle-class and a lower-middle class background. The first sub-sample of 8 infants was 12 months old at the first filming, and 14, 16, 18, 20 and 22 months at subsequent observations. The second sub-sample was 18 months at the first filming, and 28 at the last. The third sub-sample was 22 months at the first filming and 32 at the last. All children were of normal health, from intact homes, and the longitudinal sample were all first-borns. The subjects were randomly selected from birth records and recruited through letters and home visits. Each mother was paid a fee for her participation.

The mother-infant dyads were videotaped in their homes for 30 minutes at a time. Mothers were asked to behave as they usually do at that time of the day, but were asked to stay as much as possible in the same room with the child. Mothers were told that we wished to obtain a naturalistic sample of infants' interaction with their mothers, but were not told before the end of the study that either maternal or child language was to be the special focus of analysis.

The 24 mothers of the cross-sectional sample were invited to the lab within a week of the videotaping and asked to describe in detail what had happened in the videotaped interaction period, while watching the taped observation. A detailed description of the elicitation

procedure is to be found in Ninio and Wheeler (1984a).

In the present paper, the speech of all 48 mothers was examined; speech produced at all the observations was included in the analysis. Overall, these mothers were recorded interacting with young children when the latter were between 10 months and 32 months, with the ages of the children at the time of observation quite symmetrical around the middle of this range, about 20 months. The pooled speech of all 48 mothers represents a generalized data base of "caretaker speech" addressed in interaction to young children acquiring Hebrew.

Child speech

The language corpora used in this study are the first intransitive verbs in the multiword speech of 20 children acquiring Hebrew as their first language. The children were observed longitudinally. They are unrelated to the mothers' sample described above. Of the children, one, Ruti, was audiotaped by her parents twice a week between 1;6 and 2;4, for 20 minutes at a time. The recordings were immediately transcribed, and contextual information added. The other 19 children were audiotaped about once a week at home for 30 minutes, beginning at about 1;6 on the average, for 8-11 months; contextual information was added to the transcriptions. The observations were augmented by parental reports on emergent patterns. In all children the observations started before -- or just as -- word-combinations with intransitive verbs were first produced.

Table 1 presents the children' s sample with information about sex, study period, and number of observations.

Insert Table 1 about here

Data analysis

All utterances were transcribed in standard orthography. The corpora were then divided into utterances. An utterance was defined as a speech string which has a sentence intonation contour and which is separated by perceptible pauses from other locutions of the same speaker. All utterances of two words or more were included in the data base. Immediate repetitions of a single word within the same speaking turn were not taken as adding to the length of the utterance. Utterances in which a child makes a hesitation-pause between words are considered word-combinations, but not "vertical constructions" where two words belong to separate single-word turns at speech. Maternal utterances were marked with a detailed code for the communicative intent expressed by the utterance, using the Ninio and Wheeler (1994b) coding system. Excluded from the present analysis were recitations of texts (e.g. of songs or of verse) as well as elicitations of completion from the child, either of texts or of words in non-text utterances.

The maternal data base consisted of 49050 multiword utterances of two to four words per clause, in 7803 of which there was at least one intransitive verb. Longer utterances were not included, as we assume that children do not parse them at this early stage of development and that they therefore cannot function as effective input for acquisition.

Child utterances were annotated immediately after observation with detailed contextual comments regarding the communicative use of the utterance. As with maternal speech, non-spontaneous utterances, namely, immediate imitations of a previous utterance, recitations of texts and elicited completions of either texts or non-text words were excluded from the data base.

Language coding

Utterances including an intransitive verb were identified, and classified into verb-stem groups. A given verb stem was defined by its 3-consonantal root (e.g., H-L-K 'go') and by its binyan or derivational template, (e.g. pa'al, as in *halak* 'went-3sm', 'sm' indicating singular masculine). All forms of person, number and tense inflections for the same verb-stem were included in the same stem-category, including infinites (e.g., *laleket* 'to go'). The same root in another binyan was considered as a different verb stem. For example, the same consonant-root H-L-K can derive a stem in reflexive, *hithalek*, which is also an intransitive stem, meaning, roughly, 'walked around-3sm'. These two stems were considered different verbs in this study. All transitive stems derived from the same consonantal roots were excluded. For example, the root L-B-Sh derives in the reflexive binyan hitpa'el the intransitive verb *hitlabesh* 'got dressed-3sm', but in the causative binyan hif'il it derives the verb *hilbish* 'dressed (tr)-3sm', meaning put on clothes on somebody else. Only the former stem (in all its inflected forms) was included in the present study.

There is a great deal of lability in Hebrew, that is, the same derived verb-stem can be used in transitive as well as intransitive sentences. As a result, the categorization of a given stem as transitive or intransitive was not always self-evident. As a rule, a verb accepting a direct-object which is not a strict cognate object or any other obligatory complement without a preposition (except the accusative *et*) was considered a transitive verb and excluded from this study. This excluded from consideration verbs receiving predicate complements such as *haya* 'be', and other copulative verbs. It is interesting that *ha*, a verb considered to be the 'logical predicate' of the stative intransitive verb class, is, syntactically, a transitive verb.

For another reason we excluded also the verb-stem *kara* 'happen'; this verb appeared exclusively in the formulaic question *Ma kara?* 'What happened?', and it is highly questionable

whether it is anything but a frozen expression for children at the relevant age.

RESULTS AND DISCUSSION

The most frequent intransitive verbs in multiword utterances in the maternal input

The first statistics we computed to estimate the relative frequency of use of the different intransitive verb stems was the number of longitudinally observed mothers using the relevant stem in word combinations. As described in the Methods section, 24 of the mothers were observed 6 times, for an overall of 3 hours of video recordings. By contrast, the other 24 mothers were observed only once, for 30 minutes. Undoubtedly the longitudinal observations provide a much more comprehensive speech sample, with less likelihood of missing a verb due to undersampling. One the other hand, 3 hours of recording, distributed over 6 observations, can, potentially, create a ceiling-effect, so that there is a probability that many verbs will be found in the speech of all the mothers without exception at least once during this large-scale speech sampling. Thus, the proportion of the longitudinal sample using a verb stem may be a less than adequately discriminating index of relative frequency.

It is, then, significant, that of all the different intransitive stems recorded in word-combinations, only 5 were produced by all 24 longitudinally observed mothers. These were the verb-stems *ba* 'come', *halak* 'go', *nafal* 'fall', *yashav* 'sit' and *yashan* 'sleep'. These five verbs constitute, therefore, the set of most regularly used intransitive verbs in the input. Given a large enough speech sample, all mothers may be expected to be observed saying them. This characteristics distinguishes the set of five from other verbs which are quite frequent overall but which do not share this feature of "obligatoriness", such as *baka* 'cry' (produced by 23 of the 24 longitudinal mothers); *yaca* 'exit' (22), *kaav* 'hurt' (22), *amad* 'stand' (21), *nasa* 'travel' (21), *yarad* 'descend' (19) and *nikhnas* 'enter' (18).

Because of the ceiling-effect, it is impossible to order the first five verbs among themselves, and for that, and other more subtle differences among the verbs, the complete sample of 48 mothers is the better data base.

Table 2 presents the 30 most frequent intransitive verbs in word combinations in maternal speech. The major criterion for input frequency was the number of mothers, out of 48, who used the relevant verb-stem in their multiword utterances. The verbs are ordered in the table according to this measure of frequency. The table also presents the overall number of utterances produced by the sample with each verb-stem, and the cumulative percentage of all utterances accounted for by the verb and the verbs more frequent than it.

Insert Table 2 about here

The different intransitive verbs appear with a pronounced skewed distribution in the sampled corpus of maternal speech. There were 175 different intransitive verbs in the combined corpus. Of these, 12 verbs were used by more than half of the sample, more precisely, by 25 or more mothers. The other 163 verbs were used, each, by less than half of the sample, i.e., by 23 mothers or less.

The total token frequency of utterances with intransitive verbs was 7803. Of these, the 12 most frequent verbs accounted for no less than 5986, or 76.7% of all pooled tokens of intransitive verbs in combination. The less frequent 163 verbs together accounted only for 23.3% of the total pooled utterances. Because of the skewed distribution, the 30 most frequent verbs in Table 2 represent 87.6% of all utterances.

As to the most frequent verbs, the results of the complete 48-strong sample are very

similar to the results of the subgroup of 24 mothers who were recorded for 3 hours. The most frequently used verbs in multiword utterances are *ba* 'come' *halak* 'go' and *nafal* 'fall'. These verbs are used by about 90% or more of the mothers in the 48-sample. *Yashav* 'sit' is somewhat less widely used than the first three (by 83.3%). The other verbs, led by *yashan* 'sleep', are used by 60% or less of the 48-sample. Basically, of the five verbs which were used by all 24 of the longitudinal sample, only three -- 'come', 'go' and 'fall' -- retain this feature of being extremely frequent, everpresent, and near-obligatory when checked against the complete sample. The shorter recording time (30 minutes) used for 24 of the mothers separates the truly everpresent verbs from the ones that appear only with longer sampling.

The findings based on number of users are supported by the figures for number of tokens. The first five verbs, according to number of mothers using them, are also the most frequent according to the number of tokens observed. In particular, 'come' and 'go' are uttered extremely frequently, with 'sit' following, and 'fall' and 'sleep' the next most frequent. As a group, these five surpass the token frequency of any other verb. This advantage remains when the mean number of utterances are computed, relative to the number of mothers who actually produced sentences with each verb. To summarize, the first five verbs are used by more mothers and with a larger number of utterances per mother, then any other intransitive verb. The most prolific active intransitive verb in combination is *ba* 'come', and the most prolific static intransitive verb is *nafal* 'fall'.

These results support our initial hypothesis that the 'model' intransitive verbs *come/go* and *fall* (based on grammaticalization data) will be the most frequently produced verbs in maternal input. The second group of verbs that tend to be grammaticalized -- *move, stay* or *sit, stand, lie* or *sleep* -- is also relatively frequent in the input, but less so than the prototypically

grammaticalized active motion verbs *come* and *go*, and the static event or involuntary-motion verb *fall*.

Next, we shall proceed to an examination of the first intransitive verbs produced in word-combinations by children.

The first intransitive verbs in children's word-combinations

Appendix A presents the first sentence with each of the first three intransitive verbs emerging in combinations, for the 20 children of the sample. Table 3 presents the distribution of the verbs that were the first to appear in word-combinations. The table presents, for each verb stem, the number of children who had that verb as their first intransitive verb in combination; the number who had that verb among their first two verbs in combination and then among their first three verbs and their first seven verbs. Included in the table are only those verbs that at least one child used among his or her first seven verbs; this was the highest number of verbs whose order of acquisition could be documented for all 20 children in the present set of observations.

Insert Table 3 about here

There was a great deal of similarity among the initial verbs of the different children: Only 6 different verbs accounted for the verbs used by the sample of 20 children as their first intransitive verb in combination. Similarly, only 13 different verbs accounted for the 40 verbs used by the sample as their first two intransitive verbs in combination.

The one verb first produced in word combination by the largest number of children was *ba'come'*. In 12 of the 20 children, this was the first verb; in one child, the second, and in two

children, the third verb acquired. No other verb approached this strong tendency to be the initially combining intransitive verb.

The second-most likely verbs to be acquired early in combination were *nafal* 'fall' and *halak* 'go'. These verbs were especially prominent among the second and third verbs learned in combination. For example, *nafal* 'fall' was the first verb in 3 children, the second verb in 4 children, and the third verb in 4 others. These three favorites, 'come', 'fall' and 'go', together made up 34 of the 60 verbs which were the first three verbs in this group of children. The remainder is spread out among 14 other verbs.

In addition, the three favorites were almost certain to be acquired among the first 7 verbs in combination for all the children. For *ba* 'come', the probability reached 95%, namely, all but one child of the sample used it. For the other two favorites, *nafal* 'fall' and *halak* 'go', the probability of learning them among the first 7 verbs was 70-75%.

As the next most favoured verbs, *yashan* 'sleep' and *yashav* 'sit' were acquired as their very first verbs by a few children, in particular their third verb. In addition, the probability of these verbs to be among the first 7 verbs was 55%. None of the other verbs approximated this profile of preference for early acquisition.

As far as the comparison with the grammaticalization data is concerned, the first three intransitive verbs learned in combination are indeed the Hebrew forms of the three most prominent grammaticalizable intransitive verbs, *ba* 'come', *halak* 'go', and *nafal* 'fall'.

In addition, as we saw in the previous section, these three are the three most frequent verbs in the maternal input speech.

In summary, the first acquired intransitive verbs in combination are *come* and *go* (active), and *fall* (static) verbs; all three are, at the same time, the most grammaticalized, general verbs,

and also the most frequent verbs in the multiword input. The second-preferred verbs *sit* and *sleep* are also among the crosslinguistically grammaticalized verbs, and they are also relatively frequent in the input, but on both counts they are less prominent and less frequently modelled than the first three. Other, later-acquired intransitive verbs tend to be not easily grammaticalizable, and also tend to be less frequent in the maternal input.

As with other linguistic domains, the order of acquisition of combining intransitive verbs was successfully predicted by the verbs' generality, and also by their input frequency. Generality -- or perhaps cognitive simplicity -- and frequency in the input were, as usual, highly "confounded", so that it is impossible to give a single answer to the question, why are these verbs the most preferred verbs for early syntactic acquisition.

CONCLUSION

The results of this study support the suggestion initially entertained in the context of transitive verb-object combinations, that children begin the mastery of particular linguistic domains by initially learning the "best models" of that domain. These "best models" are the most general and most prototypical concrete items belonging to the domain, and as such are the best sources for analogy and transference to other, more specific, instances.

In the present study, this hypothesis was tested in yet another domain -- that of intransitive verbs in multiword combinations. We predicted that the most frequently modelled combining intransitive verbs in the linguistic input provided by mothers would be the same verbs as the intransitive verbs first produced in word-combinations by children, and that these verbs would be the translation equivalents of frequently-grammaticalized verb-stems. These predictions have been verified in the reported study.

The learning process for linguistic rules of all kinds suggested by these results is extremely

similar to concept formation: apparently while learning the features of a few concrete instances, children form what is the functional equivalent of abstract concepts applying to KINDS of objects of the same type. This concept-formation process is facilitated, or even made possible by the statistical structure of language, by which the simplest, most general, and most model-like instances in each domain are the most frequently used instances of that domain. The well-documented input-frequency effect thus has a constitutive role in language acquisition which is cardinally different from its supposedly "behavioristic" character: It makes possible the formation of abstract concepts, on the basis of prototypical or "best model" instances of a linguistic domain.

At the same time, the proposed acquisition mechanism is a proper learning procedure, without the need to invoke genetically inherited linguistic concepts or any nativist notions of this kind. Human languages are robust systems with functionally shaped features; such a system would have precisely the characteristics that makes it the easiest possible to learn from the available linguistic input. The correlation of use frequency and simplicity, generality and best-model characteristics demonstrated yet again in this study, is one of the optimal design-characteristics of language that make its learning by children a highly solvable cognitive problem.

Appendix A.

The first three intransitive verbs emerging in combinations, and the first sentence with each (N=20).

Child	Utterances	
	Hebrew	English gloss
Ruti	Ze af. Ze tas. Beca nafal.	this FLIES this FLIES (PLANE) egg FELL
Zohar	Po...lishon. Nafal li. Lalechet habayta.	here TO-SLEEP FELL to-me TO-GO home
Shay (m)	Lechi mipo. Migdal nofel. Boy Mia.	GO (FEM IMP) from-here tower FALLs COME (FEM IMP) Mia
Paz	Aba bo. Hi kama. Ze nishbar.	Daddy COME (IMP) she GETS-UP (FEM) this BROKE
Or	Marina boy. Tomer zuz. Lech mipo.	Marina COME (FEM IMP) Tomer MOVE (IMP) GO (IMP) from-here
Omer	Bo elay. Hu...koev. Habalon hitpocec.	COME (IMP) to-me he HURTs the-balloon BURST
Ayelet	Aba bo. Doda halach. Barcha balon.	Daddy COME (IMP) Untie WENT ESCAPED (FEM) balloon
Mordexai	Pkak nafal. Yored geshem. Lo roce lishon.	cork FELL DESCENDs rain not want to-SLEEP
Rotem	Al telech. Ima boy. Kol hazman nofelet.	don' €GO (IMP) Mommy COME (FEM IMP) all the-time FALLs (FEM)
Yonatan	Zizi bo. Cipor afa. Tuki nafal.	Zizi COME (IMP) bird FLIES parrot FELL

Appendix A. -- Continued

Child	Utterances	
	Hebrew	English gloss
Shira	Aba boy. Nafal li. Al tivki.	Daddy COME (FEM IMP) FELL to-me don' CRY (FEM IMP)
Matan	Galgal nafal. Kelev boxe. Navot lishon.	wheel FELL dog CRIES Navot to-SLEEP
Reut	Ima...bo. Adi zuzi. Aba shev.	Mommy COME (IMP) Adi MOVE (FEM IMP) Daddy SIT (IMP)
Bet	Ima bo. /Boy leechol. Zuzi ima. Nofelet habuba.	Mommy COME (IMP) /COME (IMP) to-eat MOVE (FEM IMP) Mommy FALLs the-doll
Merav	Kan shev. Merav yoce. Dubi yashen.	here SIT (IMP) Merav GETS-OUT Teddy-bear SLEEPs
Naama	Aba bo. Ima zuzi. Ima shelo halcha.	Daddy COME (IMP) Mommy MOVE (FEM IMP) Mommy his WENT (FEM)
Shay (f)	Mik bo. Aba lex. Shev po.	Mik COME (IMP) Daddy GO (IMP) SIT (IMP) here
Yuval	Boy shama. Nafal sus. Dubi yashan.	COME (FEM IMP) there FELL horse Teddy-bear SLEPT
Hana	Et-ze nafal. Aba halach. Sara baa.	ACC-this FELL Daddy WENT Sara CAME (FEM)
Adi	Boy...habayta. Roca...lashevet. Buba...mistovev.	COME (FEM IMP) home want to-SIT doll TURNs

REFERENCES

BARRETT, M. D. (1981). The communicative functions of early child language. Linguistics, **19**, 273-305.

BARTLETT, E. (1976). Sizing things up: the acquisition of the meaning of dimensional adjectives. Journal of Child Language, **3**, 205-220.

BERMAN, R. (1985). The acquisition of Hebrew. In D. I. Slobin (Ed.), The crosslinguistic study of language acquisition (Vol. I, pp. 255-371). Hillsdale, NJ: Lawrence Erlbaum.

BERMAN, R. (1988). Word class distinctions in developing grammars. In Y. Levy, I. Schlesinger, & M.D.S. Braine (Eds.), Categories and processes in language acquisition (pp. 45-72). Hillsdale, NJ: Lawrence Erlbaum.

BIERWISCH, M. (1967). Some semantic universals of German adjectivals. Foundations of Language, **3**, 1-36.

BLAKE, B. J. (1994). Case. Cambridge, U.K.: Cambridge University Press.

BLEWITT, P. (1982). Word meaning acquisition in young children: a review of theory and research. In H. Reese & L. Lipsitt (Eds.), Advances in child development behavior (Vol. 17, pp. 140-185). New York: Academic Press.

BLOOM, L. (1973). One word at a time: The use of single word utterances before syntax. The Hague: Mouton.

BOAS, F. (1947). Kwakiutl grammar, with a glossary of the suffixes. (Edited by H. B. Yampolsky). Transactions of the American Philosophical Society, **37**, 203-377.

BROWN, R. (1973). A first language: The early stages. Cambridge, MA: Harvard University Press.

BROWN, R. & HANLON, (1970). Derivational complexity and order of acquisition in child speech. In J. R. Hayes (Ed.), Cognition and the development of language. (pp. 11-53). New York: Wiley.

BYBEE, J., PERKINS, R. & PAGLIUCA, W. (1994). The evolution of grammar. Chicago: The

University of Chicago Press.

CAMPBELL, R. & WALES, R. (1970). The study of language acquisition. In J. Lyons (Ed.), New horizons in linguistics (pp. 242-260). Harmondsworth, England: Penguin Books.

CLARK, E. (1972). On the child' s acquisition of antonyms in two semantic fields.Journal of Verbal Learning and Verbal Behavior, **11**, 750-758.

CLARK, E. V. (1978). Strategies for communicating. Child Development, **49**, 953-959.

DE VILLIERS, J. G. (1980). The process of rule learning in child speech: a new look. In K. E. Nelson (Ed.), Children' s language(Vol. 2, pp. 1-44). New York: Gardner Press.

DIXON, R. M. W. (1980). The languages of Australia. Cambridge, U.K.: Cambridge University Press.

DIXON, R. M. W. (1982). Where have all the adjectives gone? New York: Mouton.

EILERS, R., OLLER, D. K. & ELLINGTON, J. (1974). The acquisition of word meaning for dimensional adjectives: the long and the short of it. Journal of Child Language, **1**, 195-204.

FOLEY, W. A. & OLSON, M. (1985). Clausehood and verb serialization. In J. Nichols & A.C. Woodbury (Eds.), Grammar inside and outside the clause (pp. 17-60). Cambridge, U.K.: Cambridge University Press.

FORNER, M. (1979). The mother as LAD: Interaction between order and frequency of parental input and child production. In F. R. Eckman & A. J. Hastings (Eds.), Studies in first and second language acquisition (pp. 17-44). Rowley, MA: Newbury House.

GIVON, T. (1975). Serial verbs and syntactic change: Niger-Congo. In C. Li (Ed.), Word order and word order change (pp. 47-112). Austin: University of Texas Press.

GIVON, T. (1991). Serial verbs and the mental reality of ' event' . In E. C. Traugott & B. Heine (eds), Approaches to grammaticalization (Vol. 1, pp. 81-127). Amsterdam: Benjamins.

GREENFIELD, P.M., & SMITH, J.H. (1976). The structure of communication in early language development. New York: Academic Press.

GRIFFITHS, P. (1985). The communicative functions of children's single-word speech. In M. Barrett (Ed.), Children's single-word speech (pp. 87-112). New York: Wiley.

HALLIDAY, M.A.K. (1975). Learning to mean -- Explorations in the development of language. London: Edward Arnold.

HEIDER, E. (1971). "Focal" color areas and the development of color names. Developmental Psychology, **4**, 447-455.

HOPPER, P. J. & TRAUGOTT, E. C. (1993). Grammaticalization. Cambridge, U.K.: Cambridge University Press.

INGRAM, D. (1989). First language acquisition. Cambridge: Cambridge University Press.

JOHNSON, E. G. (1977). The development of color knowledge in pre-school children. Child Development, **48**, 308-311.

LEHMANN, C. (1991). Predicate classes and participation. In H. Seiler & W. Premper (Eds.), Partizipation (pp. 183-239). Tubingen: Gunter Narr Verlag.

LI, C. N. & THOMPSON, S. A. (1981). Mandarin Chinese: a functional reference grammar. Berkeley: University of California Press.

LORD, C. (1982). The development of object markers in serial verb languages. In P. Hopper & S. Thompson (Eds.), Studies in transitivity. Syntax and Semantics 15 (pp. 277-299). New York: Academic Press.

MOERK, E. L. (1980). Relationships between parental input frequencies and children's language acquisition: a reanalysis of Brown's data.Journal of Child Language, **7**, 105-118.

NINIO, A. (1984). Functions of speech in mother-infant interaction. Final Science Report to the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel.

NINIO, A. (1993). On the fringes of the system: Children's acquisition of syntactically isolated forms at the onset of speech. First Language, **13**, 291-313.

NINIO, A. (1996). Pathbreaking verbs in syntactic development. Paper presented at the Seventh

International Congress for the Study of Child Language, Istanbul, Turkey.

NINIO, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. Journal of Child Language, In Press.

NINIO, A. & WHEELER, P. (1984a). Functions of speech in mother-infant interaction. In L. Feagans, G.J. Garvey, & R. Golinkoff (Eds.), The Origins and Growth of Communication (pp. 196-207). Norwood, NJ: Ablex.

NINIO, A. & WHEELER, P. (1984b). A manual for classifying verbal communicative acts in mother-infant interaction. Working Papers in Developmental Psychology, No. 1. Jerusalem: The Martin and Vivian Levin Center, Hebrew University. Reprinted as Transcript Analysis, 1986, 3, 1-82.

QUIRK, R., GREENBAUM, S., LEECH, G. & SVARTVIK, J. (1972). A grammar of contemporary English. London: Longman.

ROSCH, E. (1975). Cognitive representation of semantic categories. Journal of Experimental Psychology: General, **104**, 192-233.

SILVERSTEIN, M. (1972). Chinook Jargon I. Language, **48**. 378-406.

SLOBIN, D. I. (1987, January). Frequency reflects function. Paper presented at the conference on The Interaction of Form and Function in Language. University of California, Davis.

TRAUGOTT, E. C. & HEINE, B. (1991). Introduction. In E. C. Traugott & B. Heine (Eds.), Approaches to grammaticalization (Vol. 1, pp. 1-14). Amsterdam: Benjamins.

WEISENBERGER, J. L. (1976). A choice of words: two-year-old speech from a situational point of view. Journal of Child Language, **3**, 275-281.

WOLFF, J. G. (1988). Learning syntax and meanings through optimization and distributional analysis. In Y. Levy, I. Schlesinger, & M.D.S. Braine (Eds.), Categories and processes in language acquisition (pp. 179-215). Hillsdale, NJ: Erlbaum.

AUTHOR NOTES

[*] Portions of the research reported here were presented at the Seventh International Congress for the Study of Child Language, Istanbul, Turkey, July, 1996. Thanks are due to Moti Rimor for permission to use his observations of Ruti; to Naomi Bacon, Lucy Bahat, Eli Berkowitz, Adi Brill, Sarit Dag, Rachel Hadani, Noa Hocherman, Margalit Kazins, Tamar Keren-Portnoy, Chagit Magen, Liat Malhi, Sharon Marbach, Noga Meir, Ori Nahari, Ofri Tal, Ariel Wasserteil, Chamutal Yakir and Osnat Yogev, who collected the longitudinal language corpora; and to Neta Ofer for analyses of the Rimor corpus.

The research reported in this paper was supported by a grant from The Spencer Foundation. The data presented, the statements made, and the views expressed are solely the responsibility of the author.

Table 1.

Description of the longitudinal observations used in the study (N=20)

Child	Sex	Study period	No. observations
Ruti	f	1;06.00-2;04.00	125 (twice weekly)
Zohar	f	1;06.12-2;02.08	29
Shay (m)	m	1;07.21-2;03.26	28
Paz	f	1;06.22-2;03.02	24
Or	f	1;03.21-2;02.06	38
Omer	m	1;03.03-1;10.10	23
Ayelet	f	1;01.08-1;10.20	34
Mordexai	m	1;08.08-2;03.00	20
Rotem	f	1;07.19-2;03.16	31
Yonatan	m	1;06.00-2;03.29	36
Shira	f	1;04.18-2;02.16	31
Matan	m	2;01.03-2;04.03	16
Reut	f	1;05.12-1;11.25	21
Bet	f	1;10.16-2;06.20	27
Merav	f	1;06.04-1;11.16	22
Naama	f	1;08.08-2;02.24	27

Shay	f	1;06.18-2;02.05	29

Table 1 -- Continue

Child	Sex	Study period	No. observations
Yuval	m	2;01.01-2;07.27	30
Hana	f	2;01.22-2;07.11	20
Adi	f	1;07.02-2;07.22	51

Table 2.

Distribution of the most frequent intransitive verbs in maternal word-combinations

Verbs		Input Frequency of Verb Stem		
Hebrew (stem)	English gloss	Number of mothers (N=48)	Pooled Tokens	
			Utterances (N=7803)	Cumulative Frequency
ba	come	48 (100.0%)	2743	35.2%
halak	go	47 (97.9%)	1270	51.4%
nafal	fall	42 (87.5%)	307	55.4%
yashav	sit	40 (83.3%)	648	63.7
yashan	sleep	30 (62.5%)	192	66.2
' amad	stand	28 (58.3%)	168	68.3
baka	cry	28 (58.3%)	153	70.3
nasa	travel	28 (58.3%)	106	71.6
yaca	exit	28 (58.3%)	077	72.6
niknas	enter	27 (56.2%)	124	74.2
kaav	hurt	26 (54.2%)	96	75.4
yarad	descend	25 (52.1%)	102	76.7

diber	talk	22 (45.8%)	110	78.1
kam	get up	20 (41.7%)	56	78.9
naga	touch	20 (41.7%)	51	79.5

Table 2 -- Continue

Verbs		Input Frequency of Verb Stem		
Hebrew (stem)	English gloss	Number of mothers (N=48)	Pooled Tokens	
			Utterances (N=7803)	Cumulative Frequency
xika	wait	20 (41.7%)	45	80.1
histovev	turn	19 (39.6%)	53	80.8
nizhar	be careful	19 (39.6%)	52	81.4
' ala	ascend	16 (33.3%)	58	82.2
zaz	move	16 (33.3%)	47	82.8
nish' ar	stay	16 (33.3%)	34	83.2
nigmar	stop	16 (33.3%)	32	83.6
barax	escape	16 (33.3%)	29	84.0
avar	pass	15 (31.2%)	28	84.4
tiyel	take a walk	14 (29.2%)	51	85.0

nishbar	broke	14 (29.2%)	46	85.6
hirbic	hit	14 (29.2%)	30	86.0
cadak	be correct	14 (29.2%)	23	86.3
caak	yell	13 (27.1%)	32	86.7
rac	run	13 (27.1%)	30	87.1
kafac	jump	12 (25.0%)	38	87.6

Table 3.

Distribution of the first intransitive verbs appearing in word-combinations in children (N=20)

Verbs		Number of children using the verb stem among their first verbs in combination (N=20)			
Hebrew (stem)	English gloss	1st	First 2	First 3	First 7
ba	come	12	13	15	19
nafal	fall	3	7	11	14
halak	go	2	5	8	15
yashav	sit	1	2	4	11
yashan	sleep	1	1	5	11
`af	fly	1	2	2	7
zaz	move	0	4	4	7
baka	cry	0	1	2	7

kaav	hurt	0	1	1	8
yarad	descend	0	1	1	4
kam	get up	0	1	1	3
yaca	exit	0	1	1	2
tas	fly (plane)	0	1	1	2
nishbar	broke	0	0	1	3

Table 3 --- Continue

Verbs		Number of children (N=20)			
Hebrew (stem)	English gloss	1st	First 2	First 3	First 7
hitpocec	burst	0	0	1	1
barax	escape	0	0	1	1
histovev	turn	0	0	1	1
niknas	enter	0	0	0	7
kaas	be angry	0	0	0	3
naga	touch	0	0	0	2
nasa	travel	0	0	0	1
avar	pass	0	0	0	1
higi' a	arrive	0	0	0	1
' ala	ascend	0	0	0	1
tipes	climb	0	0	0	1
hitkalkel	spoil	0	0	0	1
xika	wait	0	0	0	1
saxa	swim	0	0	0	1
' amad	stand	0	0	0	1

nigmar	stop	0	0	0	1
nishpak	spill	0	0	0	1

Bibliographical note

Anat Ninio is Braun Professor of Psychology at the Hebrew University in Jerusalem. She has published extensively on pragmatic development, recently bringing out a book on the subject with Catherine Snow. Currently, she is working on the development of an empiricist theory of syntactic development in children, and conducting cross-linguistic research.