Grant Agreement no. 214668

ITALK
Integration and Transfer of Action and Language Knowledge in robots

Seventh Framework Programme (FP7)
Cognitive Systems, Interaction, Robotics (ICT-2007.2.1)

Deliverable 3.1
Constructivist grammar classifications for grounded language learning

Start Date: 2008-03-01  Duration: 48 months

Project coordinator name: Angelo Cangelosi
Project coordinator organization name: University of Plymouth
TABLE OF CONTENTS

1. Executive summary
2. Introduction
3. The usage-based model of child language acquisition
4. Empirically motivated grammar learning experiments in ITALK
5. Constructivist grammar classifications
   5.1 Introduction
   5.2 Target patterns: English basic argument structure constructions
      5.2.1 C1: Simple Intransitives
      5.2.2 C2: Intransitive Resultatives
      5.2.3 C3: Simple Transitives
      5.2.4 C4: Transitive Resultatives
      5.2.5 C5: Ditransitives
   5.3 Objectives, methods and data
      5.3.1 Methods and objectives
      5.3.2 The CHILDES database
      5.3.3 Corpus compilation
      5.3.4 Data extraction and coding
   5.4 Results
      5.4.1 Stage I
      5.4.2 Stage II
      5.4.3 Stage III
      5.4.4 Whole corpus
6. Sample stimuli for grounded learning of grammatical constructions
7. Appendix
   A Corpus contents
   B Data coding guidelines
8. References
1. Executive summary

This study presents constructivist grammar classifications derived from empirical analyses of large corpora of child-directed speech which will serve as input for ITALK grammar learning experiments. We first outline the theoretical framework on which the current grammar learning approach is built and propose a set of concrete grammar learning tasks within this framework. We then present the empirical study by means of which we developed the constructivist grammar classifications as input for automatic language learning experiments. For these analyses, we manually annotated 69,750 maternal utterances from the Childes database for the occurrence of five major argument structure constructions. These constructions were investigated for a) prototypical fillers of the verb slot, b) lexical diversity of the verb slot, and c) the semantic patterning of the attested verbs in different constructions. These distributional cues have previously been shown to facilitate child language acquisition. The empirical findings are represented in constructivist grammar classifications and serve as facilitative input in the ITALK language learning experiments.

2. Introduction

This report presents classifications of basic grammatical constructions for grounded robotic language learning experiments derived from a large-scale empirical study of child-directed speech. Section 3 gives a brief overview of the theoretical framework adopted, the usage-based model of child language acquisition. Section 4 outlines a set of empirically-motivated language learning experiments. Section 5 presents the corpus study and discusses the obtained results. Section 6 provides a sample demonstration of how these results can be used to devise concrete experimental materials for grounded robotic language learning.

3. The usage-based model of child language acquisition

Contrary to the assumptions of nativist linguistic theories, recent empiricist, ‘usage-based’ approaches to child language acquisition\(^1\) are currently amassing more and more evidence that natural languages may indeed be learnable without the need for extensive language-specific cognitive hardwiring. Although the overall puzzle is as yet far from solved, experimental research along these lines has already significantly improved our understanding of both the

---

\(^1\) A short and highly accessible introduction to this research programme can be found in Tomasello (2006).
cognitive capacities that children bring to bear on the task and the different cues with which they may be able to achieve it. Key findings of the usage-based approach relate to

- the crucial role of general cognitive skills of cultural learning and intention reading (section 3.1);
- the grounding of language in both sensorimotor embodiment and social interaction (section 3.2);
- the significance of distributional learning and the statistical structure of children’s linguistic input (section 3.3);
- the item-based nature of early child language (section 3.4);
- the gradual emergence of grammatical abstractions through processes of schematisation (section 3.5).

We briefly elaborate on these issues and point out ways in which relevant findings from cognitive linguistics and developmental psychology can inform language learning experiments in ITALK.

3.1. General cognitive prerequisites

Until recently, linguists have postulated that natural language grammars can only be acquired given substantial pre-wired language-specific cognitive constraints that form a part of children’s genetic endowment, a ‘universal grammar’ (cf. Chomsky 1981; Pinker 1994). More recently, scholars have provided evidence from a broad range of areas, such as cognitive linguistics (Langacker 1987, 2000; Goldberg 1995, 2006), developmental psychology (Tomasello 2003; Lieven et al. 2003; Carpenter et al. 1998a), language evolution (Zuidema 2005; Oudeyer 2006; Kinsella forthcoming), grammaticalization studies (Diewald & Bergs 2008; Traugott 2003, 2008), evolutionary anthropology (Deacon 1997; Tomasello et. al 2005; Herrmann et al. 2007), epigenetic robotics (Steels & Kaplan 2002; Steels 2004; Dominey & Dodane 2004; Dominey & Boucher 2005; Dominey 2006) and computational modelling (Lewis and Elman 2001; Batali 2002; Bod 2006; Kaplan et al. 2008; Chang 2008), that general cognitive mechanisms, in contrast to language-specific capabilities, may account for many aspects of language learning. In particular, research in developmental psychology has shown that language development in children builds on a number of important socio-cognitive prerequisites: at the time that they begin to engage in linguistic communication (or very early in language development), children:
• have an understanding of the triadic structure of interactions between two interlocutors and an object that is being jointly attended to (Tomasello 1988, 1995; Carpenter et al. 1998a);

• understand the behaviour of others as intentional (Behne et al. 2005a, Behne et al. 2005b, Carpenter et al. 1998b);

• are aware of the normative structure of conventional activities, including social interactions (Rakoczy et al. 2008, Casler and Kelemen 2005; Rakoczy 2007); and

• understand the cooperative logic of human communication (Große et al. 2008; Liszkowski 2005, 2006; Tomasello et al. 2007).

Thus, in contrast to positions postulating special linguistic prerequisites, such as a ‘Universal Grammar’, these more recent approaches suggest that children are equipped with a broad range of general cognitive capabilities which help them to construct a linguistic system from the input they receive.

3.2. Grounded language learning

Using language is more than the appropriate manipulation of linguistic forms; for human speakers, it is a meaningful activity (Searle 1984; Vogt 2006; Bergen & Feldman 2008). Children's linguistic categories develop in direct, meaningful interactions with their environment, and their cognitive categories co-develop with their linguistic categories. Sensorimotor experience not only plays a crucial role in the formation of categories (Feldman 2006), but neurobiological research has also demonstrated the role of body sensation for reasoning (e.g. Damasio 1994). Moreover, embodiment implies the availability of multimodal cues, which may facilitate language learning. For instance, Yu & Ballard (2006) have shown how cues provided by the caretaker’s eye gaze and gesture support word learning. Furthermore, interaction between children and their caregivers generally focuses on the here and now of the current situation (e.g. Hatch 1983), underlining that children acquire language at least initially from direct experience. Linguistic meaning, especially in the first years, is therefore experientially grounded.

Furthermore, cognitive linguistic research over the past 25 years has shown that much of language structure has to be explained in reference to embodied human experience; one case in point are spatial prepositions, such as up and down, whose polysemy, i.e. their multiple ambiguity, is based on the embodied experience of more being up, such as the level of liquids.
in a glass rising the more liquid is added. Such image schemas license uses such as *the economy is going up again* (Lakoff 1987). Other examples concern the meanings of lexical items such as *cup* (Labov 1973) or *bicycle* (Wierzbicka 1985), whose reference is crucially determined by human anatomy; embodied experience of typical situations also underlies the understanding of grammatical structures (Fillmore & Atkins 1992, see below section 5.2), and finally much of the structure of social interactions, such as the turn-taking system (Sacks et al. 1974) and preference organisation (Levinson 1983), are deeply rooted in embodied interaction.

### 3.3 Distributional learning and the statistical structure of children’s linguistic input

Usage-based language models assume that natural languages can be learned from positive evidence alone given a.) distributional pattern-finding skills as already evidenced in infants, and b.) the observable statistical properties of the linguistic input. As regards a.), statistical learning has been evidenced in infants in a variety of language-related categorisation tasks such as:

- attuning to the elements of the phonological inventory of the ambient language (Werker & Tees 1984);
- discovering the phonotactics of the ambient language (Saffran & Thiessen 2003);
- segmenting words in continuous speech (Saffran et al. 1996);
- distinguishing lexical (‘content’) from grammatical (‘function’) words (Shi et al. 1999);
- using distributional cues to bootstrap syntactic categories such as ‘noun’ and ‘verb phrase’ (Gerken et al. 2005);
- extracting, memorising and generalising regularities from miniature artificial grammars so as to be able to tell legal from illegal strings, even with a change in vocabulary between training and test (Gomez & Gerken 1999).

As for b.), the specialised linguistic input that children receive (e.g. Pine 1994, Theakston et al. 2005) facilitates such statistical learning in crucial ways: among other things, speech directed at children is considerably less variegated (i.e. it uses fewer words and constructions), it is highly stereotypical (words and constructions are used in their most common senses/functions), and also heavily redundant (i.e. it is strongly repetitive and reformulative). To illustrate with some concrete figures, empirical studies of child-directed speech in English have found that:
45% of all child-directed utterances in the entire corpus began with one out of a set of merely 17 words (Cameron-Faulkner et al. 2003);

for a given argument structure construction, a single verb covers the lion’s share of all its tokens in child-directed speech (up to 38% in the case of put and \([\text{NP}_{\text{SUBJ}} \text{ V NP}_{\text{OBJ}} \text{ NP}_{\text{OBL}}]\)), thus facilitating the association of the prototypical filler verb’s meaning with the constructional schema itself (Goldberg et al. 2004);

between 21% (Küntay and Slobin 1996: 267) and 58.6% (Onnis et al. 2008: 424) of mother's utterances in child-directed speech occurred in variation sets, i.e. in sets of reformulations.

Finally, the sheer amount of data that children are immersed in is in itself impressive: Dabrowska (2004) estimates that children hear an average of 5,000-7,000 utterances per day, thus giving an estimated 5.6 million words between year 1 and year 4.

Thus, children receive an enormous amount of data, which is characterized by repetition and reformulation, simplification in structure, content and distribution and semantic cues to grammatical structure. At the same time, children have been found to be sensitive to such distributional cues, rendering a facilitative effect for language acquisition highly probable.

3.4 The item-based nature of early child language

Broadly construed, the usage-based theory of child language acquisition distinguishes four major developmental steps: children’s first utterances (usually at around 14 months) are regarded as holophrases, rote-learned strings that may possess internal grammatical complexity from an adult standpoint (e.g. lemme-see!), but which are assumed to constitute holistic symbols for the child that can be used to express likewise holistic communicative intentions (e.g. lemme-see! – ‘show me this object’). The second type of children's utterances, emerging at around 18 months, are so-called pivot schemas which consist of at least two component words and which reflect a partitioning of the designated experiential scene into attendant semantic building blocks (e.g. Mummy gone), some of which may be schematic (e.g. more X). Next (at approx. 18-20 months) are so-called item-based constructions that make use of grammatical devices such as word order or inflectional morphology to mark semantic roles in the encoded scenes (e.g. X see Y), but do so differently for different item-based constructions. In the usage-based view, it is only after these three initial item-based phases...
that entirely abstract grammatical schemas (e.g. \( NP_{\text{SUBJ}} \ V \ NP_{\text{OBJ}} \)) will begin to emerge as generalisations over sets of partially concrete item-based constructions (e.g. \( X \ \text{see} \ Y, \ X \ \text{hit} \ Y, \ X \ \text{kiss} \ Y \) etc.), and possibly only after a certain ‘critical mass’ of such item-based schemas has been acquired (Akhtar 1999; cf. 2.3).2

Grammar learning therefore does not consist in the acquisition of grammatical rules and lexical items, but in the detection of distributional structure of meaningful patterns, i.e. the usage of linguistic forms.

3.5 The gradual emergence of grammatical abstractions through schematisation

At some, rather late, point in development, children construct abstract grammatical categories and schematic patterns of these categories that license productive language use, thereby enabling them to produce an infinite number of utterances which they have never heard before. At the moment, the details of precisely how this happens in development are still poorly understood, in particular when it comes to the factors that constrain such generalisations. On the other hand, the lack of a fully worked-out usage-based model of grammatical generalisation does not strengthen the case for nativist accounts – not just because they merely shift the explanatory burden to biology, but also because there is empirical evidence which indicates that supposedly relevant categories over which any such innate constraints would have to be defined (e.g. ‘verb’) do not exist in infants but only emerge during ontogeny (Olguin and Tomasello 1993).

Except for the constraint issue, the general usage-based vision of how grammar emerges from an inventory of holophrases is simple enough: initially, children are assumed to simply imitate chunks of unanalysed adult speech in order to accomplish particular communicative intentions. Shortly after, however, they will attempt to make sense of the structure and function of the individual phrasal constituents of these utterance-level constructions in a process which Tomasello (2003: 297) graphically calls ‘blame assignment’: using a combination of distributional and pragmatic-interactional cues, children try to work out which part of the utterance relates to which aspect of the encoded scene, thereby working towards the later formation of generalised syntactic categories such as ‘noun’ or ‘verb phrase’. The

---

2 Evidence that children’s early utterances are indeed organised around concrete individual words and phrases that are only gradually split up and decomposed is reviewed in Tomasello (2000). Evidence that the item-based representations of early child grammar do not just centre around verbs (e.g. \( X \ \text{hit} \ Y \), as suggested in Tomasello 1992), but also around pronouns and pronoun-clitic-combinations (e.g. \( \text{It’s} \ X\text{-ing, Where’s} \ X? \)) is presented in Pine & Lieven (1993), Lieven et al. (1997) and Childers and Tomasello (2001).
result of this process is an inventory of item-based constructions with functionally defined schematic slots.

In the usage-based language model in which the developmental theory of Tomasello and colleagues is couched (for example, Langacker 1987, 2000; Bybee 2006), it is assumed that speakers (including child language learners) retain memory traces of (various aspects of) fully concrete utterance tokens that they experience. At the same time, they are also assumed to extract generalisations (i.e. schemas) from these memorised units which may become entrenched routines themselves. Following Langacker (2000), such schemas can be seen as reinforced commonalities between sets of co-activated exemplars: every time a set of different units is activated in language processing because they are allrelevantly similar to a given target, those (and only those) properties of the activated exemplars that are shared will be jointly reinforced. Over time, the repetition of this process may produce a stable pattern that is not identical to, but nevertheless completely immanent in, its individual instantiations. Put differently, schemas are emergent generalisations over a set of relevantly similar instances (including lower-level schemas) that are repeatedly co-activated in processing. Second, clustering exemplars for similarity (i.e. the crucial prerequisite for later schema extraction) is assumed to be an automatic by-product of language processing: metaphorically speaking, hearers/language learners are assumed to either map an incoming target onto an already existing identical representation in their repertoire, strengthening it, or if none such unit is available, to store it ‘next to’ the most similar existing unit in their system, creating a new representation (Bybee 2006). Over time, growing exposure to expressions that are in some respect similar to already encountered ones therefore creates progressively dense exemplar clusters, with various local schemas/similarity classes emerging from the numerous individual categorisations that were involved. The resulting picture is one of a vast network growing together via slow and piecemeal abstraction, and indeed productive overgeneralisations of the type Don’t giggle me, which would indicate a stronger orientation to rule learning and generalisation, typically occur fairly late in development, i.e. between ages three and five (Bowerman 1988).4

1 Experimental evidence for the psychological reality and processing relevance of such exemplar representations is presented in Bannard and Matthews (2008) for children and Zeschel (2008) and Gurevich et al. (to appear) for adults.

4Evidence that syntactic creativity in young children involves modified re-use of memorised previous productions is presented in Lieven et al. (2003) and Dabrowska and Lieven (2005). Further support for the usage-based assumption that child grammar slowly becomes more abstract with age comes from recent computational simulations of children’s linguistic schematisation processes which identify productive slot-and-filler patterns in children’s speech at different ages (Borensztajn et al. 2008; Bannard in preparation).
4. Empirically motivated grammar learning experiments in ITALK

In the ITALK project, we aim at developing a framework for grammar learning that reflects the assumptions of the usage-based approach presented above. Automatic grammar learning in the framework assumed thus concentrates on the learning of associations between linguistic forms and meanings, where meaning consists in the internal results of the robot's interaction with its environment. Language learning experiments therefore comprise the presentation of utterance – scene pairs to the robot, where the robot's task is initially to associate holophrastic utterances with representations of the scenes observed in its environment, and in later steps to segment the utterances presented and to connect the relevant chunks with appropriate meaning representations, i.e. decompose the holophrases and build up appropriate generalisations.

Sugita & Tani (2005) have already shown that verb – noun combinations, such as *push the block, push the cup, push the ball, touch the ball, touch the cup* etc., can be successfully learned and generalised to new, unseen examples, such as *touch the block*, using grounded language learning in a remote-control scenario. This kind of learning involves the segmentation and analysis of the contributing parts and the generalisation over lexical material.

The first step in the experiments here proposed consists in a replication of Sugita's and Tani's (2005) experiments, using the humanoid robot iCub, in order to create a baseline for the investigation of the role of facilitative cues from CDS for automatic language learning. This baseline will then be extended by adding further and further constructions to the set, for example, *push the block, push the block over the table, push me the block* etc. Interesting research questions at this stage concern on the one hand the role of order in which the robot is presented with the constructions; also in child language acquisition, some constructions are acquired earlier than others, and the presence of some constructions has been suggested to support the acquisition of others (so-called 'construction conspiracy', cf. Abbot-Smith and Behrens 2006). On the other hand, the role of the degree of semantic overlap between the constructions constitutes an open issue. That is, construction learning can be attempted by presenting the robot with verb-construction associations that are unique for each construction, that comprise the same verbs across constructions, or that exhibit a certain amount of overlap, for instance, the overlap measurable in CDS. These possibilities will be investigated in different experimental conditions.
Furthermore, the more constructions will be added, the more difficult automatic grammar learning will become (see Borovsky & Elman 2006); the neural network will have to match specific aspects of more and more complex scenes connected to increasingly complex utterances it is presented with. The more options there are, the more difficult learning will be, and yet children are always presented with many constructions at the same time, albeit fewer than in communications between adults (Snow & Ferguson 1977; Pine 1994; Cameron-Faulkner et al. 2003). In order to scale up automatic language learning, hence in a third step facilitative effects extracted from distributional regularities in the input parents present their children with in child language acquisition, i.e. skewed input (cf. section 3.3), will be employed. That is, the third step consists in experimenting with those facilitating cues in the input that children are presented with as well. We therefore explore systematically distributional effects in verb-construction associations in child-directed speech (see section 5 below) and use them for the design of stimuli for learning experiments, examples of which will be given in section 6.

To sum up, the experiments will proceed in the following four steps:

1. replication of Sugita & Tani (2005) as a baseline;
2. extension from one to several constructions and investigation of the effect of acquisition order, the amount of presentation and semantic overlap between stimuli;
3. supporting generalisation by providing facilitative input cues derived from empirical input analyses, in particular, the empirically determined prototypical verb-construction associations;
4. constraining generalisation by providing facilitative input cues derived from empirical input analyses.

While the experimental simulations by Borovsky and Elman (2006) show that skewed input in the sense described in Goldberg et al. (2004) and Casenhiser & Goldberg (2005) has effects on construction learning, some of which are positive, some of which are negative, their study involved randomly chosen distributions between verbs and constructions. That is, the authors picked a random verb in each semantic category and made it “much more frequent than the other members” (Borovsky & Elman 2006: 776). In contrast, in accordance with the general theoretical assumptions of usage-based cognitive linguistic models of language acquisition, it has to be assumed that the peculiar distributions of verbs in constructions in child-directed speech are meaningful such that the meaning of the prototypical verb corresponds best to the constructional meaning of the expression (see section 5 below for details) and that thus both
the frequency of the most typical verb and the distributions of the other verbs in each construction provide helpful cues to the language learner.

The facilitative input cues we are experimenting with in ITALK therefore presuppose the detailed analysis of distributional information available to children during language acquisition that may facilitate the acquisition of the constructions to be learned. In particular, the results of the empirical analyses will provide detailed information on a) overlaps between verb-construction pairs, b) prototypical verbs per construction, and c) distributions of verbs within constructions. The next section describes the empirical study by means of which these facilitative cues in child-directed speech were analysed, allowing the development of constructivist grammar classifications used as input for learning experiments (see section 6 below).

5. Constructivist grammar classifications

5.1 Introduction

The present section reports our study of facilitative cues for grammar learning in child-directed speech. Section 5.2 introduces those aspects of the learning target that our study focuses on – the five most basic argument structure constructions of English. We discuss how the learning of argument structure relates to the overall task of grammar learning and survey the five individual target patterns that are inspected. Section 5.3 describes methods and data of our study. We introduce the CHILDES database for child language research and describe how it was used to compile a large-scale corpus of child-directed speech addressed to learners at different developmental stages. We describe data extraction and coding procedures, present the empirical objectives of our study and introduce the quantitative measures and methods used for analysing the obtained results. Finally, section 5.4 reports on the results of the corpus study, addressing each of the investigated constructions in turn.

5.2 Target patterns: English basic argument structure constructions

It is widely acknowledged that the acquisition of argument structure – the linguistic knowledge involved in mapping the semantic arguments of verbal predicates to syntactic categories like ‘subject’ and ‘object’, i.e. in signalling who did what to whom in a sentence – lies at the very heart of the acquisition problem. The developmental psychologist Steven Pinker puts it as follows: ‘Since verbs’ argument structures assume such a large burden in
explaining the facts of language, how argument structures are acquired is a correspondingly
crucial part of the problem of explaining language acquisition” (Pinker 1989: 4). In recent
years, construction-based theories of grammar (Goldberg 1995; Langacker 1987, 1991, 2008;
Croft 2001; Bybee 2006) and grammatical acquisition (Tomasello 2003; Goldberg 2006) have
introduced a new perspective on the issue that also sheds new light on how these patterns may
be learned: on the constructionist view, basic argument structure constructions such as the
intransitive (e.g. shall we draw?), the transitive (e.g. shall we draw a picture?) and the
ditransitive construction (e.g. shall we draw Daddy a picture?) are not viewed as the
products of meaningless grammatical rules that operate over contentful lexical items, but
rather as inherently meaningful schematic signs that carry functional and semantic load over
and above the lexical semantics of the inserted words. From this perspective, an important
aspect of grammatical acquisition therefore resides in learning the kinds of structural
templates that are conventionally employed to communicate about specific basic event types
linguistically. The event types in question are seen as ‘basic’ in the sense that they constitute
supposedly universal conceptual building blocks of human experience, or “humanly relevant
scenes” as postulated in Goldberg’s (1995: 39) “scene encoding hypothesis”: “Constructions
which correspond to basic sentence types encode as their central senses event types that are
basic to human experience”. The underlying assumption (borrowed from cognitive semantic
theories such as Talmy 1985 and Talmy 2000) is that humans chunk the continuous flow of
sensorimotor experience into abstract configurations of experiential primitives such as
‘things’, ‘locations’ and ‘forces’, resulting in so-called “image schemas” (Johnson 2001,
Hampe and Grady 2005) for recurrent basic event types like “something moving somewhere”,
“something causing something to change location” and “somebody transferring something to
somebody”. The configurations encoded in these experientially grounded schemas are then
extended metaphorically to all sorts of abstract notions and relations, such as the passage of
time being conceptualised as motion along a path – cf. expressions like back in the 60ties,
before noon, to postpone something – or changes of state as a change in location – cf.
expressions such as to come to life, to go furious, to fall silent. These experientially grounded
schemas thus structure the way humans make sense of the world, which is reflected in the
linguistic structure of the respective language. Goldberg’s hypothesis postulates that such
abstract image-schematic configurations do not only structure semantic extensions in the
lexicons of many languages in predictable ways (cf. Sweetser 1990), but also underlie the
semantics associated with basic grammatical structures such as the above mentioned
ditransitive construction: in Goldberg’s analysis, this structural pattern is used to encode
concrete or metaphorical transfer of an entity between two participants. To illustrate, it is in virtue of this constructional meaning that we understand the earlier example of *shall we draw daddy a picture?* to mean that *daddy* is intended to receive the object thus created, even though the verb *to draw* in itself does not signify a transfer event lexically. Instead, the transfer implication in this sentence is the result of a so-called coercion effect that arises when lexical material is reinterpreted according to the requirements of the grammatical construction that it is plugged into (Michaelis 2005): in order for a word or phrase to be compatible with a particular slot of a given argument structure construction, it has to satisfy specific semantic constraints that are imposed on this position. For instance, the ‘recipient’ argument of the ditransitive argument structure construction is conventionally required to be an animate entity. For this reason, a sentence such as *Joe sent Chicago a letter* is necessarily interpreted as meaning that the letter was sent to a group of *people* in Chicago (e.g. staff working in a local branch of Joe’s office, who are the recipients of the letter), which is yet another example of a constructionally induced reinterpretation effect (Goldberg 1995: 55).

So where do these meanings come from? On the usage-based view, the meanings of basic argument structure constructions emerge from semantic generalisations over memorised expressions of the relevant type: over time, salient implications that keep recurring across repeated uses of a given grammatical pattern come to be associated with the structural configuration itself (Goldberg 1999). Note, however, that not necessarily *all* expressions that match the structural specifications of a given argument structure construction indeed convey the same schematic meaning. Consider (1):

(1)  a. Amy asked Sam his name.
    b. He envied the prince his fortune.
    c. She forgave him his sins.
    (Goldberg 1995: 132f.)

These sentences do not designate transfer events in which the referent of the first postverbal noun phrase is the recipient of the entity denoted by the second one as in *draw daddy a picture*. Also in our CDS data, there are a number of expressions that look like ditransitives (i.e. that are headed by a verb phrase with two NP complements) but do not carry the relevant meaning:

(2)  a. you could call her ducky.
    (ENG-US\Post\tow07.cha)
    b. maybe it’ll [= peekaboo] take us longer time.
    (ENG-UK\NewEngland\07.cha)
    c. putting a helicopter on your head doesn’t make you a witch either.
    (ENG-UK\Manchester\liz18a.cha)
The semantic structure of these expressions is not ‘X causes Y to receive Z’. Still, they look exactly like regular ditransitives that normally convey this implication. How can one reconcile this observation with the previously adduced evidence that a meaningful generalisation over this pattern exists, and that it can even enforce its inherent meaning on a verb like *to draw* when used in the relevant pattern? The simple answer is that the relationship between grammatical form and meaning is not a trivial one-to-one correspondence: natural languages are *motivated* but not entirely predictable systems, and many of the puzzling irregularities, syncretisms and inconsistencies of grammar persist in the surrounding system for good independent functional reasons (cf. e.g. Bybee and Thompson 1997). Put differently, constructional generalisations such as the ditransitive are *probabilistic*, and their robustness is a function of the consistency of the subsumed exemplars. Moreover, not all argument structure constructions may be equally amenable to a unified semantic characterisation in terms of a single overarching central meaning (or, viewed differently, this commonality may be highly schematic in places).

Usage-based language models accommodate these ideas by viewing grammatical constructions as interrelated networks of locally similar types and subtypes of actually experienced structures that cluster together on different levels of abstraction. Cognitively entrenched exceptions exist alongside more regular patterns in the system, and the question of what will become a productive (and hence regular/predictable) pattern in the first place is determined by statistical properties of the input. This introduces a quantitative perspective on linguistic knowledge as accrued experience with language use that permits the formulation of insightful generalisations about grammatical structure: if quantitative aspects of language processing and learning is taken into account, it is possible to see how one could explain empirical facts such as the above mentioned coercion effects in a way that is not compromised by the odd counterexample. From an acquisition perspective, a topic of central concern therefore resides in the way that learners become acquainted with the structure and function of basic sentence types, how learners may identify *dominant* usage patterns of these structures, and more specifically, how these structures are presented to the child in child-directed speech as clues to the generalisations that these uses may afford. It is this question that is addressed in our corpus study presented in sections 5.2 to 5.5.

---

1 Like in a ditransitive construction, both postverbal NPs in (2) are (selected) complements rather than (optional) adjuncts such as *e.g.* *last week* in *We bought it last week* (cf. section 5.3.5 below for the distinction between complements and adjuncts).
Summing up, argument structures constructions are schematic templates for the probabilistic pairing of argument roles (semantic categories such as “agent”, “patient” and the like) with grammatical functions (syntactic categories such as “intransitive subject”, “transitive object” etc.). Staying with the above example of the ditransitive construction, it is possible to capture these regularities in simplified but illustrative representations like the following:

<table>
<thead>
<tr>
<th>Sem</th>
<th>CAUSE-RECEIVE</th>
<th>AGT</th>
<th>REC</th>
<th>THM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syn</td>
<td>V</td>
<td>SUBJ</td>
<td>OBJ1</td>
<td>OBJ2</td>
</tr>
</tbody>
</table>

*Figure 1* The English ditransitive construction (adapted from Goldberg 1995)

However, having argued for a prominent role of argument structure constructions for theories of grammar and its acquisition, it is necessary to stress that these constructions are nevertheless only one particular facet of the grammar of simple sentences. This is to say that the form-meaning pairing that is provided by these units is but one contribution to the overall formal and semantic properties of an utterance: apart from the individual words that must be integrated with the appropriate slots, argument structure constructions also interact with further constructions for such phenomena as e.g. tense, aspect, voice, sentence mood and information structure that also play important roles in shaping the form and meaning of a given expression. In fact, of the three basic formal devices that are available for the grammatical encoding of meaning (i.e. word order, morphological marking, intonation), actually none of the relevant properties of an utterance is directly stipulated by the argument structure construction that it happens to instantiate. The point is easy to see for morphology and intonation: the fact that English grammar requires a suffixed –s on the verb in e.g. *she draws Daddy a picture* has nothing to do with the construction being ditransitive or not, and speakers are also free to produce such ditransitive constructions with different prosodic realisations (cf. focusing *She draws daddy a PICTURE*, interrogative *She draws daddy a picture?* etc.). By contrast, the ‘Syn’-specification in figure 1 may seem to suggest that it involves an implicit stipulation of word order (subject before verb before object-1 before object-2). However, a quick look at languages like German shows that argument structure constructions are in fact underspecified in this respect:
(3)  *Mama ruft Papa.*
    mommy calls daddy
    ‘Mommy calls daddy’ / ‘Daddy calls mommy’

On both interpretations (the choice of which depends on discourse context and intonation), (3) is clearly an instance of the same transitive construction, even though the verbal arguments SUBJ and OBJ swap positions. Though this specific possibility does not exist in English, the same point can be made using the following transitive constructions from our CDS-data – in each of these, the element that is affected/effectected by the verb action semantically (the ‘proto-patient’ entity, cf. 5.2.3 below) is not found in its canonical position after the verb (marked ‘__’) but rather precedes it in the sequence:

(4)  a.  *I been smootched __ by ya.*  (ENG-US\Rollins\cb09.cha)
    b.  *oh tissue you want __.*  (ENG-US\Sachs\n02.cha)
    c.  *what sound does mister cow make __?*  (ENG-US\Rollins\me09.cha)
    d.  *that's a terribly small horse for you to ride __.*  (ENG-US\Brown\adam01.cha)
    e.  *anything else in here you want __?*  (ENG-US\Rollins\me12.cha)

Similarly, in the examples in (5), the verb does not come in its canonical position after the subject but rather precedes it; likewise, the locative expression that would otherwise follow the verb comes first:

(5)  a.  *down goes the antenna __ __.*  (ENG-US\Tardif\e25.cha)
    b.  *along came a bunny __ __.*  (ENG-US\Rollins\me12.cha)
    c.  *next goes the cow __ __.*  (ENG-US\Rollins\nb12.cha)
    d.  *out came a baby bird __ __.*  (ENG-US\Valian\04a.cha)
    e.  *ahhah off comes the hat __ __.*  (ENG-US\Rollins\im09.cha)

Without going into the details of these constructions here, the examples show that even in a strongly word order-based language like English, supposedly commonsensical folk theories of sentence structure such as “the doer of an action comes first” are at best oversimplifying misconceptions: clearly, there is more to figure out for a language learner than just a simple mapping between semantic roles like ‘agent’ or ‘patient’ and relative positions in a sequence. (3), (4) and (5) thus serve as a reminder that the grammatical mapping between form and function is nontrivial, and that argument structure is only one dimension within the complex overall system.⁶

---

⁶ As it were, examples like (3) and (4.a) also demonstrate that representations like the box diagram in figure 1 in fact only capture certain *subtypes* of the construction in question: unless one wanted to postulate separate active, passive and middle versions of each argument structure construction, it is obviously not possible to link semantic.
Nevertheless, from a construction grammar perspective, basic argument structure constructions still remain the prime learning target within the process of grammatical acquisition: if Cognitive Linguistics is correct in positing a close correspondence between humans’ conceptual chunking of experience into certain primitive scenes (‘image schemata’) on the one hand and the grammatical encoding of these scenes in terms of a restricted set of basic grammatical blueprints on the other, these patterns hold the very key to the productive encoding of open-ended conceptual content by means of finite linguistic resources.

In the following sections, we will introduce the five most basic argument structure constructions of English that constitute the focus of attention here. The constructions in question are the simple intransitive (6.a), the intransitive resultative (6.b), the simple transitive (6.c), the transitive resultative (6.d) and the ditransitive (6.e):

\[(6)\]

- a. *chickens don’t swim.* (ENG-UK\Manchester\carl22b.cha)
- b. *this little piggy went to the market.* (ENG-US\Rollins\cm12.cha)
- c. *we’ll break it.* (ENG-UK\Valian\17b.cha)
- d. *roll the ball to Toby.* (ENG-UK\Howe\nicola1.cha)
- e. *give me the kleenex.* (ENG-US\NewEngland\04.cha)

We will go through each of these in turn, offering a brief characterisation of their form and function, exemplifying them with data from our corpus of child-directed speech and pointing to more extensive treatments of the pattern in the construction grammar literature.

### 5.2.1 Pattern I: Simple Intransitives

The simple intransitive is both the simplest and the most abstract of all English argument structure constructions. Examples from our dataset are:

\[(7)\]

- a. *chickens don't swim.* (ENG-UK\Manchester\carl22b.cha)
- b. *it'll break.* (ENG-US\NewEngland\04.cha)
- c. *how do you know?* (ENG-US\Brown\adam14.cha)
- d. *it’s not raining today.* (ENG-UK\Manchester\carl34b.cha)
- e. *which of Anna's toys do you really want to play with?* (ENG-UK\Manchester\warr14a.cha)

The construction is used to encode the relation between an event and a single participant that is involved in this event. In case the event does not involve any clearly individuated participant at all (a scenario that is sometimes also referred to as ‘atransitive’, cf. van Valin and LaPolla 1997), the subject position is occupied by a semantically empty expletive (7.d). Even though the subject is the only strictly required argument of this construction, simple categories like ‘proto-agent’ to syntactic categories like ‘Transitive Subject’ directly. See section 5.3.5 for our handling of voice in the corpus study.
intransitives can still become quite complex when there are optional modifiers of the verb (cf. 7.e).

Since the construction does not contribute any arguments by itself except the subject, any valence-related coercion effects (such as the introduction of an additional recipient argument for draw in the earlier example of shall we draw daddy a picture?) that could be used as an indicator of constructional meaning would have to be of the valence-reducing type in the case of the simple intransitive. With certain verbs that are intrinsically more complex in participant structure, intransitive uses indeed convey shifted and specialised meanings (cf. time drags, details are wanting, something’s gotta give), but this is not a regular and interpretively predictable effect that could be attributed to the construction itself (cf. *she puts, *he hands; *they enjoy). The only conventional implication associated with the construction (when used in the bare form of SUBJ + V alone) is one of habituality (cf. she smokes), but the basic designated event type in such utterances is encoded by the individual verb and the respective word form rather than by the simple intransitive construction. Furthermore, there are no restrictions as to the kind of semantic (macro-)role that the single expressed participant may play in the predicated event: it can be either a proto-agent/actor (7.a), a proto-patient/undergoer (7.b) or something in between these idealised extremes (7.c). Likewise, the verb event can be of any general aspectual type including both telic (7.b) and atelic processes (7.a) as well as states (7.c).

Given the enormous heterogeneity of basic one-participant event types that would seem to qualify as ‘humanly relevant’ in some sense (and hence prone to recur in linguistic communication), it is in fact not surprising to find that the topmost structural generalisation SUBJ + V is not associated with any too specific unified semantics. This leaves two options for the grammatical description of simple intransitive constructions: the first is to aim at the most inclusive conceivable generalisation anyway, even if the ‘constraints’ thus identified are next to vacuous. In Goldberg-style notation, a representation of the construction thus identified would look as follows:

<table>
<thead>
<tr>
<th>Sem</th>
<th>THING</th>
<th>PRED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syn</td>
<td>SUBJ</td>
<td>V</td>
</tr>
</tbody>
</table>

*Figure 2* The generalised simple intransitive construction
The second option is to posit distinct semantic subtypes of simple intransitive constructions that can be characterised in greater detail, though at the expense of missing the generalisation that all of them are [SUBJ V]-patterns structurally. With regard to the semantics of the subject argument, one could for instance distinguish intransitives with theme (chickens don’t swim), patient (it’ll break) and experiencer subjects (how do you know?):

<table>
<thead>
<tr>
<th>Sem</th>
<th>THM</th>
<th>PRED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syn</td>
<td>SUBJ</td>
<td>V</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sem</th>
<th>PAT</th>
<th>PRED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syn</td>
<td>SUBJ</td>
<td>V</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sem</th>
<th>EXP</th>
<th>PRED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syn</td>
<td>SUBJ</td>
<td>V</td>
</tr>
</tbody>
</table>

*Figure 3* Semantic subtypes of the simple intransitive construction

Other options would be a distinction in terms of aspectual properties of the event (i.e. telic vs. atelic vs. stative, cf. X breaks vs. X swims vs. X knows), or a cross-classification involving both dimensions. Determining the ‘right’ level of abstraction for such generalisations (i.e. that which is psychologically real in speakers and learners) is clearly an empirical question – what is (still) required here are carefully designed experimental and/or corpus-linguistic investigations of the way in which the relevant patterns are actually processed and used.

Discussions of the simple intransitive construction from a specifically constructional/cognitive linguistic perspective can be found in Croft (2001) and in Newman and Rice (2006). Presumably because of its high degree of schematicity/internal heterogeneity, the simple intransitive has not received as much attention in the construction grammar literature as our other target patterns though.

5.2.2 Pattern II: Intransitive resultatives

The intransitive resultative is used to describe situations in which a participant undergoes some kind of transition (i.e. is moving to a particular location or state) that is not caused by an external effector. Examples from our data include:

(8) a. *this little piggie went to the market.* (ENG-US\Rollins\cm12.cha)
    b. *shall we go downstairs?* (ENG-US\Bliss\normarjo.cha)
    c. *out it comes.* (ENG-US\Manchester\carl25b.cha)
    d. *where did it go?* (ENG-US\Bernstein\alice1.cha)
    e. *the house is falling to bits.* (ENG-US\Manchester\joel18a.cha)
    f. *rabbit's fallen asleep.* (ENG-UK\Manchester\warr02a.cha)
Semantically, it is possible to distinguish spatial resultative constructions (8.a-d; also called intransitive motion constructions) from property resultative constructions (8.e-f). From a cognitive linguistic perspective, property resultatives can be analysed as metaphorical extensions of their spatial counterparts. Under spatial resultatives we not only included constructions designating motion towards a goal (8.a), but also those encoding telic motion along a path (8.b) and motion from a source (8.c).

Structurally, (8) shows that the phrase encoding the endpoint of the designated process can take a number of different shapes – it can be realised as a prepositional phrase (8.a, e), as an adverb or adverb phrase (8.b, c), as a pro-form (8.d) or as an adjective or adjective phrase (8.f). In other words, the second postverbal complement is semantically constrained to designate a path, goal or resultant state, yet underspecified syntactically. Following Goldberg and Jackendoff (2004), we will subsume these different variants under the functional cover term ‘resultative phrase’ (RP) and hence treat all of the constructions in (8) as related variants of a common pattern.

Apart from the RP and the verb, the only other required component of this construction is the subject, and usually it is this participant that is understood to be the entity that moves (literally or metaphorically) to the state encoded by the RP. However, there are also some exceptions to this rule. Again following Goldberg and Jackendoff (2004), we will use the term ‘host’ for the entity undergoing the motion or change of state and distinguish cases in which the host corresponds to the subject (9) from others where the relevant entity is left implicit (10):

(9)     a. *The pond froze solid.*
        b. *Bill rolled out of the room.*

(10)    a. *Bill sweated/bled on the floor.*
        b. *Bill ate off the floor*
        c. *Bill drank from the hose.*

(Goldberg and Jackendoff 2004: 536 ff.)

Typically, however, the host is mapped to SUBJ. Where this is the case, the intransitive resultative construction can be represented as follows:
Most construction grammar treatments of resultative constructions are devoted to *transitive* resultatives (cf. 4.2.4). Discussions that also make reference to the intransitive variant include Goldberg (1995) and Goldberg and Jackendoff (2004).

5.2.3 Pattern III: Simple transitives

The simple transitive (sometimes also called monotransitive) is the most frequent basic sentence type in our data, and also the most frequent argument structure construction in the English language at large. The simple transitive is a highly general pattern that is used to encode an interaction or relation between two participants. It can be applied to a very wide range of situations:

(11)  

a. so, making it won't kill you.  
    (ENG-US\Gleason\katie.cha)  

b. we could see what it does on the floor.  
    (ENG-US\Tardif\e25.cha)  

c. we’ll break it.  
    (ENG-UK\Valian\17b.cha)  

d. I’ll always drink coffee.  
    (ENG-US\Kuczaj\abe051.cha)  

e. I like that story.  
    (ENG-US\Higginson\june08.cha)  

f. the other button opens the drawer.  
    (ENG-UK\Manchester\nic10a.cha)  

g. old Medonald had a farm.  
    (ENG-US\Rollins\di12.cha)  

h. it attracts it.  
    (ENG-US\Suppes\nina34.cha)  

i. god bless you.  
    (ENG-US\Tardif\e10.cha)  

j. does the elephant need a diaper change?  
    (ENG-US\Valian\18a.cha)  

k. we don't practice that.  
    (ENG-US\Bernstein\alice2.cha)  

l. oh (.) he's outgrown the crib.  
    (ENG-US\Suppes\nina33.cha)  

m. that deserves a snap+shot.  
    (ENG-US\Clark\shem25.cha)  

Structurally, the transitive consists of a subject argument, a verb and an object. Both arguments can be clausal (11.a, b) or nominal in nominative/objective case (11.c-m). Semantically, events denoted by a transitive clause can be telic (11.c) and atelic processes (11.d) as well as states (11.e). The referents of either argument may be concrete or abstract as
well as individuated or unbounded. In terms of traditional semantic roles,\(^7\) the first argument can be anything from an agent (11.c) over an instrument (11.f) and an experiencer (11.e) to a possessor (11.g) and a location (11.h). The second argument can take the role of patient (11.c), stimulus (11.e), theme (11.h), beneficiary (11.i), possessed (11.g) plus a variety of less readily classifiable relations (11.j-m). Even more abstractly, either argument can also refer to actions and entire situations rather than to individual participants within such scenes (11.a, b, j).

Given this extraordinary variability, it may seem dubious that the pattern is associated with one specific constructional meaning. However, adopting the cognitive linguistic view that grammatical constructions are prototype categories that consist of networks of related structures which are connected by family resemblances, it is possible to identify particular prominent configurations of event types that are typically expressed by such clauses and that share a number of distinctive semantic features: the properties that jointly make up prototypical transitivity. On such an analysis, examples like 11.c (reproduced as 12 below) illustrate the prototypically transitive scenario:

(12) \(\text{we'll break it.}\)

In such sentences, the referent of the subject argument is animate and human, it is engaged in a dynamic volitional action that transmits energy to the referent of the object argument, it has control over of the referent of the object argument, the referent of the object argument is inanimate and stationary, and the event brings about a change in its condition. Deviations from this prototype are possible in all these dimensions, and the transitive is even found in a few contexts where actually none of the above asymmetries obtains (animate/inanimate; volitional/non-volitional; energetic/non-energetic; in control/not in control; unchanged/changed):

(13)  

a. \(\text{The fifth floor contains a library.}\)  
b. \(\text{A library occupies the fifth floor.}\)  
c. \(\text{Line A intersects line B.}\)  
d. \(\text{Line B intersects line A.}\)  
e. \(\text{Marsha resembles Hilda.}\)  
f. \(\text{Hilda resembles Marsha.}\)  

Langacker 1991: 222

\(^7\) We use these labels for illustrative purposes only. No assumption is made is that there exists a fixed universal set of discrete primitive roles.
However, the existence of a limited number of highly untypical transitives notwithstanding, the above mentioned features provide highly informative cues for the identification of scenarios that are likely to receive transitive linguistic coding, and to predict which element will be coded as grammatical subject and which as object in the resulting clause. In an influential paper, Dowty (1991) proposes the terms “proto-agent” and “proto-patient” to cover these two feature bundles, respectively. Adopting Dowty’s terminology, Goldberg (1995) proposes the following representation for the schematic transitive construction:

<table>
<thead>
<tr>
<th>Sem</th>
<th>proto-agent</th>
<th>proto-patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syn</td>
<td>SUBJ</td>
<td>OBJ</td>
</tr>
</tbody>
</table>

Figure 5   The simple transitive construction


5.2.4 Pattern IV: Transitive resultatives

As the name suggests, the transitive resultative construction (also known as the caused motion construction) is the causative equivalent of the intransitive resultative construction discussed in section 4.2.2: it is used to encode situations in which one participant acts on some other participant, thereby causing it to change state or location. Examples from our data are:

(14)  
a. roll the ball to Toby. (ENG-UK\Howe\nicola1.cha)  
b. very good, honey now set him down. (ENG-US\Tardif\e03.cha)  
c. where does Granddad take you? (ENG-US\Manchester\carl01.cha)  
d. put the duck to sleep. (ENG-US\Tardif\e21.cha)  
e. are you making her hair pretty? (ENG-UK\Howe\lucy2.cha)

As in the intransitive variant, the pattern can be subcategorised into spatial (14.a, b, c) and property resultative constructions (d, e). Again, the resultative phrase can be encoded by a prepositional phrase (a, d), by an adverb/adverb phrase (b), by an adjective/adjective phrase
Diagrammatically, the transitive resultative construction can be represented as follows:

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{Sem} & \text{CAUSE-BECOME} & \text{AGT} & \text{THM} & \text{PRED} \\
\hline
\text{Syn} & \text{V} & \text{SUBJ} & \text{OBJ} & \text{RP} \\
\hline
\end{array}
\]

Figure 6  The transitive resultative construction

Together with the ditransitive, the transitive resultative construction is the most widely discussed basic argument structure pattern in the construction grammar literature. Among others, constructionist analyses (and critiques of such analyses) can be found in Goldberg (1991a, 1991b, 1995), Goldberg and Jackendoff (2004, 2005), Boas (2003, 2005), Wechsler (2005), Müller (2006), and Steels (2005).

5.2.5 Pattern V: Ditransitives
The ditransitive construction, already introduced in section 3.2.1, is used to encode the transfer of an object between two participants. (16) gives relevant examples from the CHILDES database:

(16)  
\begin{align*}
\text{a.} & \quad \text{give me the kleenex.} & \text{(ENG-US\NewEngland\04.cha)} \\
\text{b.} & \quad \text{can you bring me some rocks?} & \text{(ENG-US\Demetras2\mot03.cha)} \\
\text{c.} & \quad \text{roll me the ball.} & \text{(ENG-US\NewEngland\70.cha)} \\
\text{d.} & \quad \text{can you throw me the ball?} & \text{(ENG-US\NewEngland\26.cha)} \\
\text{e.} & \quad \text{make me a tower.} & \text{(ENG-US\Howe\sally2.cha)} \\
\text{f.} & \quad \text{you draw Mummy a fish.} & \text{(ENG-US\Manchester\car117a.cha)} \\
\end{align*}

The construction is used to encode concrete (16.a-d) or metaphorical transfer (16.e, f). As is also case with both the intransitive and the transitive resultative construction, the ditransitive can be instantiated by a verb that has the same general meaning as the construction itself (16.a, b) or by a verb that does not signify transfer lexically and instead adds an additional implication to the overall sentence meaning (16.c-f). Constructions differ as to the type of semantic mismatch between verb and construction that they permit. In the case of the English ditransitive, attested verb classes other than transfer verbs proper include motion verbs

\footnote{Goldberg (1995: 150) argues that examples like 16.e and f are instances of metaphorical transfer due to a figurative mapping that she describes as “actions which are performed for the benefit of a person are understood as objects which are transferred to this person”}
expressing the means of the designated transfer (16.c, d) and verbs of creation that designate a precondition for the following transfer (i.e. the transferred object’s coming into existence, cf. 16.e, f).

Structurally, the pattern consists of a subject, the verb, and two object arguments. Noting the oddness of sentences like *In attempt to injure Chris, Pat threw Chris the ball, Goldberg (1995) observes that the subject argument is not simply an agent in the sense that it instigates a volitional action, but that the action must in fact be intended as a transfer event on part of the subject referent in order for the ditransitive to be felicitous. As already mentioned in section 5.2 on the example of Joe sent Chicago a letter, the referent of the first object is mapped to the role of recipient and is required to be animate. If this is not the case, the construction will coerce a figurative reinterpretation of the first object as an animate entity (where available) or otherwise produce an anomaly. Similar to the intended-transfer constraint pertaining to the agent, Goldberg (1995) notes that also the recipient must a willing recipient of the transfer, to the effect that e.g. *Bill threw the coma victim a blanket is anomalous. Finally, the second object maps to the role of theme, i.e. to that entity which is realised as a direct object in the semantically related transitive resultative construction (cf. bring me some rocks vs. bring some rocks to me). This gives the following representation:

<table>
<thead>
<tr>
<th>Sem</th>
<th>CAUSE-RECEIVE</th>
<th>AGT</th>
<th>REC</th>
<th>THM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syn</td>
<td>V</td>
<td>SUBJ</td>
<td>OBJ1</td>
<td>OBJ2</td>
</tr>
</tbody>
</table>

Figure 7 The ditransitive construction

Constructionist analyses of the ditransitive are discussed in Goldberg (1992, 1995) and in van der Leek (1996).

5.3 Objectives, methods and data

Having established the phenomena that our study is devoted to, it is now time to survey the actual aims and the design of our investigation. We begin with an account of our empirical objectives and introduce the measures that were used to operationalise the targeted variables and then subject our data to a quantitative analysis (5.3.1). Second, we introduce the
CHILDES database of child language from which these data were extracted (5.3.2). Section 5.3.3 describes the criteria and procedure for selecting our material (clause-level maternal utterances directed at children at different developmental stages) from the overall database. Section 5.3.4 describes the data extraction procedure for each investigated construction. Section 5.3.5 summarises the main analytical decisions involved in coding the data according to the interests of our analysis.

5.3.1 Objectives and methods
In sections 2 and 3.2, we have already hinted at what might help children learn the semantics of abstract argument structure constructions: it is assumed that the meanings of these patterns arise through gradual semantic generalisation over specific encountered expressions, and that caretakers’ use of these (and other) linguistic forms shows a strong bias for typical uses of the structure in question, thus facilitating the discovery of the right generalisation. As briefly mentioned in section 2.1 already, a previous study by Goldberg et al. (2004) has already indicated that the usage of basic argument structure constructions in child-directed speech is strongly centred on particular verbs that are lexically synonymous with the construction in question (such as put for the transitive resultative and give for the ditransitive), which may help children to interpret novel instances of the pattern (e.g. VERB X into Y) on the analogy to a strongly entrenched prototype (e.g. put X into Y). We continue and expand this line of research by considering more constructions, more data and more sophisticated quantifications. With the target constructions already introduced and data issues covered in sections 5.3.2 – 5.3.4, the present section focuses on the operationalisations and quantitative measures involved in analysing our dataset of child-directed speech. We focus on the following properties of basic argument structure constructions in CDS directed at learners at different developmental stages:

- lexical diversity of the verb slot
- prototypical fillers of the verb slot
- the semantic patterning of the attested verbs

To begin with, the lexical the diversity of a construction is quite simply measured by type frequency counts (in our case of the verb slot). Entrenched prototypical instances of a given construction (‘pathbreaking verbs’) are identified through a collostructional analysis of the

9 In the Goldberg et. al (2004) study, three constructions were considered (intransitive motion, caused motion and ditransitive), the analysis is based on a single corpus out of the entire CHILDES database (ENG-US/Bates), and privileged verb-construction combinations were identified through textual raw frequency counts/percentages.

Constructivist grammar classifications (Deliverable 3.1)
verb slot. Collostructional analysis is a family of quantitative corpus-linguistic methods that is specifically geared at investigating the interaction of words and grammatical constructions as conceived in usage-based construction grammar (Stefanowitsch and Gries 2003; Gries and Stefanowitsch 2004b, Stefanowitsch and Gries 2005): in the terminology of Stefanowitsch and Gries, all grammatical patterns are associated with a range of significant collexemes (i.e. lexemes collocating with a particular slot of the construction). The affinity between the construction and the range of attested collexemes can be ranked in terms of individual collostruction strengths (the term being a blend of ‘collocation’ and ‘construction’), i.e. the association strength of the particular collexeme with the constructional slot in question. In experimental studies with adult native speakers (Gries et al. 2005), collostruction strength has been found to outperform token frequency (of verbs in a given construction) as a predictor of speaker’s processing behaviour in sentence completion tasks, thus suggesting that it should be the preferred measure of cognitive entrenchment for verb-construction combinations as compared to raw token frequency. The reason is that in contrast to raw frequency counts of the target measure alone (i.e. occurrences of a particular word in a particular construction), collostructional analysis also takes the overall frequencies of the word and the construction into account, thus eliminating distortions of the association ranking that are simply due to chance.\footnote{To illustrate, suppose that 99\% of all occurrences of a word A are found in one particular construction C (i.e., the construction is a highly typical environment for this word to occur). Now suppose that there is a word B in the corpus that is 500 times more frequent than word A, yet only 1\% of its occurrences are attested in construction C (which is therefore not a particular typical environment for B). If the attested instances of the construction are simply ranked by raw corpus frequency, the untypical combination B-C will take a share more than five times the size of the highly typical combination A-C, thus obscuring the existence of a significant connection between A and C.} Collostruction strength (henceforth CxStr) will be reported directly in the form of the $p$-value of the association strength computation here – in other words, the lower the figure, the stronger the association. In extreme cases, the figure may be too low to compute, in which case an output of ‘0’ appears in the results. The computation was performed using a script for R for Windows, coll.analysis 3, that automates the procedure (Gries 2008).

Collostructional analysis requires that all instances of a given construction in the corpus are included in the computation, so that the results do not return the distributionally identified protoype alone but rather an exhaustive ranking of all attested fillers of a given slot. The resulting table can therefore also be used to identify groups of similar verbs that are statistically attracted to a given construction, thus permitting the identification and quantitative weighting of different semantic subtypes of a given construction (e.g. ditransitives with transfer, communication and creation verbs etc.). We use this method to
determine the lexical diversity in the verb slot of a construction and to identify the prototypical fillers in each construction.

5.3.2 The CHILDES database
The CHILDES database is an international database comprising annotated transcripts of interactions with children in first and second language acquisition in about twenty different languages. The project was initiated by Brian MacWhinney in the 1980ies, and besides the huge database, it offers tools for annotation and transcription as well as for search and data manipulation. All interactions have been transcribed in a uniform format, and the transcriptions are freely available for interested scholars to download. Because of its impressive size and easy availability, the CHILDES database has developed into the major resource for child language and language acquisition research.

5.3.3 Corpus compilation
At the time of investigation, the English section of the CHILDES database comprised 5211 transcribed dialogs (1432 files in the ENG-UK archive, and 3779 in the ENG-US archive). In a first step, files from both archives were collapsed into a single repository. However, since the analysis aimed only at maternal utterances directed at children younger than three years that came with full morphosyntactic annotation and were transcribed according to the most widely used annotation conventions within the system, a number of files had to be removed from this initial set. First of all, analyses were restricted to those 4590 transcripts (88.1%) in which the file header identified a specific “Target_Child” with an explicitly declared age (or, in fact, more than one target child with an explicitly specified age). Second, for ease of data extraction, analyses were restricted to transcripts in which the target children interacted with their mother (at least as one among other interaction partners), and where the mother was explicitly identified as ‘*MOT’ in the participant declaration of the file header (4066 transcripts, or 88.6% of the remaining files). Third, the analysis focused on interactions between mothers and target children younger than 3;0.0. Again for ease of data extraction, all transcripts in which the age specification of the target child(ren) was either missing or merely implicit in the information provided in the file header were removed.\footnote{In the latter type, the header specified date of target child birth and date of recording (rather than age of target child at day of recording, thus leaving the age implicit).} 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. This left a total of 2015 mother-child interactions in the dataset. Fourth, target children’s age declarations in the file headers were checked against the corpus documentation in order to weed out possible coding errors in the

\footnote{In the latter type, the header specified date of target child birth and date of recording (rather than age of target child at day of recording, thus leaving the age implicit).}
transcripts. Fifth, transcripts from corpora that did not possess full morphological (%mor) and grammatical (%xgra) annotation at the time of investigation were excluded from consideration, thereby reducing the number of target interactions to 1957. Finally, any of the remaining transcripts in which the mother was declared as a participant in the file header but did not actually produce any linguistic utterance in the course of the interaction were likewise removed.

From the remaining 1920 files (36.9% of the entire English section of the database), 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. Given that the present study is specifically geared at the relation between verbs and argument structure constructions, two further adjustments were made as a last preparatory step: first, all verbless utterances in the dataset (i.e. lines without an element tagged either ‘V’ or ‘AUX’) were removed. Second, seeing that argument structure constructions are a clause-level phenomenon, all utterances containing more than one clause (including utterances with tag questions) were likewise removed in order to facilitate a maximally error-free automatic extraction of the targeted argument structures. With these restrictions in place, the

---

12 Two such errors were found: “rosie08.cha” in the ENG-UK Wells corpus and “monmt2.cha” in the ENG-USA HSLLD corpus involved erroneous age declarations. Both errors have been corrected since.

13 At the time of investigation, most of the English language corpora provided full morphosyntactic annotation. However, annotations were still missing for some entire corpora (e.g. UK-Belfast and ENG-USA:McCune) as well as for certain individual files within otherwise parsed corpora (e.g. ENG-USA:Bates/Free28/george28.cha).

14 Restricting the corpus to mono-clausal utterances ensured that an element identified as a specific type of complement (e.g. an object) was in fact a dependent of the (automatically extracted) first lexical verb identified in the given transcript line, rather than the complement of a verbal predicate embedded somewhere further downstream in the utterance (as in the case of e.g. *I think you’ll like it*, where *it* functions as the object of *like* rather than of *think*). In principle, it would of course be possible to use the information provided in the %xgra-tier to correctly identify grammatical dependencies in complex sentences, too. However, since the present investigation is not concerned with syntax beyond the clause level and a recovery of the complete dependency relations in the corpus would have been considerably more costly computationally, the analysis was restricted to mono-clausal utterances for ease of data extraction. Grammatical relations to be removed on these grounds (or, rather, utterances involving these relations) include ‘CSUBJ’ (i.e. finite clausal subjects of another clause, e.g. *That Eric cried, moved Bush*), ‘XSUBJ’ (i.e. non-finite clausal subjects of another clause, e.g. *Eating vegetables is important*), ‘COMP’ (i.e. finite clausal complements of a verb, e.g. *I think that was Fraser*), ‘XCOMP’ (i.e. non-finite clausal complements of a verb, e.g. *stop throwing the blocks*), ‘CPRED’ (i.e. finite clausal predicates of verbs such as *be* and *become*, e.g. *This is how I drink my coffee*), ‘XPRED’ (i.e. non-finite clausal predicates of verbs such as *be* and *become*, e.g. *My plan is to win the competition*), ‘CJCT’ (i.e. finite adjunct clauses, e.g. *I remember when we had macaroni*), ‘XJCT’ (i.e. non-finite adjunct clauses, e.g. *She’s outside sleeping in the carriage*), ‘CMOD’ (i.e. finite clausal nominal modifiers or complements, e.g. *That’s what I do*), ‘XMOD’ (i.e. non-finite clausal nominal modifiers or complements, e.g. *This is what you are doing* and ‘CPZR’ (i.e. complementisers as tokens of clausal subordination, e.g. *Wait until the noodles are cool*). Of these, ‘CSUBJ’ and ‘XSUBJ’ did not actually occur in the presently assembled corpus. Finally, we excluded utterances involving serial verb constructions (e.g. go play with your toys), coordinated VPs (Go and get your telephone) and any remaining instances of the above structures with more than one lexical verb per utterance that were still included in the corpus due to grammatical tagging errors.

15 Recall and precision of an automated extraction of relevant utterances from the corpus of course depend on both tagging accuracy and analytical coding decisions on part of the corpus creators (e.g. the decision to tag both have and V in modal ‘have to V’-constructions of the type John has to drink milk as main verbs, thus resulting in their exclusion from the present study).
remaining dataset still comprised 254,414 morphosyntactically annotated maternal utterances directed at children between ages 0;6 and 2;11 from 1,797 transcripts in 25 different English CHILDES corpora.

Since one of the issues to be investigated in the present study is the way in which facilitative cues change with increasing learner competence, these ~250k utterances were then stratified according to stage of learner development. Unfortunately, the precise nature of the connection between learners’ estimated capabilities and speakers’ resultant communicative fine-tuning is not very well understood. On the one hand, the very fact that a simplified register like CDS is being used clearly testifies to a certain special recipient design on part of the speaker, and it is intuitively obvious that the relevant conception will be different for an addressee that is, say, five years old and normally developing as compared to a preverbal infant, or an aphasic peer. On the other hand, little is known about the interplay of specific individual factors that may contribute to shaping this particular conception: are estimates primarily based on learners’ perceived level of understanding (as inferable from their behavioural responses to a given communicative act), or are they (also) informed by certain properties of learners’ own productions? If, as seems fair to assume (at least from a certain stage of development onwards), both comprehension and production capabilities play a role, then specifically which aspects of learners’ productions will lead to which kinds of adaptations on part of the speaker? For instance, is it possible to correlate individual indicators of child language development such as MLU (mean length of utterance) with specific fine-tuning effects in the input that a given child receives? Previous research suggests that the relationship between fine-tuning and production is in fact more complex than this, or at any rate does not involve such a direct coupling of purely linguistic properties alone (Shatz 1982; Bohannon & Marquis 1977; Bohannon & Warren-Leubecker 1988; Snow 1994; Behrens 2006; Huttenlocher et al. 2008).

Therefore, we did not stratify the dataset according to one specific corpus-derivable indicator such as MLU. By contrast, we assume that caretaker’s conception of their addressee is shaped by tacit assessments of both linguistic and more general socio-cognitive capabilities of the child, and that both dimensions involve more than just a single variable. Specifically, following Tomasello (2003) and colleagues, we assume that language acquisition presupposes a certain set of pre-established social-cognitive foundations (cf. section 2), and that in normally developing children, there is a typical timetable for the emergence of both these social-cognitive prerequisites and certain important developmental breakthroughs on the way for instance, in terms of linguistic cues, tacit assessments of learner competence may be based in part on MLU, but there are also many other straightforward cues on all linguistic levels – think of phonetic abilities, mastery of allophony patterns, correct use of inflectional morphology etc. – that probably conspire to yield an integrated overall impression of the linguistic competence of a child.
to language that will build on them. Slightly simplifying this timetable, our study thus distinguishes the following broad developmental stages on the basis of target children’s age:

- **Stage 1: Preverbal stage.** According to Tomasello (2006), children’s first holophrastic utterances typically appear at the age of approximately 13 months. The first section of the corpus thus consists of maternal utterances directed at prelinguistic children with developing general communicative (social cognitive) capabilities up until the age of 1;1;
- **Stage 2: 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) and (2006), this section of the corpus consists of utterances directed at children between ages 1;2 and 1;8;
- **Stage 3: Item-based constructions.** From the age of approx. 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 corpus consists of CDS directed at children between ages 1;9 and 2;11.

Table 1 gives an overview of the distribution of the dataset over these three classes:

<table>
<thead>
<tr>
<th>Stage</th>
<th>Age Range</th>
<th>Utterances</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>0 – 1;1</td>
<td>7780</td>
</tr>
<tr>
<td>II</td>
<td>1;2 – 1;8</td>
<td>30985</td>
</tr>
<tr>
<td>III</td>
<td>1;9 – 2;11</td>
<td>215649</td>
</tr>
<tr>
<td>TOTAL</td>
<td>0 – 2;11</td>
<td>254414</td>
</tr>
</tbody>
</table>

*Table 1* Distribution of source dataset across developmental stages

Not surprisingly for a child language resource, interactions with preverbal infants are somewhat underrepresented in the database as compared to dialogs with older children. Since we did not want to limit the investigation to merely 8000 utterances per stage, we therefore matched only stages II and III in terms of number of included utterances and simply took everything that we could get for stage I. Specifically, the corpus that forms the basis of the present investigation comprises

- all 7780 utterances to stage I children from the original dataset;
- all 30,985 utterances to stage II children from the original dataset;
- 30,985 randomly selected utterances to stage III children from the original dataset.

Summing up, the corpus thus compiled consists of

- 69,750 morphosyntactically annotated maternal utterances;
- directed at children between ages 0;6 and 2;11;
- stratified into three sections according to stage of target child language development.
5.3.4 Data extraction

In spite of the recent addition of syntactic dependency annotation to almost the entire English section of the database, we subjected our data to full manual (post)coding. This was necessary for three reasons. First, the dependency analysis provided by the CHILDES system does not map neatly to distinctive tag configurations for each of our five target patterns. Consider the resultative constructions in (17):

\[(17)\]
\[a. \quad \text{make} \quad \text{your} \quad \text{knees} \quad \text{dirty.}\]
\[v|\text{make} \quad \text{pro:poss:det|your} \quad \text{n|knee-PL} \quad \text{adj|dirty} \]
\[1|0|\text{ROOT} \quad 2|3|\text{MOD} \quad 3|1|\text{OBJ} \quad 4|1|\text{PRED} \]
\[(\text{ENG-UK}\Manchester\gail23a.cha)\]
\[b. \quad \text{put} \quad \text{the} \quad \text{duck} \quad \text{to} \quad \text{sleep.}\]
\[v|\text{put}\&\text{ZERO} \quad \text{det|the} \quad \text{n|\lowercase{d}uck} \quad \text{prep|to} \quad \text{n|\lowercase{sleep}} \]
\[1|0|\text{ROOT} \quad 2|3|\text{DET} \quad 3|1|\text{OBJ} \quad 4|1|\text{LOC} \quad 5|4|\text{POBJ} \]
\[(\text{ENG-US}\Tardif\e21.cha)\]
\[c. \quad \text{uhm (.)} \quad \text{you} \quad \text{frighten} \quad \text{them} \quad \text{all} \quad \text{away.}\]
\[\text{fil|uhm} \quad \text{pro|you} \quad \text{v|frighten} \quad \text{pro|them} \quad \text{qn|all} \quad \text{adv|away} \]
\[1|3|\text{COM} \quad 2|3|\text{SUBJ} \quad 3|0|\text{ROOT} \quad 4|3|\text{OBJ} \quad 5|4|\text{JCT} \quad 6|5|\text{JCT} \]
\[(\text{ENG-UK}\Howe\barry2.cha)\]

These examples show that the construction-defining resultative phrase RP is not only underspecified structurally (i.e. can be instantiated by different types of phrasal constituents such as PP, AP, AdvP, NP), but that relevant elements also receive rather heterogeneous coding in the function-based CHILDES annotation scheme: in some cases, the resultative phrase is tagged as a predicative complement (‘PRED’, cf. 17.a), in some as a subcategorised locative complement (‘LOC’, cf. 17.b) and in yet others as an adjunct (‘JCT’, cf. 17.c). This means that the annotation provided by the ‘xgra’-tier could only be used to narrow down the set of potentially relevant concordance lines rather than uniquely identify relevant utterances directly.

The second reason is tagging accuracy. Given our focus on the quantitative attraction between verbs and different argument structure constructions, it was important to correct wrongly classified utterances like those in (18) manually:

\[(18)\]
\[a. \quad \text{eat} \quad \text{it} \quad \text{Rosie (3.).}\]
\[v|\text{eat} \quad \text{pro|\lowercase{t}} \quad \text{n|\lowercase{prop|Rosie}}. \]
\[1|0|\text{ROOT} \quad 2|1|\text{OBJ} \quad 4|1|\text{OBJ2} \]
\[(\text{ENG-UK}\Wells\rosie06.cha)\]
\[b. \quad I \quad \text{used} \quad \text{it} \quad \text{last} \quad \text{night.}\]
\[\text{pro|I} \quad \text{v|use-PAST} \quad \text{pro|\lowercase{t}} \quad \text{adj|last} \quad \text{n|\lowercase{night}}. \]
\[1|2|\text{SUBJ} \quad 2|0|\text{ROOT} \quad 3|2|\text{OBJ} \quad 4|5|\text{MOD} \quad 5|2|\text{OBJ2} \]
\[(\text{ENG-UK}\Wells\geoffr03.cha)\]
In these examples, a vocative (18.a), a temporal adjunct (18.b) and an apposition (18.c) are erroneously classified as the second object of a ditransitive construction. If one simply relied on the pre-established annotation, one would thus get the puzzling and distorting result that e.g. eat, use and see are collexemes of the ditransitive. Sagae et al. (2007) report an accuracy of 92.2% for their automatic parsing of adult corpus sentences as compared to a manually annotated gold standard. However, to stay with the example, out of the 724 utterances containing both an element tagged “OBJ” and one tagged “OBJ2” in our dataset, manual postcoding revealed that only 460 (63.6%) were in fact instances of the ditransitive construction. Conversely, the extracted dataset also contained an additional 203 ditransitive utterances that were not identified as double object constructions by the CHILDES annotation scheme. An example is (19):

(19) **did rabbit give Warren a kiss?**

Finally, a third argument for manual postcoding of the data is the notorious ambiguity of grammar: many inappropriate parses in the dataset were not in fact technically wrong (i.e. inconsistent with the rules of English grammar), but still not the correct analysis of the given utterance. Consider (20):

(20) **I know you like the telephone.**

The dependency analysis in (20) translates into the phrase structure tree in (21.a) where you functions as the direct object of know and like the telephone is analysed as an adjunct PP, thus giving a reading that can be paraphrased as ‘I know you like I also know the telephone’. Even though we did not consider the target utterances in their actual discourse context during postcoding, it seems safe to assume that like is a verb here instead and you like the telephone is an embedded clause as diagrammed in (21.b):
(21) a. S
   | NP  VP
   |   |   
   | NP  PP
   |   |   | NP
   |   | N  V  N  P
   |   |   |   | Det  N
   | I  know  you  like  the  telephone

b. S
   | NP  VP
   |   |   
   | S
   |   NP  VP
   |   |   NP
   |   | N  V  N  V
   |   |   | Det  N
   | I  know  you  like  the  telephone

Nevertheless, the grammatical annotation provided by the ‘xgra’-tier was still very helpful to us in that it served to relieve the amount of necessary handcoding substantially: even though the five targeted patterns constitute the most common argument structure constructions of English (and the dataset was restricted to verbal utterances from the outset), the corpus nevertheless still contains a number of highly frequent patterns that are not within the focus of attention here (for instance copular constructions such as that’s him or you look great) which would otherwise have to be weeded out manually. In order to avoid this, sets including all
potentially relevant utterances were extracted for each of the five targeted patterns by applying the following procedure:

- Potential simple intransitives were identified by removing all utterances containing either an object (OBJ, e.g. try this one) or a predicative complement (PRED, e.g. that’s nice) from the original dataset. In addition, all utterances involving subcategorised locative phrases (LOC) were removed. Finally, all remaining utterances with a form of to be as the main verb were removed (mostly misclassified utterances with predicative complements, e.g. it’s over there)

- Potential intransitive resultatives were identified as follows: first, all utterances involving an object (OBJ) were removed from the original dataset. Next, all utterances with the main verb to be used in conjunction with a predicative complement (PRED) were removed. This way, copular constructions like That’s her bottle could be eliminated, whereas relevant types of predicative complements (e.g. it’s getting tall) were retained. Among the remaining set of intransitive utterances, potential resultatives were identified by extracting all utterances in which a verb combined with a predicative complement (PRED) or some kind of locative expression. The latter included combinations with elements tagged as locative adverbial adjuncts (adv:loc|JCT, e.g. how did it get there?) or prepositional complements (identified via their object POBJ, e.g. do they go on your feet). As pointed out above, supposedly intransitive utterances with subcategorised locatives (LOC) were in fact all wrongly coded transitive resultatives and hence discarded right away

- Potential simple transitives were identified by extracting all utterances containing an object and removing those with an additional lexically selected complement (i.e. OBJ2, IOBJ or LOC)

- Potential transitive resultatives were identified by extracting all utterances containing a direct object (OBJ) in conjunction with at least one of the following categories: an indirect object (‘IOBJ’, i.e. show it to Mama), a predicative complement (e.g. oh that makes you mad), a subcategorised locative (e.g. put that on) or a (putative) locative adjunct (can you put the crayons in the box for me)

- Potential ditransitives were identified by extracting all utterances containing both an object (OBJ) and an element identified as the second object of a ditransitive clause (‘OBJ2’)

Applying this procedure, the size of the original corpus (69,750 utterances) was reduced by 13% to a total of 61,022 potential instances of our five target constructions. Moreover, since the extraction criteria for simple and resultative constructions were overlapping (in order to compensate for the unreliable identification of adjuncts in the annotation), weeding out duplicates prior to postcoding reduced the dataset further to 46,207 utterances, thus relieving the overall coding load by one third as compared to a full manual annotation of the entire corpus. Table 2 gives an overview of the data extraction results (where C1 = utterances matching the extraction criteria for simple intransitives, C2 = intransitive resultatives, C1-C2 = matching both simple intransitives and intransitive resultatives, C3 = simple transitives, C4

17 All of these were misclassified utterances involving the complex transitive verb put.
= transitive resultatives, C3-C4 = matching both simple transitives and transitive resultatives, C5 = ditransitives, and C4-C5 = matching both transitive resultatives and ditransitives):

<table>
<thead>
<tr>
<th>Type</th>
<th>C1</th>
<th>C2</th>
<th>C1-C2</th>
<th>C3</th>
<th>C4</th>
<th>C3-C4</th>
<th>C5</th>
<th>C4-C5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens</td>
<td>8977</td>
<td>1492</td>
<td>8923</td>
<td>18103</td>
<td>5346</td>
<td>724</td>
<td>62</td>
<td>46207</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Raw data collection results for the five target constructions in our corpus

Manual postcoding of the data of course does not guarantee a 100% error-free annotation either: also for a trained human analyst, there remain many difficult individual analytical decisions to be made, and also a number of very general questions to be settled (for instance, regarding the treatment of ungrammatical, fragmented or elliptical utterances in the corpus). The following section gives an overview of the most important of these decisions.

5.3.5 Coding

Maternal utterances that were both elliptical/ungrammatical and incomprehensible out of context were removed from the dataset:

(22) 

a. *mummie go that at a restaurant.* (ENG-UK\Manchester\gail23b.cha)
b. *I can't cold [*] it .* (ENG-UK\Manchester\liz09a.cha)
c. *cut baby Sarah.* (ENG-US\brown\eve08.cha)

Utterances that were fragmented, elliptical or in some respect ungrammatical but still understandable were retained in the dataset provided that their argument structure pattern was unambiguously discernible. This was the case with subjectless utterances with complete VPs (23.a, b), with utterances in which the ellipsis or breach of grammaticality affected an optional adjunct of the VP (c, d), and with utterances where a particular type of verbal complement was unambiguously given, even though its internal structure was somehow fragmented or disrupted (e, f):

(23) 

a. *can’t see?* (ENG-UK\Manchester\anne17b.cha)
b. *first make the pool.* (ENG-US\Bloom70\peter13.cha)
c. *put it somewhere safe while* [//] where +/- (ENG-UK\Manchester\anne25a.cha)
d. *it can only tow things with the +//.* (ENG-UK\Manchester\aran09a.cha)
e. *put your xxx here.* (ENG-US\NewEngland\13.cha)
f. *I'll give you all they [= sweets] (1.)* (ENG-UK\Wells\ellen05.cha)
Sometimes, a particular complement whose presence would be expected given the norms of adult written language was not fragmented but absent from the utterance altogether. Such ellipses are not unusual in spoken discourse when the element that would be encoded by the ‘missing’ complement is pragmatically inferable, thus making its explicit mention in the utterance redundant. We included such uses in our study, and counted them as instances of the structure that they actually instantiated in the given utterance – provided this structure was one of our target patterns (i.e., if a semantically ditransitive verb was used with only one instead of two objects as in 24.a, the example was counted as a simple transitive utterance). The motivation for this is that child language learners are faced with the same problem of implicit arguments and pragmatically induced valency changes, i.e. they have also heard verbs occurring in a number of non-canonical linguistic environments a certain number of times. Even if such uses cloud an otherwise more consistent relationship between a specific verb and a given range of argument structure constructions, they should therefore not be eliminated from the data in advance simply because they are felt to be unusual out of context:

(24)  
\[ \text{a. give a hug.} \quad \text{(ENG-US\Rollins\t12.cha)} \]
\[ \text{b. tell about the circus.} \quad \text{(ENG-US\Brown\sarah031.cha)} \]
\[ \text{c. you put in the corner?} \quad \text{(ENG-US\Tardif\e22.cha)} \]

Moreover, there is no clear-cut demarcation between cases like (24) and those in (25) where the verbs are also intrinsically more complex (regarding the number of actual participants in the encoded event), yet the valency-reduced variant is much more conventional than in (24):

(25)  
\[ \text{a. the baby goat is eating.} \quad \text{(ENG-US\Suppes\nina03.cha)} \]
\[ \text{b. I promise you.} \quad \text{(ENG-UK\Manchester\domin19b.cha)} \]
\[ \text{c. oh perhaps you've forgotten.} \quad \text{(ENG-UK\Manchester\john31a.cha)} \]

To repeat, however, elliptical utterances like (24) and (25) were of course only included in the dataset if they still matched one of our five target patterns. In short, argument structures were annotated on an as-is basis and not mediated by our own grammaticality intuitions (as long as the structural choice was unambiguous and it was sufficiently clear what the utterance was supposed to mean). By contrast, utterances that were explicitly marked as interrupted or trailing off in the transcription at a point before the argument structure of the clause could be established were excluded:

(26)  
\[ \text{a. are you making +...} \quad \text{(ENG-US\Rollins\ds12.cha)} \]
\[ \text{b. what does he +/.} \quad \text{(ENG-UK\Manchester\anne26a.cha)} \]
\[ \text{c. boy we can even +...} \quad \text{(ENG-US\NewEngland\28.cha)} \]
Elliptical utterances without lexical verbs (often echo questions) were likewise removed:

(27) a. can you? (ENG-US\Rollins\ds09.cha)
b. should we? (ENG-US\Demetras2\mot04.cha)
c. it sure did. (ENG-US\Rollins\cr12.cha)

In case an utterance in the dataset still contained more than one clause (usually because one of the verbs had been mistakenly tagged as a noun) and either clause instantiated one of the target patterns, we annotated the first clause in the transcript line (e.g. the underlined string in 28.a). Where only one of the two clauses was non-elliptical, we picked this one irrespective of sequential order (b). If only one of the two clauses instantiated a relevant construction, we picked this element (c):

(28) a. I'll give you Raggedy Ann (.) you give me the blanket . (ENG-US\Sachs\n62.cha)
b. can you give him a [//] are you shaking his hand . (ENG-US\Rollins\ds06.cha)
c. <you're crazy> [//] watch your head honey . (ENG-US\NewEngland\31.cha)

Linear order of the constituents was not seen as crucial. As long as the argument structure could be unambiguously identified, both nominally embedded VPs (29.a-c) and clauses with extraposed complements (d-f) were included in the data:

(29) a. that's a board for putting lego on. (ENG-UK\Manchester\aran16b.cha)
b. a skill nobody needs any more. (ENG-US\Tardif\e01.cha)
c. five pieces to find (8.). (ENG-UK\Wells\iris05.cha)
d. where are you taking the horse? (ENG-UK\Manchester\aran18b.cha)
e. one more red brick I can find in here. (ENG-UK\Manchester\anne29b.cha)
f. out teddy gets . (ENG-UK\Manchester\carl01b.cha)

Among the remaining set of non-elliptical/non-fragmented utterances, all samples that did not match the structural specifications of our five target patterns were excluded. The most common among these were copular constructions (29.a-c) and constructions with verbally selected clausal (d-f), quotative (g-i) or reflexive complements (j-l):

(29) a. now it's closed. (ENG-US\Rollins\cb12.cha)
b. sounds a good idea . (ENG-UK\Manchester\warr20b.cha)
c. it was next to it . (ENG-UK\Manchester\joel33a.cha)
d. show me where your headache is. (ENG-UK\Manchester\ruth32a.cha)
e. I expect hospitals do lovely lunches.  
   (ENG-UK\Manchester\warr17b.cha)

f. maybe you can let Nina play with it for a while.  
   (ENG-US\Suppes\nina05.cha)

g. can you say <my name's Joel and I am two> ['"]?  
   (ENG-UK\Manchester\joel14b.cha)

h. did you tell Fraser goodbye?  
   (ENG-US\Brown\eve18.cha)

i. can you sing Rock_a_bye_baby?  
   (ENG-UK\Manchester\becky06.cha)

j. well () the heater turns on itself ()  
   (ENG-US\Brown\eve19.cha)

k. you've been hurting yourself quite a bit this morning.  
   (ENG-US\Sachs\n07.cha)

l. I'm enjoying myself.  
   (ENG-UK\Manchester\nic30b.cha)  all

Less common patterns that were removed include the configurations verb + (object) + predicative conjunctional phrase (30.a-c), verb + clausal pro-form (d-f), and verb + particle + prepositional phrase (g-i):

(30)

a. will this do as a towel?  
   (ENG-UK\Manchester\anne17b.cha)

b. Child, I see that as a helicopter on top of the landing strip.  
   (ENG-US\Valian\08b.cha)

c. that [//] we can count that as a nightgown.  
   (ENG-US\Suppes\nina31.cha)

d. I think so.  
   (ENG-UK\Wells\nevill02.cha)

e. I hope not.  
   (ENG-US\Suppes\nina03.cha)

f. I suppose so.  
   (ENG-UK\Manchester\becky27b.cha)

g. are you going to give up on those?  
   (ENG-US\Tardif\e08.cha)

h. can you hang on to that.  
   (ENG-US\Rollins\im06.cha)

i. Caroline'll have to put up with you weeing.  
   (ENG-UK\Manchester\gail23b.cha)

Further non-basic constructions that were removed from the dataset include serial verb constructions (31.a-c), non-clause integrated verb reduplications (d-f), periphrastic aspectual constructions of the type go + N and go + V (g-i) and a variety of more marginal clause-level constructions (j-m):

(31)

a. where did you go play with Leo?  
   (ENG-US\Suppes\nina32.cha)

b. can you go hide behind one of the bean bags?  
   (ENG-US\Bates\olivia20.cha)

c. should we go talk to the camera?  
   (ENG-US\NewEngland\06.cha)

d. pop pop!  
   (ENG-US\NewEngland\13.cha)

e. ahhah knock knock knock.  
   (ENG-US\Rollins\tt12.cha)

f. jump jump jump jump jump!  
   (ENG-US\NewEngland\49.cha)

g. did the block go boom?  
   (ENG-US\Bernstein\alice3.cha)

h. can you go clap clap clap?  
   (ENG-US\Rollins\im12.cha)

i. she's going night(ie)+night?  
   (ENG-US\Bates\ruth28.cha)

j. does Mummy's foot need washing?  
   (ENG-UK\Manchester\warr18b.cha)

k. why don't you have him come  
   (ENG-US\Bates\keith20.cha)

Constructivist grammar classifications (Deliverable 3.1)
over here?

l.  <you’re> [←] not to touch the milk.  (ENG-UK\Wells\harrie06.cha)

m.  you’d better stop now.  (ENG-UK\Higginson\june09.cha)

With all these more or less uncontroversially different structures removed, the dataset was reduced to utterances that matched one of the nine structural templates in (32):

(32)  
a.  \[VP V\]            listen         (simple intransitive)
b.  \[VP V [AP A]\]      fall asleep    (intransitive resultative)
c.  \[VP V [Adv P Adv]\] come out      (intransitive resultative)
d.  \[VP V [PP P [NP N]]\] go to bed     (intransitive resultative)
e.  \[VP V [NP N]\]      leave it       (simple transitive)
f.  \[VP V [NP N] [AP A]\] make it dry (transitive resultative)
g.  \[VP V [NP N] [Adv P Adv]\] bring it here (transitive resultative)
h.  \[VP V [NP N] [PP P [NP N]]\] put bunny to sleep (transitive resultative)
i.  \[VP V [NP N] [NP N]\]      give me a kiss (ditransitive)

Note, however, that the match between the structural surface configurations on the left and the constructions on the right is not perfect. For instance, the following sentences match the structural description of different possible realisations of the resultative construction as given in (32), yet none of them conveys the constitutive semantic implication that something is moving/being moved to a new state or location:

(33)  
a.  you’d never pass for an Irish Ma.  (ENG-UK\Wells\gerald03.cha)
b.  they belong to Samantha.  (ENG-US\Suppes\nina34.cha)
c.  you’re interfering with my building.  (ENG-UK\Manchester\domin28b.cha)
d.  you’re doing wonderful.  (ENG-US\Bernstein\dale1.cha)
e.  yeah (.) he’s acting goofy.  (ENG-US\Bates\gloriast.cha)
f.  chicken would do nicely.  (ENG-UK\Manchester\liz28b.cha)
g.  tell me about it.  (ENG-US\Rollins\hp12.cha)
h.  take the chicken for a ride.  (ENG-US\Bates\ivy20.cha)
i.  what are you feeding them on?  (ENG-UK\Manchester\aran30a.cha)
j.  take it easy.  (ENG-US\Valian\13b.cha)
k.  serves him right.  (ENG-UK\Manchester\domin32a.cha)
l.  Mummy’s got it wrong.  (ENG-UK\Manchester\aran22a.cha)

To complicate things further, such isomorphisms may arise both with phrases that are lexically selected (i.e. complements of the given verbs/verb readings) and with structures in which the second postverbal constituent is an optional adjunct. In practice, distinguishing what is a ‘selected complement’ of a verb/verb reading and what is not is often very difficult to tell, and it was not in fact possible to obtain a consistent and non-arbitrary binary classification of our data in terms of this distinction. Intuitively, this may seem unfortunate since a sensible
analysis should recognise that e.g. the following utterances are all of the same basic structural type:

(34)  a.  don’t make such a fuss.
     b.  don’t make such a fuss about it.
     c.  don’t make such a fuss about it again.
     d.  don’t make such a fuss about it again next time etc.

In other words, adding the underlined adjunct phrases in (34.b-d) does not change the basic sentence type of (34.a): it always remains a simple transitive clause, which suggests that an identification of adjuncts is crucial for a correct assignment of an utterance to a given argument structure construction pattern. On the other hand, we do not in fact assume that children are equipped with innate categories for grammatical functions, and hence a pre-existent differentiation between complements and adjuncts: instead, child learners of argument structure have to discover the meanings of verbs and master the structure and function of the different phrasal patterns that co-occur with these verbs in their input on different levels of complexity (such as the noun phrase construction, the prepositional phrase construction, the verb phrase construction or the transitive argument structure construction). Given an understanding of the semantic frame associated with a specific verb and the basic constructional means available for encoding such foundational notions as “location”, “motion”, “possession”, “causation” etc., they will at some point also arrive at an understanding of which elements of a clause are structurally required for encoding a particular construal of a given event and which are merely optional embellishments (for instance, they will realise that prepositional phrases can be used to supply additional information about the spatial, temporal or causal circumstances of an event encoded in a simple transitive construction). Assuming that they do not start out with such information from birth but have to discover it through learning of course complicates a purely distributional identification of meaningful patterns like the resultative, since this opens the door for a wide variety of semantically heterogeneous V + AP/AdvP/PP adjunct structures that enter the input space for patterns II and IV. On the other hand, it is equally important to remember that children bring more to the task of language understanding than a capacity for structural pattern recognition, which means that they will not sort everything into the same category that happens to instantiate the same string of (adult) syntactic categories in a sequence. For instance, going back to the earlier example of the misanalysed putative ditransitive eat it Rosie in (18), the addressee will know that Rosie is her own name and understand that it is added to a simple
transitive construction in order to address her, rather than construct the outlandish ditransitive interpretation that she herself is being requested to eat an entity called Rosie for the benefit of some unspecific other entity it.

This means that some sort of compromise is needed between two conflicting desiderata: on the one hand, the data should not be restricted to a narrow semantic definition of the target structures from the outset, since in this case the statistics would do little more than illustrate the analyst’s preconception of which occurrence of a given string should count as a particular construction and which others should not. On the other hand, the effect of semantic and interactional grounding on children’s grammatical categorisations must not be underestimated either, which means that learner’s sorting of the input into categories that will enter into a particular distributional analysis is not determined by (adult) syntactic category information alone. Here, we will make the following rather conservative assumptions about the semantic and interactional aspects of language understanding that influence children’s categorisation of utterances into argument structure patterns:

- children are able to understand that a single referential entity may be identified by a linguistically complex appositive structure in discourse (e.g. bambi the deer), and conversely use such referential information to disambiguate relevant N + N sequences as a single noun phrase;
- children understand that spoken discourse may be interspersed with elements such as vocatives (e.g. some things you just can't fix Peter) and disjuncts (you need the gates sorry at the same height) that are meaningful components of the larger communicative interaction scenario though not syntactically integrated into the surrounding clause;
- children are able to distinguish prepositions and adverbs that (potentially) encode paths and locations (e.g. in, to, into) from others where this is not the case (e.g. for, with, yesterday).

Staying with the example of the resultative construction, note that these criteria leave a number of expressions in the dataset that will enter the distributional calculations for patterns II and IV even though they are not in fact resultatives:

(35)  a. look at that! (ENG-US\Bernstein\kay1.cha)
b. oh are you dancing to the music? (ENG-US\Rollins\nb12.cha)
c. I couldn't tell her in advance. (ENG-UK\Wells\abiga02.cha)
d. what did you win at the party? (ENG-UK\Manchester\joel06a.cha)
e. we have those at home. (ENG-US\Rollins\me12.cha)

On the other hand, these criteria still ensure that even less closely related utterances like (36) are excluded from consideration:
a. you look for the baby cow. (ENG-UK\Manchester\john19a.cha)
b. perhaps we'd better not bother with that table. (ENG-UK\Manchester\anne20a.cha)
c. can you think of something else? (ENG-US\Valian\12b.cha)
d. you don't care about the top huh? (ENG-US\NewEngland\85.cha)
e. you'd never pass for an Irish Ma. (ENG-UK\Wells\gerald03.cha)

In contrast to the difficult complement/adjunct-distinction, our handling of the problem of grammatical voice is straightforward: given the constructionist focus on surface generalisations – grammatical generalisations over surface forms that do not recognise derivational or transformational relationships between structures (cf. Goldberg 2006) – we did not count active and passive sentences as different realisations of the same underlying structure. Instead, passive occurrences of simple transitive, transitive resultative and ditransitive constructions like those in (37) were discarded from the dataset:

(37) a. the other ones have been eaten. (ENG-US\Manchester\becky31.cha)
   b. they've been dumped down there. (ENG-US\Manchester\joel33a.cha)
   c. are they given a cuddle? (ENG-UK\Manchester\ruth01.cha)

A final question concerns the status of idioms (i.e. prefabs and set expressions) such as the following:

(38) a. thank you. (ENG-US\Rollins\cc12.cha)
   b. oh beg your pardon. (ENG-UK\Howe\oliver1.cha)
   c. there you go. (ENG-US\Post\lew03.cha)

Clearly, these are preassembled formulas whose specialised pragmatic functions attach directly to the lexicalised composite structure. At the same time, in contrast to ‘extragrammatical idioms’ such as by and large, they are still instances of some of the most regular syntactic patterns of English. Since usage-based construction grammar maintains that elements of linguistic knowledge can be stored and analysable at the same time, we decided that a syntactic analysis should therefore classify them syntactically rather than simply discard them as ‘long words’. Unless they were lexicalised to the degree of downright fusion (e.g. peekaboo!, lookit!), idioms and prefabs that were instances of one the investigated argument structure constructions were therefore included in the analysis on the same basis as their non-specialised counterparts in the relevant pattern. This also applied to lexicalised verb-particle combinations like those in (39.a-f):
As long as the adverb/particle had the required directional or locative meaning (outside the bleached verb-particle combination, that is), such combinations were all included in the study.

An advantage of this is that it was not necessary to make more fine-grained distinctions between different readings of one and the same verb-particle combination:

(39)  
\[ \begin{align*} 
    a. & \quad \text{don't wake the baby up.} \quad \text{(ENG-US\NewEngland\66.cha)} \\
    b. & \quad \text{come on then.} \quad \text{(ENG-UK\Howe\lucy1.cha)} \\
    c. & \quad \text{shall I finish off the spider?} \quad \text{(ENG-UK\Manchester\john01a.cha)} \\
    d. & \quad \text{are you going to calm down?} \quad \text{(ENG-UK\Manchester\nic29b.cha)} \\
    e. & \quad \text{oh shut up.} \quad \text{(ENG-UK\Wells\harrie02.cha)} 
\end{align*} \]

Drawing to a close, (41) demonstrates the application of our coding scheme on the example of the verb \textit{to look}. Five out of these nine examples were included in the dataset: (a) and (b) are simple intransitives (in the case of (b) involving an incorporated/fused object \textit{it}), (c) and (d) match the structural template of the intransitive resultative (even though neither of them is in fact a resultative construction) and (e) represents a non-canonical, yet nevertheless included transitive use of \textit{look}. Excluded were the uses with non-locative/directional PPs in (f) and (g), the copular use in (h) and the variant with clausal complementation in (i):

(40)  
\[ \begin{align*} 
    a. & \quad \text{come on Laurie.} \quad \text{(ENG-US\NewEngland\56.cha)} \\
    b. & \quad \text{the refrigerator came on.} \quad \text{(ENG-US\Brown\eve15.cha)} 
\end{align*} \]

(41)  
\[ \begin{align*} 
    a. & \quad \text{Corinna look!} \quad \text{(ENG-US\NewEngland\16.cha)} \\
    b. & \quad \text{oh (.) lookit!} \quad \text{(ENG-US\Bernstein\amelia2.cha)} \\
    c. & \quad \text{look out.} \quad \text{(ENG-US\NewEngland\36.cha)} \\
    d. & \quad \text{look at that!} \quad \text{(ENG-US\Bernstein\kay1.cha)} \\
    e. & \quad \text{what're you looking?} \quad \text{(ENG-US\Brown\adam18.cha)} \\
    f. & \quad \text{are you looking for the baby?} \quad \text{(ENG-UK\Manchester\carl05b.cha)} \\
    g. & \quad \text{farmer's looking after that lamb.} \quad \text{(ENG-UK\Manchester\anne28b.cha)} \\
    h. & \quad \text{do I look a bit funny?} \quad \text{(ENG-UK\Manchester\ruth05a.cha)} \\
    i. & \quad \text{look what we have here Anne.} \quad \text{(ENG-US\NewEngland\28.cha)} 
\end{align*} \]

Applying these criteria, 37,460 out of the 46,207 automatically extracted utterances (81.1\%) were retained in the dataset as instances of one of the five targeted patterns. The following section presents the results of our analysis of the categorised data.
5.4 Results

Table 3 gives an overview of the target pattern frequencies for all three investigated stages:

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Stage I</th>
<th>Stage II</th>
<th>Stage III</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>350</td>
<td>1834</td>
<td>1931</td>
<td>4115</td>
</tr>
<tr>
<td>II</td>
<td>1138</td>
<td>4515</td>
<td>3679</td>
<td>9332</td>
</tr>
<tr>
<td>III</td>
<td>1961</td>
<td>6954</td>
<td>7388</td>
<td>16303</td>
</tr>
<tr>
<td>IV</td>
<td>787</td>
<td>3203</td>
<td>3019</td>
<td>7009</td>
</tr>
<tr>
<td>V</td>
<td>88</td>
<td>276</td>
<td>337</td>
<td>701</td>
</tr>
<tr>
<td>Total</td>
<td>4324</td>
<td>16782</td>
<td>16354</td>
<td>37460</td>
</tr>
</tbody>
</table>

Table 3 Raw pattern frequencies across developmental stages

Figure 8 translates these results into a graphic representation of proportionate frequencies (i.e. the proportion of the total number of utterances per stage that is taken up by each pattern). Even though at least some of them are in fact more inclusive than their names suggest, we will refer to these patterns by the names of their targeted constructional subtypes for ease of exposition (e.g. pattern V = ‘ditransitive’ etc.).

![Figure 8: Proportionate constructional token frequencies across developmental stages](image-url)
The graph not only shows how the token frequencies of our five target patterns compare to one another, but also that the overall frequency of each of the investigated constructions remains more or less constant in CDS across developmental stages. This is not too surprising since these patterns invariably constitute the simplest choice for encoding a particular event type linguistically. However, the picture changes when we have a look at the internal structure of the five categories in our data:

![Figure 9](image)

**Figure 9** Verbal type frequencies in each pattern across developmental stages

Even though their relative token frequencies remain stable, all five categories show a steady increase in verbal type frequency, i.e. become much more diversified internally. Again, such lexical fine-tuning effects are not unexpected: from the perspective of the speaker, there is certainly reason to be confident that a normally developing child between two and three (roughly, our stage III) can be trusted to have mastered more concepts and words than a preverbal infant. This result reflects recent findings by Huttenlocher et al. (2008) who find no quantitative differences in mothers' speech with respect to number of words, sentences or sentence length over time, but report on significant differences with respect to the diversity
and complexity of the linguistic structures for child age. For a detailed look at how the maternal usage patterns of these five constructions attune to this development, we will consider the results for each of our three subcorpora in separate sections.

5.4.1 Stage I
Figure 10 offers an informative visualisation of distributional peaks in the patterning of attested verbs across the investigated constructions. The graph shows that in speech directed at preverbal infants, three of the five target patterns are typically used with exactly one specific verb: go in the case of intransitive resultative (40.1%), put (43.8%) in the case of transitive resultative, and give in the ditransitive (63.6%). These results confirm similar findings by Goldberg et al. 2004 (though the exact percentages differ). As anticipated, the more schematic simple intransitive and simple transitive construction show less extreme lexical biases. However, also the dominant simple intransitive verb look makes up for roughly a quarter of all constructional tokens in the corpus, and in the case of the simple transitive, the three top frequent items (see, get, do) still add up to more than a third of all attestations.

Figure 10  Proportionate collexeme frequencies (stage I)
Zooming in on the semantic structure of the five categories, table 4 reports the top 30 verb–construction associations for each pattern in stage I-CDS:

For each construction, table 4 lists the frequency of the top attracted verbal collexemes in the pattern (‘f’) and their collostruction strength score (‘CxStr’, i.e. significantly attracted combinations have a score <5E-02). We will briefly go through each of the patterns in turn.

By far the most common verb in pattern I is used for the coordination of attention (look, further down on the list are also listen and see). Most other verbs in the pattern likewise refer to aspects of a joint attention scenario, either by providing descriptions of what participants and/or salient discourse referents of the scenario are doing (this includes toys and agents in picture books etc., e.g. you play with the ball, their phone doesn’t ring, the little bunny that hops around?) or by encoding directives (don’t chew, clap with Big Bird, sit). Consequently, most verbs occurring in the pattern denote concrete activities (play, fall, ring, squeak, dance, draw, clap, chew, wave, sit, bounce, quack, push, sing, fly, hop, rock, touch, catch). A second discernible group denotes mental states and perceptions of the interaction participants (look, know, peek, listen, understand, remember). Many of the descriptions in this section of the corpus refer to palpable perceptual qualities (do birds quack?, the duck squeaks, did somebody whistle?), whereas utterances with verbs for more abstract intentional (you missed,
should we try a little bit) or functional or relational concepts (that won’t work, will it fit) are comparably rare. In terms of semantic role types, more than half of the 30 top attracted verbs in our data have unambiguously agentive subjects (e.g. play, dance, clap), but there are also experiencers (look), themes (bounce) and patients (fall). On the whole, then, stage I uses of pattern I do not prefigure a specific constructional meaning apart from signalling the relationship between an event and a single participant (with an underspecified role).

This is different in pattern II, the intransitive resultative. As indicated in figure 10, this structural configuration is much more strongly centred on three individual verbs – go, come and look – which taken together make up for more than 80% of the tokens. Dominant uses of the pattern in stage I CDS thus already prefigure the three most important constructions in the construction grammar sense that instantiate this pattern: the first two are the intransitive motion construction (come here) and the intransitive resultative construction (is he going to sleepie) that were analysed as metaphorically related subtypes of a single overarching construction in section 5.2.2. The third one is the so-called conative construction which encodes that an agent directs an action (including abstract processes such attending to something) at some other participant (look at that!). Most other verbs in our data pattern neatly into these three categories: top attracted items are to a large extent motion verbs (go, come, fall, walk, run, crawl, ride, roll, hop, rock) and conative verbs (look, talk, point, write, listen, growl, draw, speak). However, children are also exposed to deviant uses of the pattern that cannot be subsumed under the above generalisations from the beginning. The first type of these are locative adjunct uses with stative and posture verbs that do not convey any sense of directedness or orientation toward a result (stand, sit, lie, hide). A second class ‘deviant’ uses are verb-particle constructions and verbs with prepositional complements. Such expressions form a continuum from transparently motivated (grow up, break out) to ever more arbitrary combinations (watch out > happen to > come on). From a usage-based perspective, it is predicated that frequent exposure to structures without an obvious semantic motivation for the locative complement will work against the discovery of the resultative generalisation. The data show that children are exposed to these ‘counterexamples’ from the very beginning, albeit many of them are quasi-lexicalised and presumably learned as word-like entrenched exceptions (come on, watch out). The existence of such cases notwithstanding, pattern II shows a marked distributional centering on three prototypical usages of the structure that facilitate learner’s discovery of the appropriate generalisation.

Moving on to the simple transitive construction, the first thing to note in the results is the discrepancy between the verb raw frequency ranking in figure 10 and the statistical
association strength ranking reported in table 4. Specifically, the third most frequent collexeme do only comes sixth in the association strength ranking, and like the two top collexemes of the class, the three less frequent but more distinctive verbs that come before it are not what one would think of as prototypically transitive items (thank (you) is a lexicalised exception, and like and want are experiencer verbs whose implicit arguments are not effected by the verb event). Indeed, especially among the most strongly attracted items, prototypically transitive verbs are comparably rare in stage I CDS: none of the top 5 and less than half of the top 30 collexemes refer to strongly asymmetrical interactions between prototypically agentive subjects and prototypically patient-like objects. Pragmatically, of course, the ranking in table 4 it is not surprising (i.e. the communicative need to talk about seeing, getting, wanting and having things in mother-child interaction is greater than talking about e.g. hitting or kicking them). Like the simple intransitive, then, the collexeme ranking of the simple transitive in stage I CDS is strongly shaped by extralinguistic factors, does not invite obvious generalisations from salient instances to a shared/dominant constructional meaning and is not markedly centered on a single pathbreaking verb either.

The transitive resultative is different again: not only is there a single prototypical verb again (put, occurring almost five times as often as the second most frequent collexeme), but also a clear semantic patterning of the remaining items, almost all of which are verbs of caused motion (take, throw, bring, pick (up), get, roll, knock (over), pull, stack, set, peel (off), sling, tip (over), wipe (off), drop). The few exceptions illustrate stative uses of the pattern (you want it in this hand maybe?, it has lots of colors on it, leave it here) and a few complex transitive verb-particle constructions/prepositional complementation patterns with various degrees of semantic bleaching of the locative complement (open X up, close X down, check/test X out, feed X on Y). On the whole, maternal usage of this pattern in stage I CDS provides a good basis for arriving at a coherent semantic generalisation about the typical meaning/function of this structure.

This is even more strongly the case with the ditransitive construction. Here, the prototypical filler verb give accounts for near two thirds of the entire tokens. The remaining verbs (only 11 other types in stage I) are either like give in that they refer to situations of concrete physical (hand, bring, get) or abstract metaphorical transfer (tell, read, show, wish, sing). A prominent subtype in the latter pattern are verbs of communication. The remaining items are metonymically related to the central constructional meaning by encoding preconditions (bake) or means (throw) of the designated transfer event. The only type that is not transparently
motivated is a fixed expression involving the light verb do (i.e. do me a favour). Still, all types invite the generalisation that the structure is used for encoding a transfer scenario.

5.4.2 Stage II
In Stage II, we get a similar picture as in Stage I, yet the comparison reflects the increasing variability we have seen in Figure 9; Figure 11 illustrates the distribution of verbs across patterns in stage-II CDS:

![Figure 11: Proportionate collexeme frequencies (stage II)](image)

Even though there are characteristic distributional peaks again, it is interesting to see that they are less marked than in stage I: among the strongly skewed patterns, only the transitive resultative remains constant with put slightly over 40 percent of all tokens. By contrast, the top frequent collexeme of the intransitive resultative drops by more than ten percent, and the top ditransitive one even by more than 20 percent. Table 5 reports the top collexemes for each pattern in this stage:
Constructivist grammar classifications (Deliverable 3.1)

In the simple intransitive, more than half of the most strongly attracted verbs were also among the top collexemes in stage I. Most changes in the ranking reflect the addition of new activities to typical interaction scenarios in this age group (e.g. color, write), whereas the marked jump of touch from rank 28 to rank 2 reflects the fact that children are now also beginning to explore the environment on their own (don’t touch!). As in stage I, the top combinations also contain a few mental predicates, but on the whole, the pattern shows a clear association with concrete activities (most often with agentive subjects). The changes also reflect the children’s linguistic development; for instance, while draw is significantly attracted to the intransitive resultative in stage I, as in draw here, in Stage II (and III) it occurs significantly more often in the transitive construction, as in draw an X. The example illustrates that the verb-construction associations are functionally determined.

As regards pattern II, about half of the most strongly attracted collexemes are new in the top 30 as compared to stage I. However, the new items are from classes that already dominated the picture in stage I: half of the new ones are other motion or posture verbs (climb, jump, lean), the rest are verb-particle constructions and prepositional verbs such as hang on, belong to and clean up.

In pattern III, the highest ranking verb is now also a light verb: do comes in almost twice as often as the second most frequent item, see. With the exception of watch and smell, all new
items among the top instances now have agentive subjects (*use, close, draw, touch, fix, crack, pat, miss*). In other words, more typical transitives are beginning to encroach on the pattern in stage II, but other uses still remain highly salient (cf. X see/want/have Y).

Reminiscent of pattern II, about half of the top verbs in pattern IV have changed, yet the new additions again pattern neatly into two already established classes: motion and posture verbs (*push, turn, stick*) and verb-particle constructions, many of them with completive *up* (*drink up, clear up, line up*). Here, the exceptional status of the pathbreaking verb is most pronounced: *put* is nearly seven times as frequent as the second most frequent pattern-IV verb, *get*.

Also in pattern V, the first exceptional (i.e. non-ditransitive) uses of the structure now appear in the data (*to call sb. sth., to make sb. sth., to leave sb. sth, to take sb. some time*). Apart from that, new additions to the list are primarily verbs of creation (*build, draw, play, write*), but there are also other kinds of benefactives (*feed, find, save*) and new concrete (*roll*) and abstract transfer verbs (*find*). With only four deviant items and all other subtypes transparently related, the motivation for semantic generalisations over the data remains strongest in pattern V.

Noteworthy is also the much higher variability of the ditransitive construction; as Figure 12 reveals, the verb *give* accounts for 'only' 42.8% in Stage II, in comparison with the 64.5% in Stage I. Furthermore, as Table 5 shows, in Stage II already 21 different verbs are associated with the construction, compared to the only eight types in Stage I.

**5.4.3 Stage III**

Turning to stage III-CDS, the following verbs emerge as most frequent in the five target patterns:
These figures are very similar to stage II: with the exception of know, the peaks are still on the same verbs, and also the proportions remain more or less constant: the count for give is quasi-identical, put drops by a few percent and go goes up by roughly the same amount. Also in the collexeme ranking reported in table 6, there are no drastic changes as compared to the trends already identified in the earlier stages:

![Proportionate collexeme frequencies (stage III)](image)

**Figure 12** Proportionate collexeme frequencies (stage III)
Table 6  Collexeme analysis for all five patterns in stage III

Seeing that there are no major new developments, we will not go through each individual construction again. Instead, as a final step, we will have a look at the data of all three subcorpora combined, which gives an approximation of the cumulated experience with argument structure patterns that a child has acquired by the age of three years.

5.4.4 Whole corpus

Figure 13 and table 7 offer the by now familiar overview of top frequent/attracted verbs in each pattern:

![Proportionate collexeme frequencies (whole corpus)](image)

Figure 13  Proportionate collexeme frequencies (whole corpus)
Constructivist grammar classifications (Deliverable 3.1)

Our five target patterns. These 52 types add up to 31,656 occurrences or 84.5% of the entire text. They occurred more than 100 times in the corpus and were significantly attracted to at least one of the most frequent core elements of mothers’ vocabulary. For this, we identified all verbs that occurred at least 30 times, and each structural configuration occurs with many more than just 30 verbs in the corpus. Table 7 of course only presents a small subset of the results: with the exception of pattern V (+CxStr), we show the most frequent verb for all other patterns.

Table 7

Collexeme analysis for all five patterns in the whole corpus

Table 7 of course only presents a small subset of the results: with the exception of pattern V (+CxStr), each structural configuration occurs with many more than just 30 verbs in the input. In all, we found 593 different verb types occurring in our five target patterns. We do not give a complete matrix of the distribution here for reasons of space but concentrate on the most frequent core elements of mothers’ vocabulary. For this, we identified all verbs that occurred more than 100 times in the corpus and were significantly attracted to at least one of our five target patterns. These 52 types add up to 31,656 occurrences or 84.5% of the entire tokens. Table 8 reports the distribution of these core items across the investigated patterns and indicates statistically significant associations where applicable:
Table 8 shows also that the 52 most frequent verbs in CDS are never unique to a particular construction, which reflects the highly reformulative character of CDS, discussed in section 3.3. That is, as Table 8 reveals, mothers present the verbs in their utterances in several different distributional patterns, providing their children with variation sets (Küntay & Slobin 1996; Onnis et al. 2008). In particular, 88.5% of the verbs are attested in three or more different constructions. At the same time, however, 92.3% of these verbs are significantly associated with only one or two of these constructions.

Figure 14 presents a visualization of the distribution of the 52 most frequent verb types across different argument structure constructions, as well as their significant attraction to particular constructions. Thus, while verbs typically occur in several different argument structure constructions, they exhibit distributional biases towards only one or two of these constructions.
Finally, switching from the verb to the pattern perspective, Table 9 reports the quantitative proportions between the 50 most frequent collexemes for each pattern, giving the raw frequency of each type and the corresponding percentage among the pattern’s overall tokens.

<table>
<thead>
<tr>
<th>Verb</th>
<th>f</th>
<th>%</th>
<th>Verb</th>
<th>f</th>
<th>%</th>
<th>Verb</th>
<th>F</th>
<th>%</th>
<th>Verb</th>
<th>f</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>look</td>
<td>411.00</td>
<td>9.99</td>
<td>go</td>
<td>3116</td>
<td>33.39</td>
<td>do</td>
<td>2478</td>
<td>15.20</td>
<td>put</td>
<td>2751</td>
<td>39.25</td>
</tr>
<tr>
<td>know</td>
<td>370.00</td>
<td>8.99</td>
<td>come</td>
<td>2207</td>
<td>23.65</td>
<td>get</td>
<td>1423</td>
<td>8.73</td>
<td>get</td>
<td>602</td>
<td>8.59</td>
</tr>
<tr>
<td>play</td>
<td>349.00</td>
<td>8.48</td>
<td>look</td>
<td>1336</td>
<td>14.32</td>
<td>see</td>
<td>1360</td>
<td>8.34</td>
<td>take</td>
<td>384</td>
<td>5.48</td>
</tr>
<tr>
<td>see</td>
<td>178.00</td>
<td>4.33</td>
<td>sit</td>
<td>470</td>
<td>5.04</td>
<td>have</td>
<td>1282</td>
<td>7.86</td>
<td>have</td>
<td>309</td>
<td>4.41</td>
</tr>
<tr>
<td>go</td>
<td>165.00</td>
<td>4.01</td>
<td>get</td>
<td>376</td>
<td>4.03</td>
<td>want</td>
<td>1200</td>
<td>7.36</td>
<td>throw</td>
<td>282</td>
<td>4.02</td>
</tr>
<tr>
<td>happen</td>
<td>130.00</td>
<td>3.16</td>
<td>fall</td>
<td>221</td>
<td>2.37</td>
<td>like</td>
<td>870</td>
<td>5.34</td>
<td>want</td>
<td>234</td>
<td>3.34</td>
</tr>
<tr>
<td>touch</td>
<td>127.00</td>
<td>3.00</td>
<td>stand</td>
<td>117</td>
<td>1.25</td>
<td>make</td>
<td>577</td>
<td>3.54</td>
<td>push</td>
<td>194</td>
<td>2.77</td>
</tr>
<tr>
<td>remember</td>
<td>122.00</td>
<td>2.96</td>
<td>happen</td>
<td>108</td>
<td>1.16</td>
<td>eat</td>
<td>384</td>
<td>2.36</td>
<td>turn</td>
<td>190</td>
<td>2.71</td>
</tr>
<tr>
<td>draw</td>
<td>96.00</td>
<td>2.33</td>
<td>stay</td>
<td>98</td>
<td>1.05</td>
<td>find</td>
<td>336</td>
<td>2.06</td>
<td>pick</td>
<td>147</td>
<td>2.10</td>
</tr>
<tr>
<td>come</td>
<td>87.00</td>
<td>2.11</td>
<td>tall</td>
<td>91</td>
<td>0.98</td>
<td>try</td>
<td>246</td>
<td>1.51</td>
<td>bring</td>
<td>130</td>
<td>1.85</td>
</tr>
<tr>
<td>cry</td>
<td>76.00</td>
<td>1.85</td>
<td>hang</td>
<td>65</td>
<td>0.70</td>
<td>turn</td>
<td>227</td>
<td>1.39</td>
<td>give</td>
<td>124</td>
<td>1.77</td>
</tr>
<tr>
<td>catch</td>
<td>70.00</td>
<td>1.70</td>
<td>fit</td>
<td>52</td>
<td>0.56</td>
<td>need</td>
<td>198</td>
<td>1.21</td>
<td>leave</td>
<td>114</td>
<td>1.63</td>
</tr>
<tr>
<td>color</td>
<td>66.00</td>
<td>1.60</td>
<td>turn</td>
<td>47</td>
<td>0.50</td>
<td>draw</td>
<td>186</td>
<td>1.14</td>
<td>see</td>
<td>112</td>
<td>1.60</td>
</tr>
<tr>
<td>fit</td>
<td>56.00</td>
<td>1.36</td>
<td>write</td>
<td>46</td>
<td>0.49</td>
<td>hold</td>
<td>162</td>
<td>0.99</td>
<td>pull</td>
<td>97</td>
<td>1.38</td>
</tr>
<tr>
<td>sleep</td>
<td>54.00</td>
<td>1.31</td>
<td>belong</td>
<td>38</td>
<td>0.41</td>
<td>know</td>
<td>159</td>
<td>0.98</td>
<td>do</td>
<td>78</td>
<td>1.11</td>
</tr>
<tr>
<td>try</td>
<td>52.12</td>
<td>1.26</td>
<td>climb</td>
<td>37</td>
<td>0.40</td>
<td>say</td>
<td>157</td>
<td>0.96</td>
<td>make</td>
<td>76</td>
<td>1.08</td>
</tr>
<tr>
<td>fall</td>
<td>50.00</td>
<td>1.22</td>
<td>play</td>
<td>34</td>
<td>0.36</td>
<td>open</td>
<td>151</td>
<td>0.93</td>
<td>knock</td>
<td>67</td>
<td>0.96</td>
</tr>
<tr>
<td>hurt</td>
<td>49.00</td>
<td>1.19</td>
<td>draw</td>
<td>34</td>
<td>0.36</td>
<td>take</td>
<td>148</td>
<td>0.91</td>
<td>roll</td>
<td>58</td>
<td>0.83</td>
</tr>
<tr>
<td>eat</td>
<td>47.00</td>
<td>1.14</td>
<td>hide</td>
<td>34</td>
<td>0.36</td>
<td>touch</td>
<td>142</td>
<td>0.87</td>
<td>stick</td>
<td>40</td>
<td>0.57</td>
</tr>
<tr>
<td>write</td>
<td>46.00</td>
<td>1.12</td>
<td>live</td>
<td>30</td>
<td>0.32</td>
<td>think</td>
<td>135</td>
<td>0.83</td>
<td>hold</td>
<td>35</td>
<td>0.50</td>
</tr>
<tr>
<td>work</td>
<td>45.00</td>
<td>1.09</td>
<td>run</td>
<td>29</td>
<td>0.31</td>
<td>throw</td>
<td>133</td>
<td>0.82</td>
<td>keep</td>
<td>35</td>
<td>0.50</td>
</tr>
<tr>
<td>move</td>
<td>44.00</td>
<td>1.07</td>
<td>jump</td>
<td>27</td>
<td>0.29</td>
<td>wait</td>
<td>125</td>
<td>0.77</td>
<td>like</td>
<td>34</td>
<td>0.49</td>
</tr>
<tr>
<td>hide</td>
<td>43.00</td>
<td>1.04</td>
<td>push</td>
<td>27</td>
<td>0.29</td>
<td>push</td>
<td>117</td>
<td>0.72</td>
<td>let</td>
<td>34</td>
<td>0.49</td>
</tr>
<tr>
<td>sit</td>
<td>41.00</td>
<td>1.00</td>
<td>lie</td>
<td>27</td>
<td>0.29</td>
<td>play</td>
<td>111</td>
<td>0.68</td>
<td>eat</td>
<td>32</td>
<td>0.46</td>
</tr>
<tr>
<td>dance</td>
<td>40.00</td>
<td>0.97</td>
<td>hold</td>
<td>26</td>
<td>0.28</td>
<td>read</td>
<td>111</td>
<td>0.68</td>
<td>find</td>
<td>29</td>
<td>0.41</td>
</tr>
<tr>
<td>read</td>
<td>40.00</td>
<td>0.97</td>
<td>sleep</td>
<td>25</td>
<td>0.27</td>
<td>hear</td>
<td>111</td>
<td>0.68</td>
<td>open</td>
<td>28</td>
<td>0.40</td>
</tr>
<tr>
<td>do</td>
<td>39.00</td>
<td>0.95</td>
<td>walk</td>
<td>24</td>
<td>0.26</td>
<td>remember</td>
<td>103</td>
<td>0.63</td>
<td>stand</td>
<td>28</td>
<td>0.40</td>
</tr>
<tr>
<td>think</td>
<td>38.00</td>
<td>0.92</td>
<td>ride</td>
<td>24</td>
<td>0.26</td>
<td>watch</td>
<td>102</td>
<td>0.63</td>
<td>move</td>
<td>26</td>
<td>0.37</td>
</tr>
<tr>
<td>listen</td>
<td>36.00</td>
<td>0.87</td>
<td>see</td>
<td>23</td>
<td>0.25</td>
<td>use</td>
<td>96</td>
<td>0.59</td>
<td>tip</td>
<td>26</td>
<td>0.37</td>
</tr>
<tr>
<td>peek</td>
<td>31.00</td>
<td>0.75</td>
<td>listen</td>
<td>19</td>
<td>0.20</td>
<td>tell</td>
<td>93</td>
<td>0.57</td>
<td>need</td>
<td>24</td>
<td>0.34</td>
</tr>
<tr>
<td>open</td>
<td>29.00</td>
<td>0.70</td>
<td>move</td>
<td>16</td>
<td>0.17</td>
<td>build</td>
<td>92</td>
<td>0.56</td>
<td>stack</td>
<td>23</td>
<td>0.33</td>
</tr>
<tr>
<td>wait</td>
<td>28.00</td>
<td>0.68</td>
<td>chew</td>
<td>16</td>
<td>0.17</td>
<td>brush</td>
<td>88</td>
<td>0.54</td>
<td>close</td>
<td>21</td>
<td>0.30</td>
</tr>
<tr>
<td>run</td>
<td>25.00</td>
<td>0.61</td>
<td>pop</td>
<td>16</td>
<td>0.17</td>
<td>catch</td>
<td>87</td>
<td>0.53</td>
<td>drink</td>
<td>19</td>
<td>0.27</td>
</tr>
<tr>
<td>throw</td>
<td>25.00</td>
<td>0.61</td>
<td>point</td>
<td>15</td>
<td>0.16</td>
<td>close</td>
<td>85</td>
<td>0.52</td>
<td>read</td>
<td>17</td>
<td>0.24</td>
</tr>
<tr>
<td>walk</td>
<td>25.00</td>
<td>0.61</td>
<td>want</td>
<td>13</td>
<td>0.14</td>
<td>break</td>
<td>80</td>
<td>0.49</td>
<td>build</td>
<td>16</td>
<td>0.23</td>
</tr>
<tr>
<td>bounce</td>
<td>24.00</td>
<td>0.58</td>
<td>wake</td>
<td>13</td>
<td>0.14</td>
<td>love</td>
<td>72</td>
<td>0.44</td>
<td>wipe</td>
<td>14</td>
<td>0.20</td>
</tr>
<tr>
<td>sing</td>
<td>24.00</td>
<td>0.58</td>
<td>fly</td>
<td>12</td>
<td>0.13</td>
<td>thank</td>
<td>72</td>
<td>0.44</td>
<td>say</td>
<td>13</td>
<td>0.19</td>
</tr>
<tr>
<td>drive</td>
<td>23.00</td>
<td>0.56</td>
<td>stick</td>
<td>12</td>
<td>0.13</td>
<td>wear</td>
<td>70</td>
<td>0.43</td>
<td>color</td>
<td>13</td>
<td>0.19</td>
</tr>
<tr>
<td>fly</td>
<td>21.00</td>
<td>0.51</td>
<td>love</td>
<td>12</td>
<td>0.13</td>
<td>hurt</td>
<td>68</td>
<td>0.42</td>
<td>send</td>
<td>13</td>
<td>0.19</td>
</tr>
<tr>
<td>jump</td>
<td>21.00</td>
<td>0.51</td>
<td>eat</td>
<td>11</td>
<td>0.12</td>
<td>move</td>
<td>66</td>
<td>0.40</td>
<td>wind</td>
<td>13</td>
<td>0.19</td>
</tr>
<tr>
<td>speak</td>
<td>21.00</td>
<td>0.51</td>
<td>throw</td>
<td>11</td>
<td>0.12</td>
<td>pull</td>
<td>64</td>
<td>0.43</td>
<td>drop</td>
<td>12</td>
<td>0.17</td>
</tr>
<tr>
<td>stop</td>
<td>21.00</td>
<td>0.51</td>
<td>pull</td>
<td>11</td>
<td>0.12</td>
<td>drop</td>
<td>63</td>
<td>0.39</td>
<td>set</td>
<td>12</td>
<td>0.17</td>
</tr>
<tr>
<td>miss</td>
<td>20.00</td>
<td>0.49</td>
<td>roll</td>
<td>11</td>
<td>0.12</td>
<td>fix</td>
<td>63</td>
<td>0.39</td>
<td>cover</td>
<td>12</td>
<td>0.17</td>
</tr>
<tr>
<td>crash</td>
<td>19.00</td>
<td>0.46</td>
<td>make</td>
<td>11</td>
<td>0.12</td>
<td>help</td>
<td>60</td>
<td>0.37</td>
<td>cut</td>
<td>11</td>
<td>0.16</td>
</tr>
<tr>
<td>help</td>
<td>19.00</td>
<td>0.46</td>
<td>hurry</td>
<td>11</td>
<td>0.12</td>
<td>sing</td>
<td>58</td>
<td>0.36</td>
<td>hang</td>
<td>11</td>
<td>0.16</td>
</tr>
</tbody>
</table>
Using the information in these tables, it is now possible devise empirically motivated experimental stimuli for grounded language learning that reflect important statistical properties of naturalistic CDS such as

- distributional biases for individual ‘pathbreaking’ verbs;
- lexical overlaps between the investigated patterns;
- frequency rankings within the internal structure of a given category.

### 6. Sample stimuli for grounded learning of grammatical constructions

On the basis of the empirical results on verb-construction associations just presented, experimental stimuli for language learning with iCub will be developed.

All stimuli are constrained by the following criteria:

- the stimuli represent natural language utterances;
- they reflect the findings of the empirical study in content and distribution;
- iCub is capable of carrying out the actions described;
- the meaning of the action can be learned by remote-control (for details on the experimental set-up and learning architecture, see Sugita & Tani 2005).


research question:

replication of Sugita & Tani (2005) using the humanoid child-like robot iCub constructions learned:

transitive construction

verb types:

\[ push, hit, touch^{18} \]

---

\(^{18}\)The actions and corresponding verbs chosen here are two of the three used by Sugita and Tani (2005), \textit{push} and \textit{hit}, while \textit{point} cannot be used in transitive constructions in English; thus, instead we suggest to use \textit{touch} instead.
verb tokens:

\textit{VERB(imp)-the-NOUN.}

stimuli:

18 (14 for training and four for testing)

sample stimuli:

push-the-block.
push-the-cup.
push-the-doll.
hit-the-block.
hit-the-cup.
hit-the-doll.
touch-the-block.
touch-the-cup.
touch-the-doll.

6.2. Sample stimuli for experiment II: Extensions

research questions:

- role of order of presentation (construction conspiracy)
- role of constructional overlap
- role of presentation of stimuli for each construction individually, sequentially (in different orders) or all constructions at once

constructions learned:

intransitive construction
intransitive resultative construction
transitive construction
transitive resultative construction
ditransitive construction

verb types:

- Condition I: 100\% construction overlap

\textit{throw, draw, push}^{19}

---

\textsuperscript{19}The verbs used in all five constructions in CDS (see table 7) often necessitate abstract nouns (for instance, \textit{play the game}) or actions iCub is not able to carry out (\textit{read, write}). Thus, the list was adapted to the experimental set-up.
• Condition II: no construction overlap/unique verbs per construction
  intransitive construction
    *nod, smile, blink*
  intransitive resultative construction
    *move, point, turn*
  transitive construction
    *push, hit, touch*
  transitive resultative construction
    *throw, roll, take*
  ditransitive construction
    *bring, show, slide*

• Condition III: empirically determined construction overlap/variation sets (see Table 8)
  *get* (transitive, transitive resultative, and ditransitive construction)
  *take* (transitive and transitive resultative construction)
  *throw* (transitive resultative and ditransitive construction)
  *draw* (intransitive, intransitive resultative and transitive construction)

verb tokens:
  intransitive construction
    *VERB(imp).*
  intransitive resultative construction
    *VERB(imp)-RP.*
  transitive construction
    *VERB(imp)-the-NOUN.*
  transitive resultative construction
    *VERB(imp)-the-NOUN-RP.*
  ditransitive construction
    *VERB(imp)-the-NOUN-the-NOUN.*

stimuli:
  18 per construction (14 for training and four for testing)

sample stimuli:
  *nod.*
  *smile.*
  *move-to-the-right.*
  *move-to-the-left.*
point-to-the-right.
point-to-the-right.
push-the-block.
push-the cup.
hit-the-block.
hit-the-cup.
throw-the-block-to-the-right.
throw-the-block-to-the-left.
roll-the-block-to-the-left.
roll-the-block-to-the-right.
bring-the-teddy-the-cup.
bring-the-teddy-the-block.
show-the-teddy-the-cup.
show-the-teddy-the-block.

6.3. Sample stimuli for experiment III: Facilitating construction learning using frequency information

research question:

replication of Experiment II using empirically established prototypes

constructions learned:

as in Experiment II

verb types (see Table 4 for Stage I prototypes):

intransitive construction

play, smile, wave

intransitive resultative construction

look, point, turn

transitive construction

see, have, get

transitive resultative construction

put, take, roll

ditransitive construction

give, show, pass

verb tokens:

as in Experiment II
stimuli:

18 per construction (14 for training and four for testing)

sample stimuli:

play.
smile.
wave.
look-to-the-right.
look-to-the-left.
see-the-block.
see-the-cup.
put-the-cup-to-the-left.
put-the-block-to-the-right.
give-the-teddy-the-cup.
give-the-teddy-the-block.

6.4. Sample stimuli for experiment IV: Constraining generalisation using frequency information

research question:

- role of frequency information in constructional overlap
- learning semantic distinctions using with prototypes and prototypical distributions

constructions learned:

transitive resultative construction
ditransitive construction

verb types (see Table 7):

transitive resultative construction

put, take, throw
ditransitive construction

give, show, throw

verb tokens:

VERB(imp)-the-NOUN-RP.
VERB(imp)-the-NOUN-the-NOUN.
stimuli:

- 50 stimuli (40 for training, 10 for testing)
- distribution of verbs per construction according to their distribution in CDS (Table 9)²⁰

<table>
<thead>
<tr>
<th>Verb</th>
<th>Count</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>put-the-NOUN</td>
<td>20</td>
<td>put-the-NOUN-RP.</td>
</tr>
<tr>
<td>take-the-NOUN</td>
<td>3</td>
<td>take-the-NOUN-RP.</td>
</tr>
<tr>
<td>throw-the-NOUN</td>
<td>2</td>
<td>throw-the-NOUN-RP.</td>
</tr>
<tr>
<td>give-the-NOUN</td>
<td>22</td>
<td>give-the-NOUN-the-NOUN.</td>
</tr>
<tr>
<td>show-the-NOUN</td>
<td>3</td>
<td>show-the-NOUN-the-NOUN.</td>
</tr>
<tr>
<td>throw-the-NOUN</td>
<td>2</td>
<td>throw-the-NOUN-the-NOUN.</td>
</tr>
</tbody>
</table>

sample stimuli:  

- put-the-block-to-the-left.
- put-the-cup-to-the-right.
- take-the-block-to-the-left.
- take-the-block-to-the-right.
- throw-the-block-to-the-teddy.
- give-the-the-teddy-the-block.
- show-the-teddy-the-block.
- throw-the-teddy-the-block.

6.5. Conclusion and future work

Recent developments in cognitive linguistics, developmental psychology, cognitive science and various other disciplines have contributed to a new perspective on language acquisition which assumes that children construct their linguistic systems on the basis of the massive input they receive, relying on the one hand on their cognitive and social prerequisites and on facilitating cues provided in their caretakers' speech on the other. In this study, we analysed how mothers present their children with five basic argument-structure constructions in a large corpus of child-directed speech, detailing the potential role of frequency and skewed input in verb-construction associations for construction learning. On the basis of these findings, we proposed grammar learning experiments investigating the role of these potentially facilitating effects for automatic language learning.

²⁰The distributions were calculated by assuming that these verbs together (104.8 according to Table 9) constitute 100% of the data (104.8 = 100%, yielding for give, for instance, 45.9 = 43.8%, which translates into 44 stimuli). In order to keep the amount of stimuli for this experiment manageable, the resulting numbers were divided by two (yielding 22 stimuli for give).
Our study was restricted in numerous respects; for instance, many utterances directed towards the children do not contain verbs at all and thus do not instantiate argument-structure constructions. Yet, these utterances contribute to language learning (e.g. isolated words have been shown to facilitate segmentation and word learning, cf. Brent & Siskind 2001). We also focused on a limited set of argument-structure constructions and ignored numerous aspects of linguistic form, such as grammatical mood, tense and aspect, as well as potential fillers in the noun and the RP slots, although the item-based nature of children's speech suggests that also the concrete realisation of these slots may play a crucial role (see also Stefanowitsch & Gries 2005 on the analysis of skewed input with respect to the possessive construction, which exhibits effects similar to those investigated here for the verb-construction associations).

Furthermore, frequency is only one, though crucial, feature characterizing child-directed speech; other facilitative properties still need to be investigated and included into the experiments.

In order to address some of the shortcomings of the given study, we propose at least the following extensions of the experiments outlined:

- One possible extension of the experiments presented here concerns the sentence types in which the linguistic material is presented to the robot, which could correspond to the inclusion of different modes of learning; for instance, the robot could learn from observation instead of from being remote-controlled. For instance, the input used in the experiments described above consists of utterances in the imperative; one extension could be to generalise the linguistic structures acquired to other sentence types, for example, to the declarative.

- The categories used will be extended to more realistic variability with respect to both form and meaning; for instance, the RP may comprise, besides to-the-NOUN phrases used above also adverbs, like away, down, or left or prepositions, like up, and they may comprise, besides the typical target locations, also descriptions of the source, as in she drank water from the tap, allowing the investigation of polysemy effects of the resultative (for discussion, see 5.2.2 and 5.2.4).

- We also intend to develop experiments on the basis of input created from the three developmental stages investigated separately, using the increasing variability mothers produce for their children to extend the robot's linguistic knowledge in naturalistic ways.

- Furthermore, as we have seen in section 3.3, the verb-construction associations are only one type of facilitative cue the child is presented with; another central point is the
repetitive and reformulative nature of CDS, providing the child with rich distributional information. Recently, Onnis et al. (2008) have shown that variation sets can facilitate automatic language learning as well. While they have focused on artificial language learning, the aim of our experiments will be to employ variation sets in natural language learning experiments.

Thus, several further extensions of the experiments described here are conceivable. To conclude, in spite of its limitations, the study presented constitutes a first step into grounded language learning experiments that rely on empirically established constructivist grammar classifications. While previous work in grammar learning has shown that the effects studied here may facilitate language learning in principle, the empirical basis of the classifications developed here allow the experimentation with natural language learning on the level of both content and distribution of different readings and through time on a solid, empirical basis.

7. References


Abbot-Smith, K., Lieven, E. V. M. & Tomasello, M. 2004. Training 2;6-year-olds to produce the transitive construction: The role of frequency, semantic similarity and shared syntactic distribution. Developmental Science 7, 48-55.


