DE GRUYTER MOUTON DOI 10.1515/gcla-2013-0008 — GCLA 2013; 1: 129–150

Arne Zeschel

Constructional 'scene encoding' and acquisition: Mothers' use of argument structure constructions in English child-directed speech[1]

Abstract: Construction-based language models assume that grammar is meaningful and learnable from experience. Focusing on five of the most elementary argument structure constructions of English, a large-scale corpus study of child-directed speech (CDS) investigates exactly which meanings/functions are associated with these patterns in CDS, and whether they are indeed specially indicated to children by their caretakers (as suggested by previous research, cf. Goldberg, Casenhiser and Sethuraman 2004). Collostructional analysis (Stefanowitsch and Gries 2003) is employed to uncover significantly attracted verb-construction combinations, and attracted pairs are classified semantically in order to systematise the attested usage patterns of the target constructions. The results indicate that the structure of the input may aid learners in making the right generalisations about constructional usage patterns, but such scaffolding is not strictly necessary for construction learning: not all argument structure constructions are coherently semanticised to the same extent (in the sense that they designate a single schematic event type of the kind envisioned in Goldberg's [1995] 'scene encoding hypothesis'), and they also differ in the extent to which individual semantic subtypes predominate in learners' input.

Keywords: argument structure, constructional meaning, child-directed speech, collostructional analysis

Arne Zeschel: Institut für Deutsche Sprache. E-mail: zeschel@ids-mannheim.de

1 This research was supported by the ITALK project (EU ICT grant n. 214668). I thank Kerstin Fischer and the audiences at DGKL 4 in Bremen and at the conference 'Construction Grammar: New perspectives for the study of German and English' in Kiel for helpful comments and discussion.

1 Introduction

According to Goldberg's (1995: 39) 'scene encoding hypothesis', argument structure constructions such as the ditransitive or certain complex transitives encode 'humanly relevant scenes' like 'successful transfer' (1a) or 'something causing something to change location' (1b):

(1) a. *Sally baked her sister a cake.*
b. *Frank sneezed the napkin off the table.*
(Goldberg 1995)

Goldberg argues that constructional meanings like the 'transfer' interpretation of (1a) and the 'caused motion' meaning of (1b) are learned by generalising over concrete memorised exemplars of relevant expressions: over time, salient implications that keep recurring across a great many uses of the construction in question come to be associated with the schematic structural template itself (Goldberg 1999).

Furthermore, Goldberg and colleagues argue that learners are *aided* in finding the right semantic generalisations by the way these structures are presented to children during language acquisition. Specifically, in a corpus study of child-directed speech (CDS), Goldberg et al. (2004) found that *go* accounted for 39% of mothers' uses of the intransitive motion construction ([SUBJ V LOC]), *put* for 38% of mothers' uses of its transitive counterpart ([SUBJ V OBJ LOC]), and *give* for 20% of mothers' uses of the ditransitive construction ([SUBJ V OBJ-1 OBJ-2]) when speaking to their children. A parallel analysis of the children's speech in the same corpus revealed that the exact same three verbs were also the top frequent choices in learners' own productions of these constructions. Apart from the close input-output correspondence, what is particularly interesting about these findings is that the one verb which mothers used disproportionately often in each construction was always the closest *lexical* semantic equivalent to the central meaning of the construction itself on Goldberg's (1995) analysis ('motion', 'caused motion' and 'transfer', respectively). On the basis of these findings, the authors suggest that "the dominance of a single verb in the construction facilitates the association of the meaning of the verb in the construction with the construction itself, allowing learners to get a 'fix' on the construction's meaning" (Goldberg et al. 2004: 308).

On the other hand, a unified 'central' constructional meaning may not always be obvious even to competent adult speakers (be they parents or linguists). For instance, Goldberg characterises the central sense of the simple transitive construction as a scenario involving "a volitional actor affecting an inanimate patient – a causative event". At the same time, she acknowledges that there also exist "clusters of cases which are not instances of the general semantic template" thus defined (Goldberg 1995: 118). Indeed there are a number of very common transitive constructions (for instance transitives with verbs of possession, cognition and perception) that do not coerce the ascribed interpretation on speakers' construal of the encoded scene. Moreover, Sethuraman and Goodman (2004) found that caretakers' use of the transitive construction in CDS does not follow the characteristic pattern described above either: in their data, mothers' top frequent transitive verb *do* occurs mainly in expressions like *do it* and *do what*? that are not exactly typical transitives from a semantic point of view; the proportional frequency of *do* in the transitive construction is markedly lower than that of *go* and *put* in the intransitive/caused motion construction; the next most frequent items on the list are not transparently related to *do* such that they would support a single overarching generalisation; and finally, children's own productions of the transitive construction do not mirror the frequency ranking of transitive verbs in the input. In short, both the elusive semantics and the quantitative usage patterns of transitive constructions in CDS are more difficult to reconcile with the above vision of inherently meaningful constructions and the way these meanings are indicated to children.

What is more, also those constructions that *are* intuitively more coherent semantically are not in fact restricted to the assumed central meaning. Rather, just like in lexical semantics, the formal pole of these constructions may be associated with a more or less complex category of distinct senses that are related through family resemblance. For instance, (2b–e) illustrate that the English ditransitive construction is not limited to the meaning of concrete/physical transfer that it coerces on the verb *throw* in (2a):

(2) a. X threw Y a ball.
 b. X drew Y a picture.
 c. X told Y a secret.
 d. X cost Y a fortune.
 e. X asked Y a favour.

(2) shows that the assumed central meaning (a) coexists with a number of motivated semantic variants (2b–d) and also apparently unmotivated exceptions (2e). This suggests that also for the three constructions analysed by Goldberg et al. (2004), there is more to be said about their meanings and the way they are indicated to learners in CDS than just the identification of a single schematic central sense and the single most frequent verb in the input.

The present study addresses these issues and continues the line of inquiry pursued in Goldberg et al. (2004) and Sethuraman and Goodman (2004). Like these two earlier studies, it seeks to characterise children's experience with some of the most basic grammatical patterns of English in a construction-based format. However, it also differs from them in that it investigates a wider range of constructions, considers more data for each pattern, presents a different quantification of the results and also investigates the larger semantic structure of each category. Specifically, the study is devoted to the following five constructions:

(3) a. Simple intransitive: *Listen.*
b. Complex intransitive: *Come here.*
c. Simple transitive: *Don't touch it!*
d. Complex transitive: *Put it back.*
e. Ditransitive: *Give Mummy a kiss.*

These constructions were investigated in a large sample of CDS extracted from 25 different English language corpora on the CHILDES database (MacWhinney 2000). In contrast to the abovementioned studies, the relative importance of individual verbs for generalisations about constructional meanings was not assessed on the basis of verbal raw frequency but of collostruction strength (Stefanowitsch and Gries 2003), a quantitative corpus-linguistic measure of the statistical association between grammatical constructions and their lexical fillers. And finally, the study broadens the focus of attention from an identification of just the single dominant filler of the verb slot to the larger semantic category structure of each construction. This is achieved by grouping verbs that encode the same general type of event (e.g. 'motion', 'perception', 'creation' etc.) and take the same configuration of semantic participant types (e.g. <agent>, <patient>, <stimulus> etc.) together to a semantic category that is contrasted with other such categories that are attested within the respective construction. Following this procedure, the study seeks to uncover quantitatively salient constructional usage patterns in the input on which child language learners could base the aforementioned generalisations about constructional meanings in a bottom-up manner.

2 Methods

2.1 Materials

The study is based on 23,340 semi-automatically analysed maternal utterances to children under 3;0 from 1,652 transcripts in 25 different English language CHILDES corpora. Table 1 gives an overview of the corpora and the number of transcripts and utterances included from each source:

Table 1: Materials used in the study

Corpus	Documentation	Transcripts	Utterances
Bates	Bates et al. (1988)	98	441
Bernstein	Bernstein (1982)	27	902
Bliss	Bliss (1988)	1	5
Bloom70	Bloom (1970)	18	88
Brown	Brown, 1973	76	985
Clark	Clark (1978)	13	46
Cornell	cf. MacWhinney (2000)	10	317
Demetras2	Demetras (1989)	11	164
Feldman	cf. MacWhinney (2000)	9	29
Gleason	Gleason (1980)	7	167
Higginson	Higginson (1985)	21	803
Howe	Howe (1981)	31	538
Kuczaj	Kuczaj (1977)	45	132
MacWhinney	MacWhinney (2000)	14	59
Manchester	Theakston et al. (2001)	791	5,810
NewEngland	Ninio et al. (1994)	52	4,062
Peters	Peters (1987)	2	6
Post	Demetras et al. (1986)	30	673
Rollins	cf. MacWhinney (2000)	50	4,383
Sachs	Sachs (1983)	52	264
Suppes	Suppes (1974)	36	577
Tardif	cf. MacWhinney (2000)	25	1,427
Valian	Valian (1991)	43	433
Warren	Warren-Leubecker and Bohannon (1984)	9	86
Wells	Wells (1981)	181	943
Total		1,652	23,340

2.2 Corpus compilation

At the time of the investigation, the English section of the CHILDES database comprised 5,211 transcribed dialogs. Since the analysis aimed only at maternal utterances directed at children under 3;0 that come with full morphosyntactic annotation and that are transcribed according to the most widely used annotation conventions within the system, a number of files were removed from this initial set. First, analyses were restricted to those transcripts in which the file header identified a specific 'Target_Child' with an explicitly declared age. Second, analyses were restricted to transcripts in which the target children interacted with their mother, and where the mother was explicitly identified as '*MOT' in the participant declaration of the file header. Third, the analysis focused on interactions between mothers and target children younger than 3;0. All transcripts in which the age specification of the target child(ren) was either missing or merely implicit in file header information were removed. Transcripts of interactions with more than one target child at the same time were included in the dataset if at least one of the target children had the required age. Fourth, transcripts from corpora that did not possess full morphological ('%mor') and grammatical ('%xgra') annotation at the time of the investigation were excluded. From the remaining files, all maternal utterances and their associated morphosyntactic annotation (i.e. main lines, %mor- and %xgra-tiers) were extracted, tagged for the respective source corpus file and merged into a single file. Since the study is specifically geared at the relation between verbs and clause-level argument structure constructions, two further adjustments were made: first, all verbless utterances in the dataset (i.e. lines without an element tagged either 'V' or 'AUX') were removed. Second, for ease of analysis, the investigation was restricted to utterances consisting of a single clause. With these restrictions in place, the remaining dataset still comprised 254,414 morphosyntactically annotated maternal utterances directed at children between 0;6 and 3;0 from 1797 transcripts in 25 different English CHILDES corpora.

From this dataset, a sample of approximately 10% was drawn. The sampling procedure sought to compensate possible finetuning effects in mothers' formulation preferences: on the one hand, the database contains many more interactions with older children (here: learners between 2;0 and 3;0) than conversations with younger toddlers or even prelinguistic infants (not surprisingly for a child language resource, of course). On the other hand, it is conceivable that the lexical variability of mothers' speech may increase with the child's growing language skills, thus making a sample in which conversations with older children predominate potentially less representative for the targeted population as a whole (speech to children in the first three years of life). To properly investigate such possible finetuning effects, mothers' lexical variability would have to be related to an indication of the given child's state of language development, such as for instance mean length of utterance (MLU). Since children's own utterances were not considered and the finetuning issue is not in itself of interest here, however, a simpler strategy was adopted to avoid a possible bias in the data: following Tomasello's (2006) account of typical developmental milestones on the way to language, the original dataset was stratified according to target children's age. Specifically, the data were partitioned into the following three broad strata:

- Stratum I: utterances to prelinguistic children. According to Tomasello (2006), children's first holophrastic utterances typically appear at the age of approximately 14 months. The first section of the sample thus consists of speech to children with developing general communicative (social cognitive) capabilities up until 1;2;
- Stratum II: utterances to children producing first words and word combinations. Combining the holophrastic and pivot schema-phases, this stage covers the typical timeframe for early linguistic productions before the onset of grammatical marking. Adopting the age thresholds suggested in Tomasello (2003, 2006), this section of the sample consists of speech to children between 1;2 and 1;9;
- Stratum III: utterances to children with item-based constructions. From the age of approximately 20 months on, children begin to build an increasingly complex repertoire of item-based constructions with initially idiosyncratic and progressively ever more regular grammatical marking. This section of the sample consists of speech to children between 1;9 and 3;0.

Table 2 reports the distribution of utterances across the three strata:

Table 2: Sample stratification

Stratum	Child age range	Utterances	Included	% Included
I	0–1;2	7,780	7,780	100.0
II	1;2–1;9	30,985	7,780	25.1
III	1;9–3;0	215,659	7,780	3.6
Total	0–3;0	254,414	23,340	9.2

From each of the three strata, an equal number of utterances was sampled and combined to form the final dataset. With the smallest stratum containing 7,780 utterances in total, this procedure gave a sample of 7,780 × 3 = 23,340 utterances.

2.3 Data extraction and coding

Data extraction proceeded semi-automatically with full manual postcoding. Structurally, the five target patterns exemplified in (3) above can be characterised as follows:

- simple intransitives are utterances involving a lexical verb and no dependents except the subject in non-imperatives (e.g. *you've won*);
- complex intransitives are a family of related constructions that comprise intransitive locatives (consisting of a subject, a lexical verb and a locative phrase headed by either a preposition or an adverb, e.g. *Santa's not coming to you*) on the one hand and intransitive resultatives on the other (consisting of a subject, a verb and a 'resultative phrase' in the sense of Goldberg and Jackendoff (2004) that can be headed by either an adjective, a preposition or an adverb/particle; e.g. *rabbit's fallen asleep*);
- simple transitives are utterances consisting of a subject, a verb and a nominal direct object of the verb with no other dependent (e.g. *he's biting his hand*);
- complex transitives include 'caused motion' and transitive resultative constructions that consist of a subject, a verb, a nominal direct object and a locative/resultative phrase that can be headed by either an adjective, a preposition or an adverb (e.g. *that'll keep Eve busy*);
- ditransitives are double object constructions consisting of a subject, a verb and two nominal objects of the verb (e.g. *can you tell me the story?)*

Data extraction took advantage of the recent addition of syntactic dependency annotation to the English section of the database (Sagae et al. 2010). Hence, in order to identify e.g. potential simple transitives in the data, all utterances were extracted in which the verb combined with exactly one element that was tagged as an object of this verb. Potential complex transitives, by contrast, were utterances that met the criteria for simple transitives and furthermore included an element that could be analysed as an instance of the construction-defining locative/resultative phrase. This means that there was overlap between the extraction criteria for the five target patterns, and the automatically identified initial candidate set had to be analysed further to eliminate duplicates and resolve ambiguities. This last step was done by hand. On the one hand, this was necessary in order to compensate for tagging errors. On the other hand, manual postcoding was also useful in view of the pervasiveness of syntactic ambiguity: many inappropriate parses in the dataset were not in fact technically *wrong*, but still not the contextually adequate analysis of the given utterance (e.g. *I know you like the telephone* analysed as a single clause with *you* as the direct object of *know* and *like the telephone* as an adjunct PP). Importantly, however, care was taken not to restrict the focus of attention to just those utterances that instantiated the ultimately 'correct' (i.e. adult grammar) generalisation from the outset: seeing that the goal of the analysis was to assess whether a particular grammatical pattern is really preferentially associated with a certain (range of) interpretation(s) or function(s) in maternal speech, it was of course crucial not to discard counterexamples to the hypothesised generalisation(s) prior to the actual analysis. To illustrate, consider the following pairs:

(4) a. *draw <u>a kitty</u>.* (ENG-US/NewEngland/38.cha)
b. *wait <u>a minute</u>.* (ENG-UK/Wells/benjam02.cha)

(5) a. *I got you <u>a pig</u>.* (ENG-US/Post/she02.cha)
b. *They call it <u>a truck</u>.* (ENG-US/NewEngland/14.cha)

Though superficially similar, the underlined elements in (4) and (5) have a different grammatical status in each case. But if children do not come to the task of language acquisition with innate knowledge of categories like 'argument' and 'adjunct' or 'object' and 'predicative complement' that are needed to differentiate them, how should they know? For present purposes, two modest assumptions were made about the capabilities that children bring to bear on this task: on the formal side, the ability to recognise patterns of phrase structure categories such as NP, AP and PP (which can be learned distributionally from the input; cf. Redington et al. 1998), and on the semantic side, the ability to discriminate denotata of the type 'person'/'thing' from denotata of the type 'location'/'state'. The preliminary characterisation of the five target patterns given above can therefore be fleshed out as follows:

Table 3: Possible realisations of the five target patterns

Pattern	Possible realisations	Example
1	$[_{VP}\ V]$	*listen*
2	$[_{VP}\ V\ [_{AP}\ A]_{LOCATION/STATE}]$ $[_{VP}\ V\ [_{AdvP}\ Adv]_{LOCATION/STATE}]$ $[_{VP}\ V\ [_{PP}\ P\ [_{NP}\ N]_{LOCATION/STATE}]]$	*fall asleep* *come out* *go to bed*
3	$[_{VP}\ V\ [_{NP}\ N]_{PERSON/THING}]$	*leave it*
4	$[_{VP}\ V\ [_{NP}\ N]_{PERSON/THING}\ [_{AP}\ A]_{LOCATION/STATE}]$ $[_{VP}\ V\ [_{NP}\ N]_{PERSON/THING}\ [_{AdvP}\ Adv]_{LOCATION/STATE}]$ $[_{VP}\ V\ [_{NP}\ N]_{PERSON/THING}\ [_{PP}\ P\ [_{NP}\ N]_{LOCATION/STATE}]]$	*make it dry* *bring it here* *put bunny to sleep*
5	$[_{VP}\ V\ [_{NP}\ N]_{PERSON/THING}\ [_{NP}\ N]_{PERSON/THING}]$	*are you giving me the towel?*

This means that in the case of (4), only the (a) variant was recognised as a potential input to Pattern 3, $[_{VP}$ V $[_{NP}$ N$]_{PERSON/THING}]$, whereas in the case of (5), both variants were recognised as potential inputs to Pattern 5, $[_{VP}$ V $[_{NP}$ N$]_{PERSON/THING}$ $[_{NP}$ N$]_{PERSON/THING}]$, even though (5b) is not in fact a ditransitive. In other words, the data were coded in a strongly surface-oriented way, irrespective whether the utterance did in fact instantiate the actually targeted adult construction from a linguistic point of view or not. Linearisation was not seen as crucial, to the effect that observations with re-ordered constituents (*out teddy gets*) were included in the dataset. Passives, however, were removed.[2] Sentence mood was ignored, and utterances marked as interrupted or trailing off were removed. Utterances consisting of fixed expressions that are presumably learned as 'long words' like *thank you* and *peek a boo!* were likewise removed. Manual postcoding along these lines reduced the automatically extracted original candidate set to a total of 13,496 instances of the five target patterns.

2.4 Quantification

When the data were coded, collexeme analysis (Stefanowitsch and Gries 2003) was applied to evaluate all verb-construction combinations in the dataset for statistical significance. Collexeme analysis is a variant of collostructional analysis, a group of corpus-linguistic methods that extend the analysis of lexical collocations to the interplay of lexical items with grammatical constructions. The purpose of collexeme analysis is to uncover semantic restrictions on the targeted constructions by identifying lexical items ('collexemes') that are significantly 'attracted to' a given constructional slot. Four frequencies are needed to calculate a given collexeme's degree of attraction to a grammatical construction: first, the frequency of the collexeme in the construction itself (e.g. all occurrences of *draw* in the simple transitive construction); second, the frequency of all other words in the relevant slot of this construction (e.g. the number of all simple transitives in the corpus minus the frequency of transitive *draw*); third, the frequency of the target word in all other (paradigmatically alternative) constructions in the corpus (e.g. the number of all occurrences of the verb *draw* with other argument structures than the simple transitive); and fourth, the number of all other words occurring in all other (paradigmatically alternative) constructions in the corpus. The fourth number is obtained in two steps: first, the frequency of the target construction is subtracted from the number of all other constructions of the same type in the

2 cf. Goldberg (2006: 22) for arguments (on the example of the English ditransitive construction) for the treatment of actives and passives as separate constructions (with distinct linking and information structure properties).

corpus (e.g. the number of all single-clause transcript lines in the dataset minus the number of simple transitive clauses within this set). From the resulting total frequency of all other single-clause argument structures in the corpus, the frequency of the target word in these other constructions is subtracted, thus giving the frequency of all other words in all other argument structure constructions in the corpus. Given these four frequencies, it is then possible to compare the observed frequency with which a given word occurs in a particular construction to the frequency expected by chance. Finally, the difference is evaluated for statistical significance using the Fisher-Yates exact test. All calculations reported in this paper were done with an R-script for collostructional analysis kindly provided by Stefan Gries (Gries 2004). Results were corrected using the Bonferroni correction (significant results are highlighted in the tables). A value of ‘0’ indicates a probability of error that is so low that it could not be computed anymore.

The fact that collostruction strength rather than token frequency was used to identify privileged exemplars of a construction in CDS is not to downplay the significance of token frequency. Learners are clearly sensitive to frequency effects on various linguistic levels, and as indicated above, collostruction strength is a frequency-derived measure itself. Still, the question remains whether it is the absolute frequency or rather the *relative likelihood* with which a verb appears in a construction (i.e. the measure that lies at the heart of collostruction strength) that helps learners form the right generalisations about which verbs to use in which environments. Following arguments presented in Stefanowitsch (2008), it was here assumed that speakers/learners indeed form word-construction associations on the basis of their experienced deviations from (subconscious) co-occurrence expectations, although the issue is contested and as yet in need of further empirical research (cf. Bybee 2010: 97).

2.5 Semantic classification

Finally, the resulting collexeme tables were systematised semantically. For this, each verb-construction combination was analysed for the types of participants involved in the encoded event as well as for the type of the event itself.

Methodologically, this was clearly the most questionable part of the analysis. The most pressing problem is that there is no fixed and generally applicable role inventory that such classifications can draw on. Since the aim of the analysis was to capture all supposedly relevant generalisations about event participants as precisely as possible (rather than to impose some ill-fitting general scheme from the literature), reference to existing role typologies was eclectic: where possible and appropriate, the classification employed widely used general labels like 'agent', 'patient', 'goal' and the like; where none of these was applicable or judged as precise enough to capture a particular generalisation, more fitting labels were employed as appropriate. For instance, in order not to lose the generalisation that *cry*, *squeak* and *clap* are all verbs of sound emission (and that this property sets them apart from other agentive verbs like e.g. *dance*, *play* and *walk*), the subject argument of utterances like *do birds quack?* was annotated as a 'stimulus producer' and not as an 'agent'. For reasons of space, classification was arbitrarily restricted to the twenty most strongly attracted verbs per pattern (i.e. no special significance attaches to the number '20').

A second problem is that a purely *form-based* collexeme analysis treats all realisations of the target collexeme as tokens of the same type. However, in lexicography, verbs are usually analysed as complex categories with (potentially many) different readings that some researchers might want to distinguish for such an analysis (including their uses as part of larger 'idiomatic' phrases such as e.g. *go on, go crazy* and *how did it go?* for *go*). Again, however, the question is not yet settled on the theoretical level, either, since there is disagreement about the correct framing of the relationship between a verb's meanings and its syntagmatic contexts. As in previous work on collostructional analysis, the approach taken here was to treat verbs as semantically unified on a certain (and sometimes rather schematic) level. For instance, *go* was analysed as encoding a change of location (be it or concrete or metaphorical). This means that the analysis was based on the quantitatively dominant (and possibly fairly abstract) 'core' meaning of a verb alone and any more finegrained semantic differentiations in its usage patterns were ignored. Following this procedure, each of the top 20 collexemes of each target pattern was assigned to a semantic frame like the following:

(6) a. cause-receive <agent> <recipient> <mover>
(e.g. X bring Y Z)
b. cause-receive <stimulus producer> <experiencer> <stimulus>
(e.g. X tell Y Z)
c. intend-receive <creator> <recipient> <effected object>
(e.g. X bake Y Z)

3 Results

3.1 Pattern 5

The highest degree of semantic consistency is found in Pattern 5 (ditransitive):

Table 4: Collexeme analysis, Pattern 5

Rank	Collexeme	f	p_{FYE}	Rank	Collexeme	f	p_{FYE}
1	*give****	124	5.67E–210***	11	*wish*	1	2.15E–02
2	*tell****	36	5.84E–41***	12	*paint*	1	3.20E–02
3	*show****	15	2.36E–18***	13	*cost*	1	4.25E–02
4	*bake****	6	6.59E–10***	14	*teach*	1	4.25E–02
5	*bring****	10	1.19E–08***	15	*send*	1	8.32E–02
6	*throw****	12	1.03E–05***	16	*make*	6	1.09E–01
7	*read****	7	1.16E–05***	17	*call*	3	1.48E–01
8	*pass**	2	1.69E–03*	18	*buy*	1	1.69E–01
9	*sing**	4	1.78E–03*	19	*find*	3	2.38E–01
10	*hand*	2	3.98E–03	20	*draw*	2	3.57E–01

The semantic groups that stand out in Table 4 are scenarios of physical transfer with the frame structure ‘cause-receive <agent> <recipient> <mover>’ (*give*, *bring*, *throw*, *pass*, *hand*, *send*), communication scenarios with the frame structure ‘cause-receive <stimulus producer> <recipient> <stimulus>’ (*tell*, *show*, *read*, *teach*), and events of creation for the benefit of an intended recipient with the frame structure ‘intend-receive <creator> <recipient> <effected object>’ (*bake*, *make*, *sing*, *draw*, *build*).[3] More remote similarities are found between ditransitive *wish*, *buy* and *find*, all of which signify that the subject referent intends the referent of the first object to ‘have’ the referent of the second object (*she is wishing you a happy birthday*, *you find me some more blue bricks*, *what did Daddy buy you yesterday?*). None of these three verbs designates transfer to a third party lexically,[4] so that the recipient role of their ditransitive use is inherited from the construction. What makes them stand out from other such verbs in the construction

3 This illustrates that the labels employed for the classification were not necessarily mutually exclusive between categories: a ‘stimulus producer’ is by definition also a ‘creator’, and both types of arguments are also an ‘agent’. Still, permitting categories from different taxonomic levels to coexist in the analysis was useful to capture more finegrained differentiations in the data where they were discernible.

4 The verb *buy* of course does signal transfer lexically, but the recipient role in ditransitive *buy* is nevertheless contributed by the construction (i.e. the final recipient is not the subject referent).

(such as *throw*) is that their lexical meaning furthermore modulates the constructional meaning from ‘cause-receive’ to ‘intend-receive’. Apart from that, however, the lexical meanings and semantic roles of these verbs are too heterogeneous to define a discernible class. Of the two remaining items, *cost* is a ‘verb of future non-having’ (Pinker 1989) and as such also related to the constructional meaning of change of possession, so that only *call* defies the semantic generalisation that expressions of the type $[\mathrm{NP}]_{\mathrm{PERSON/THING}}$ V-s $[\mathrm{NP}]_{\mathrm{PERSON/THING}}$ $[\mathrm{NP}]_{\mathrm{PERSON/THING}}$ are about change of possession.

3.2 Pattern 4

A similar degree of consistency is also found in Pattern 4 (complex transitive):

Table 5: Collexeme analysis, Pattern 4

Rank	Collexeme	f	p_{FYE}	Rank	Collexeme	f	p_{FYE}
1	*put*	1055	0.00E+000***	11	*pull*	29	4.04E−13***
2	*take*	148	1.60E−78***	12	*roll*	23	2.84E−11***
3	*throw*	114	6.95E−45***	13	*give*	46	5.81E−07***
4	*pick*	48	2.59E−31***	14	*set*	7	8.40E−06**
5	*bring*	52	2.85E−30***	15	*stack*	10	2.76E−05**
6	*leave*	38	2.07E−26***	16	*hook*	5	1.16E−04*
7	*get*	236	7.37E−25***	17	*tuck*	5	1.16E−04*
8	*push*	66	2.64E−23***	18	*wind*	5	1.16E−04*
9	*turn*	60	3.14E−16***	19	*want*	116	1.40E−04*
10	*knock*	22	3.59E−13***	20	*keep*	13	3.05E−04*

Again, the class of uses in which the verb matches the assumed central meaning of the construction (‘cause-move <agent> <mover> <goal>/<path>’) is highly prominent among the top 20 types (*put*, *take*, *throw*, *pick*, *bring*, *get*, *push*, *pull*, *roll*, *give*, *set*, *stack*). The remaining verbs illustrate four other usages of Pattern 4, at least three of which are again closely related to the central sense: first up are events of ‘caused posture’ with the frame structure ‘cause-assume <agent> <mover> <posture>/<orientation>’ (*turn*, most commonly used with (*a*)*round*, and *knock*, most commonly used with *over*). Second is a scenario of maintained rather changed location (*leave* and *keep*, i.e. ‘not cause-move <agent> <mover> <location>’), and third a modal extension to contexts of intended rather than directly caused change of location (*want*: intend-move <experiencer> <mover> <location>, e.g. *you want that one off?*). The only two remaining items are particle verbs (*tuck in* and *wind up*), which can also be argued to denote a special kind of caused motion (to a resulting state).

In contrast to the ditransitive, however, the top 20 collexemes in Table 5 are really only the tip of the iceberg: all in all, the pattern is attested with 154 different verbs in the data, and the pattern also subsumes many more specialised constructions such as the *get/have-* passive (*are you having your hair cut this morning?*), the ‘*What’s* X *doing* Y’ construction (*what’s your foot doing on the table?*) as well as predicative constructions with ‘depictive’ complements (*you need it harder?*). However, the collexeme analysis clearly reveals the heavy predominance of ‘caused motion’-uses of the pattern in CDS.

3.3 Pattern 2

Table 6 reports the top 20 collexemes of Pattern 2 (complex intransitive):

Table 6: Collexeme analysis, Pattern 2

Rank	Collexeme	f	p_{FYE}	Rank	Collexeme	f	p_{FYE}
1	*come*	800	0.00E+000***	11	*lie*	10	2.92E–09***
2	*go*	942	5.95E–284***	12	*fit*	23	1.17E–08***
3	*look*	569	7.70E–182***	13	*point*	6	7.59E–06***
4	*sit*	162	4.90E–103***	14	*climb*	10	1.03E–05**
5	*fall*	81	1.17E–42***	15	*happen*	27	1.41E–05**
6	*stay*	40	4.67E–30***	16	*chew*	13	1.37E–04*
7	*stand*	40	2.61E–23***	17	*live*	5	2.87E–04*
8	*hang*	24	3.02E–15***	18	*ride*	10	8.66E–04
9	*talk*	35	2.47E–14***	19	*write*	14	1.23E–03
10	*belong*	15	1.56E–13***	20	*hide*	12	1.55E–03

Here, the class of verbs that are more or less synonymous with the constructional meaning ‘move <mover> <location>/<path>’ again includes the two most strongly attracted types, but it is already substantially smaller (*come*, *go*, *fall*, *climb*, *ride*). However, it is still the largest group in the table, even if it shares this rank with the group of verbs of posture and spatial orientation (e.g. *what’s hanging onto the magnet?*), represented in Table 6 by *sit*, *stand*, *hang*, *lie* and *fit*. Many occurrences of Pattern 2 with these verbs are not instances of the intransitive motion/resultative construction but simple intransitives combining with a locative adjunct (e.g. *sit right there*, *you’re kneeling on my calf*; but also: *stand up*, *lie down*). In addition, as was also the case in the corresponding transitive pattern, the structure is also found with verbs that signal ongoing locatedness rather than change of location/state (*stay* and *live*). Apart from these uses, there is only one other notable group in the data, and this is again unrelated to the central change of location/state-frame: the frame 'direct-action-at <agent> <goal>', exemplified in Table 6 by *look*, *talk* and *point* (*at* <goal>), and arguably also *write* and *chew* (*on* <goal>). Semantically, such expressions are related to caused-motion constructions, but the element/action that is being directed at the goal is implicit in the lexical semantics of the verb (*look*: 'direct gaze at'; *talk*: 'direct utterance to'; *point*: 'direct gesture at' etc.). Apart from that, the verbs in this group are rather heterogeneous semantically (i.e. they denote rather different types of events involving different types of participant categories), but their intended interpretation will nevertheless occur as similar to a learner when they are combined with a <goal> argument. Two of the three remaining verbs in Table 6, *happen* and *belong,* commonly occur with *to*-PPs, but the argument in the PP is not a goal but rather an experiencer or patient (in the case of *happen*) or a possessor (*belong*). The last remaining verb, *hide*, does not group with any of the other verbs – like the group of posture and orientation verbs, it just happens to be used with locative adjuncts rather often in the data. All in all, then, the data for Pattern 2 do not encourage the generalisation that structures of the type $[\text{NP}]_{\text{PERSON/THING}}$ V-s $[\text{PP/AdvP}]_{\text{LOCATION/STATE}}$ are typically about motion events to the same extent that similar inferences are warranted for Patterns 5 and 4.

3.4 Pattern 3

This is all the more true in the case of Pattern 3 (simple transitive):

Table 7: Collexeme analysis, Pattern 3

Rank	Collexeme	f	p_{FYE}	Rank	Collexeme	f	p_{FYE}
1	*see*	547	2.07E−167***	11	*hear*	53	1.69E−25***
2	*get*	603	1.74E−115***	12	*try*	82	7.67E−24***
3	*want*	460	3.28E−110***	13	*open*	56	4.66E−20***
4	*have*	451	1.62E−85***	14	*need*	57	1.17E−18***
5	*like*	332	1.45E−73***	15	*hold*	72	1.25E−16***
6	*do*	808	2.08E−63***	16	*love*	45	7.75E−15***
7	*make*	201	3.39E−52***	17	*use*	31	1.32E−13***
8	*find*	128	3.59E−50***	18	*wear*	27	7.77E−13***
9	*eat*	139	6.54E−48***	19	*watch*	32	8.70E−10***
10	*brush*	58	3.52E−29***	20	*drop*	24	1.77E−09***

In this pattern, the most strongly attracted collexeme again represents a recurrent frame among the top 20 verbs ('perceive <experiencer> <stimulus>': *see*, *hear*, *watch*), but this is by no means the only obvious grouping in the data, and most other patterns are quite different semantically. Still closely related to the first class are instances of the construction denoting events of volition and emotion (*want*, *like*, *need* and *love*). However, already the second highest ranking verb *get* represents a semantically unrelated frame that is about having or gaining access to/contact with an object (*get*, *have*, *find* and *hold*, as well as possibly also *wear*). Somewhat unexpectedly in view of standard notions of prototypical transitivity (cf. Section 1), none of these relatively high-ranking verbs in Table 7 encodes a strongly asymmetrical interaction of the type ‘act-on <agent> <patient>’ with a subsequent change of state of the patient argument. And this is still true for the next highest-ranking verb on the list which does not fall into any of the previous groupings, *do*: here, the verb and the object nominal *jointly* designate an underspecified process (most frequent usage: *do it*), in this comparable to *try* and *use*. Supposedly prototypical transitives in which an agent either effects (*make*) or affects a patient argument (*eat*, *brush*, *open*, *drop*) only come fairly late in the table, which casts doubt on the idea that they embody the central sense of the construction (or, in fact, the idea that such central senses are reliably evidenced by acquisition/CDS data). Further down on the list, Pattern 3 also contains a number of highly unprototypical transitives (from a semantic point of view): uses like *that deserves a snapshot* or *he’s outgrown the crib* which once more underscore that Pattern 3 is semantically much less consistent than the previously considered environments.

3.5 Pattern 1

The last remaining category is Pattern 1 (simple intransitive):

Table 8: Collexeme analysis, Pattern 1

Rank	Collexeme	f	p_{FYE}	Rank	Collexeme	f	p_{FYE}
1	*know*	119	1.41E–78***	11	*work*	16	9.13E–14***
2	*play*	101	8.25E–78***	12	*colo(u)r*	19	5.44E–12***
3	*happen*	43	1.95E–31***	13	*draw*	29	1.09E–11***
4	*touch*	42	4.17E–28***	14	*sleep*	13	2.13E–11***
5	*cry*	24	2.41E–24***	15	*squeak*	9	6.27E–11***
6	*catch*	32	2.44E–21***	16	*fit*	16	3.52E–09***
7	*peek*	24	3.76E–19***	17	*walk*	11	7.44E–08***
8	*remember*	32	3.97E–18***	18	*clap*	8	1.15E–07***
9	*dance*	17	5.06E–16***	19	*hide*	12	2.06E–07***
10	*look*	137	5.97E–15***	20	*pretend*	7	2.08E–07***

Here, the semantic heterogeneity of the results is yet more marked, and it is quite difficult to discern even mere candidates for a category prototype. The highest ranking verb, *know*, has only one other member of its semantic class in the table (*remember*), and both are in fact inherently *transitive* – it is just that the information contributed by their complement can often be inferred in discourse (*I know*, *you remember?*). Pragmatic licensing of the simple intransitive construction for verbs that 'normally' select for additional complements also accounts for the appearance of *touch*, *catch*, *colo(u)r*, *draw* and *pretend* in the table. Among the usages that do not require special pragmatic licensing, verb groups are never larger than two (<mover> V-s: *dance*, *walk*; <perceiver> V-s: *peek*, *look*) or three items (<stimulus producer> V-s: *cry*, *squeak*, *clap*), and the encoded event types are quite different in each of these cases. For the remaining verbs *play*, *happen*, *work* (as in *That won't work*), *sleep* and *hide*, no specific semantic class (over and above general *Aktionsart* categories) readily suggests itself, so that it is not possible to say that the simple intransitive preferentially encodes a certain schematic event type in the data. Rather, it appears that the verbs in Table 8 are prominent in Pattern 1 simply because they happen to denote states and events that are commonly talked about in CDS for non-linguistic reasons.

4 Discussion

The extent to which the data provide evidence for coherent constructionally encoded event types ('humanly relevant scenes') in child-directed speech was found to differ strongly between the patterns under consideration. In line with previous research on mothers' use of argument structure constructions in CDS, especially the data for Pattern 5 (ditransitive) and Pattern 4 (complex transitive) indicate that the structure of the input may aid learners in forming the right semantic generalisations about constructional usage patterns. Corroborating findings by Goldberg et al. (2004), maternal usage of both constructions was heavily dominated by a single verb that may serve as a lexical 'anchor' from which the prototypical meaning of the construction can be abstracted. What is more, both items are transparently related to the majority of other attracted collexemes in the construction, and also where the most common variants of these constructions differ in some finer semantic detail, they are still strongly coherent on a more schematic level: ditransitive structures in the data are typically about changes of possession (whether the possession be literal or metaphorical, and the change be caused, intended, prevented etc.), and complex transitives are typically about changes of location (again both in the literal and metaphorical sense, with the change being caused, intended or refrained from). To a weaker extent, a similar tendency is also observable for Pattern 2 (complex intransitive): the two verbs that are most characteristic for the pattern are motion verbs, and again they are members of the largest semantic class within the top attracted items. However, there are also other salient usages of the pattern $[\text{NP}]_{\text{PERSON/THING}}$ V-s $[\text{PP/AdvP}]_{\text{LOCATION/STATE}}$ in the data that work against the generalisation that such structures encode motion. Also, in contrast to the study by Goldberg et al. (2004), the hypothesised verbal ‘anchor’ of the construction was *come* rather than *go* in the present data. Incidentally, this is a consequence of the different approach to quantification in the two studies: if the present study had also used raw frequency rather than collostruction strength to rank the verbs, *go* would have come out on top here, too (cf. Table 6).[5] Apart from that, however, the present results do not differ fundamentally from Goldberg et al.’s frequency-based findings.

A more notable difference between the studies is the thrust of the conclusions that are drawn from the results. Goldberg and colleagues argue “that the input is structured in such a way as to make the generalization from verb islands to argument structure constructions straightforward” (Goldberg et al. 2004: 307). It must be added, however, that their study focused exclusively on three constructions whose degree of semantic coherence is in fact rather exceptional. For these three constructions, it is indeed possible to discern a schematic prototypical meaning in CDS (and other) data. For the remaining two constructions that were investigated in the present study, by contrast, it is much more difficult to say what their schematic meaning might be and how it is indicated to children. This is not to suggest that they are somehow fundamentally different from the former: it is just that the latter two cover a semantic space that is too large to be subsumed under a single sufficiently coherent generalisation about the encoded event type. Nor is it to imply that there are no generalisations whatsoever to be extracted here: in the case of the English simple transitive construction, for instance, it is clearly essential to learn which of two given semantic arguments tends to map to the syntactic subject function and which to the object function. Likewise, when learning the simple intransitive construction, it is important to figure out which kinds of states of affairs are in fact expressed verbally in the respective language in the first place and which are encoded by other kinds of one-place predications (cf. *the man sleeps*/*is asleep* vs. *the apple* **greens*/*is green*). However, the fact remains that the latter two constructions do not inherently specify a certain default interpretation/scene type that is fitted onto instantiating utterances.

---

5 The same explanation also accounts for the apparent mismatch with Sethuraman and Goodman’s (2004) finding that *do* is the most prototypical filler of the simple transitive construction in CDS (cf. Table 8).

Summing up, the CDS data presented in Section 3 suggest that learners may be aided in figuring out constructional generalisations by the structure of the input. However, it appears that such cues are not strictly *necessary* in order to master the conventional usage of argument structure constructions. Even among the semantically more coherent types, individual verbs' occurrence or non-occurrence (or relatively high/low frequency) within a given construction is simply not fully predictable on semantic grounds alone, which shows that individual verb-construction associations are ultimately *memorised*. Some constructions then warrant more or less robust generalisations over the many individual 'verb-islands' that they subsume, and it is these constructions that begin to accrue an inherent 'scene' meaning of their own (as it shows up in the famous coercion effects that figure prominently in the construction grammar literature). Other constructions, by contrast, never reach that stage because they are simply too wide in semantic scope to begin with. Although this qualifies the 'scene encoding hypothesis' to some extent and may seem to conflict with the postulate of meaningful constructions 'all the way down' at first sight, this observation is actually not at odds with the general theoretical premises of usage-based construction grammar. Quite possibly, the answer is just that 'the' simple intransitive and 'the' simple transitive construction simply do not exist over and above the clusters of their heterogeneous lower-level subtypes – except of course, of all things, as 'taxonomic artefacts' (Chomsky 2002: 95) of linguistic theorising.

References

Bates, Elizabeth, Inge Bretherton & Lynn Snyder. 1988. *From First Words to Grammar: Individual Differences and Dissociable Mechanisms*. Cambridge, MA: Cambridge University Press.

Bernstein, Nan. 1982. *Acoustic Study of Mothers' Speech to Language-Learning Children: An Analysis of Vowel Articulatory Characteristics*. Unpublished doctoral dissertation, Boston University.

Bliss, Lynn. 1988. The development of modals. *The Journal of Applied Developmental Psychology* 9. 253–261.

Bloom, Lois. 1970. *Language Development: Form and Function in Emerging Grammars*. Cambridge, MA: MIT Press.

Brown, Roger. 1973. *A First Language: The Early Stages*. Cambridge: Harvard University Press.

Bybee, Joan L. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Chomsky, Noam. 2002. *On Nature and Language*. Cambridge: Cambridge University Press.

Clark, Eve. 1978. Awareness of language: Some evidence from what children say and do. In Anne Sinclair, Robert J. Jarvella & Willem Levelt (eds.), *The Child's Conception of Language*, 17–43. Berlin: Springer.

Demetras, Marty. 1989. Changes in parents' conversational responses: A function of grammatical development. Paper presented at ASHA, St. Louis, MO.

Demetras, Marty, Kathryn Post & Catherine Snow. 1986. Feedback to first-language learners. *Journal of Child Language* 13. 275–292.
Gleason, Jean B. 1980. The acquisition of social speech and politeness formulae. In Howard Giles, W. Peter Robinson & Philip M. Smith (eds.), *Language: Social Psychological Perspectives*, 21–27. Oxford: Pergamon.
Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
Goldberg, Adele E. 1999. The emergence of the semantics of argument structure constructions. In Brian MacWhinney (ed.), *The Emergence of Language*, 197–212. New Jersey: Lawrence Erlbaum Associates.
Goldberg, Adele E. 2006. *Constructions at Work. The Nature of Generalization in Language*. Cambridge: Cambridge University Press.
Goldberg, Adele E., Devin M. Casenhiser & Nitya Sethuraman. 2004. Learning argument structure generalizations. *Cognitive Linguistics* 15. 289–316.
Goldberg, Adele E. & Ray Jackendoff. 2004. The English resultative as a family of constructions. *Language* 80. 532–568.
Gries, Stefan Th. 2004. Coll.analysis 3. A program for R for Windows.
Higginson, Roy P. 1985. *Fixing: Assimilation in Language Acquisition*. Unpublished doctoral dissertation, Washington State University.
Howe, Christine. 1981. *Acquiring Language in a Conversational Context*. New York: Academic Press.
Kuczaj, Stan. 1977. The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior* 16. 589–600.
MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk*, 3rd ed. Mahwah, NJ: Lawrence Erlbaum Associates.
Ninio, Anat, Catherine Snow, Barbara Pan & Pamela Rollins. 1994. Classifying communicative acts in children’s interactions. *Journal of Communications Disorders* 27. 157–188.
Peters, Ann. 1987. The role of imitation in the developing syntax of a blind child. *Text* 7. 289–311.
Pinker, Stephen. 1989. *Learnability and Cognition: The Acquisition of Argument Structure*. Cambridge: MIT Press.
Redington, Martin, Nick Chater & Steven Finch. 1998. Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science* 22. 425–469.
Sachs, Jaqueline. 1983. Talking about the there and then: The emergence of displaced reference in parent-child discourse. In Keith E. Nelson (ed.), *Children’s Language*, Vol. 4, 1–28. Hillsdale: Lawrence Erlbaum Associates.
Sethuraman, Nitya & Judith C. Goodman. 2004. Children’s mastery of the transitive construction. Paper presented at the Stanford Child Language Research Forum, April 2004.
Sagae, Kenji, Eric Davis, Alon Lavie, Brian MacWhinney & Shuly Wintner. 2010. Morphosyntactic annotation of CHILDES transcripts. *Journal of Child Language* 37. 705–729.
Stefanowitsch, Anatol. 2008. Negative entrenchment: A usage-based approach to negative evidence. *Cognitive Linguistics* 19(3). 513–531.
Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243.
Suppes, Patrick. 1974. The semantics of children’s language. *American Psychologist* 29. 103–114.

Theakston, Anna L., Elena V. M. Lieven, Julian M. Pine & Caroline F. Rowland. 2001. The role of performance limitations in the acquisition of verb-argument structure: an alternative account. *Journal of Child Language* 28. 127–152.
Tomasello, Michael. 2003. *Constructing a Language. A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
Tomasello, Michael. 2006. Construction Grammar for Kids. *Constructions* SV1–11/2006. http://elanguage.net/journals/constructions/article/view/26. Last accessed: 31-12-2012
Valian, Virginia. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40. 21–81.
Warren-Leubecker, Amye & John N. Bohannon. 1984. Intonation patterns in child-directed speech: Mother-father speech. *Child Development* 55. 1379–1385.
Wells, C. Gordon. 1981. *Learning Through Interaction: The Study of Language Development.* Cambridge: Cambridge University Press.