Learner corpora research: Acquisition of ditransitive constructions

Proroković, Jakov

Master's thesis / Diplomski rad

2018

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zadar / Sveučilište u Zadru**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:162:671160

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-03-29**

Image not found or type unknown

Repository / Repozitorij:

University of Zadar Institutional Repository

zir.nsk.hr

Image not found or type unknown

Sveučilište u Zadru

Odjel za anglistiku

Diplomski sveučilišni studij engleskog jezika i književnosti; smjer: znanstveni (dvopredmetni)

Jakov Prorokovi

Learner corpora research: Acquisition of ditransitive constructions

Diplomski rad

Zadar, 2018

Sveučilište u Zadru

Odjel za anglistiku
Diplomski sveučilišni studij engleskog jezika i književnosti; smjer: znanstveni (dvopredmetni)

Learner corpora research: Acquisition of ditransitive constructions

Diplomski rad

Student/ica:
Jakov Proroković

Mentor/ica:
Doc.dr.sc. Lidija Štrmelj

Zadar, 2018



Izjava o akademskoj čestitosti

Ja, Jakov Proroković**,** ovime izjavljujem da je moj diplomski rad pod naslovom **Learner corpora research: Acquisition of ditransitive constructions** rezultat mojega vlastitog rada, da se temelji na mojim istraživanjima te da se oslanja na izvore i radove navedene u bilješkama i popisu literature. Ni jedan dio mojega rada nije napisan na nedopušten način, odnosno nije prepisan iz necitiranih radova i ne krši bilo čija autorska prava.

Izjavljujem da ni jedan dio ovoga rada nije iskorišten u kojem drugom radu pri bilo kojoj drugoj visokoškolskoj, znanstvenoj, obrazovnoj ili inoj ustanovi.

Sadržaj mojega rada u potpunosti odgovara sadržaju obranjenoga i nakon obrane uređenoga rada.

Zadar, 2018

Table of contents

1. Introduction 1
2. Two paradigms within language acquisition research 2
3. Ditransitives 8
4. CHILDES corpora 13
5. Methodology 16
6. Research Questions 21
7. Results and discussion 22
   7.1. Internal structure of ditransitives across age groups 22
   7.2. Frequency of ditransitive construction use 25
   7.3. Verb selection (hierarchy and range) 31
8. Conclusion 40
9. Bibliography 42

1. Introduction

The studies of language acquisition and the mechanisms behind it still remain one of the most extensive ones within the field of linguistic research. Corpus related research has been present within the field of language acquisition from the very beginnings, but it is only due to the technological development that the research has become available to the wider range of scholars interested in the field. This thesis utilizes the data from the CHILDES corpus and serves as a small contribution to the ever-growing field of language acquisition research. It could be said that the thesis combines the language acquisition research and corpus linguistics with focus on the acquisition and syntax of ditransitive constructions in the English language.

After the introduction, section 2 outlines the two dominant paradigms within language acquisition research and shortly discusses the nature vs. nurture paradigm. The section offers a brief introduction into language acquisition research within the paradigms in order to ease the interpretation of the data. In section 3, the thesis turns to ditransitive constructions and their definition, since they remain in the focus of this study and the main object of analysis. The fourth section elaborates on the CHILDES corpora, the most famous corpus in language acquisition research and the source of data in this study. The fifth section deals with the methodological intricacies of the study, sampling details, elicitation and the selection of data. Research questions and hypotheses are noted in the sixth section. The main part of this thesis is found in section 7 under results and discussion. The section is further subdivided in three parts according to the data found and analyzed: internal structure of ditransitives across age groups, frequency of ditransitive construction use, the distribution of verbs across age groups (their hierarchy and range). In the final section 8, the main points and results found in this study are yet again summarized and shortly discussed.

2. Two paradigms within language acquisition research

The theory of language acquisition can roughly take on two different forms. The feud between the two dominant linguistic paradigms is also known as that of nature vs. nurture. One account argues for the innate linguistic competence, the idea that children are equipped with the core linguistic knowledge from the very birth. This account was made famous by Noam Chomsky in his reply to Skinner (1959), where the main reason for the rapidity by which children's acquisition of language occurs was seen in the innateness argument, a proposal that human race is exclusively equipped with a particular device, one to be later known as language acquisition device (LAD), that facilitates the process of acquisition. The second theoretical strain sought to explain the language acquisition by stressing the general cognitive capacities of children and the power of the input. What was once a philosophical dispute between empiricists Locke and Hume and nativist Kant, has quickly taken over the field of modern linguistics and most of its fundamental theoretical concerns. In linguistics, the two opposing theoretical strains have adopted the names of linguistic nativism, the most famous representatives of which are Chomsky (1959, Hauser et al. 2002) and Pinker (2002 & 2003), and cognitive linguistics, comprised of the names such as Lakoff (1987, Lakoff & Thompson 1975), Langacker (1987), Goldberg (1995) and Tomasello (2000 & 2003).

Naturally, the rapid acquisition of language that children demonstrate was not the only argument for nativism. The rapidity, of course, is a matter of perception, assessment of which can only be discussed in relative terms, since there is nothing to compare it to. The main underlying argument for nativism has always been the supposed existence of "linguistic universals", the backbone of the Universal Grammar theory (UG) (Cook and Newson 2007). Indeed, it does appear that the existence of such universals could only be explained by a genetically dictated shared linguistic competence. However, this claim for linguistic universals was never fully confirmed, and was even disputed to a certain degree (Evans and

Levinson 2009). The extreme and literal interpretations of nativism are always problematic as they seem to imply the complete linguistic competence of a toddler, which remains unperformed until certain age not because it is lacking, but because the toddler has yet to develop its motor skills and receive enough input to apply its hardwired grammatical knowledge onto it.

In a relatively modern era of philosophy, John Locke rehabilitated the term "tabula rasa" (1690) to infer that the human mind was completely subjectable to whatever input it might receive, without having any biologically hardwired rules on how this data should be processed. In linguistics, this has come to be known as cognitive linguistics, with Construction Grammar theory (CxG**)** emerging as the main opposition to nativist tradition. Unlike UG, CxG treats form-meaning correspondences as the basic language units. UG, instead, assumes the existence of fundamental linguistic structure and grammar shared between languages. Meanwhile, CxG recognizes the distinction between the "core grammar" and the constructions that fall out of this category, thus concerning itself with all linguistic phenomena. Because of this, the "non-core" cases are usually of main interest for the CxG, not just because they constitute a significant part of the corpus data, but because if the theoretical account of such non-canonical constructions succeeds to account for the canonical ones as well, it would ultimately prove as the complete and unifying theory of language (Goldberg 1995: 6).

The theoretical position in language acquisition related to the philosophical "tabula rasa" model and possibly the most prominent theory within CxG is a Usage-based model. The leading authority on the Usage-based model has arguably been Michael Tomasello. Perhaps the term "tabula rasa" is not necessarily a suitable marker of the approach, since the approach does emphasize the role of a child's general cognition in the process. The crucial cognitive skills for processing the input are pattern finding (grammatical role) and intention-reading

(functional role) (Tomasello 2008: 69). Nevertheless, it is one thing to argue these cognitive skills to be central for language acquisition and completely another thing to regard cognition as something separate from the supposed innate linguistic capacity, the claim of the nativists. With the dismissal of the innateness hypothesis, the most prominent role is attributed to the input (Tomasello 2003 & 2008). An important claim by Tomasello is the one which still distances his model from the "tabula rasa" approaches (which are just as irrational in their extreme forms as nativist fundamentalism), is the one where he claims intention-reading to be species specific. This theoretical assumption is important for the understanding of ditransitives, as they imply intentional action, which is discussed further in the continuation of the paper. However, this claim is also problematic since it cannot be put through a test of falsifiability. Intention-reading implies understanding the meaning of intention, even the capacity for the intentional action, which, after all, requires a greater cognitive level, the one unrelated to language comprehension and processing.

Linguists working within the usage-based linguistic framework are particularly interested in verb acquisition, the idea behind this being that the acquisition of certain word classes proves more challenging than that of others. The acquisition of nouns appears to be less challenging, given that they are conceptually less vague, especially in terms of early acquisition where the learned nouns represent objects from the adjacent surroundings (Tomasello and Brandt 2009).[1] This is only logical when one acknowledges the VP as the core of the sentence, determining the syntactic properties of the sentence. Since the paper deals with the acquisition of ditransitive constructions, children's understanding of the verbs' meaning becomes central to the interpretation of results. The evidence related to this is often discussed in terms of flexibility and productivity. Flexibility in verb usage is evidenced in the degree of variety of verbs used and the variety of the communicative context in which they

[1] Here, it is possible to distinguish between the sortal nouns (those denoting more tangible concrete phenomena such as toy, chair, pillow etc.) and the relational nouns which are more conceptually vague and probably represent a similar level of difficulty to verbs when it comes to their acquisition.

are found. Productivity, on the other hand, is the usage of certain verbs in a totally new, if not radically different way in terms of syntactic and morphological conventions. Otherwise, the flexibility of usage can always be credited to the power of input, i.e. reiteration of the utterances heard by the parents or others.

From a nativist perspective, the acquisition of verbs is often discussed in terms of syntactic bootstrapping, the idea where children are endowed with the knowledge of grammatical categories from their birth (Brown: 1957). The innate part in this account refers to the ability of children to recognize the syntactic categories of the words in question, such as nouns, adjectives etc. It is not clear, however, why is it necessary to presume that any innate categories are needed. In his further research on first language acquisition (1973), that is now a part of the CHILDES corpora, Brown takes a less nativist stance, assuming the children's capacity to recognize meanings is a product of children's sensorimotor capacities:

> "Linguists and psycholinguists when they discover facts that are at all general have, nowadays, a tendency to predict that they will prove to be universal and must, "therefore," be considered innate. The Stage I meanings have proved to have some generality in a sampling of child speech studies, and I do feel tempted to hypothesize universality. But not innateness. Not innateness because, though I have not worked out the relation in any detail, it is my impression that the first meanings are an extension of the kind of intelligence that Jean Piaget calls sensorimotor." (Brown 1973: 198)

However, he does highlight on several occasions that it is hard to separate grammatical relations from semantic roles, such as agent, object, location, also referred to as deep cases in Fillmore's transformational grammar (TG) (1968), precisely because they are often used interchangeably (Brown 1973: 201). In the end, the idea is that the capability for linking the

syntactic and semantic roles would provide the children with syntactic guidance in the acquisition of verbs. Again, it is not clear why it is plausible to argue that children can deduce meanings on a sensorimotor level, while at the same time not being able to deduce the grammatical relations using just the extralinguistic cues. Another problem is that the concept of grammatical relation is highly abstract, and it is questionable whether its "assumed" innateness can be put to a test. An alternative explanation for the recognition of these semantic categories, which correspond to now more elaborated thematic roles of the arguments, and possibly grammatical ones, is credited to Saffran et al. and their proposal for statistical learning as a way of language acquisition (1996). Although the "statistical learning" explanations were mostly used for accounting lexical acquisition, the same reasoning was attempted in the field of syntactical acquisition (Gomez & Gerken 2000). Distributional learning, a term used interchangeably with statistical learning, is a learning mechanism by which children can access all the instances in which they heard a particular word in its respective syntactic and semantic contexts, which ultimately allows them to infer their meaning. The same logic applies to syntax, where grammatical structures and categories are learned only after the children have stored the information from countless instances, deducing syntactical patterns from the analogous contexts in which the words that belong to the same grammatical category appeared.

Here, the distinction between distributional learning and syntactic bootstrapping becomes rather clear. Within the nativist approaches, the necessity to account for the acquisition of syntax is often found redundant, the implication being that there is no need to account for the acquisition of something that is already considered acquired from birth. On the other hand, for the alternative approach, it is not enough to account just for the acquisition of verbs, as it also needs to account for the acquisition of syntax, which already puts these alternatives in somewhat inferior starting positions in construction of theories. This research

deals with the acquisition of ditransitive constructions, rather than ditransitive verbs. Naturally, the concept of a "ditransitive verb" implies the company of two noun phrases (NPs) (the omission of one of the arguments is possible, but unlikely), and thus the ability to acquire the verb should imply the ability to acquire the construction. However, if one disregards the cases where one of the arguments was omitted from the construction because it could have been deduced from the context, one is still left with the cases where the child might have acquired the verb but was simply misusing it (e.g. *Mama give lunch* instead of *Mama, give me lunch*). In such cases, it is even justifiable to question the acquisition of the verb itself, since the question remains whether syntactically and semantically improper demonstration of a verb, can still be considered acquired. Nonetheless, the important thing to note is that only the fully demonstrated constructions will be targeted in this corpus research, rather than simply verbs on their own.

3. Ditransitives

Construction Grammar is a theory that rejects an absolute division of syntax and the lexicon. It is only generative in so much as it attempts to account for the infinity of conceivable grammatically unconstrained expressions. This, however, does not mean that the Constructional Grammar dismisses the existence of those constraints, whether morphological or syntactic, but rather that there are the cases where the boundary of semantics and syntax remains unclear. If anything, CxG is attempting to account for the potential infinity of both grammatically constrained and unconstrained constructions, without positing any latent syntactic or semantic forms (Goldberg 1995: 7). Ditransitive constructions are often subjected to analysis within CxG, as they seem to grant a pathway to understanding of this intricate syntax-semantics interface.

The argument structure of ditransitive constructions is defined in terms of the predicate being accompanied with the subject and two objects (direct and indirect one), with the thematic relations of agent, recipient and theme. The first argument of the construction is the agent argument (A), the second being the recipient (R) and the third being the theme (T) (Malchukov et al. 2010). In English, two forms most commonly express the three-argument structure of ditranstives:

(1) a) Agent →Recipient → Theme

"Jack gave John a ball."

b) Agent → Theme → Recipient

"Jack gave a ball to John."

The shift from a) to b) is referred to as "dative alternation" or "dative shift", where the verb takes the noun phrase (NP) and a (PP), instead of two NPs. It is also possible to use the terms "double-object dative" (VP → NP1 → NP2) for constructions comparable to the example a), or "prepositional dative" (VP → NP → PP) for those comparable to example b). The English

language belongs to the minority of languages where these dative alterations can be found. It has been argued in previous studies that there are approximately 6% of such languages with analogous alterations (Siewierska 1998: 179). The Croatian language also allows alterations in ditransitive constructions with the change of objects' cases. However, the dative alternation in Croatian does not call for the need of preposition "to", and its double object constructions in the same case do not allow for such alterations (Silić & Pranjković 2005). To illustrate, Zovko Dinković (2007: 65) gives an example for Croatian "alternated constructions":

(2) a) Lena je poslužila gostima čaj i kekse (guest – DAT) (tea and biscuits – ACC)

Lena served tea and biscuits to the guests.

b) Lena je poslužila goste čajem i keksima. (guest – ACC) (tea and biscuits – INSTR)

Lena served the guests tea and biscuits.

Although the verbs allowing this type of alteration vary across languages, they generally involve a transfer characteristic for ditransitive constructions, relocation of theme from agent towards the recipient. The existence of the prepositional phrase in English changes the argument structure in the sense of its "core arguments", It is often argued that the English dative alternation has only two "core arguments" ("Lena" and "tea and biscuits" in 1b), whereby the instances in Croatian retain the three "core arguments" in both examples (Zovko Dinković 2007: 66). For the purposes of this research, the focus is solely on the acquisition of ditransitive constructions with three "core arguments".

Another possible occurrence when it comes to ditransitive constructions is the omission of an object, most typically the indirect one. There are not many different verbs which tolerate such omission in English, the examples being the verbs such as *show*, *tell*, *pay*.

There are two main reasons often stated to account for the lack of an indirect objects in such constructions, and these include the irrelevance to designate the object (theme) or the possibility that it can be salvaged from the context (Carnerero 2007). In these cases, we label those expressions as "dimonotransitives". Similarly to other derived ditransitives, these constructions will also remain excluded from the analysis.

Goldberg (1995: 141) uses an example of a verb *bake* being used ditransitively: *Sally baked her sister a cake*. Her argument is that the construction necessarily implies intention, which is a semantic construct. To bake something for someone means to do it not only intentionally, but with a very specific intention – intention of giving it away to that someone. The implication is not only the one of giving, but also the one of receiving. Pinker (2013: 56-57) also recognizes the common semantic property of dativizable verbs. The prospect of a particular object occupying the second position within the construction being possessed by the object occupying the first position is characteristic for ditransitives. He distinguishes the semantic property of the first object being the possessor, rather than goal to which the transferred entity is *given* or *sent* to, even when this possession is not literal as in the expressions, *He told her the story*, *He asked her a question*, and *She showed him the answer*. Certain authors even argue that there is a semantic difference between the alternating ditransitive constructions. For instance, *teaching French to the students* carries a different set of connotation than *teaching the students French*, with the former expression implying that the students did not necessarily master French, and the latter entailing that the process was successful (Green, 1974). Similarly, *throwing John a ball* implies that John is probably able to catch it, while *throwing a ball to John* only invites the inference of John being the target, without the requisite skillset to receive it (being dead or asleep) (Pinker 2013: 97-98). In a double-object ditransitive construction the recipient is bound to exist, while its presence is much less overt in the prepositional dative.

The thematic relations that we find in ditransitive constructions are nothing but very specific semantic constraints. Goldberg uses the terms "volitional agent and a willing recipient" to describe the semantic structure of a ditransitive (1995: 141). Ditransitives are the focus of this research precisely because they, like many other, embody the disputes not only between the two approaches to grammar, but between the two opposing theoretical paradigms in language acquisition theory. For the Chomskyan framework, which assumes the strict separation of syntax and lexicon, every idiomatic expression proves to be problematic, especially if these constructions prove to be productive or semi-productive. To illustrate, Boas gives the example of an expression "*Joe laughed his head off?*", where laugh, normally designated as intransitive verb, behaves in a transitive manner (Boas 2008: 116-117). The question is whether the generative framework can provide the general syntactic rules that derive such expressions and how to restrict the application of these rules on these specific idiosyncrasies. Many of the ditransitives discussed in the following chapters are considered to be idiosyncratic within traditional grammar, and yet they prove to be extremely productive forms with the capacity to incorporate an infinite range of arguments.

The immediate constituents of a ditransitive construction seem to be the VP→NP→NP (excluding the dative alternation). Naturally, such definition is inadequate for the proper description of a construction. For instance, in cases with verbs such as "call" and "name", as in *Jack called/named his boy John*, the construction does not merely share the same constituents, but almost the same underlying logic. Ditransitive structures are defined in terms of having an agent, a recipient and a theme-like argument. One could possibly argue that the semantic roles of the expression above are fairly similar to those which one might find in a ditransitive construction, and yet the expression belongs to the realm of complex transitive constructions. This becomes especially important in understanding the difference between benefactives and ditransitives, where the boundary is sometimes blurred, or when the

overlap is almost complete (Kittilä 2005). How do we decide whether "*She brought me coffee*" translates to "*She brought a coffee for me*" or "*She brought coffee to me*" (Malchukov et al. 2010: 3). This is why the discussion on difference between the complex transitive "*Jack called John a fool*" and a benefactive "*Jack called John a taxi*" are suited for this paper. Theoretically, it is possible to interpret the former expression as a benefactive as well, i.e. "*Jack called a fool for John*". Of course, Jack's intention was certainly not the one of looking around for a fool in order to present it to John, but rather the one of regarding John himself a fool. This is the same reason why it is difficult to separate the thematic relations of a "beneficiary" and a "recipient". Even in such scenario, the semantic structure invites us to consider the label "fool" as theme and John as a possible recipient of it. The same applies to the thematic relations of "John" and "taxi", perhaps even more so. For instance, if we consider the expression "*Jack named his boy John*", its dative alternation proves nonsensical ("*Jack named John to his boy*"). However, if one acknowledges the fact that the definition of the verb "to name" is literally "to give a name", the semantic likeness of a ditransitive construction "*Jack gave the name John to his boy*" invites one to deem a complex transitive as a ditransitive that has undergone ellipsis and a word-formation process of conversion.

Nonetheless, the argument of this paper is certainly not the one of labelling such constructions as ditransitive ones or challenging the already solid theoretical foundations, which the labelling of such constructions depend on. Here, the consideration is introduced solely to highlight the entanglements of methodological and theoretical concerns the researches face. In this particular corpus analysis, the focus is exclusively on ditransitive constructions. Needles to say, benefactives and other types of constructions, such as causatives or applicatives, sometimes dubbed derived ditranstives, were also eliminated from the analysis.

4. CHILDES corpora

Learner corpora research have recently become an invaluable asset in language acquisition studies. There are a lot of morphological and syntactic data that can be exploited in this way and used to make general inferences about language acquisition. The research of this kind was utilized from the very beginnings in the studies of language acquisition in general, including both second and first language acquisition. Corpora studies that use this type of data as a source for making inferences about first language acquisition, fall under the domain of what is called *Cognitive Corpus Linguistics*, especially in case of using the Construction Grammar framework (Arppe 2010: 1). To think about language acquisition necessarily implies thinking about cognition, how language is stored, how it is processed, and this is precisely where the cognitively oriented Construction Grammar appears fit for this type of data.

This study exploits the data provided by the CHILDES corpus, corpus that was initially founded with the idea of resolving the mysteries surrounding the first language acquisition. The initial idea of the creators, Brian MacWhinney and Catherine Snow, was the one of gathering data in naturally occurring situations by capturing the spontaneous speech of both children and adults. The recorded speech is mostly that between children and their interlocutors, extracted from the recordings of their spontaneous interactions. Most of these conversations are between parents and their children, but there are also recorded situations of child-child interactions and kindergarten classroom interactions. With time, the corpora grew bigger with the help of more than 100 researchers, ultimately resulting in great data diversity, including different language impairments, aphasics and even some second language learner data (MacWhinney 2015: 10-11). Although the majority of data is elicited from monolingual spontaneous speech, the CHILDES corpora also merges the so-called narrative corpora - e.g. including a “frog story” narrative, where a picture book was used to obtain speech from

younger children (Gillis 2014). Naturally, this implies that some of the semantic patterns will repeat themselves, perhaps compromising the syntactic ones as well. What is sometimes seen as its main strength, remains the main fault of this particular research, since the entire corpus was used, without the cherry-picking of data and excluding certain sub-corpora. However, in quantitative terms, this excess of data falls under the scope of standard error, especially if dispersed across different age groups. Another thing to consider is that the inclusion of a small portion of linguistically impaired people actually reflects the everyday language better, since language impairment is a normal societal occurrence. The converging body of data doesn't automatically imply that the same cognitive structures governing language processing will be reflected in the data, or that the linguistic data will be identical. The use of converging evidence is justified by the fact that the recurring results across different methods and data types prove to be the best indicator of their gravity. In fact, such studies may even yield results superior to those studies which deal with only one type of data sets (Arppe 2010: 4-5). Nonetheless, the results presented in this research should only be considered for research purposes, and any serious or conclusive generalizations about first language acquisition would require more careful sampling and control over data inclusion.

Given that the data for all age groups is extracted from the same corpus, where the recorded instances include the conversations between parents and their children, it is possible to reflect on the "poverty of the stimulus" argument by comparing the speech of adults and children. However, there are severe methodological liabilities with this endeavor, and any reference that is made to the "poverty of the stimulus argument" needs to be taken into consideration for further critical evaluation. The idea behind this particular argument is that the children display linguistic abilities that are not in correspondence with what they could have extracted from the input. For any conclusive evidence, the researcher would have to have separate data on the children and the entire input that they receive, not only from parents

but from others as well, which is still inconceivable because of both ethical and practical reasons. Again, this type of study is to be seen as a study of “big data” with its restrictions in relation to controlling of the conditions. The poverty of the stimulus argument is mentioned on a side note, where any similarity between the speech of children and adults is brought up for the sake of provoking further language acquisition research.

5. Methodology

This study relies on *Sketch Engine* as one of the leading corpus tools in linguistic research. The term applies to both software and the web service, the difference being that the web service incorporates "pre-loaded" corpora, the core software, and the necessary tools for creating one's own corpora (Kilgariff et al 2014: 8). The basic tool for operating the corpora, and the main tool in carrying out this research, was the concordancer, which enables one to access the raw data. As already stated, the studied evidence is from the CHILDES corpora, which contains the data gathered from more than a 100 studies. The gathering of data was carried out during April 2018.

The first step in the research was to make sure that all of the ditransitive constructions get targeted by concordancer, besides the prepositional dative, which was not included in the analysis. Necessary for this task was the use of part-of-speech tagging (POS tagging) annotation and the query program, capable of identifying word classes within the text of the corpora. Once the list of possible class variation was established for the double-object ditransitive constructions, the same queries would be made for different age groups. For simplicity sake, one could possibly group the constructions in the following way:

1. Ditransitive constructions without any determiners

a) Verb → Noun →Noun (e.g. *X told James something*)

Sketch Engine query: [tag="V.*"][tag="N.*"][tag="N.*"]

(note that *something*, although pronoun, is targeted as a noun by POS tool)

b) Verb → Pronoun → Noun (e.g. *X told him something*)

Sketch Engine query: [tag="V.*"][tag="PP"][tag="N.*"]

c) Verb → Noun → Pronoun (e.g. *X gave James it*)

Sketch Engine query: [tag="V.*"][tag="N.*"][tag="PP"]

d) Verb → Pronoun → Pronoun (e.g. *X gave him it*)

Sketch Engine query: [tag="V.*"][tag="PP"][tag="PP"]

2. Ditransitive constructions with one determiner/adjective in the second object postion

a) Verb → Noun → Determiner (e.g. *X gave James the Y*)

Sketch Engine query: [tag="V.*"][tag="N.*"][tag="DT"]

b) Verb → Pronoun → Determiner (e.g. *X gave him the Y*)

Sketch Engine query: [tag="V.*"][tag="PP"][tag="DT"]

c) Verb → Noun → Adjective (e.g *X gave James good Y*)

Sketch Engine query: [tag="V.*"][tag="N.*"][tag="J.*"]

d) Verb → Pronoun → Adjective (e.g. *X gave him good Y*)

Sketch Engine query: [tag="V.*"][tag="PP"][tag="J.*"]

3. Ditransitive constructions with one determiner/adjective in the first object position

a) Verb → Determiner → Noun → Noun (e.g. *X gave the man something*)

Sketch Engine query: [tag="V.*"][tag="DT"][tag="N.*"][tag="N.*"]

b) Verb → Determiner → Noun → Pronoun (e.g. *X gave the man it*)

Sketch Engine query: [tag="V.*"][tag="J.*"][tag="N.*"][tag="PP"]

c) Verb → Determiner → Adjective → Noun → Noun (e.g. *X gave the white man something*)

Sketch Engine query: [tag="V.*"][tag="DT"][tag="J.*"][tag="N.*"][tag="N.*"]

d) Verb → Determiner → Adjective → Noun → Pronoun (e.g. *X gave the white man it*)

Sketch Engine query: [tag="V.*"][tag="DT"][tag="J.*"][tag="N.*"][tag="PP"]

4. Ditransitive constructions with determiners/adjectives in both object positions

a) Verb → Determiner → Noun→ Determiner (e.g. *X gave the man the Y*)

Sketch Engine query: [tag="V.*"][tag="DT"][tag="N.*"][tag="DT"]

b) Verb → Determiner → Noun→ Adjective (e.g. *X gave the man good Y*)

Sketch Engine query: [tag="V.*"][tag="DT"][tag="N.*"][tag="J.*"]

c) Verb → Determiner → Adjective→ Noun→ Determiner (e.g. *X gave the white man the Y*)

Sketch Engine query: [tag="V.*"][tag="DT"][tag="J.*"][tag="N.*"][tag="DT"]

d) Verb → Determiner → Adjective → Noun → Adjective (e.g. *X gave the white man good Y*)

Sketch Engine query: [tag="V.*"][tag="DT"][tag="J.*"][tag="N.*"][tag="J.*"]

Since there are various other constructions targeted by these queries, the final list of ditransitive constructions needed to be stripped from these cases. The elimination of redundant data (mistakenly targeted constructions) was carried out within the spreadsheet to which all of the queries above were exported from Sketch Engine.

The inclusion/exclusion of certain cases was problematic from semantic reasons. For instance, let us consider one example from CHILDES Manchester corpora (Theakston et al. 2001): "…me not building a house . panda get upset . me not **building panda house** . but you are building panda's house . What ? are…". This particular example was targeted by the query [tag="V.*"][tag="N.*"][tag="N.*"] for the 0-3 age group. Naturally, the question is whether the child wants to express *I am building a panda's house* or *I am building panda a house* (in other words, a house for panda). In this case, the construction is a benefactive one, but the point remains the same. The assumption of the interlocutor, as one can see from the expression following the targeted sequence is that the child's intended words were panda's house. This is an understandable presumption since the expression *I am building a panda's house* does sound more natural than the other variation. Even if it does not, there is a greater likelihood of the child omitting an inflectional affix than expressing the ditransitive construction in this case. Another possibility is that of a child forming a compound on purpose. Nevertheless, the point is that even the context does not provide enough evidence for researcher to reach a conclusive verdict on the character of this construction. In such cases, one is forced to make an "educated guess" about whether or not to include the expressed

construction in their analysis. In the end, if one does decide to include the construction, then one needs to do the same with other similar dubious cases across all language groups in order to "balance the scales".

The CHILDES corpora on Sketch Engine differentiates 6 age groups: 0 to 3 year olds (0-3), 4 to 6 year olds (4-6), 7 to 9 year olds (7-9), 10 to 12 year olds (10-12), 13 to 17 year olds (13-17) and 18 or older (18+). These language groups are maybe not ideal for the precise tracking of language development because most of the acquisition whether phonological, lexical, morphological or syntactic is already nearing its end by the age of 3 (Bloom 1998). However, although certain structures might have been acquired, the flexibility and the productivity levels might be different. Although the initial idea of the research was to do a cross examination between all of the age groups stated, the problem was the misbalance between the sizes of these subcorpora. More accurately, it is not so much the misbalance as it is the lack of input in the corpora containing the language of children and adolescents from 6 to 18 years of age. The table 1 below indicates the sizes of these subcorpora at the time when the data was collected:

Table 1. Subcorpora sizes

Age group	The number of tokens in the subcorpus	The number of words in the subcorpus
0-3	4 580 207	3 525 724
4-6	1 874 520	1 442 956
7-9	326 970	251 693
10-12	109 008	83 911
13-17	68 774	52 940
18+	20 571 114	15 835 110

It is obvious that the age groups 7-9, 10-12 and 13-17 are underrepresented in the CHILDES corpora and cannot be used for comparative purposes, or even research purposes in this case. Because of the sizes of these corpora, certain patterns of ditransitive constructions were not located within these age groups, regardless of being noted in the younger stages of linguistic development. This was clearly a consequence of the diminished corpora sizes, not allowing this research to pursue any comparative work. While these age groups were only included in the general overview of the evidence collected, the data of 0-3, 4-6 and 18+ age groups was used for comparative purposes. Needless to say, more tokens in the 4-6 age group would prove advantageous for this study, but the data is still reasonably large for research purposes.

6. Research Questions

1. Is there a difference in the character of ditransitive constructions used across different age groups with regards to the type and number of constituents?

2. Is there a difference in the frequency of ditransitive usage across different age groups?

3. Is there a difference in verb selection between different age groups when it comes to usage of ditransitive constructions?

Hypotheses

1. More elaborate ditransitive constructions (constituent-wise) will occupy significantly larger portions of subcorpora with the progression of age.

2. The number of ditransitive constructions used will be grow with the progression of age.

3. a) The hierarchy of usage with regard to verb selection will not differ across age groups as the early usage of ditransitives should reflect the input adult language provides.

3. b) The range of lexical verbs expressed by different age groups will differ since the development of lexicon is a function of time more than linguistic ability itself.

7. Results and discussion

7.1. Internal structure of ditransitives across age groups

The considerations about the complexity of ditransitive constructions are stated in the first research question. Naturally, there is not a lot of variation in the sense of an underlying syntactic pattern of a ditransitive construction (VP-complex NP-complex NP). The following table indicates the number of times certain constructions appeared in the specific age group.

Table 2. Distribution of constructions within age groups

Constructions	0-3	Percentage	4-6	Percentage	18+	Percentage
V-PRO-comNP	1634	64.95%	1156	60.62%	20960	69.94%
V-PRO-NP	414	16.45%	371	19.45%	2950	9.84%
V-PRO-PRO	176	7.00%	188	9.86%	1848	6.17%
V-NP-comNP	117	4.65%	47	2.47%	2398	8.54%
V-NP-NP	101	4.01%	50	2.62%	537	1.79%
V-comNP-comNP	38	1.51%	54	2.83%	731	2.44%
V-NP-PRO	27	1.07%	26	1.36%	336	1.12%
V-comNP-NP	9	0.36%	10	0.52%	174	0.58%
V-comNP-PRO	0	0.00%	5	0.26%	34	0.11%
Total count	2516	100.00%	1907	100.00%	29968	100.00%

Legend: V-verb, NP-bare noun, PRO-pronoun, comNP- complex noun phrase

The categories in Table 2 represent the merged data of constructions targeted by this research (pg. 13-15). Basically, the arguments, when summarized, are either represented by a standalone pronoun, a standalone noun (marked as NP) or a noun-phrase (marked as complex

NP), which covers a wide range of situations, typically involving adjectives, determiners etc. For example, V – NP – comNP incorporates search queries such as Verb → Noun → Determiner (e.g. *X gave James the Y*) and Verb → Noun → Adjective (e.g *X gave James good Y*). Sentence examples can be found in Table 3.

Table 3. Sentence examples

Constructions	Sentence examples
V-PRO-comNP	X gave him a ball/ X gave him (a) beautiful ball etc.
V-PRO-NP	X gave him cappuccino
V-PRO-PRO	X gave him it
V-NP-comNP	X gave James a ball/beautiful ball/big beautiful ball etc.
V-NP-NP	X gave James cappuccino
V-comNP-comNP	X gave the customer a big bowl of soup
V-NP-PRO	X gave James it
V-comNP-NP	X gave the customer cappuccino
V-comNP-PRO	X gave the customer it

The question, of course, is whether this classification can prove as a good measure for the estimation of construction complexity. Even when further dissected, it remains unclear whether the sentence such as *X gave James the ball* is any less complex than the sentence *X gave James a good beating*. The number of constituents does not necessarily imply a greater complexity in terms of grammar. While it might potentially reveal differences in lexicon, any judgment on their complexity should invite serious reservations. However, the omission of determiners might prove as a good indicator of "linguistic immaturity" and possibly a good pointer for determining structural complexity of the constructions in question.

From Table 2 we can easily see that there are some discrepancies in the portion of particular constructions found across different age groups (Pearson Chi-square = 604.174, $p<0,05$). The percentage indicates the fraction a particular ditransitive construction takes out of all ditransitive constructions found within one age group. Although the difference is statistically significant, the similarity between the trends is still rather conspicuous. The greatest disparity across age groups can be found under the constructions V-PRO-NP, V-PRO-comNP and V-NP-comNP. However, it is clear that the patterns V-PRO-NP and V-PRO-comNP still take up most of the ditransitive constructions under each age group. The common denominator here is the pronoun following the VP immediately. It is not surprising that the expressions like *tell him something* or *tell him a story* will be used more frequently than *tell Jack a story*. What is surprising is that these constructions are 15-20 times more frequent than the ones with the noun following the VP. One would expect that the use of pronouns would be more frequent in the younger age groups (for instance, the self-centered use of *me*) and although there is a statistical difference, the data indicates a very similar pattern in all three subcorpora.

The difference in the frequency between the ditransitives that lack determiners, especially V-PRO-NP and V-NP-NP, across different age groups clearly indicates that the dropping of determiners happens more frequently at younger ages. Probably, a part of the explanation for this rests on the fact that children are more prone on making grammatical mistakes. Research (Soja 1994) shows that the crucial age for mastering the determiners, which relates to distinguishing normal count nouns from NP-type nouns (her label for singular count nouns that do not require determiner such as *church*), is the age between 2 and 3;6 years. It is only after that period that the determiner omission starts declining rapidly. This somewhat corresponds to the results of this research, although the period does extend to 6 years old, but this might be an effect of the subcorpora (the age distinction is not as precise).

7.2. Frequency of ditransitive construction use

Syntactically and semantically complex ditransitive constructions should occupy significantly larger portions of subcorpora with the progression of age. The assumption is that the ditransitive constructions will be demonstrated more frequently as the age progresses. This might seem self-evident with the comparisons between 0-3 age and 18+ age group, especially if one considers that the crucial age for the acquisition of ditransitive constructions and the flexible use of more than one verb seems to be around three (Gropen et al. 1989 & Tomasello 2001). Nevertheless, it might be interesting to see the differences between the 4-6 age groups and adults. In the continuation, we look at the frequencies of particular constructions and the total frequencies observed. The total number of ditransitives can be found in the final row of Table 2. Once again, the size of the subcorpora is significantly different, which is why the number of observed frequencies is clearly different. This is why it is necessary to calculate the ratio to find out the true portion of the subcorpora occupied by the ditransitive constructions.

One way to do it is to divide the number of frequencies with the number of verbs in the respective subcorpus. The other way to do it is to divide the number of observed frequencies with the number of words in the same subcorpus. However, none of these options proves ideal for the analysis. Ideally, the number of ditransitive constructions should be divided by the number of constructions in general (in the corresponding corpus), but given that this information is not available, one must opt for one of the stated options. The option with verbs seems to be more plausible, given the assumption that the proper syntactic constructions in English are bound to contain a verb phrase. The problem is that VPs quite often imply more than one verb with all its auxiliaries, meaning that the number of constructions does not really correspond with the number of verbs targeted by the corpus tools. Hence, the following calculation is to be considered with certain reservations about it. In the following table, the observed frequency of ditransitive uses is divided by the number of

verbs found in the corresponding corpus and multiplied by a million, indicating the hypothetical usage of ditransitives per million verbs.

Table 4. Usage of ditransitives per million verbs

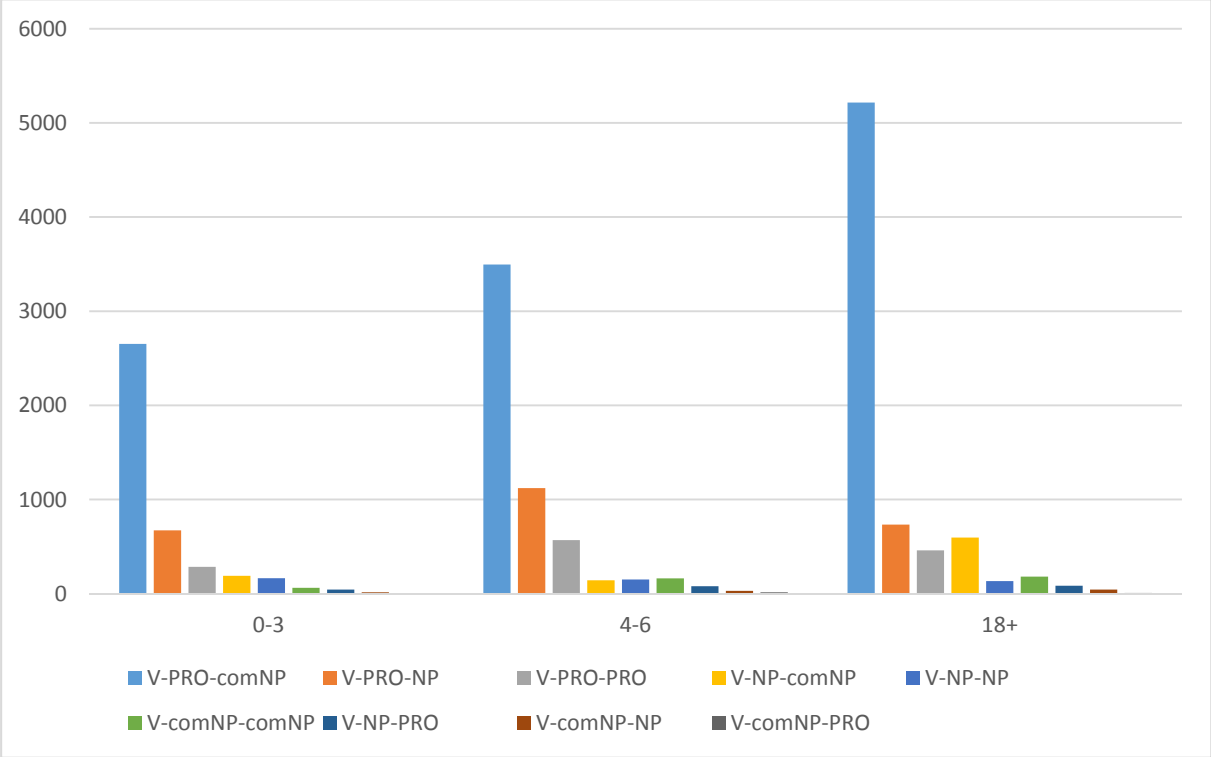
Constructions	Normalized frequency		
V-PRO-comNP	2653	3496	5217
V-PRO-NP	672	1122	734
V-PRO-PRO	286	569	460
V-NP-comNP	190	142	597
V-NP-NP	164	151	134
V-comNP-comNP	62	163	182
V-NP-PRO	44	79	84
V-comNP-NP	15	30	43
V-comNP-PRO	0	15	8
Total count	4086	5767	7459

Legend: V-verb, NP-bare noun, PRO-pronoun, comNP- complex noun phrase

Another problem with this ratio lies in the fact that different age groups might be using different tense and aspect constructions, which implies different span of auxiliaries across the expressed VPs. For instance, present perfect continuous or past perfect continuous contain at least one auxiliary more on average when compared to other tenses, and yet it is unlikely that the frequency of using these tenses is the same across different age groups. Besides, these results would carry even more weight with the implication that the number of auxiliary use rises with the age (meaning that the usage of ditransitives – gap would be even larger with the

data on the number of sub-constructions[2] in the corresponding subcorpora). When illustrated (Figure 1), the data clearly shows diverging points in the frequency of usage between different age groups.

Figure 1. Usage of ditransitives per million verbs



Legend: V-verb, NP-bare noun, PRO-pronoun, comNP- complex noun phrase

Before progressing with the ratios of particular sub-constructions, it is worth noting that the chi-square for the total count of ditransitive sub-constructions (once put into comparable ratios) showed significant difference between age groups (Chi-square= 985,776, df=2, $p<0,01$). The results of the analysis indicate a clear difference in the portion of corpora occupied by the ditransitive constructions. The progression is rather linear and obvious, with the difference of almost 3400 uses of ditransitives (per million verbs) between the children

[2] The term sub-construction is used in the thesis for different realizations of ditransitive constructions, which ultimately differ depending on their constituent structure.

younger than 3 and adults. In this respect, the 4-6 age group resembles the adult speech more than the 0-3 age group, which is to be expected. Again, the differences in the sub-constructions such as V-PRO-NP do not follow the pattern completely, but they may easily be a consequence of the subcorpora sizes (the size of the 4-6 subcorpus is relatively good, but not ideal).

The study of ditransitives in four-year-old native speakers of Kannada suggests that the children at this stage already possess adult-like representations of ditransitive constructions (Viau & Lidz 2011: 682). They consider this age to be the youngest point at which the children might start to understand the contexts in which the ditranstives are used. According to the data, their knowledge is independent of particular verbs found in these structures, while being only partly conditioned by the surface word order. While the data on the Kannada-speaking four-year-olds shows the children's demonstration of ditransitives to be adult-like, it remains unclear how they accomplish this (Viau & Lidz 2011: 701-702). The question is how children manage to map particular strings of words accurately to the suitable constructions? It needs to be highlighted that these strings of words are not always and necessarily mapped appropriately by children. If anything, in the subcorpus of early speech we often find semantically and syntactically dubious examples such as the one from CHILDES Manchester corpora (Theakston et al. 2001):

(3) "me like the fork out . there . yeah ! **make tea a baby** please . **make tea the baby** too . **make tea the baby** . **make tea the baby** . **make tea a big baby** . I don't know what you're saying . **make the tea a baby** . make the little baby ? **make tea a little baby** . make tea for the little baby ."

This example clearly shows a child demanding a tea, probably from his parents. Although the correct word order should be *make the baby tea*, the child fails to accurately map the utterance to the ditranstive construction. Therefore, one mustn't fall into the trap of interpreting the early demonstrations of ditranstives as a consequence of innate knowledge of these constructions. Children's control of adult-like representations is hardly immediate, and the incremental mastering of these constructions is to be seen as a "practice makes perfect" process.

There are other peculiarities that correspond to the results of Table 2. For instance, one might look at the sub-construction found in Table 4, V-NP-NP, one which was particularly difficult for filtering out the proper ditransitive constructions from those which only imitate them. The data in Table 4 does not indicate a great difference when it comes to that particular sub-construction across age groups although in general, ditransitive constructions occupy a significantly larger portion of the adult subcorpus. As stated previously, the explanation for this might be found in the omission of determiners characteristic for younger ages. While Table 2 indicated the percentage out of all the ditransitives used, in Table 4 it becomes clear that this specific ditransitive sub-construction appears as frequently within the speech of adults and the two younger generations. Considering that the ditranistives are generally much more used among adults, this indicates that the similarity in this category must surely be attributed to the omission of determiners as presumed. Let us consider an example from the famous Brown corpus (1973) incorporated into CHILDES and the utterance targeted by the query [tag="V.*"][tag="N.*"][tag="N.*"]:

(4) "…Perro . where Perro go ? . did you say you ? say you . **giving Perro ride**. you're giving Perro a ride ? does he like it ?"

The example (4) was produced by Adam, a child almost as equally famous in the field of language acquisition as the scholar who studied his language. There are many other examples like this in the 0-3 subcorpus, all of them sharing the same characteristic of determiner omission. On the other hand, in adult speech the frequency of ditransitive constructions lacking determiners appears lower when compared to other expressions targeted by the query, which contain either mass nouns or pronouns such as *something*, *anything* (which are targeted as nouns by the POS tools).

After evening out the ratios across the three age groups, the answers to the first and second research question seem apparent and somewhat expected. The results clearly indicate a significant rise in the use of ditransitive constructions with the progression of age, especially those containing determiners and adjectives. Considering that extreme nativism grants full syntactic competence to young children, these data proves more fitting for the cognitive approaches to language acquisition. It appears that children have indeed mastered certain uses of ditransitive constructions with certain verbs. However, their acquisition mirrors the input when it comes to basic constructions, indicating that the acquisition is probably item-based at this point. The context in which the verb appeared in the input is probably remembered in terms of the three arguments surrounding the verb, while the rest (determiners, adjectives) are acquired subsequently. While the nativist claims would not necessarily reject the possibility of ditransitive constructions being used less frequently in younger ages, they would still imply that once the ditransitives were indeed acquired, the complexity of these constructions would vary just like in the adult speech. This is not because it would mirror the input but because it would mirror their own capacity for producing the syntactically complex constructions with all its variations. For further confirmation of this reasoning, the paper turns to the prototypical verbs used in these constructions.

7.3. Verb selection (hierarchy and range)

The following part of the paper deals with the frequency of particular verbs across different age groups. The data on verb frequency might reveal whether the children truly possess the abstract knowledge about these constructions or whether this acquisition is item-based and merely reflects the input.

Table 5. Distribution of ditransitive verbs across age groups

Age	0-3	Proportion	4-6	Proportion	18+	Proportion
give	1073	42.65%	798	41.85%	11162	37.25%
tell	233	9.26%	367	19.24%	6295	21.01%
get	374	14.86%	207	10.85%	2654	8.86%
show	157	6.24%	124	6.50%	1885	6.29%
make	186	7.39%	71	3.72%	1405	4.69%
buy	91	3.62%	92	4.82%	950	3.17%
bring	75	2.98%	40	2.10%	867	2.89%
read	68	2.70%	37	1.94%	600	2.00%
find	22	0.87%	5	0.26%	407	1.36%
ask	5	0.20%	20	1.05%	385	1.28%
pour	12	0.48%	6	0.31%	376	1.25%
do	8	0.32%	8	0.42%	270	0.90%
draw	28	1.11%	6	0.31%	245	0.82%
pass	2	0.08%	12	0.63%	220	0.73%
send	5	0.20%	6	0.31%	210	0.70%
teach	11	0.44%	8	0.42%	175	0.58%

throw	6	0.24%	3	0.16%	159	0.53%
hand	10	0.40%	11	0.58%	148	0.49%
build	14	0.56%	4	0.21%	148	0.49%
sing	3	0.12%	5	0.26%	139	0.46%
feed	28	1.11%	14	0.73%	137	0.46%
write	24	0.95%	1	0.05%	81	0.27%
bake	7	0.28%	0	0.00%	80	0.27%
offer	0	0.00%	0	0.00%	78	0.26%
leave	7	0.28%	1	0.05%	76	0.25%
wish	2	0.08%	1	0.05%	70	0.23%
fix	6	0.24%	3	0.16%	67	0.22%
put	3	0.12%	1	0.05%	47	0.16%
cook	8	0.32%	0	0.00%	45	0.15%
sell	0	0.00%	2	0.10%	39	0.13%
promise	1	0.04%	1	0.05%	38	0.13%
cut	3	0.12%	0	0.00%	37	0.12%
save	0	0.00%	0	0.00%	36	0.12%
serve	1	0.04%	1	0.05%	33	0.11%
TOTAL	2473	98.29%	1855	97.27%	29564	98.65%

Table 5 indicates the number of instances in which these particular verbs were found across different age groups. The frequencies are ordered in a descending order beginning from the most frequent ones according to the 18+ category. These are not all of the verbs found in ditransitive constructions although they constitute around 98% of each corpus as visible in the bottom of Table 5. Numerous other verbs have also appeared in ditransitive constructions,

with greater variety in adult language, which may not appear so considering the total percentages occupied by the verbs included in Table 5, but it needs to be noted that the size of the 18+ subcorpus is significantly larger than the other two. Other verbs include *lend, blow, play, pay, owe, cost, pick, peel, deliver, keep, paint, fetch, wipe, color, want, catch, brush, charge, owe, trade* etc. Verbs like *blow, owe, cost, pick, peel, keep, paint, fetch, color, catch, owe, pack, trade, charge, post, mail, pop, comb, wash, answer, order, grate* and many others were found almost exclusively under the 18+ subcorpus in a ditransitive format, with the exception of few which were located in 4-6 age group (*owe, cost, color, pack*).

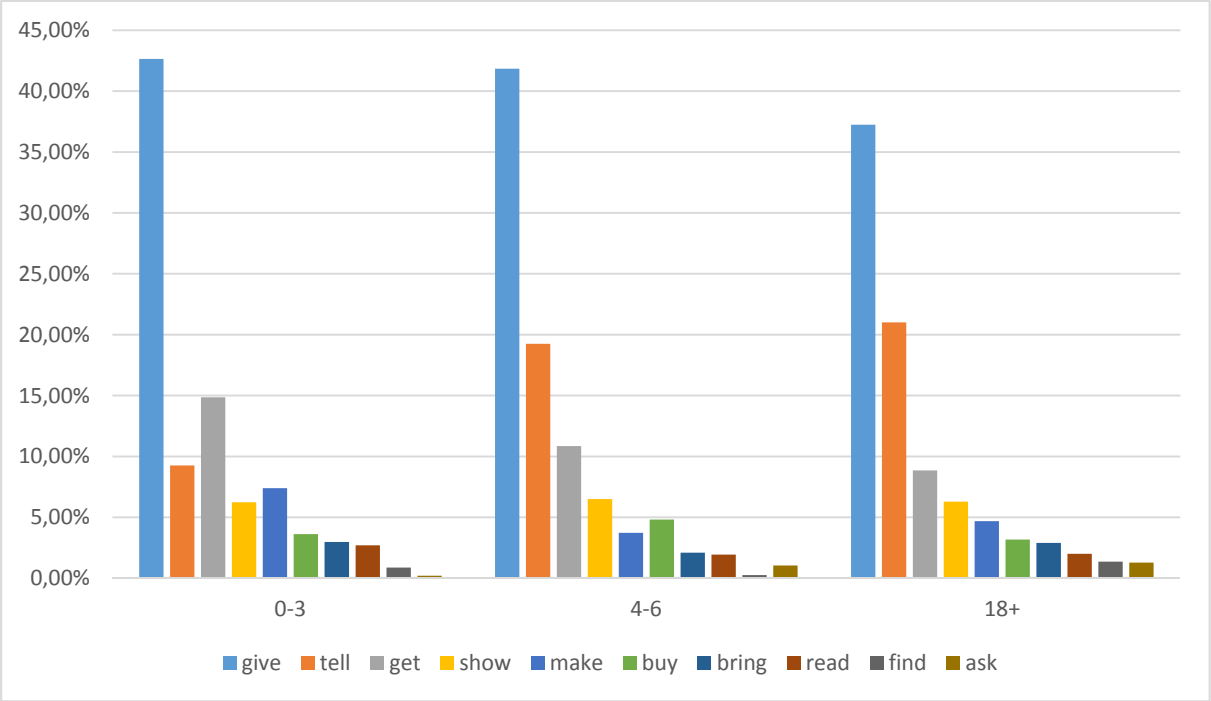
Some of the ditransitive constructions found in the speech of children aged below three seem suspicious, considering the observed frequency and the relatively challenging semantics of the verb in question. For instance, verbs *promise* and *serve* can be found in a ditransitive construction on one occasion in the age groups 0-3 and 4-6. It is unclear whether these constructions were truly produced by children, or whether this was a mistake made in the process of annotating the corpora ("I told him not to go to work . give me my lolly . you ***promised me a lolly*** when my come back from Sunday school .") (Cruttenden 1978). Here, it is a relatively safe assumption that the person uttering the sentence is indeed a child, given that the requested object is a lollypop. However, in other cases, the issue might not be so clear, but it would be equally hazardous for the research to throw away these cases, even if isolated, without the unambiguous proof in favor of the alternative reasoning.

As in Tables 2 and 4, the verbs found in the speech of young children, and the ratios in which they are used, appear strikingly similar to the ones found in adult speech. There is some discordance between the ratios. In all three age groups, the verb *give* easily occupies the first place when it comes to frequency of usage. However, while *tell* is the verb following in 4-6 and 18+ age groups, the second position in the 0-3 subcorpus is occupied by *get* rather than tell. If one were to make an educated guess about why this might be the case, one might

attribute this to the nature of child-parent relationship, whereby children are always in need of assistance. This explanation would also account for the verb *make* and the greater portion of the subcorpus it occupies in the youngest of stages (0-3: 7,39%, 4-6: 3,72%, 18+: 4,69%). Even the verb *give* shows a slight decline with the progression of years, but this difference may be attributed to the greater variety of verbs used in ditransitive constructions among adults.

Learning of ditransitive constructions might be item-based, but the difference in frequency of verbs such as *get* and *make* indicate that the early language, very much like the adult language, is situationally dictated. This is not necessarily decisive for any of the accounts of language acquisition, but it does provoke debate on the involvement of cognitive mechanisms in the early language. For instance, one might wonder whether the early use of these particular verbs takes up a larger portion of the ditransitive uses than it should because of the instinctive behavior (demanding food or drinks) or is it because the children in their early age rely on these communicative tools, not necessarily because of the survival as such, but because of their ability to recognize their incapacity to feed themselves and the capacity of others to fill this skill-gap once provoked by the linguistic cue that calls upon assistance. Naturally, the cognitive skill of comprehending situation and its requirements needs not be strictly separated from the survival instinct or instinctive behavior. The hierarchy of verb usage can be seen more clearly in Figure 2.

Figure 2. Distribution of ditransitive verbs across age groups



In the second hypothesis the assumption is that the verb selection across different age groups will differ with regard to variety and hierarchy of usage. When it comes to variety, the hypothesis is somewhat confirmed, given that the range of verbs used in ditransitive format is significantly greater than the ones used in younger stages. At first, it might appear that the reason for this is the restricted vocabulary, given that its size is a function of time and age. However, many of the verbs not found in ditransitive constructions were mastered by the children aged 0-3, such as *play* (5 208 times), *blow* (579 times), *sell* (35 times), *deliver* (29 times) etc. This indicates that the use of ditransitive constructions might be verb-based, and that that overgeneralization in the speech of children, at least when it comes to ditransitive constructions, is not present.[3]

[3] Here, it needs to be noted that an expression *Jack played me a toy*, would not be deleted from the data collected even though the semantics of the sentence seem rather awkward. This was done precisely for this reason, so that one could possibly compare whether children truly overgeneralize, even when these constructions prove semantically invalid. This casts serious doubt on children's productivity as praised by the nativists.

Demonstration of a verb does not necessarily imply understanding, especially in terms of its semantic and pragmatic constraints. The demonstration of a certain verb relies heavily on the input, on the repetition of what was said in a particular instance. This does not imply a supreme linguistic ability, but rather remembering the instance in which somebody used the particular verb, if not the entire construction, and then repeating the same in hope that the uttered words will yield the same results. As Tomasello puts it:

> "However, beginning language learners are not creative or productive with their language in some other basic ways. For example, they do not use a verb in a sentence frame in which they have not heard it used. Thus, on the basis of hearing just *The window broke* (and no other uses of this verb) they cannot go on to produce *He broke it* or *It got broken*, even though they are producing simple transitive and passive utterances with other verbs. This lack of productivity suggests that young children do not yet possess abstract and verb-general argument structure constructions into which different verbs may be substituted for one another as needed, but rather they are working more concretely with verbs as individual lexical items whose syntactic behavior must be learned one by one" (2000: 210).

Another assumption of the third hypothesis is that the verbs will not differ according to the hierarchy of usage. Regardless of the minor differences found in the verb frequency (*give, get, make*), the general pattern is strikingly similar. In fact, children's speech seems to mirror the input completely, regardless of the syntactically simplified constructions, which are in alignment with what is to be expected from the early syntactic competence. Here, it seems that we can separate two main factors in determining the reason behind such similar distributions across age groups:

a) The input seems to be the main source and the guide for the acquisition of verbs.

b) The nature of children's early behavior and the situational requirements are bound to influence the choice of lexical items first learned and demonstrated.

Nevertheless, the input seems to maintain its primacy when it comes to children's productivity, or better to say non-productivity. Such mirroring of the input when it comes to the hierarchy of usage invites one to characterize the language acquisition process to be parrot-like. The situational requirements bear only minor influence on verb usage, enough to account for the small differences found under the verbs such as *give, get* and *make*, but not enough to account for the entire acquisition process or to undermine the usage-based account of language acquisition. This is especially interesting when one looks at the earliest uses of dative constructions and the verbs acquired first. The data on the emergence of dative constructions in Brown's children was taken from Tables 2 and 3 in Gropen et al. (1989: 213-214) and integrated into Table 6:

Table 6.

Verbs in double-object use	Adam's age	Eve's age
bring	3;1	1;10
buy	3;3	/
draw	3;4	/
get	2;4	2;0
give	2;3	1;9
hand	2;6	/
make	3;5	/
read	4;2	1;8
show	3;0	1;9

sing	3;0	/
tell	3;0	/

It is clear from Table 6 that the first verbs used in double-object form are the same ones that are found to be the most frequent ones in the input. Indeed, according to Campbell, Aimee & Tomasello (2001: 262) and their examination of Brown's children, the first verbs used in ditransitive fashion were: *give*, *show*, *bring, feed, send, read, get, make*. One of the suggestions on how children acquire complex constructions is that they take certain verb from the input, a "light" verb that leads the acquisition process of the rest of the syntactic constructions in which it may be found. The acquisition of syntactic constructions as something relying on prototypical verbs which constitute the majority of these constructions was asserted by many scholars coming from cognitive approaches to language acquisition (Goldberg et al. 2004 & Boyd and Goldberg 2009). According to them, the acquisition of ditransitive constructions, and other syntactically idiosyncratic constructions, relies on several verbs found the majority of these constructions in the input. The prototypical relationship of verbs and its arguments associates itself with the meaning of the verb in question, which ultimately allows the learners to make generalizations about it through time.

In dative constructions, the verb *give* is implied to be the "light" one (Campbell et al 2001: 254). These verbs were often labelled as "pathbreaking" verbs (Ninio 1999), as they pave the way for the acquisition of not only constructions, but other verbs in general. Once the "pathbreaking" verb is mastered, the rest should follow. The implication of the "pathbreaking" verbs is that they are learned first, but the question remains why they would be learned first and if so, what would be best way to prove this assumption. When it comes to proving the latter, a number of longitudinal studies would prove more suitable for this

endeavor. When it comes to the reason as to why they are acquired first, it is possible to argue that this primacy is a consequence of the input. However, Campbell et al. (2001) noted that not all of the children began with the verb *give*, which might be considered "pathbreaking", but with other verbs listed here. Out of the four children who did begin with the verb *give*, three of them started using another verb within a very short time frame, i.e. within a month.

Some studies (Thothathiri & Snedeker 2008, Conwell & Demuth 2007) indicate that three- and four-year old children possess knowledge about dative constructions that is verb-independent. This seems to be the crucial age for the acquisition of ditransitives. The results on the verb-independent comprehension on datives does seem to deviate from the results gained from this research. However, it needs to be noted that their conclusion about the abstract knowledge of dative constructions can still be in accordance with the acquisition of dative constructions via prototypical verbs. The prototypicallity hypothesis assumes that the abstract knowledge will follow eventually, but the initial demonstration of these constructions will rely on several verbs. The difficult task of the linguists is to locate the precise point at which the demonstration of dative constructions based on prototypical verbs ends and the acquisition based on abstractness begins.

8. Conclusion

The first hypothesis may be regarded as partially confirmed. Similarly to verbs, there is a striking similarity when it comes to the dispersion of constructions and their internal structure between age groups. Nevertheless, as expected, ditransitive constructions containing adjectives and determiners do constitute a larger portion of constructions in adult language.

The second hypothesis is confirmed as ditransitive constructions proved to take up a greater part of syntactic constructions in general with the progression of age. It would be interesting to see whether this progression in usage corresponds perfectly with all age groups, but the size of the age 6-18 subcorpora was unfortunately too small for valid calculus.

Finally, the third research question is concerned with the verb selection in ditransitive constructions where we can differentiate between the hierarchy of usage and the range of lexical verbs used. Surprisingly enough, the alignment of verbs is strikingly similar across the examined groups. The differences found are almost insignificant on such scale and thus inconsequential for any other conclusion but the one where the input holds primacy on language acquisition. Yes, the range of verbs used is different, but this is clearly a mark of reduced vocabulary in younger stages, thus confirming both of the hypotheses related to the third research question. The important observation here is that the first 15 or so verbs (the so called “light” verbs) are almost evenly distributed across groups.

The results of this study can be reduced to the question of what can this evidence tell us about language acquisition? While this research provides us with definitive answers on the distribution, usage and structure of ditransitives in different age groups, the empirical data is still insufficient to reach any verdicts on the nature of language acquistion. However, the data does seem to fit the cognitive models of acquisition theory more than it does the nativist accounts. Although the research is restricted to ditransitives, the results have unequivocally shown that the syntax, like lexicon, evolves through time. It appears that there are two general

inferences that can be drawn from the study. The first is that the acquisition of grammar is a long-term process, the outcome of which relies on repetition and cognitive abilities. The second is that the input holds primacy over the acquisition of both lexicon and syntactic constructions such as ditransitive ones.

9. Bibliography

Arppe, A., Gilquin, G., Glynn, D., Hilpert, M., Zeschel, A. (2010). Cognitive Corpus Linguistics: Five Points of Debate on Current Theory and Methodology, *Corpora* 5(1), 1-27.

Boas, H. C. (2008). Determining the structure of lexical entries and grammatical constructions in Construction Grammar, *Annual Review of Cognitive Linguistics* 6(1), 113-144.

Boyd, J. K., Goldberg, A. E. (2009). Input Effects within a Constructionist Framework." *The Modern Language Journal* 93(3), 418–429.

Bloom, L. (1988). Language acquisition in its developmental context, *Handbook of Child Psychology Volume 2: Cognition, Perception and Language 5 ed*., edited by D. Kuhn & R. Siegler, New York: Wiley, 309-370.

Brown, R. (1973). *A first language: The early stages*. Harvard U. Press.

Brown, R. (1957). Linguistic determinism and the part of speech, *The Journal of Abnormal and Social Psychology* 55(1), 1–5

Campbell, A. L., Tomasello, M. (2001). The acquisition of English dative constructions, *Applied Psycholinguistics* 22(2), 253-267.

Carnerero, C. A. (2007). On "Dimonotransitive" Structures in English, *Conference on Ditransitive Constructions*, 23-25 November 2007, Leipzig (Germany), Max Planck Institute for Evolutionary Anthropology.

Chomsky, N. (2007). A Review of B.F. Skinner's Verbal Behaviour, in *Language* 35(1) (1959), 26-58.

Cook V.J., Newson, M. (2007). *Chomsky's Universal Grammar*. Blackwell Publishing.

Conwell, E., Demuth. K. (2007). Early syntactic productivity: Evidence from dative shift, *Cognition* 103, 163–79.

Cruttenden, A. (1978). Assimilation in child language and elsewhere, *Journal of Child Languag, 5*, 373–378.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Vojtěch, K., Michelfeit, J., Rychlý, P., Suchomel, V. (2014). The Sketch Engine: ten years on, *Lexicography* 1(1), 7-36.

Evans, N., Levinson, S. (2009). The Myth of Language Universals: Language diversity and its importance for cognitive science, *Behavioral and Brain Sciences* 32(5), 429-448.

Fillmore, C.J. (1968). The Case for Case". In Bach and Harms (Ed.): *Universals in Linguistic Theory*. New York: Holt, Rinehart, and Winston, 1-88.

Gillis, S. (2014). Child Language Data Exchange System, http://www.cnts.ua.ac.be/~gillis/pdf/2014_CHILDES.pdf. Accessed 1 May 2018.

Goldberg, A.E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Goldberg, A.E., Casenhiser, D.M., Sethuraman, N. (2004). Learning argument structure generalizations, *Cognitive Linguistics* 15(3), 289–316.

Gómez, R.L., Gerken, L. (2000). Infant artificial language learning and language acquisition, *Trends in Cognitive Sciences* 4(5), 178–186.

Green, G. M. (1974). *Semantics and syntactic regularity*. Indiana University Press.

Gropen , J., Pinker, S., Hollander, M., Goldberg, R., Wilson, R. (1989). The learnability and acquisition of the dative alternation in English, *Language* 65, 203-257.

Hauser, M.D., Chomsky, N., Fitch, W.T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?, *Science* 298. 5598, 1569 – 1579.

Kittilä, S. (2005). Recipient-prominence vs. beneficiary-prominence, *Linguistic Typology* 9(2), 269–297, DOI: https://doi.org/10.1515/lity.2005.9.2.269.

Lakoff, G., (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. University of Chicago Press.

Lakoff, G., Thompson, H. (1975). Dative questions in cognitive grammar, *Parasession on Functionalism*, 337-50.

Langacker, R.W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*. Vol. 1. Stanford university press.

Locke, J. (1690). *An Essay Concerning Humane Understanding*. 1st ed. 1 vols. London: Thomas Bassett.

Malchukov, A., Haspelmath, M., Comrie, B. (2010). Ditransitive constructions: a typological overview, *Studies in ditransitive constructions: A comparative handbook*, 1-64.

MacWhinney, B. (2015). Tools for Analyzing Talk–Electronic Edition Part 1: The CHAT Transcription Format.. The CHILDES Project: Tools for Analyzing Talk. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates, https://talkbank.org/manuals/CHAT.pdf

Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity, *Journal of Child Language 26*, 619–654.

Pinker, S (2013). *Learnability and cognition: The acquisition of argument structure*. MIT press.

Pinker, S. (2002). *The blank slate: The modern denial of human nature*. Penguin Group.

Pinker, S. (2003). *The language instinct: How the mind creates language*. Penguin UK.

Saffran, J.R., Aslin, R.N., Newport, E.L. (1996). Statistical learning by 8-month-old infants, *Science* 274.5294, 1926-1928.

Siewierska, A. (1998). Languages with and without objects: The functional grammar approach, *Languages in contrast* 1(2), 173-190.

Silić, J., Pranjković, I. (2005). *Gramatika hrvatskoga jezika za gimnazije i visoka učilišta*. Zagreb, Školska knjiga.

Soja, N.N. (1994). Evidence for a distinct kind of noun, *Cognition* 51, 267-284.

Theakston, A.L., Lieven, E.V.M., Pine, J.M., Rowland, C.F. (2001). The role of performance limitations in the acquisition of verb-argument structure: an alternative account, *Journal of Child Languag, 28*, 127-152.

Thotathiri, M., Snedeker, J. (2008). Syntactic priming during language comprehension in three-and four-year-old children, *Journal of Memory and Language* 58, 188–213.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Harvard University Press.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition 74,* 209-253.

Tomasello, M. (2008). The Usage-based theory of language acquisition, in Bavin, E, (Ed.) *The Cambridge handbook of child language*, 69-88.

Tomasello, M., Brandt, S. (2009). Flexibility in the Semantics and Syntax of Children's Early Verb use, *Monographs of the Society for Research in Child Development* 74(2), 113-26.

Viau, J., Lidz, J. (2011). Selective learning in the acquisition of Kannada ditransitives, *Language* 87(4), 679-714.

Zovko Dinković, I. (2007). Dative alternation in Croatian, *Suvremena lingvistika* 63(1), 65-83.

Learner corpora research: Acquisition of ditransitive constructions

Abstract

Language acquisition remains one of the most mysterious processes that modern linguistics has yet to fully understand and explain. This thesis utilizes the data from the CHILDES corpus by looking at the production of ditransitive constructions across age groups. The data compared includes 0-3, 4-6 and 18+ age groups (the size of other subcorpora was too small for comparative purposes). This research targeted only "double-object dative" ditransitives (VP → NP1 → NP2) and did not include the "prepositional dative" variation (VP → NP → PP). The main research questions are concerned with the structural characteristics of ditransitive constructions across different age groups, frequency of ditransitive usage and verb selection in ditransitive constructions across different age groups. First, the results show a high similarity in the dispersion of constructions and their internal structure between age groups. Secondly, the use of ditransitive constructions rises with the progression of age. Thirdly, the range of verbs used across age groups differed (due to obvious vocabulary differences), but the hierarchy in verb selection again showed a striking similarity (the same verbs occupied almost equal portions of ditransitive uses in all three age groups). The research has shown that the input holds primacy over language acquisition, which appears to be an incremental process relying on cognitive abilities.

Key words: language acquisition, learner corpora research, CHILDES, ditransitive constructions

Istraživanje učeničkog korpusa: Usvajanje ditranzitivnih konstrukcija

Sažetak

Usvajanje jezika je i dalje jedan od najzagonetnijih procesa kojeg moderna lingvistika još treba razumjeti i objasniti. Ovaj diplomski rad koristi podatke iz CHILDES korpusa promatrajući produkciju ditranzitivnih konstrukcija u različitim dobnim skupinama. Proučavani podatci uključuju dobne skupine od 0-3, 4-6 i 18+ (veličina ostalih podkorpusa bila je premala za komparativne svrhe). U istraživanje su uključene samo ditranzitivne konstrukcije s dva objekta u dativu (VP → NP1 → NP2), dok je varijacija ditranzitiva s prijedložnim dativom (VP → NP → PP) izuzeta iz konačnog rezultata. Glavna istraživačka pitanja odnose se na strukturne karakteristike ditranzitivnih konstrukcija u različitim dobnim skupinama, učestalost korištenja ditranzitivnih konstrukcija te izbor i raspon glagola u ditranzitivnim konstrukcijama u različitim dobnim skupinama. Rezultati su pokazali visoku sličnost u disperziji konstrukcija i njihovoj unutarnjoj strukturi među dobnim skupinama. Osim toga, uporaba ditranzitivnih konstrukcija povećava se sa starenjem. Nadalje, raspon glagola korišten u dobnim skupinama varira (zbog očitih razlika u vokabularu), ali hijerarhija u odabiru glagola je ponovo pokazala nevjerojatnu sličnost (isti su glagoli zauzimali jednak udio u korištenju ditranzitiva u sve tri dobne skupine). Istraživanje je pokazalo da input okoline igra ključnu ulogu u usvajanju jezika, te da se radi o postepenom procesu koji se oslanja na kognitivne sposobnosti.

Ključne riječi: usvajanje jezika, korpusno istraživanje, CHILDES, ditranzitivne konstrukcije