Article:

KEREN-PORTNOY, T. orcid.org/0000-0002-7258-2404 and KEREN, M. (2011) The dynamics of syntax acquisition: facilitation between syntactic structures. Journal of Child Language. pp. 404-432.

https://doi.org/10.1017/S0305000909990559

J. Child Lang. **38** (2011), 404–432.
doi:10.1017/S0305000909990559

The dynamics of syntax acquisition: facilitation between syntactic structures*

TAMAR KEREN-PORTNOY

Department of Language and Linguistic Science, University of York

AND

MICHAEL KEREN

Department of Economics, Hebrew University of Jerusalem

(*Received* 26 *March* 2008 – *Revised* 19 *January* 2009 – *Accepted* 14 *December* 2009 – *First published online* 16 *July* 2010)

ABSTRACT

This paper sets out to show how facilitation between different clause structures operates over time in syntax acquisition. The phenomenon of facilitation within given structures has been widely documented, yet inter-structure facilitation has rarely been reported so far. Our findings are based on the naturalistic production corpora of six toddlers learning Hebrew as their first language. We use regression analysis, a method that has not been used to study this phenomenon. We find that the proportion of errors among the earliest produced clauses in a structure is related to the degree of acceleration of that structure's learning curve; that with the accretion of structures the proportion of errors among the first clauses of new structures declines, as does the acceleration of their learning curves. We interpret our findings as showing that learning new syntactic structures is made easier, or facilitated, by previously acquired ones.

INTRODUCTION

This paper sets out to show how facilitation between different clause structures operates over time in syntax acquisition. The phenomenon of facilitation within given structures has been widely documented, yet

[*] We would like to thank the editors and anonymous reviewers, Anat Ninio, Izchak Schlesinger and Marilyn Vihman, for their thoughtful comments. We are very grateful to Yoel Haitovsky for statistical advice and to Nadav Benzeev for important statistical help. We also thank all the families who took part in the research reported here, Sarit Dag, David Heller, Osnat Yogev and Motti Rimor for help with the data collection, and Ronny Alony, Tamar Parush, Doron Magen and Osnat Yogev for assistance with the transcription and coding.

inter-structure facilitation has not, to the best of our knowledge, been reliably reported so far. Our findings are based on the naturalistic production corpora of six toddlers learning Hebrew as their first language. We use regression analysis, a method that has not, as far as we are aware, been used to study this phenomenon.

Like many researchers who espouse an empiricist–emergentist perspective to language development, we interpret our results using an exemplar-based language learning approach. The main tenet of exemplar-learning models is that learning need not involve extraction of rules which refer to categories that are more abstract than the items of knowledge themselves (for clear definitions of such theories, see Hahn & Chater, 1998). This approach sees generalization-like behaviour as the result of similar exemplars being stored close to each other, thus creating dense clumps of exemplars, whereas items which are less similar to others inhabit more sparsely connected parts of the network. Knowledge remains associated with specific items rather than with categories which are abstracted from them, and resides mostly in the connections between different exemplars. These connections get strengthened with recurring co-activation or weakened otherwise. Such models were originally suggested for category learning in general (e.g. Medin & Schaffer, 1978; Nosofsky, 1988). In relation to language, this view is now held by many researchers studying morphology and phonology (e.g. Miller, 1994; Bybee, 2001) as well as syntax (e.g. Tomasello, 1992; 2000; Ninio, 1999a; 1999b). In the context of the early stage in syntax development which is the focus of this paper, the exemplars of interest to us are verbs, along with information about the arguments each takes.

FACILITATION is a process whereby previously known 'old' exemplars, by having already been learned, FACILITATE or make the learning of new exemplars easier. Facilitation operates between similar items: unfamiliar exemplars are processed like familiar exemplars that are similar though not identical to them, by way of analogy. When a new item is met, the child searches for a similar known item, and this becomes easier as the stock of learned items grows. As more exemplars from the same category are learned, the basis for their similarity becomes clearer as in the comparison process the weights assigned to their features are adjusted and refined. Consequently comparing new exemplars to old ones, learning and assimilating them into the existing knowledge structure becomes progressively easier.

Facilitation has been shown to occur both in non-linguistic learning (Homa & Chambliss, 1975) and in language acquisition, specifically in syntax learning (Abbot-Smith & Behrens, 2006; Keren-Portnoy, 2006; Kiekhoefer, 2002; Ninio, 1999a; Vihman, 1999). Most of the work on facilitation in syntax has focused on facilitation between different verbs within a syntactic structure. Abbot-Smith & Behrens (2006) were the first to look for facilitation between different syntactic structures and have shown that the acquisition of later acquired structures may be supported (or at times hindered) by previously acquired ones. They focus on the role of semantics and lexical overlap as determinants of facilitation. Our research complements theirs and shows, using a different methodology, facilitation operating between early and later acquired structures. When discussing facilitation between different structures, the carriers of knowledge are again the same specific exemplars, the verbs. What we would like to show is that the effect of previously learned verbs in previously learned structures extends not only to other verbs sharing a similar structure, but also to other structures. Facilitation within a structure is usually demonstrated by showing a gradually diminishing time lag between consecutively learned items (e.g. verbs) as learning progresses. Although the first items are learned slowly and effortfully, learning new items becomes gradually easier, and new items start to be learned at a faster pace. Old items are said to facilitate the learning of subsequent items. When measuring the cumulative number of items learned per period of time, the growing pace of learning is evidenced in an accelerating learning curve.

Keren-Portnoy (2006) suggested that practice which takes place during the early stages of learning a structure can explain (at least in part) the phenomenon of facilitation. During this practice phase the child develops the skill that is involved in producing certain kinds of combinations, simply by using them over and over again and comparing her output to the input she hears. Word-order problems – determining the location of the different arguments relative to the verb – are also worked out through 'learning-by-using', and errors can be shown to diminish as learning progresses (Keren-Portnoy, 2002). It is natural to assume that these details, once they have been mastered for one or more structures, should help to smooth and accelerate further learning. In this scenario, early learning in one structure makes later learning in a different structure easier. We therefore see it as facilitation between structures. We do not know which dimensions of similarity among structures or utterances are most relevant and beneficial for such facilitation. It may seem that having learned to use a specific verb in one structure may be beneficial for learning to use that very same verb in another structure, or that structures which contain certain argument types (such as subject or direct object) would be especially beneficial for learning further structures which contain the same argument types. Thus learning verb–direct object (VO) structures or subject–verb (SV) structures, or both, should catalyze the learning of subject–verb–object (SVO) structures. Similarly, it may be that learning is facilitated by semantic similarity, so that items tend to be learned which are semantically similar to those learned before them. There is, however, evidence that none of these three kinds of similarity (verb identity, overlap in argument structure and semantic similarity) operates in syntax acquisition in a simple and predictable manner. Regarding specific verb identity, Ninio (2003) has shown that the acquisition of later learned structures does not necessarily begin with verbs already learned in previous structures. In the English corpus she looked at, 40 percent of the earliest verbs produced in SVO have not been previously produced in either VO or SV. She argues that facilitation is not, in general, carried by specific verbs, and that it is not the case that, once an argument is combined with a given verb, the child will learn to produce new structures by gradually combining more arguments with that same verb. As for the possibility that specific argument types are the vessels of facilitation – this may usually be the case, but Keren-Portnoy (in preparation) reports a counter-example. She shows that cases can be found in which mastering the word order for a multi-argument structure (e.g. SVO) did not induce the correct word order for its component structures (e.g. VO), and errors in the latter continued after having ceased to occur in the multi-argument structure. Regarding semantic similarity, Ninio (2005a; 2005b) found no evidence that semantic similarity plays a role in mediating syntactic learning. In the current paper we have restricted ourselves to MEASURING facilitation between the nth and the $(n+1)$th acquired structure, and no attempt will be made to identify the agents of facilitation.

There are two potential objections to our interpretation of an increase in the ease of use as evidence for facilitation driven by previous learning. The more extreme objection, first suggested to us by Izchak Schlesinger and previously tackled by Ninio (1999a), questions whether the development involved is at all linguistic. The other possible counter-claim accepts that we are indeed dealing with language learning, but questions the usefulness of conceptualizing the learning process as one of learning new STRUCTURES.

The first objection may be phrased thus: children's cognitive and memory faculties mature, becoming faster and more efficient over time, and learning anything new becomes easier. Hence, how justified is the claim that facilitation is due to what has already been learned, rather than to simple biologically driven maturation? We contend that maturation, i.e. the improvement of the brain, the hardware, which permits faster reception, retention and processing of information, is not the whole story, and that actual early syntactic knowledge supports and scaffolds later knowledge. Murphy, McKone & Slee (2003) show that implicit memory (as evidenced in priming effects) develops with age in domains in which the knowledge base develops. They thus demonstrate that automaticity and processing speed are not simply a result of maturation, but of development in the structure of the database, not through learning new items, but primarily through the creation of new connections between items. If knowledge progresses by strengthening some connections and weakening others, then the two accounts for the improvement in processing, that of the improvement of hardware and that of the increasing density of the database, can be seen as two facets of the same process. In such a system, previous syntactic knowledge shapes the brain, enabling it to work faster and more efficiently. However, hardware improvement resulting from changes in the database will be much more local than improvement which results from biologically determined brain growth.

The second objection may be that what facilitation data actually show is that combining words becomes gradually easier, due not particularly to learning new structures but rather to learning to produce multiword utterances. In essence, the claim is that specific syntactic structures do not play an essential role in syntax learning.

Both of these claims, that of general maturation and that of general word-combining knowledge, can only explain across-the-board changes. The timing of mastery of different structures will therefore be investigated: if different structures used at the same time seem to point to different levels of maturation or knowledge (some being mastered well while others are not), then neither of these two general accounts can fully explain the data. If found, such patchy knowledge would point to structures (as exemplified in specific verbs) being the likely unit of knowledge which is acquired.

The main variables and hypotheses of our study will now be presented, followed by the methods and a description of the data. Since the statistical methods used in this study are not customary in this type of research, a fairly detailed description is provided before presenting the regression results. Following these results, evidence is presented that the learning of previous structures, rather than more general mechanisms, is the factor responsible for the facilitated acquisition of later learned structures. We conclude with a brief discussion of the implications of our findings.

PARAMETERS AND HYPOTHESES

We now come to the type of evidence for facilitation that will be adduced from naturalistic production data. If later learned structures actually build on knowledge that has been acquired for previously acquired structures, they should exhibit fewer signs of struggle and fewer traces of search for solutions than earlier ones. We focus on two properties in the acquisition of new structures: (1) the number of errors made in the learning process, and (2) the trajectory of learning – the rate at which verbs join the structure.

(1) *Word-order errors*. Keren-Portnoy (2006) has shown that word-order errors in the earliest clauses constructed in a structure can be taken as signs of trial and error in the early stages of learning a new structure (cf. Braine's 1976 'groping patterns'). Such errors seem to be the result of a search for the correct location of different arguments, for

the solution to the puzzle of how those arguments 'go together' with the verb. If later learned structures are indeed learned more easily on the basis of what is already known, and if they involve less of a need for problem solving, we would expect children to produce fewer word-order errors among the earliest clauses constructed in those structures.

(2) *The learning trajectory*. It has been shown that the great majority of structures exhibit a learning curve characterized by a slow start and a gradual acceleration (Ninio, 1999a; 2005a; 2005b; Vihman, 1999; Kiekhoefer, 2002; Keren-Portnoy, 2006). If later structures are indeed easier to learn due to some previous experience, they should exhibit learning trajectories in which the initial overcoming of hurdles followed by a 'take-off' is less pronounced as more structures are mastered. Later structures will therefore be expected to be learned more evenly, that is, to have a less-accelerating learning curve. At the extreme one may get a linear, non-accelerating curve, as reported by Abbot-Smith & Behrens (2006).[1]

The degree of acceleration and the tendency to make word-order errors (see above) are both taken to signify the degree of difficulty. We therefore expect them to show a positive correlation, supporting the notion that they are both affected by a common underlying variable, DIFFICULTY.

We summarize the predictions to be tested in the following hypotheses:

Hypothesis 1: Errors and degree of acceleration as measures of difficulty: the smaller the proportion of errors among a structure's earliest clauses, the less accelerating its learning curve.

Hypothesis 2 (facilitation I): the later a structure is acquired, the smaller the proportion of errors among its earliest clauses.

Hypothesis 3 (facilitation II): the later a structure is acquired, the lower the acceleration of its learning curve.

To challenge the suggestion that facilitation, if found, is the result of the state of the entire mental, syntactic or cognitive system of the child and to underline the benefit of describing syntactic knowledge as being channelled through structure learning, we will investigate whether, during the very

[1] A reviewer raised the possibility that less accelerating trajectories are actually ones which show less successful overcoming of the initial difficulties, and therefore are signs of learning which has never actually 'taken off' and requires more, not less, effort. However, evidence from Keren-Portnoy (2006), which is based on the same corpora as those used in the current paper, shows that this is unlikely to be the case. Keren-Portnoy compared structures with accelerating trajectories to those with linear, non-accelerating trajectories (a minority). The structures whose trajectories were linear had significantly fewer errors and more different verbs used among the first clauses created in them, signifying that these structures were easier, not harder to learn, relative to accelerating structures. Linear trajectories can be taken to be an extreme case of reduced acceleration.

TABLE 1. *The corpora*

Child's name	Age at first recording	Age at last recording	Number of recordings
Bareket	1;1.8	1;10.20	34
Lior	1;7.16	2;3.5	25
Naomi	1;6.25	2;7.22	51
Ofer	1;6.14	2;6.16	45
Shuli	1;5.25	2;4.8	125[a]
Tal	1;7.22	2;3.27	28

a: Twice weekly. Source: Table 1 in Keren-Portnoy (2006).

same period in a given corpus, late structures may be at their slow stage of learning while earlier ones are already in their accelerating period. Note that a similar argument has been made by Ninio (1999a).

THE METHODS AND THE DATA

Participants and corpora

Our basic dataset consists of the production corpora of six children acquiring Hebrew as their first language. One child received some English input as well, but all his documented productions are in Hebrew. Data collection started before any word combinations were produced and continued for eight to thirteen months. The children were audio-recorded in naturalistic interaction with a parent. Five children were audio-recorded weekly for about half an hour. One girl was recorded for twenty minutes twice a week. The average age at the first session was 1;5.29 and at the last session 2;4.1 (see Table 1 for a general description of the corpora). Three of the corpora (those of Naomi, Ofer and Shuli) are much richer than the other three in terms of the number of recordings made and, as a result, in the number of clauses available for analysis. The observers transcribed the recordings in standard Hebrew orthography, often noting (in very broad phonetic transcription) the phonetic realization of the verbs. The recordings were supplemented by parents' written reports of some utterances heard outside the recording sessions; these were excluded for one child due to too great a divergence between the verbs reported by the parents and those recorded by the observer. In addition, a couple of observers documented additional utterances heard outside the recording session, usually after the audiotape had been turned off.

Primary data

For a detailed description of the data and the forms which were analyzed, see Keren-Portnoy (2002; 2006). The development of each individual structure in each corpus was followed. All clauses containing a verb with some or all of its arguments were used in the analysis. Utterances containing hesitation or pauses are considered unitary utterances, unless another speaker's turn intervenes. Each clause was analyzed separately, regardless of whether it contained a finite or a non-finite verb. However, if one clause served as an argument of another, only the main clause was analyzed. Thus the following were regarded as separate clauses: coordinated clauses, relative clauses, adverbial clauses (when not serving as obligatory adverbial arguments). Only utterances which were uninterrupted, intelligible, comprehensible and spontaneous were used in the analysis

The arguments examined were:

Subject (S)
Direct object (O) (including sentential complements)
Indirect object (I) (including all datives and obliques)
Obligatory adjuncts (A) (This category included mostly adjuncts indicating goal, source, location and, in rare cases, time or manner.)

The fifteen structures that were investigated were SV, VO, VI, VA, SVO, SVI, SVA, VOI, VOA, VIA, SVOI, SVOA, SVIA, VOIA, SVOIA. Table 2 lists, for each of these structures, an example of a clause constructed by one of the children. For most of our statistics each structure in each corpus provides a datapoint. Had all six children produced all the potential structures, there would have been ninety datapoints. In actual fact for the analyses, which necessitate constructing learning curves (see below), only forty structures can be used; in those that involve error statistics, even fewer. This is due to the fact that only three children had twelve or more structures represented in their corpora, and not in all cases were there enough valid points in a structure for it to be used in our statistical analysis. As for error analysis – only four children had any errors recorded, and only their corpora could therefore be used for this purpose. Each clause was coded as belonging to a single structure. Consequently, clauses of the form SVO, for example, were not coded as instances of SV or VO as well.

Variables associated with the structures

Word-order errors. Each clause is coded as having a canonical or a non-canonical word order.[2] Word order is considered canonical if the subject

[2] We do not go into issues of morphology is this paper. Hebrew is a language with a root and pattern (or *binyan*, in the case of verbs) morphology, and a verb's pattern allows one to predict, at least to some extent, the argument structure of the verb. However, this study is not aimed at understanding how children learn which argument structure each verb has, but rather, once this has been learned, we are interested in following the RATE of learning.

TABLE 2. *Examples of clauses constructed in each of the structures*[a]

1	**subject+verb** Bareket, age 1;4.9: *Aba halax*, Daddy go-3SG-MS-PT 'Daddy went'	**SV**
2	**verb+direct object** Lior, age 1;11.1: *Rotse et ze*, want-SG-MS-PR ACC this '[I] want this'	**VO**
3	**verb+indirect object** Tal, age 1;10.28: *Ten le-maya*, give-2SG-MS-IMP to-maya 'Give Maya'	**VI**
4	**verb+obligatory adjunct** Tal, age 1;9.10: *Lexi mi-po*, go-2SG-FM-IMP from-here 'Go away'	**VA**
5	**subject+verb+direct object [non-canonical order: VSO]** Ofer, age 2;1.2 *Oxel Bobo regel Ofer*, eat-SG-MS-PR Bobo (doll's name) leg Ofer 'Bobo is eating Ofer's leg'	**SVO**
6	**subject+verb+indirect object** Naomi, age 2;0.24: *Ima taazor lax!*, Mommy help-3SG-FM-FUT to-you 'Mommy will help you!' (a request for help from Mommy)	**SVI**
7	**subject+verb+obligatory adjunct [non-canonical order: VSVA]** Ofer, age 1;11.4: *Halax aba … halakh … haxutsa*. Go-3SG-MS-PT Daddy go–3SG-MS-PT outside 'Daddy went outside'	**SVA**
8	**verb+direct object+indirect object** Bareket, age 1;9.28: *Tni li lehikanes*, let-2SG-FM-IMP to-me enter-INF 'Let me enter'	**VOI**
9	**verb+direct object+obligatory adjunct** Shuli, age 2;0.4: *Lasim et ze kan*, put-INF ACC this here 'Put this here'	**VOA**
10	**verb+indirect object+obligatory adjunct** Shuli, age 2;1.14: *Koev li kan ba-yadayim*, hurt-SG-MS-PR to-me here in-the-hands 'My hands hurt here'	**VIA**
11	**subject+verb+direct object+indirect object [non-canonical order: SIVO]** Shuli, age 1;11.10: *Tami Shuli natan sukarya*, Tammy Shuli give-3SG-MS-PT candy 'Tammy gave Shuli a candy'	**SVOI**
12	**subject+verb+direct object+obligatory adjunct** Ofer, age 2;3.25: *Ha-ish hixnis yad letox ha-helikopter bifnim*, the-man put-in-3SG-MS-PT hand into the-helicopter inside 'The man put his hand inside the helicopter'	**SVOA**
13	**subject+verb+indirect object+obligatory adjunct** Naomi, age 2;3.12: *Ze koev li po*, this hurt-SG-MS-PR to-me here 'This hurts here'	**SVIA**
14	**verb+direct object+indirect object+obligatory adjunct** Naomi, age 2;4.10: *Nasim lo trufa ba-rosh*, put-1PL-FUT to-him medicine in-the-head 'We'll put medicine on his head'	**VOIA**
15	**subject+verb+direct object+indirect object+obligatory adjunct** Naomi, age 2;4.10: *Ima yavi li oto, et haxalav, hena*, Mommy bring-3SG-MS-FUT to-me ACC-3SG-MS, ACC the-milk, to-here 'Mommy will bring me it, the milk, over here'	**SVOIA**

a: Source: Table 2 in Keren-Portnoy (2006). The orthography is conventional and does not attempt to be a faithful phonetic description of the children's pronunciation. Inflected verb forms merely approximate a possible target form.

precedes the verb and the other arguments follow the verb, with the internal order among them immaterial. Immediate repetition of a word does not affect canonicity, but repetition of an argument or of the verb in two or more different locations in the clause is considered non-canonical. Although word order in Modern Hebrew is basically SVO (Ben-Horin, 1976; Berman, 1990; 1994; Dromi & Berman, 1986; Glinert, 1989; Ravid, 1995; Ziv, 1976; 1995), it may be altered due to pragmatic considerations (topicalization, focalization and presentation), causing, for example, the verb to precede the subject (Giora, 1981; Givón, 1976; Glinert, 1989; Ravid, 1995). We assume (as do, for instance, Lieven, Behrens, Speares & Tomasello, 2003) that such pragmatically driven variation in word order is not mastered at the early stage on which we are focusing, and clauses with non-canonical word order in the corpora, at least in the early stages, are authentic errors. Keren-Portnoy (2002) documented a decline in the proportion of non-canonical word-order errors as learning progresses. Such a decline would only be expected if, indeed, most cases of early non-canonical word order were a result of lack of knowledge, rather than a result of more sophisticated knowledge. Consequently all non-canonical clauses were coded as cases of word-order errors. Presentational clauses which have a VS word order even in very early child productions (as evidenced in our data), and which contain a specific set of verbs (e.g. *nafal* 'fell') were excluded from analyses concerning word-order errors.[3] A more detailed description of the coding schema and problems in coding can be seen in Keren-Portnoy (2002; 2006).

Age of acquisition. The age of acquisition (AoA below) of each VERB in a given structure is the age at which the first clause was produced by the child using that verb in the given structure. The age of acquisition of a STRUCTURE is the age at which the first clause was produced in the structure. This differs from the definition of age of acquisition used in Keren-Portnoy (2006), where the age of acquisition was defined as the age at which the first clause with a CANONICAL WORD ORDER was produced. This criterion was changed in order not to create a dependency between the rate of learning and number of errors in a structure (see also discussion of the results for Hypothesis 1). AoA as defined here marks the beginning of the acquisition process. This by no means implies that learning is complete and acquisition has ended by the production of a single clause.

[3] In children's corpora, unlike in adults' usage, it is mostly verbs denoting 'accidents' (falling, breaking, etc.) which are used in this manner. This is a well-defined set of verbs that often participate in VS sentences in child (Berman, 1982) as well as in adult Hebrew.

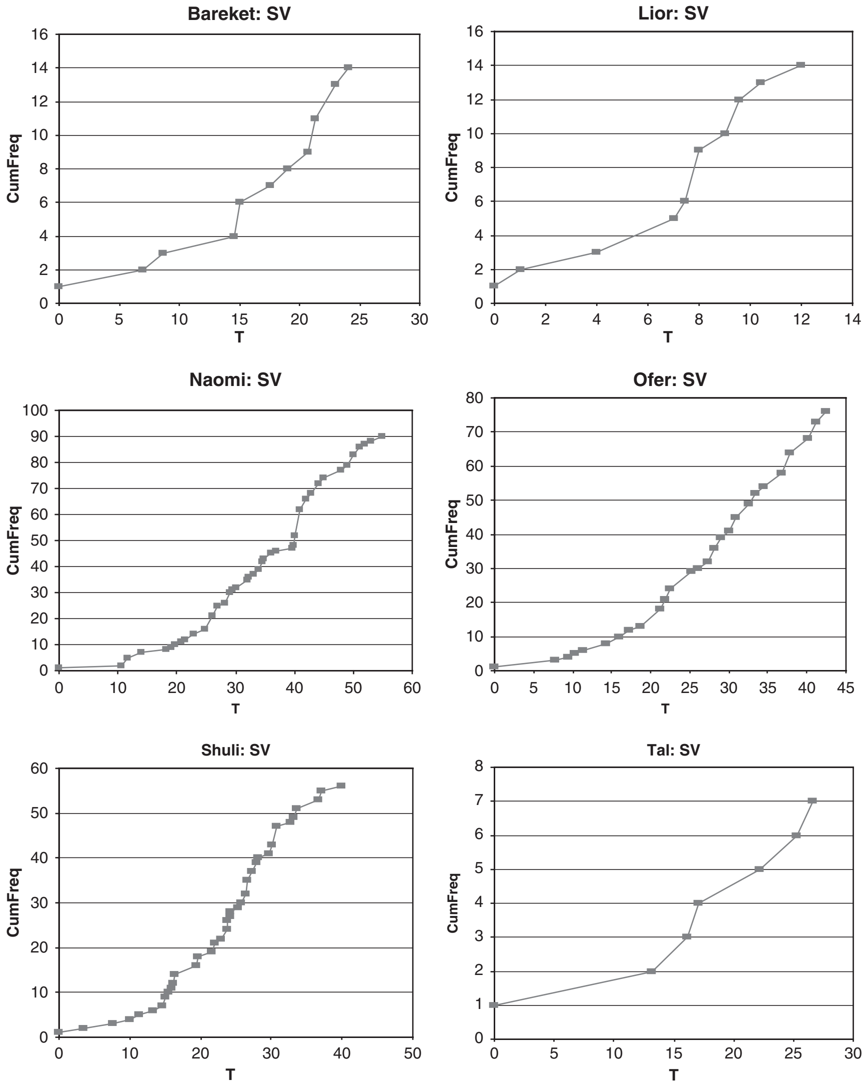

Fig. 1. Learning curves for the structure SV in all of the corpora.

Secondary data

In the primary data described above, each case is a clause produced by a child. Based on this primary database, we created a SECONDARY DATABASE, whose cases are the forty structures in the six corpora for which a learning curve was constructed (see Figure 1 and Figure 2 for examples in two

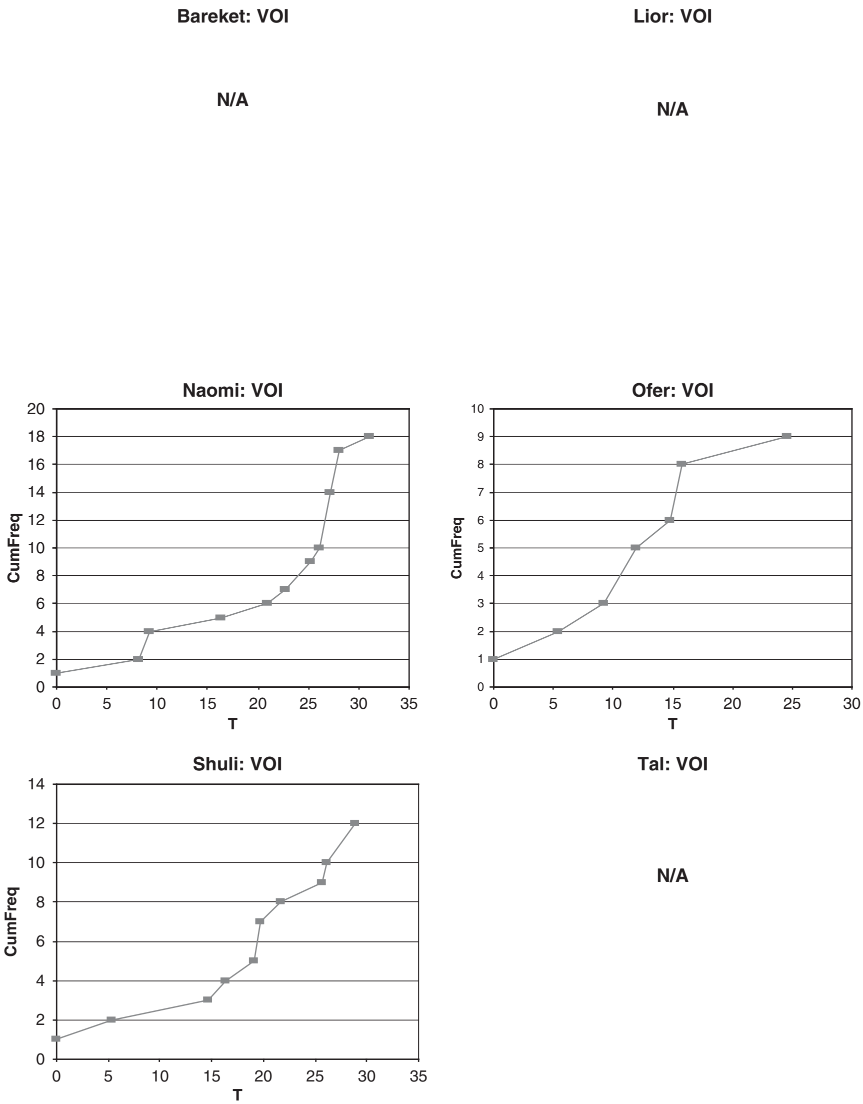

Fig. 2. Learning curves for the structure VOI. (NOTE: The VOI structure was not found in three of the corpora.)

structures), plotting the cumulative number of verbs acquired by the child in that structure as a function of the age of acquisition. Regressions using primary data are referred to as PRIMARY REGRESSIONS. Regressions using secondary data, i.e. in which the data are values obtained from entire structures, are referred to as SECONDARY REGRESSIONS (all secondary regressions reported in this paper, i.e. the regressions reported in Table 3, Table 4 and Table 6 are weighted regressions, see below for definition. The primary regressions reported in Table 5 are unweighted). To estimate a structure's degree of acceleration using a regression at least four points, four different dates in which new verbs were learned in that structure, are needed. The estimate of the degree of acceleration, the quadratic coefficient of the regression fitted to the learning curve, labelled *B2*, is also interpreted as a measure of the difficulty of acquisition of a structure. This is explained below. Table A1 in Appendix A presents the secondary data in full.

STATISTICAL METHODS

The acceleration estimates

The estimates of the quadratic coefficients of the primary regressions described above ('Secondary data') (see Table A1 in Appendix A) were obtained from the regressed learning curves of each of the structures produced by each child. These regression equations have the following form:

$$y = b_0 + b_1 T + b_2 T^2 + \varepsilon, \qquad (1)$$

where y is the cumulative number of verbs produced in the structure, T the number of weeks that have elapsed from the start of the structure's acquisition (see the definition of T below) and ε is a random disturbance or error. The estimator of b_2 is the measure of acceleration, which, we claim, indicates the difficulty of acquiring a given structure. When used as a variable in a secondary regression, the estimator of b_2 is denoted *B2*. The other parameters are of no interest and not determinate since they depend on the exact definition of time: if, instead of T as defined below, calendrical time were used, b_1 and b_0 would adjust, but b_2 would remain unchanged. For a description of the mathematical structure of a typical secondary regression, see below.

A basic problem arises with respect to the *B2* variable: different *B2*s are based on samples of very different sizes. Only a small selection of each child's utterances was sampled for this study, and data collection was terminated at a stage that left some structures, especially those that appear late in our corpora, poorly represented. The significance and reliability of the statistical generalizations that we make is lower for these structures than for earlier ones. Thus SV and VO often produce significant *B2*s, while estimators of later structures are often not statistically significant. Observe that sample size, though affecting significance, should not affect a learning curve's rate of acceleration, neither from a mathematical point of view nor from an empirical one. In fact, a previous study concerned with the acceleration of VO learning curves used only the earliest six verbs from each corpus, and its results still show quite strong acceleration (Ninio, 2005a).

As a result of the differing sample sizes, the samples from which we obtained our estimates are of unequal variance. The technical term for this phenomenon is HETEROSCEDASTICITY. Heteroscedasticity leads to inefficient estimates in those regressions in which *B2* is a predicted variable (Greene, 2000: 499ff.). Weighted regressions, a tool used in other disciplines which face similar problems, e.g. econometrics or meteorology, are therefore used in all secondary regressions.

The statistical model for the secondary regressions

The following is a typical secondary regression:

$$B2 = Const. + \beta_z * z + \beta_n Child + v, \qquad (2)$$

where *B2* is the estimator of b_2 from regression (1), and z any other variable that may be included in the regression, such as *Structure-AoA*, the age (in weeks) in which the structure was first used in the corpus. *Child* is a so-called dummy variable explained below. Thus the betas indicate the contribution of the predictor variables to the degree of acceleration, represented by *B2*.

The use of weighted regressions

The regression model assumes that all observations are drawn from the same distribution, and in particular, that if there are errors in the points for which the regression is computed, all these errors belong to the same distribution and the variance of all the errors at each point is equal. This, as mentioned above, does not hold for the secondary equations, if only because the estimated *B2*s were obtained from samples of highly varied sizes.[4] Although ordinary least squares (OLS) regressions run on such data are unbiased, they are inefficient, that is, their variance is larger than the data would permit, and the estimates of the standard errors are likely to be erroneous (Greene, 2000: 503). To equalize the variances and obtain efficient estimators we use weighted regression. The value of the degrees of freedom was used as the weight in secondary regressions reported below. This is based on the assumption that all these estimators belong to the same distribution, whose variance depends on the inverse of the degrees of freedom of the relevant estimator (Greene, 2000: 514f.). To equalize the variance we

[4] A critical analysis of regression when variances are unequal can be found in Ryan (1997: 60ff.). Ryan states that weighted regression should be used in such cases, but points out that various ways of weighting may be unsatisfactory when the samples are small.

therefore have to multiply by the degrees of freedom, that is, where *B2* is involved, by $n-3$, where n is the number of dates in which new verbs were learned in the structure (the number of datapoints in the learning curve), and three are the degrees of freedom used by Regression (1). An alternative method, weighting by the precision of the estimates, which is the inverse of the variance, $precision = 1/(SE)^2$, seems to us less reliable, due to the poor accuracy of the estimated standard error (SE) for the regressions with few points. Furthermore, with a small numbers of observations per structure, any form of weighting may in some instances produce a worse outcome than OLS, as an anonymous reviewer commented. We have therefore run all three types of regression, weighted by df and precision as well as unweighted OLS, but only the df-weighted ones are reported in this paper. Full results can be obtained from the authors.

The child-corpus dummy variables

An additional problem in the data is that the structures were not sampled independently, and structures originating in the same corpus may be more similar to each other than to those sampled from other corpora. In all the secondary regressions reported in the text, dummy variables which stand for the corpus in which the structures originated are therefore used. A dummy variable is a device that is often used when groups of datapoints have different sources, thus accounting for the lack of independence in sampling within each group. In our case the corpus variable *Child*, where *Child* = *Bareket*, *Lior*, …, *Tal* represents the idiosyncratic vertical shift of the regression curve of each child. Thus, if one child progresses more slowly and acquires syntax at an older age, her coefficient would be positive, while the vertical shift for a child with a faster learning rate would be negative.

The variables used in the regressions

Variables used in several regressions are listed below. Variables used in a single regression are explained where the relevant regression is introduced.

Time related variables:

- *Structure-AoA*: The age, in weeks, at which the first verb has been used in a given structure. This variable is a measure of the length of time that the structure has been in use in a child's corpus.
- *T*: *T* = *Verb-AoA*–*Structure-AoA*, where *Verb-AoA* is the age, in weeks, in which a given verb has joined a structure. Thus *T* is the time in weeks between the acquisition of a given verb in a structure and the acquisition of the very first verb in that same structure in a given child's corpus. For the first verb in any structure $T = 0$.

- T^2:T squared. The coefficient of T^2 in the learning curve, b_2, is the measure of acceleration of the learning curve, which is shown below to be a measure of the difficulty of acquiring a structure.

Other variables:

- *Error proportion*: The proportion of non-canonical clauses out of the first twenty clauses in a structure. For structures in which there were fewer than twenty, the proportion of non-canonical clauses out of all clauses. Two corpora in which no errors were recorded as well as all structures with fewer than four clauses were excluded from the analyses concerned with errors. This is one of the variables which we take to indicate degree of difficulty of acquisition.
- *Cumulative frequency*: Cumulative frequency of different verbs used in a given structure at a given point of time. The learning curves are based on the primary data and plot *Cumulative Frequency* as a function of T.
- *B2*: The estimator of the rate of acceleration of the learning curve of each structure, which we take as an indicator of its difficulty of acquisition. *B2* is the estimator of b_2, the coefficient of the quadratic term T^2, in the primary regressions of type (1).
- *Child*: Child-corpus dummies. See above.
- *Structure-rank*: The rank order of acquisition of a structure in a child's corpus, an alternative measure for the length of time that the structure has been in use in a child's corpus.

REGRESSION RESULTS AND DISCUSSION

We begin with a short description of the acceleration data, and then move on to report the results of the regression analyses.

Table A1 in Appendix A presents the estimates of the *B2*s, the acceleration parameters of the learning curves. Only forty structures contain enough datapoints for regression analysis. Of these, thirty-one or 78% have a positive *B2*, and nine, 23%, are non-accelerating.[5] Structures containing an Adjunct (of which VA, SVA, VOA and SVOA appear in the table) have a much lower proportion of significantly accelerating structures (42%) than the other structures (93%). The reason for this is still unclear, and merits further investigation. The significance of *B2* declines as the number of

[5] Since speech was sampled for only thirty minutes each week, it is possible that some important points in a sparse part of the curve were missed. It is the early period of acquisition of any structure that contains very few points, and if all of these are missed, an event of low yet positive probability, a curve that is convex may appear straight. This may be a possible explanation for the fact that some of the learning curves were judged non-accelerating.

verbs per structure declines: 81% of the *B2* values for structures with an average of over twenty-nine verbs per structure are significant, compared to only 51% of structures with fewer verbs. Three corpora have only three estimated learning curves each, all accelerating and all but two significantly so. Naomi's corpus includes eleven curves, eight of which are significantly accelerating, and three of which are not significantly non-accelerating. Shuli's nine include eight accelerating curves of which five are significantly accelerating, and a single significantly non-accelerating one. The outlier is Ofer, of whose eleven curves six are accelerating and five non-accelerating, three of each significantly so; even VO, an early structure with many verbs, is significantly non-accelerating in Ofer's case.

Hypothesis 1 – Errors and acceleration

Our first task is to establish that the acceleration of regression (1), measured by *B2*, is indeed positively correlated with the proportion of errors among the earliest clauses in the structure, thus justifying the use of these two variables as measures of the degree of difficulty encountered in the acquisition of a structure. *B2* was regressed on *Error proportion* as the independent variable. Note that these two variables are not only conceptually independent of each other, but that each of them originates from a different set of clauses or datapoints: the learning curve, whose acceleration is given by *B2*, is based on the first use of each new verb in the given structure over the whole recording period; the proportion of errors is computed for the first twenty clauses in each structure, where any verb may appear more than once (see details in previous section). Thus the number of errors cannot affect the shape of the learning curve and *Error proportion* and *B2* are statistically independent.

Table 3 presents the results of the regression. The proportion of errors is a significant explanatory variable of *B2*.[6] This supports the assumption that both are affected by the same latent variable, the degree of difficulty of acquiring a structure. We thus have statistical support for the claim that acceleration, like errors, is a symptom of friction, or difficulties at the start of the acquisition of a new syntactic structure, and that the rate of acceleration can serve as a measure of the degree of difficulty.

Hypothesis 2—The evolution of errors (Facilitation I)

Hypothesis 2 refers to the evolution of errors over time. The dependent variable in Table 4 is *Error proportion*, the proportion of errors in the first

[6] Observe that R^2 is just over 20%, i.e. that the included independent variables (which include also the unreported *Child-proxy dummies*) explain but a small fraction of the total variance. Nevertheless, Error proportion is a significant predictor, with $p = 0{\cdot}03$.

TABLE 3. *Proportion of errors and acceleration*[a]

Dependent variable: B2	
Error proportion	0·054** (2·269)
Child corpus dummies	Yes
N	34
R^2	0·212

a: df-weighted OLS. Variables omitted from table: *Child corpus dummies*. Reported values are the regression coefficients of the relevant variables; *t*-values in parentheses; ** significant at 0·05.

TABLE 4. *Structure's acquisition age and proportion of errors* (*df weighted*)[a]

Dependent variable: Error proportion		
Structure-AoA	−0·009*** (−3·289)	
Structure-rank		−0·021** (−2·634)
Child corpus dummies	Yes	Yes
N	34	34
R^2	0·337	0·266

a: df-weighted OLS. Variables omitted from table: *Child corpus dummies*. Reported values are the regression coefficients of the relevant variables; *t*-values in parentheses; ** significant at 0·05, *** at 0·01.

clauses produced in each structure. The independent variables are, alternatively, *Structure-AoA* or *Structure-rank*. The results are striking: the age or order at which a structure is learned are highly significant and negative explanatory variables.[7] In other words, the later a structure enters the child's vocabulary, the lower the probability of word-order errors among its earliest clauses. This decline in errors is all the more impressive as it is detected despite a probable overestimation of the errors in later structures – see above in 'Variables associated with the structures', subsection 'Word-order errors'.

Hypothesis 3—The evolution of acceleration (*Facilitation II*)

According to Hypothesis 3, acceleration should decline over time. We test this hypothesis using two complementary tests. The first test was run on the

[7] See footnote 6: a larger part of the variance in explained by these regressions (over a third or a quarter), and they are more significant ($p < 0·001$ and $p < 0·013$, respectively) for Structure-AoA and for Structure-rank than the regression of Table 3.

TABLE 5. *Time of acquisition of new verbs and acceleration (primary data)*[a]

Dependent variable: Cumulative frequency	Bareket (1)	Lior (2)	Naomi (3)	Ofer (4)	Shuli (5)	Tal (6)	All (7)
T	0·100	0·326	0·094	−0·120	0·494***	−0·058	0·161
	(0·993)	(0·942)	(0·472)	(−0·471)	(3·329)	(−0·650)	(1·482)
T^2	0·018***	0·030*	0·026#	0·034#	0·023#	0·023#	0·025#
	(4·120)	(2·019)	(6·593)	(3·982)	(6·110)	(4·982)	(7·966)
Structure-AoA	0·140**	0·192	0·165	−0·435	0·013	1·033**	0·025
	(2·476)	(0·756)	(1·338)	(−1·313)	(0·269)	(2·801)	(0·256)
Structure-rank*T2	0·001	0·007	−0·002#	−0·001	−0·003***	−0·005***	−0·001**
	(0·487)	(0·594)	(−2·365)	(−0·330)	(−3·140)	(−3·186)	(−2·403)
Child corpus dummies	na	na	na	na	na	na	Yes
N	30	28	198	132	156	21	565
R^2	0·941	0·917	0·982	0·931	0·977	0·951	0·964
D-W	1·894	1·699	1·945	1·850	1·781	1·280	1·930

a: Generalized Least Squares regression. Reported values are the regression coefficients of the relevant variables; *t*-values in parentheses; * significant at 0·10; ** at 0·05, *** at 0·01, # at 0·001; na not applicable. D-W the Durbin-Watson statistic, measures serial correlation, all of which, except Tal's are above d_u, i.e. no indication of serial correlation. Tal's lies between d_U and d_L, not ruling out serial correlation.

primary data, for all structures in each corpus and also for all structures in all corpora combined together. For this purpose we add to T^2 an interaction variable, used only in the regressions reported in Table 5:

- *Structure-rank*T^2*: the interaction term between T^2 and *Structure-rank*. This term assigns weight to the quadratic variable, T^2, in proportion to the order of entry of the structure to a corpus. It measures the change in acceleration of a learning curve in relation to the order in which a structure is learned: the greater the (absolute) value of the coefficient of this interaction variable, the greater is the change in slope as acquisition rank increases. When acceleration declines over time, it is negative.

The regressions of Table 5 were run on the primary data, using all the datapoints which form the learning curves (i.e. all the dates at which new verbs joined a structure), with *Cumulative frequency* as the dependent variable (altogether 565 datapoints were available from all six corpora).[8]

[8] We had originally used ordinary least squares (OLS) to estimate this equation, but discovered that this leads to inefficient estimates due to heteroscedasticity and serial correlation. Inefficiency means that the estimators of goodness-of-fit, such as *t*-values, are overstated. Heteroscedasticity – variance of the deviations from the fitted curve

TABLE 6. *Order or normalized age at structure's acquisition and acceleration: secondary data*[a]

Dependent variable: B2		
	(1)	(2)
Structure-AoA/100	−0·075* (−1·932)	
Structure-rank/100		−0·183 (−1·593)
Child corpus dummies	Yes	Yes
n	40	40
R^2	0·308	0·285

a: df-Weighted OLS. Reported values are the regression coefficients of the relevant variables; *t*-values in parentheses; * significant at 0·10.

We initially focus on the coefficient of T^2. It is positive everywhere, signifying that the learning curves are accelerating, thus supporting the general finding that facilitation (of later verbs by earlier ones) occurs within structures. The interaction term *Structure-rank*$*T^2$ is significantly negative in the regression which combines data from all children (column (7)) and is negative in most regressions describing individual children's data, including all three large corpora, and positive in only two regressions, both for small corpora (containing three structures each). We thus obtain support for the hypothesis that acceleration declines the later a structure is learned. Since the interaction terms in the more reliable regressions, those based on the three larger samples, are negative (significantly so in two of the three), we consider this a strong support for the hypothesis of facilitation across structures (the third, based on Ofer's corpus, is negative, but not significantly so. However, Ofer is an outlier in other respects too).

The behaviour of *B2*, the measure of acceleration, as an indicator of the degree of difficulty of learning as acquisition progresses is of special interest. Therefore a second test, which complements the first, was run on secondary data, using *B2* as the predicted variable (see Table 6). The independent variables are two alternative measures of the time of entry of each structure: *Structure-AoA* and *Structure-rank*. The coefficients of both variables are negative and in the case of *Structure-AoA*, the coefficient is also marginally

that is not constant, e.g. increasing as *T* is rising (Greene, 2000: Ch. 12), and serial correlation – deviations that are correlated (Greene, 2000: Ch. 13), can be eliminated by generalized least squares (GLS) and adjustment for first-order autoregression – AR(1). These methods have produced the estimates in the table. Only the estimate for Tal (column 6) MAY still have some serial correlation, and its *t* values may still be overstated.

TABLE 7. *Acceleration among early and late structures*

	Number of structures				
		Early structures		Late structures	
Child	All structures	Total	Accelerating	Total	Accelerating
Bareket	3	1	1	1	1
Lior	3	1	1	1	0
Naomi	11	5	3	5	4
Ofer	10	5	2	5	3
Shuli	9	4	4	4	3
Tal	3	1	1	1	1
Total	39	17	12	17	12

significant ($p=0{\cdot}062$). Again, the hypothesis of acquisition becoming easier gains support as acceleration declines over time.

FURTHER RESULTS: THE FACILITATOR – MATURATION, SYNTACTIC KNOWLEDGE OR KNOWLEDGE OF STRUCTURES?

The evidence presented below shows that most new structures, at least during the period of early acquisition, begin with a slow phase, and the rate of acceleration of a structure's learning curve depends, at least partly, on previous experience with the specific structure in question. In other words, general maturation or general state of knowledge alone cannot explain the pattern of our results.

New structures enter with a slow start

As mentioned above (in 'Regression results and discussion'), in 31 (78%) out of the 40 structures in the corpora for which regressions could be run, the learning curve accelerates (see Table A1 in Appendix A). This tendency continues with later learned structures. Each child's structures were divided into two groups: those learned first and those learned last (in case of an uneven number of structures, the middle structure was not counted). Table 7 tabulates the number of accelerating structures among the earliest and the later learned in each corpus. In all the corpora the number of accelerating structures is very similar among the early and late structures. For two children the number is higher by one for the early structures, and for two children the number is higher by one for the later structures. Altogether there are 12 accelerating curves out of the 17 early structures, and 12 out of the 17 late structures. The tendency for newly learned structures to accelerate does not diminish as learning progresses, although, as shown in the results for Hypothesis 3, the RATE of acceleration does decline.

TABLE 8. *Time from first to twentieth verb in each structure in Naomi, Ofer and Shuli's corpora*[a]

Child		Age (weeks) at:				Duration of acquisition		
		1st verb	10th verb	11th verb	20th verb	First 10 verb	Next 10 verbs	20 verbs[a]
Naomi	SV	83	103	104	109	20	5	26
	VO	84	104	104	111	20	7	27
	SVO	94	110	112	119	16	7	25
	VA	95	112	112	131	17	19	36
	VI	105	116	117	130	11	13	25
	SVI	108	125	125	134	17	9	26
	Mean	–	–	–	–	17	10	28
	SD	–	–	–	–	3	5	4
Ofer	SV	90	106	107	112	16	5	22
	VO	104	106	107	112	3	5	9
	SVO	105	116	116	121	11	5	17
	SVA	101	116	116	133	15	17	32
	Mean	–	–	–	–	11	8	20
	SD	–	–	–	–	6	6	10
Shuli	SV	83	98	99	105	15	6	22
	VO	86	103	105	107	16	2	20
	SVO	91	105	105	110	14	5	19
	Mean	–	–	–	–	15	4	20
	SD	–	–	–	–	1	2	2

a: When there is a lag between the week in which the 10th and the 11th verb were learned, the time it took to learn twenty verbs is longer than the sum of the times it took the first and the second ten verbs.

The rate of learning is a function of the stage of acquisition of a structure

In order to bring all structures onto a comparable scale only structures in which twenty or more verbs have been acquired were used for the following demonstration. As three of the children did not have a sufficient number of verbs in any of their structures, data from only three corpora – Naomi's, Shuli's and Ofer's – were used. For each structure, two subperiods were defined, the first during which the first ten verbs were acquired and the second during which the second ten were acquired. The first period in an accelerating learning curve is characterized by a gentle incline, while the second is faster and steeper (see Table 8 and Figure 3).

The majority of structures in Naomi's corpus, all the structures in Shuli's and half of those in Ofer's took longer to acquire their first ten verbs than their second ten verbs (see the second and third columns from the right in Table 8).[9] Importantly, the rate of entry of verbs into newly acquired

[9] Note that in those structures which do not fit this pattern both periods tend to be of similar lengths, i.e. learning is constant, neither accelerating nor decelerating.

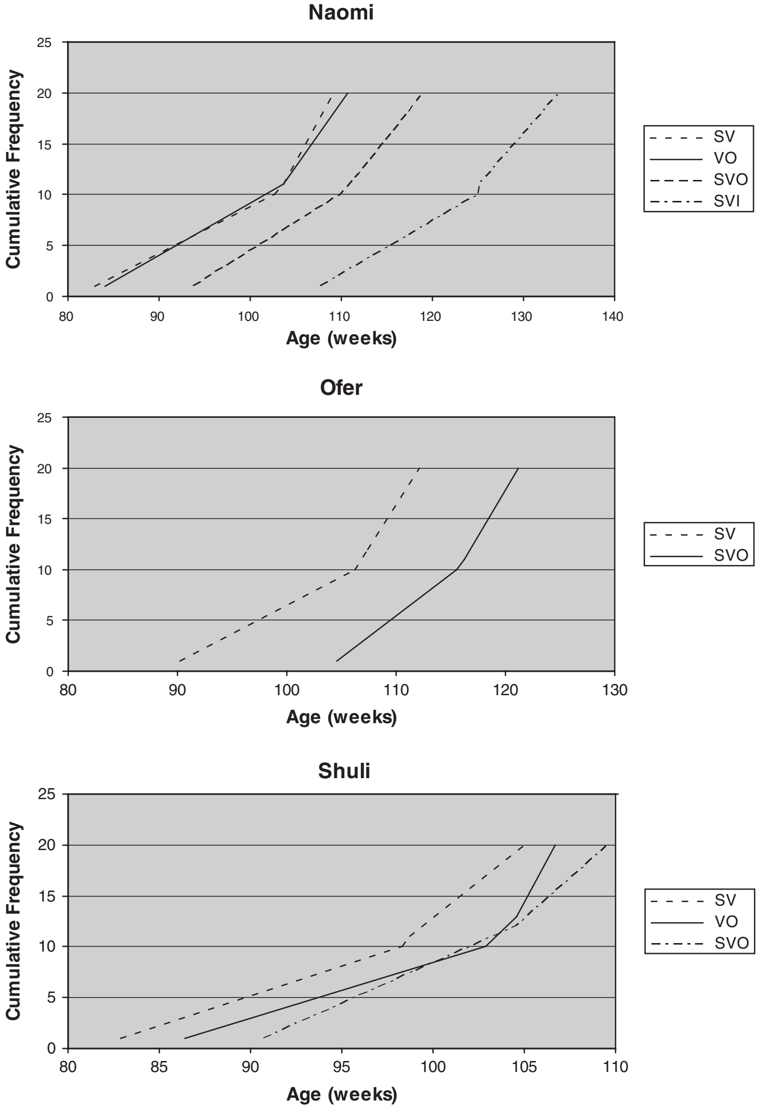

Fig. 3. Schematized learning curves for structures with at least twenty verbs.

structures is slower than the rate of accretion of verbs in previously acquired structures DURING THE SAME TIME PERIOD. This can be seen most clearly in Figure 3, which omits all non-accelerating learning curves.

What Figure 3 shows is that at the very same time that the earlier learned structures are accelerating, new structures are entering at a much slower pace. In Naomi's corpus, during the very same period that one structure is undergoing the slow learning phase, others are already at the fast stage: the slow stage for SV and VO ends at 103 or 104 weeks, at which point the acquisition of new verbs in these structures speeds up, while SVI, a little later at 108 weeks, begins its slow period; it takes 5 and 7 weeks to learn the second ten verbs in SV and VO respectively, and more than twice the time, 17 weeks, for the first ten verbs in the new structure SVI. Figure 3 underscores this finding: the slopes of the second halves of the SV and VO curves are considerably steeper than that of the first part of the SVI curve. The same applies to Ofer's two structures: the SVO structure starts its slow phase almost at the same time as the earlier SV structure starts its second, quicker stage (both around age 105 or 106 weeks). In Shuli's corpus the learning curves of the different structures are much closer together, yet here, too, one can see the same phenomenon: curves beginning at different times show very similar slopes.[10]

In summary, for accelerating structures in all three corpora, one can find later structures which undergo their slow stage while earlier structures have already begun to accelerate. These data suggest that the rate of learning new verbs in a structure is a function of the stage of acquisition of that particular structure rather than of the stage of syntactic development in general or the cognitive system (or the brain) in general. Even for structures whose acquisition starts late, when the child is already capable of fast learning, new structures present new hurdles which must be overcome.

CONCLUSIONS AND GENERAL DISCUSSION

Our main finding in this paper is that facilitation in language acquisition is not limited to intra-structure learning but that powerful inter-structure facilitation can also be documented. We find support for facilitation in the decline of symptoms of difficulty in the learning of new structures as the process of language acquisition progresses. The first symptom is the gradual reduction in word-order errors in early stages of acquisition of new structures as more structures are learned. This is an especially robust finding: non-canonical clauses are still produced in later learned structures, but they are not necessarily the result of a lack of word-order knowledge. Rather, some of them may be the result of pragmatic considerations which the child may have mastered by later stages of acquisition. Yet in spite of these new word-order options available to the children, we still see a fall in

[10] A referee has pointed out that the difference between the acceleration rate for Naomi's VO, SVO and SVI seems marginal, belying the finding that acceleration declines in later structures. This is, of course, true. The claim is not that the difference is easily discernable, but that a difference exists that can be documented by statistical methods, namely regression analysis.

non-canonical word order. This makes the decline even more impressive: it is clear that children base their knowledge of word order in later acquired structures on what they have previously learned through the use of earlier structures.

Our second line of evidence relies on the shape of the learning curves. We find that learning curves become significantly less accelerating over the acquisition process. The change from the slow pace of verb acquisition at first to a more rapid rate later is much more pronounced in the first learned structures.

Furthermore, the two phenomena are related. The rate of acceleration of a structure's learning curve is correlated with the number of errors at the onset of learning a structure. This lends support to the claim that the rate of acceleration, like the proportion of errors, indicates the degree of difficulty at the onset of acquisition of a structure.

Yet the signs of difficulty do not disappear altogether, even for the latest structures. Even later learned structures tend to start their acquisition with accelerating rather than straight learning curves, indicating that each new structure presents new challenges before learning can take off at a more rapid rate. In addition, it has been shown that when, within any corpus, the period of rapid learning for an early acquired structure coincides with the slow period of later structures, the former exhibits a faster rate of growth than the latter.

This finding allows us to respond to the two potential objections to the interpretation of the facilitation phenomenon mentioned in the 'Introduction'. The first suggested that facilitation may be due to general maturation rather than to the accrual of specific linguistic knowledge. The second was that although combining words becomes gradually easier, this has nothing to do with learning new structures but rather with having learned to produce multiword utterances, or, in other words, that structures do not play an essential role in syntax learning.

We claim that maturation is not the whole story, and that actual early knowledge supports and scaffolds later knowledge. The evidence shows that a new hurdle is encountered and some new learning must take place when a new structure is met, regardless of what has already been achieved in others. New structures 'take off' more slowly than older structures which are their coevals, showing that the actual structure is a factor affecting the ease of learning, over and above the possible contributions of age or general syntactic or cognitive advance at the time at which it is learned. Time, the proxy for general development, cannot be the only explanatory element, and variables related to the time or order of entry of specific structures were also significant explanatory factors. These data justify the analysis of learning as a process which to some extent occurs within structures, in addition to showing the propagation of knowledge from previous structures to new ones.

These results are in accordance with a general developmental picture in which newly acquired items are integrated into networks of previously learned information, rather than remaining isolated (see also Ninio, 2003). Things that are similar are stored together and are connected or associated with each other, thus forming a system. That does not mean that there is a preordained or pre-planned system into which items must fit, but rather that analogy is a very strong method of organization. Facilitation between early and late structures fits well into this view of a system that is constantly evolving.

REFERENCES

Abbot-Smith, K. F. & Behrens, H. (2006). How known structures influence the acquisition of new structures: The German periphrastic passive and future structures. *Cognitive Science* **30**, 995–1026.

Ben-Horin, G. (1976). Aspects of syntactic preposing in spoken Hebrew. In P. Cole (ed.), *Studies in Modern Hebrew syntax and semantics: The transformational-generative approach*, 193–207. Amsterdam: North-Holland Publishing Company.

Berman, R. A. (1982). Dative marking of the affectee role: Data from Modern Hebrew. *Hebrew Annual Review* **6**, 35–59.

Berman, R. A. (1990). On acquiring an (S)VO language: Subjectless sentences in children's Hebrew. *Linguistics* **28**, 1135–66.

Berman, R. A. (1994). Developmental perspectives on transitivity: A confluence of cues. In Y. Levy (ed.), *Other children, other languages*, 189–241. Hillsdale, NJ: Lawrence Erlbaum Associates.

Braine, M. D. S. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development* **41** (1, Serial No. 164).

Bybee, J. (2001). *Phonlogy and language use*. Cambridge: Cambridge University Press.

Dromi, E. & Berman, R. A. (1986). Language-specific and language-general in developing syntax. *Journal of Child Language* **13**, 371–87.

Giora, R. (1981). Sentence ordering as a text dependent phenomenon. In S. Blum Kolka, Y. Tobin & R. Nir (eds). *Studies in discourse analysis*, 263–302. Jerusalem: Academon [in Hebrew].

Givón, T. (1976). On the VS word order in Israeli Hebrew: Pragmatics and typological change. In P. Cole (ed.), *Studies in Modern Hebrew syntax and semantics: The transformational-generative approach*, 153–181. Amsterdam: North-Holland Publishing Company.

Glinert, L. (1989). *The grammar of Modern Hebrew*. Cambridge: Cambridge University Press.

Greene, William H. (2000), *Econometric analysis*, 4th edn. Upper Saddle River, NJ: Prentice Hall.

Hahn, U. & Chater, N. (1998). Similarity and rules: Distinct? Exhaustive? Empirically distinguishable? *Cognition* **65**, 197–230.

Homa, D. & Chambliss, D. (1975). The relative contributions of common and distinctive information on the abstraction from ill-defined categories. *Journal of Experimental Psychology: Human Learning and Memory* **104**, 351–59.

Keren-Portnoy, T. (2002). The development of the verb category in the grammars of children learning Hebrew as their first language. PhD dissertation, Hebrew University of Jerusalem.

Keren-Portnoy, T. (2006). Facilitation and practice in verb acquisition. *Journal of Child Language* **33**, 487–518.

Keren-Portnoy, T. (in preparation). Long is easy, short is hard.

Kiekhoefer, K. (2002). The acquisition of the ditransitive structure. Paper presented at the 9th International Congress for the Study of Child Language and the Symposium on Research in Child Language Disorders, Madison, WI, July.

Lieven, E., Behrens, H., Speares, J. & Tomasello, M. (2003). Early syntactic creativity: A usage-based approach. *Journal of Child Language* **30**, 333–70.

Medin, D. L. & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review* **35**, 207–238.

Miller, J. L. (1994). On the internal structure of phonetic categories: A progress report. *Cognition* **50**, 271–85.

Murphy, K., McKone, E. & Slee, J. (2003). Dissociations between implicit and explicit memory in children: The role of strategic processing and the knowledge base. *Journal of Experimental Child Psychology* **84**, 124–65.

Ninio, A. (1999a). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language* **26**, 619–53.

Ninio, A. (1999b). Model learning in syntactic development: Intransitive verbs. *The International Journal of Bilingualism* **3**, 111–31.

Ninio, A. (2003). No verb is an island: Negative evidence on the Verb Island Hypothesis. *Psychology of Language and Communication* **7**, 3–21.

Ninio, A. (2005a). Testing the role of semantic similarity in syntactic development. *Journal of Child Language* **32**, 35–61.

Ninio, A. (2005b). Accelerated learning without semantic similarity: Indirect objects. *Cognitive Linguistics* **16**, 531–56.

Ninio, A. (2006). *Language and the learning curve: A new theory of syntactic development*. Oxford: Oxford University Press.

Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **14**, 54–65.

Ravid, D. (1995). *Language change in child and adult Hebrew: A psycholinguistic perspective*. Oxford: Oxford University Press.

Ryan, Thomas P. (1997). *Modern regression methods*. New York: John Wiley & Sons, Inc.

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.

Tomasello, M. (2000). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences* **4**, 156–63.

Vihman, M. M. (1999). The transition to grammar in a bilingual child: Positional patterns, model learning, and relational words. *The International Journal of Bilingualism* **3**, 267–301.

Ziv, Y. (1976). On the reanalysis of grammatical terms in Hebrew possessive structures. In P. Cole (ed.), *Studies in Modern Hebrew syntax and semantics: The transformational-generative approach*, 129–52. Amsterdam: North-Holland Publishing Company.

Ziv, Y. (1995). A special VS structure in Modern Hebrew. *Hebrew Linguistics* **39**, 29–39 [in Hebrew].

APPENDIX A: B2 BY CORPUS AND STRUCTURE

TABLE A1. *Age at acquisition, quadratic coefficients and number of verbs learned in each of the structures*

Child's name	Structure	Error proportion	Structure-AoA	Structure-rank	B2	SE of B2	B2 significance	No. of verbs acquired	df = Number of dates − 3
Bareket	SV	0·15	72·1	1	0·027#	0·004	$p<0·001$	14	8
	VA	0	77·0	2	0·023***	0·004	0·002	11	6
	VO	0·05	80·9	3	0·022	0·020	0·333	9	4
Lior	VO	n/a	100·0	1	0·063***	0·015	0·003	15	8
	SV	n/a	106·0	2	0·073*	0·035	0·078	14	7
	SVO	n/a	109·0	3	0·000	0·055	0·999	4	1
Naomi	SV	0·15	83·0	1	0·029#	0·003	$p<0·001$	90	37
	VO	0·35	84·1	2	0·008#	0·002	$p<0·001$	59	34
	SVO	0·25	93·9	3	0·025#	0·002	$p<0·001$	51	26
	VA	0·05	94·7	4	−0·003	0·003	0·369	20	12
	VOA	0·15	100·0	5	−0·002	0·002	0·452	7	4
	VI	0	104·9	6	0·016**	0·007	0·025	32	14
	SVA	0	106·7	7·5	−0·001	0·002	0·809	15	8
	SVOI	0·05	106·7	7·5	0·032#	0·002	$p<0·001$	13	5
	VOI	0·05	106·9	9	0·026***	0·008	0·01	18	8
	SVI	0	107·7	10	0·018#	0·003	$p<0·001$	25	16
	SVOA	0	108·7	11	0·058**	0·002	0·025	4	1
Ofer	SV	0·5	90·3	1	0·041#	0·003	$p<0·001$	76	24
	VA	0·05	99·7	2	−0·024#	0·004	$p<0·001$	16	9
	SVA	0·2	100·6	3	0·000	0·007	0·976	20	8
	VO	0·05	103·6	5	−0·035#	0·004	$p<0·001$	46	16
	VI	0	103·6	5	0·013	0·007	0·116	13	6
	VOI	0·1	103·6	5	−0·001	0·006	0·841	9	4
	SVO	0·3	104·6	7	0·043***	0·010	0·001	47	14
	SVI	0	105·3	8	0·012**	0·004	0·024	19	9
	SVOI	0·1	106·3	9	0·003	0·003	0·386	11	7
	VOA	0·4	110·7	10	−0·114	0·078	0·381	6	1
	SVOA	0	121·1	11	−0·091*	0·007	0·052	6	1

TABLE A1. (*Cont.*)

Child's name	Structure	Error proportion	Structure-AoA	Structure-rank	B2	SE of B2	B2 significance	No. of verbs acquired	df = Number of dates − 3
Shuli	SV	0·15	82·9	1	0·030#	0·005	$p < 0.001$	56	34
	VO	0	86·4	2	0·020#	0·005	$p < 0.001$	38	21
	SVO	0	90·7	3	0·033#	0·005	$p < 0.001$	35	19
	VOI	0·1	93·9	4	0·014***	0·003	0·002	12	7
	VI	0·05	99·1	5	0·021***	0·004	0·001	15	9
	VA	0·15	99·9	6	0·008	0·006	0·219	10	7
	SVI	0·35	100·7	7	0·026	0·022	0·254	18	9
	SVOI	0·05	101·4	8	0·005	0·016	0·739	17	9
	SVA	0·15	102·9	9	−0·007**	0·002	0·016	11	8
Tal	VO	0	92·9	1·5	0·016#	0·002	$p < 0.001$	12	6
	VA	0	92·9	1·5	0·011*	0·003	0·076	6	2
	SV	0	94·7	3	0·008**	0·002	0·014	7	4

Legend: * significant at 0·10; ** at 0·05, *** at 0·01, # at 0·001.