

Weight effects in the acquisition of English

Evidence from first and second language acquisition

Masterarbeit

vorgelegt von

Daniel Bürkle

an der

Universität Konstanz

Geisteswissenschaftliche Sektion

Fachbereich Sprachwissenschaft

1. Gutachter: PD Dr. Markus **Bader**
2. Gutachter: Prof. Dr. Janet **Grijzenhout**

Konstanz 2011

Acknowledgements

I am very grateful to Dr. Markus Bader and Prof. Janet Griizenhout, my advisors, for providing much needed focus to this initially very fuzzy thesis. They set me a challenge, but were always happy to help with theoretical, methodological, and other issues and thus motivated me to do the best work I could.

Further thanks are owed to Prof. Josef Bayer, Dr. Heidi Altmann, Dr. Andreas Trotzke, Dr. Antje Lahne, Iris Bräuning, and Annika Nitschke for readily giving up valuable time in their classes to my questionnaire study. I also thank all their students for filling out the questionnaire without any compensation whatsoever.

Ursula Scholz, Mariya Kharaman, and Michael Merz also shaped this thesis in many small ways and made writing it rather more enjoyable. All remaining errors and mistakes are entirely my own. Finally, I am thankful that T.F., Peter, Jochen, Christoph, and all the Andis reminded me once or twice of the world beyond publications, prepositions, and p -values.

Contents

1	Introduction	1
2	Reordering phenomena	6
2.1	Heavy NP shift	6
2.2	The dative alternation	11
2.3	Particle verbs	19
3	Theories of syntactic weight	24
3.1	Weight in isolation	24
3.2	Relative weight	26
3.3	The multifactorial approach	31
4	Corpus study	34
4.1	Corpus	34
4.2	Results and discussion	35
4.3	Statistically modelling the dative alternation	37
4.3.1	The model	40
4.3.2	Validation	42
4.3.3	Discussion	44
4.4	Summary	49
5	Questionnaire study	51
5.1	Participants	51
5.2	Stimuli	52
5.3	Predictions	53
5.4	Results	54
5.5	Discussion	56
5.6	Summary	58
6	Conclusion	60
	References	62
A	Sample questionnaire	70

Sentence first, verdict afterwards.

Queen of Hearts, *Alice's Adventures in Wonderland*

1 Introduction

While it is common knowledge that the word order of modern English is relatively fixed, there are some well-attested phenomena where two different constituent orders are possible. These phenomena are affected by what have been called ‘syntactic weight effects’, wherein different orderings of constituents become available or unavailable according to features of the constituents in question. Arguably the most common theory for this is that these effects are caused by the syntactic structure or length of the constituents, two measures which of course correlate with each other. It is only natural to think of larger constituents as ‘heavier’, and so those theories certainly influenced the term ‘weight effects’. The structural idea of weight also extends well to the phonological structure of constituents, which arguably also correlates with syntactic structure and length (see Scaglione 1979:238–239). Additionally, the phenomenon has been studied from a pragmatic perspective. Whether a constituent is new in context can conceivably also affect its weight and therefore the structure of the clause that contains it.

Three English phenomena that exhibit weight effects on constituent order are heavy noun phrase (NP)¹ shift, the dative alternation, and certain particle verbs. Heavy NP shift, as in (1), is the phenomenon that a postverbal NP, usually an object, can occur after another constituent if and only if the NP is ‘heavy’ in some way.

¹I will use the term ‘NP’ here. It is established in the literature, and it is not less helpful in explaining the phenomenon than any alternative (like ‘DP’).

- (1) a. ?I saw some angry people who all needed a shower and a stiff
drink on the ship.
- b. I saw on the ship some angry people who all needed a shower
and a stiff drink. NP IS 'SHIFTED'
- (2) a. Simon gave Malcolm the money.
DOUBLE OBJECT CONSTRUCTION
- b. Simon gave the money to Malcolm.
PREPOSITIONAL CONSTRUCTION

The term 'dative alternation' means the simple fact that some English verbs that take two objects (and thus one dative object²) allow their objects to be either two NPs, as in (2a), or one NP and one prepositional phrase (PP), as in (2b). The PP in the second construction is typically the dative object, so it is the dative object that alternates between two possible phrases. The two possible constructions are often called the double object construction (or dative) and the prepositional construction (or dative) respectively, and I will use these terms here. The dative alternation is relevant to the discussion of weight effects because it shows the opposite effect to the one in heavy NP shift: whereas a heavy constituent *licenses* the different order in heavy NP shift, it *blocks* one order in the dative alternation. Of course, this blocking goes both ways, so that a heavy goal argument blocks the double object construction (as in (3a) and (3b)) and a heavy theme blocks the prepositional construction (as in (3c) and (3d)).

²Of course, I do not mean to categorically state that non-pronominal NPs in English can have dative case; again, I use an established term for simpler reference.

- (3) a. ? Simon gave the tired and suspicious Malcolm, who could not stand the sight of him at all, the money.
- b. Simon gave the money to the tired and suspicious Malcolm, who could not stand the sight of him at all.
- c. Simon gave Malcolm a lot of money, which he had earned through years of hard work and would rather have kept.
- d. ? Simon gave a lot of money, which he had earned through years of hard work and would rather have kept, to Malcolm.

English is rich in particle verbs, and so it is no surprise that not all of them behave exactly alike. The difference I will be interested in here is of course the possibility of different object-particle orderings: some transitive particle verbs, as in (4), always have to occur in the V–Prt–NP order; others allow both this and the V–NP–Prt order, as in (5). Interestingly, (6) shows that the latter order becomes odd as the object NP becomes heavy.

- (4) a. Simon tried out for the rowing team.
- b. * Simon tried for the rowing team out.
- (5) a. Malcolm tried the ship out.
- b. Malcolm tried out the ship.
- (6) a. ? Malcolm tried the ship he had bought the day before out.
- b. Malcolm tried out the ship he had bought the day before.

A number of different explanations and formalizations for these and other weight effects have been proposed. As I will show, the most successful ones of these are not binary, but scalar—that is, they employ a graded concept of weight instead of a simple ‘yes or no’ definition. Corpus and experimental

data provide strong support for these gradient explanations. Furthermore, it is intuitive: while the asterisk and question mark (and any combinations thereof) as labels of grammaticality are not rigorously defined, it is obvious to any competent speaker of English that the above examples of weight effects, particularly (3a) and (3d), are neither fully grammatical nor fully ungrammatical. Gradient theories capture this ‘gray area’ and thus quite rightly show weight effects to be examples of gradience in syntax.

After first discussing some theories of the three phenomena mentioned above (section 2) and of weight effects (section 3), I will report on two studies of syntactic weight effects. The first is a corpus study of some of the English corpora in the Child Language Data Exchange System (CHILDES; MacWhinney 2000), presented in section 4. The aim of this study is to validate some theories concerning weight effects against first language acquisition data. To achieve additional results that are both quantifiable and comparable, a subset of this study is then used to construct a statistical regression model of the constituent ordering.³ Based on the results of these analyses, I designed and carried out a questionnaire study among young adults in Germany with English as a second language. Section 5 presents this study. The aim of this is to compare first and second language acquisition, or first language acquisition and second language *learning* if one distinguishes naturalistic acquisition from academic learning. I will not be making that distinction here, but further research on weight effects in naturalistically acquired second language would no doubt be insightful.

As mentioned above, the results of these two studies show theories that go beyond a simple categorical definition to be the most descriptive as well as

³All statistics and computations reported here were run in the R environment (R Development Core Team 2011).

the most predictive. This thesis thus contributes some data to the discussion surrounding the causes and development of syntactic weight effects in English. These data do not prove any theory of syntactic weight effects to be correct. They do not conclusively refute any theory, or any of the counter-arguments that have been made against the theories discussed below. They do not show how children acquire the various possibilities of constituent ordering in English, or how adult learners learn them. What they *do* do is support some theories much better than others.

2 Reordering phenomena

Before discussing the theories of weight effects and then testing them, a description of the phenomena that exhibit them is in order. This thesis does not focus on the structure of sentences with weight effects, and it is limited to Modern English only. Nevertheless, it is necessary to discuss some syntactic descriptions as well as crosslinguistic accounts of weight-related phenomena because, as will be seen later, the more descriptive and predictive theories are based on the structure of constituents as well as the ideal of an universal approach to the study of human languages. Indeed, all linguistic theorizing must be aware of phenomena which are similar to the ones under discussion, but are found in other languages. To this end, this section presents some of the theoretical work on heavy NP shift, the dative alternation, and particle verbs (following Arnold et al. 2000's overview).

2.1 Heavy NP shift

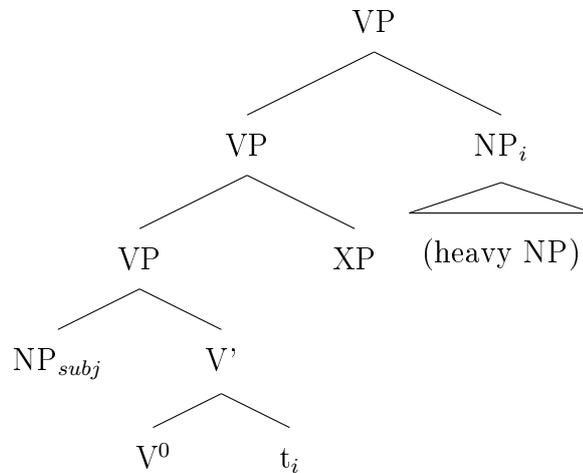
In English, heavy NP shift means that a heavy constituent is in a position that is to the right of the canonical position for that type of constituent, as in (1). The same applies in other languages where full NP objects are canonically placed after the verb, such as French (de Wind 1999), Croatian (Zovko Dinković 2007:68), or Spanish (Lozano and Mendikoetxea 2010:48). In languages with a different canonical ordering, heavy NP shift works differently: for example, the head- and verb-final Japanese has heavy NP shift that places the heavy constituent in a position before its canonical one. This supports theories which explain weight effects by positing that constituent heads should be as close together as possible. In head-initial languages like

English, this means that the longer constituents are put to the right edge; in head-final languages like Japanese, this means that longer constituents are put to the left edge (Chang 2009). Chang fed both English and Japanese sentences into a computational model capable of learning and then made it produce further sentences based on what he calls ‘messages’, which are semantic frames encoding an action, its participants, and their relative importance. The sentences that were produced showed heavy NP shift with the heavy phrase to the right in English and to the left in Japanese. He concludes that heavy NP shift is an emergent phenomenon that arises out of the pragmatic and semantic features of the ‘message’ that is to be conveyed.

While these crosslinguistic considerations are instrumental in teasing out the triggers of heavy NP shift, its structural aspects must also be considered in order to arrive at a reasonable theory of the phenomenon. Three types of syntactic analyses for heavy NP shift structures have been proposed: one sees it as an instance of rightward movement, another posits leftward movement, and the third argues that heavy NPs remain in situ. This section will briefly discuss these analyses in general before turning to one that views heavy NP shift as being outside of syntax.

Saito and Fuku (1998)’s analysis of heavy NP shift ties into their larger theory of generative syntax. In that view, heavy NP shifting is “treated on a par with scrambling” (Saito and Fuku 1998:445), which means that the heavy NP is moved out of its base position into a adjunct position on the right edge. The simplified tree in (7) illustrates the essential structure of this theory.

(7)

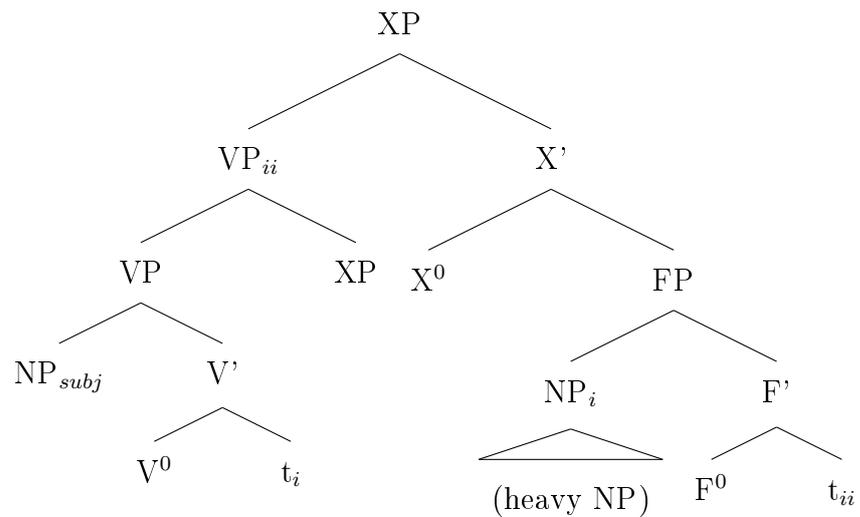


Saito and Fuku elaborate on this and find that the Move operation is not strictly necessary in a Bare Phrase Structure approach. Therefore, heavy NP shift, like other movement operations, is created by a special interpretation of Merge. Headedness rules apply, of course, and therefore a heavy NP must always be merged on the non-head side of the larger phrase. Assuming that other constituents are merged before the heavy NP, this works well in a variety of languages, as will be seen below. However, it also removes the motivation for heavy NP shift, and (1) shows that weight effects can make one of two possible orderings much less preferred than the other. Thus, heavy NP shift has a reasonably well-established trigger and is clearly not completely free and optional—two points that this rightward movement analysis misses entirely.

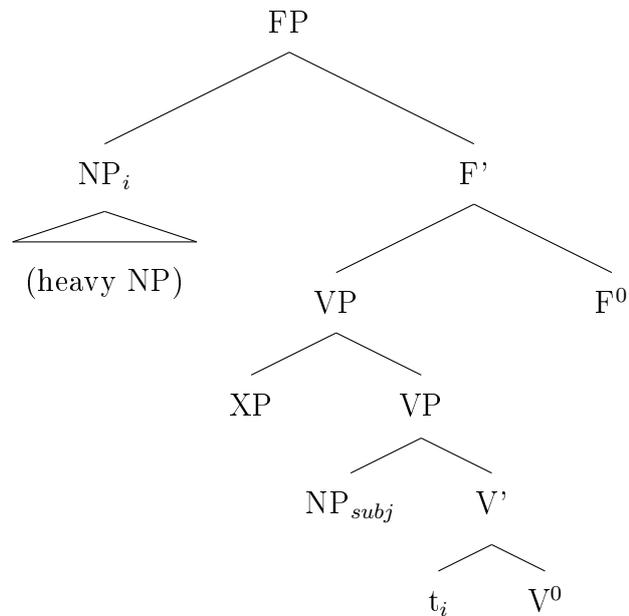
For English and other head-initial languages, the leftward movement analysis is naturally somewhat more complex: typical English heavy NP shift has all constituents in their canonical order, except the heavy NP that is to the right of all other constituents. Such theories are therefore less intuitive than the rightward movement account and also more complex, as they need more than one operation to arrive at the shifted order. In the leftward movement

analysis, as for example Takano (2003) or Rizzi (2006) implement it, the heavy NP first moves out of and to the left of the VP; then, the VP (and other material, as per Kayne 1998:section 4.5) moves leftward over the NP (see the simplified (8a)).

(8) a.



b.



While not a complete Duke of York derivation, this account is more complex than the intuitive rightward movement. However, as Takano (2003) points out, this double leftward movement deals with some theoretical problems; and additionally, it is unintuitive only in languages where heavy NPs end up to the right of their base position (compare (8b) to (7)). Languages where heavy NPs end up on the left edge make this approach seem much more natural. English speakers may intuitively reject the leftward movement analysis, but that is nothing more than a coincidence of headedness.

Finally, de Wind (1999) argues that heavy NPs are not moved at all, but simply merged in their base position, thus forcing the θ -role to percolate to them. This, according to de Wind, can only happen if and when there is exactly one θ -role and one θ -less NP left, which in turn only happens if the other arguments move around the heavy NP to receive their θ -roles first. He bases this on aspects of Standard and Canadian French, but argues that it should generalize to other languages with weak Case features and other similarities to French. As I pointed out above, working towards generalization is commendable; there are other problems with this approach, however. To conform to independent assumptions, de Wind assumes that the heavy NP and its features can be separated and merged in different positions. This also allows him to explain the syntactic differences in heavy NP shift between Standard and Canadian English: he argues that the weight of an NP must be established before syntactic operations begin. Because of this, he must “assume that heaviness is expressed by some feature in the Numeration” (de Wind 1999:95–96). However, he does not explain what this feature might be and on which criteria it might be assigned. As it stands, this analysis therefore leaves heavy NP shift completely without a cause or trigger and multiplies the entities involved without much need.

Opposing all these syntactic analyses, Lozano and Mendikoetxea (2010) posit that heavy NPs do not move at all because heavy NP shift is not a syntactic operation. Observing that the proposition of an utterance is not affected by whether a heavy NP is shifted or not, Lozano and Mendikoetxea argue that heavy NP shift is not present at LF and thus not in syntax. The only possibility, then, is PF, where heavy NP shift is caused by the mere length of the heavy NP and not by any syntactic features. In a corpus study of native and learner English, Lozano and Mendikoetxea find that postverbal subjects are often heavy. While they complicate their analysis by using a nominal scale that is closer to some early definitions of weight (see section 3.1), the grades on this scale correlate acceptably with length in words. Interestingly, NP weight also correlates with focus. Lozano and Mendikoetxea thus find that a postverbal subject is more likely to be heavy and new than a preverbal one. These findings are in line with previous research on weight and information structure, as will be seen, but they do not automatically mean that heavy NP shift is not a syntactic phenomenon. Some of the problems solved by the syntactic analyses remain unexplained in this theory, as they do in the analysis that uses partial spell-out to reconcile weight effects on PF with syntactic structure building (Shiobara 2008) or in the minimalist approach (Akasaka and Tateishi 2001). Furthermore, excluding all morphosyntactic features from an analysis of heavy NP shift means excluding some strong and helpful predictions. To avoid this, I will consider *all* available information in my approach to weight effects. The critical question of just what information is actually available to a speaker's language faculty as it makes ordering decisions remains unanswered, however.

2.2 The dative alternation

While the dative alternation has been claimed to exist in languages other than English, the features of the two alternating constructions can be quite different. Dutch, being closely related to English historically and typologically, has very similar structures to those in (2) (Coleman and de Clerck 2009:6). The similarities between the English and Greek dative alternations are already more abstract (Cuervo 2003:119). The comparison becomes even murkier when involving languages that are not Indo-European: it has been argued that Basque has an equivalent to the dative alternation (Ormazabal and Romero 2010) and that it does not (Oyharçabal 2010); and Chinese has verbs that alternate between two dative-equivalent structures and verbs that alternate between three of them (Liu 2006).

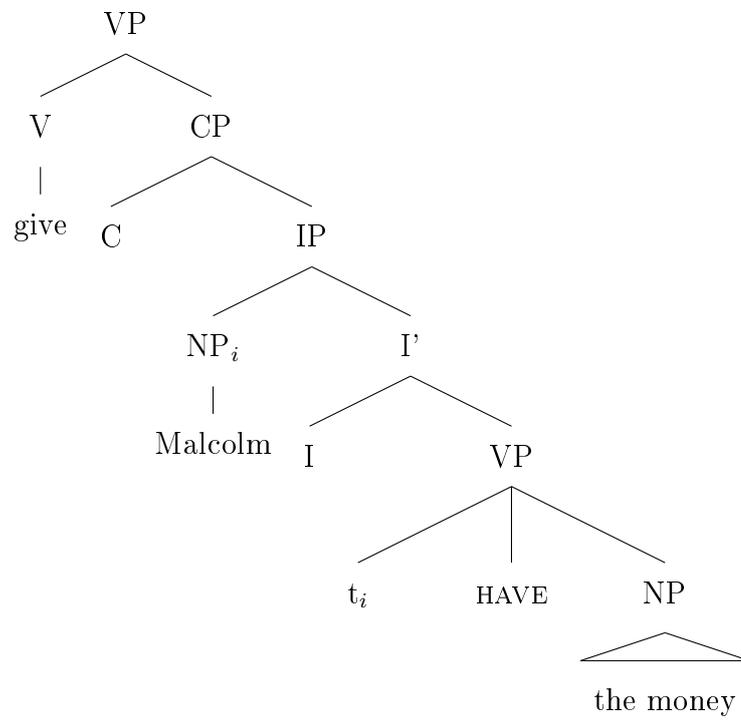
Spanish, then, provides a good example of dative alternation in languages other than English: it is generally agreed to have a dative alternation, but that is sufficiently different from the English one to show only the fundamental similarities. The two alternating structures in Spanish appear to differ only by a dative clitic. The construction with the clitic present “corresponds to the [English] double-object configuration” (Cuervo 2003:124) and the clitic-less alternative corresponds to the prepositional dative. Cuervo argues that quite different derivations underlie the two: while the clitic-less construction is derived straightforwardly, in accordance with its surface order of ACC–DAT, the clitic construction is derived by merging the two objects in the reverse order and then raising and thus preposing the accusative object. Cuervo argues that this structure explains supposedly idiosyncratic *c-command* as well as a crucial restriction in meanings. As she shows, the clitic can only be used when the dative argument can be interpreted as a

benefactive or possessor while the clitic-less constructions goes together with any meaning containing movement towards a goal.

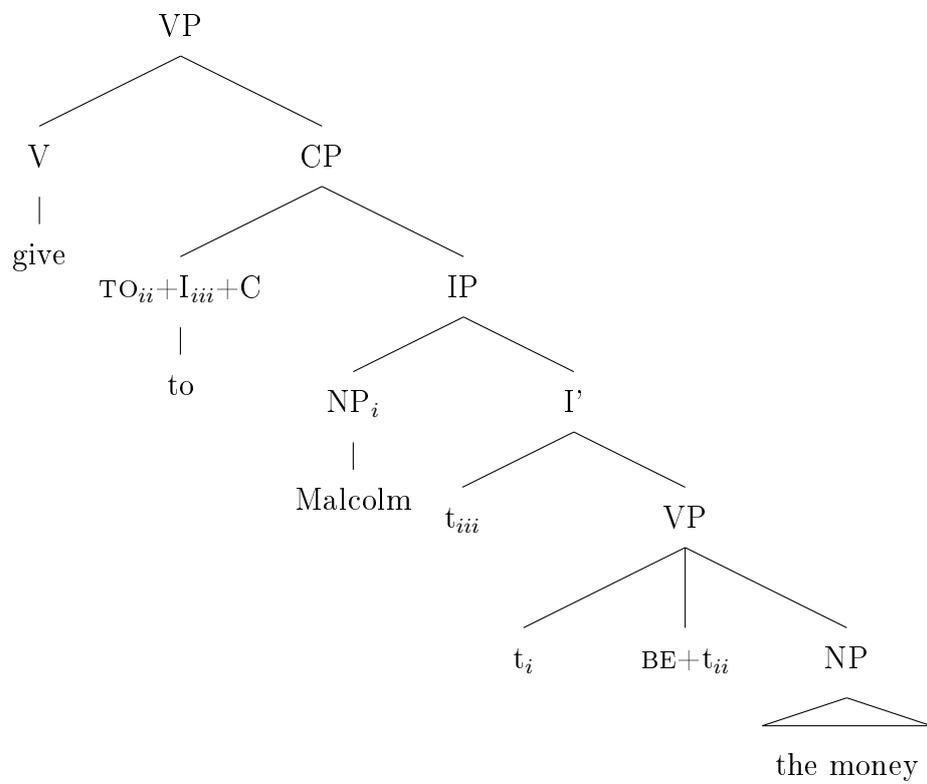
A similar approach to the English dative alternation would instinctively seem quite straightforward as well: both the double object and the prepositional construction are examples of a verb with two objects, and both should be easily generated in syntactic theories. However, it has been argued that the prepositional construction reflects the basic underlying structure (Ormazabal and Romero 2010:221), which would mean that there has to be a derivation path from that to the double object structure. I will briefly discuss such approaches before turning to proposed triggers for one construction or the other.

It has been argued that the double object construction is the basic one and that the prepositional construction is derived from it (Zovko Dinković 2007:66). For the English dative alternation, these theories must of course explain the appearance of the preposition in this derivation. (9) shows a greatly simplified representation of how Ogawa (2008:108) accomplishes this: like Cuervo (2003), Ogawa argues that the double object construction means possession and the prepositional construction means mere spatial movement. The possession is then encoded in a small clause by an invisible ‘verb’ HAVE and the change of location by BE. Crucially, Ogawa assumes that HAVE contains a TO-element that can be excorporated from this invisible ‘verb’ and moved to the empty C^0 head, where it is subsequently phonologically realized.

(9) a.



b.



In the prepositional structure shown in (9b), the theme NP (like a few other elements) is later moved out of the small clause into a higher position, of course. However, such small clause analyses are problematic: even assuming that only the prepositional construction contains a small clause, as Bruening (2010) does, the supposed small clause still fails some of the tests for small clauses as laid out in Farrell (2005:109–117).

- (10) a. * What I want you to try to give is the money to Malcolm.
b. * The money to Malcolm is difficult to give.
c. ?? The money to Malcolm, I can hardly give.
d. * The money to Malcolm could only be given by someone like
you. (after Farrell 2005:111)

The examples in (10) show most clearly how the proposed prepositional small clause is not like typical small clauses. While it could of course be explained as a special type of small clause, it has been argued that the small clause analysis is not really necessary to account for the dative alternation, at least in English (Snyder and Stromswold 1997). As it is problematic, the small clause approach using the double object construction as the basic one should therefore be abandoned for reasons of parsimony alone.

As in Spanish, the most prominent argument for the prepositional construction as the basic one in English appears to be that the two dative constructions encode slightly different meanings, with the prepositional structure expressing motion towards a goal and the double object construction expressing the bringing about of possession. This seems reasonable, given the clearly acceptable (11c) and the clearly unacceptable (11d).

- (11) a. John sent a package to his brother.
b. John sent a package to Chicago.
c. John sent his brother a package.
d. * John sent Chicago a package. (all Gropen et al. 1989:234)

Gropen et al. (1989) use these examples to show for English that the double object construction is generally acceptable if the goal argument can be understood as a possessor and unacceptable if it cannot. They do find some double object constructions that are far from perfectly acceptable to all speakers, and even those have potential possessors as goal arguments. Gropen and colleagues note that spatial transfer is a relatively general semantic frame and that transfer of possession is a more specific subtype of spatial transfer. Thus, they conclude that the spatial prepositional construction is the basic one and the dative alternation can only affect verbs that allow a transfer of possession in their meaning. Rappaport Hovav and Levin (2008:143) further elaborate on how this is apparently inherent in certain types of verbs and argue that “the meaning of *to* is less restricted than that of the first object”. It follows that the prepositional construction is to be preferred as the basic structure of dative alternation.

This view of the prepositional as the basic structure is relatively traditional, as Scaglione (1979)’s historical research shows. He argues that the prepositional structure in its linearization shows “the normal sequence . . . in the SVO languages, like English”, while the double object construction is “a remnant from the days when English was still predominantly SOV” (Scaglione 1979:236). Support for this comes from crosslinguistic comparison: it is commonly assumed that German is an SOV language, at least underlyingly, and Scaglione notes that it has only an equivalent to the double object construc-

tion, but no prepositional dative construction.⁴ Therefore, Scaglione appears to be correct in connecting the double object construction to SOV languages and the prepositional construction to SVO languages. Interestingly, this would also explain why the dative alternation is relatively rare among the languages of the world: there are presumably not that many SVO languages with a recent SVO history.

While recent research has contested the idea that the alternating dative constructions express different types of events (Coleman and de Clerck 2009), it can therefore be concluded that the prepositional construction is the basic dative construction, at least in English. However, this long-standing theory is still facing some internal problems (see Ormazabal and Romero 2010). Ormazabal and Romero therefore decide to abandon it. Rather than dealing with every problem individually, they derive the double object construction from the prepositional dative by incorporating the preposition into the verb. This constructivist approach explains the seemingly reversed antecedent command mentioned above as well as the counterintuitive fact that double object idioms can (in certain conditions) appear with the prepositional structure. Interestingly, it also allows the the *for*-benefactive structure to be derived from the double object benefactive rather elegantly by having it ‘tag along’ the same derivational path. While this approach thus solves some problems, it does not address the fundamental question of dative alternation structure:

⁴The benefactive alternation with *für* (or English *for*) is not relevant to this point, as will be seen shortly. While double object construction as in (i) is canonical in German, the prepositional alternative in (ii), though extremely ‘odd’, is parseable.

- (i) Sie gab ihm einen Schraubenzieher.
She.NOM gave him.DAT a.ACC screwdriver.
- (ii) ?? Sie gab einen Schraubenzieher an ihn.
She gave a.ACC screwdriver to him.ACC.

I therefore also assume that German does not have a dative alternation.

why must one construction be derived from the other?

Other than theory-internal reasons, I see no obvious answer to this question. Therefore, I will close this section by opposing these generative-transformational theories and considering the data in support of a *choice* between two equal constructions (see also the Optimality Theoretic account in Bresnan and Nikitina 2003). Of course, the two structures must still be syntactically derived somehow; I merely propose that constructing intricate derivations from one structure to the other is less interesting than investigating the apparent triggers for one construction over the other.

A CHILDES-based corpus study similar to the one presented in section 4 below showed that length and syntactic features have an effect on the choice of construction in the English dative alternation (de Marneffe et al. 2007). Furthermore, preceding dative constructions affected the choice as well: de Marneffe and colleagues found that prepositional constructions make the prepositional construction more likely in following dative structures, and that double object constructions prime the double object construction in a similar fashion. Comparing this study to a corpus study of adult English, they concluded that “child speech only differs from adult speech in degree, not in kind” (de Marneffe et al. 2007:13). It must be noted, however, that their child speech corpus was rather small, considering the size of CHILDES. This means that their conclusions, though valid, are not as strongly supported as those in other corpus-based research. With my studies, I intend to add somewhat more substantial evidence to this body of data. If further research should discover conclusive proof of triggers for one dative construction or the other, this would undoubtedly inform and advance the theories concerned with the structure of the dative constructions.

2.3 Particle verbs

Different ordering possibilities in constructions with particle verbs appear to be rare: other than English, they have been reported in Dutch, Norwegian (both Kremers 2009), and Swedish (Sigurd 1995). They are not found in Spanish (Gilkerson 2006:62) or German, for example. As only this peculiar subset of modern Germanic languages allows different object-particle orderings, I will not consider the crosslinguistic differences here. This section only presents and discusses some structural theories of particle verbs (following Farrell 2005's comprehensive overview).

Their entries in any dictionary of English show that particle verbs intuitively appear to be single lexical items (as in (4)). In usage, however, some of them can be separated, which makes them seem to be two items (as in (5)). This duality is fundamental to much of the theoretical work on particle verbs, such as the question of semantic compositionality. As the meanings of most particle verbs do not follow simply from the meaning of the verb and the preposition or particle, it has generally been assumed that they are idiomatic (for example in Sigurd 1995). Research based on this assumption has found some support for it (Gilkerson 2006), but it has been shown that the principle of compositionality can apply to particle verbs as well: Lüdeling (1999:chapter 4) argues for this view, and merely giving the particle more than one narrow locative meaning allows Lechler and Roßdeutscher (2009) to reconstruct compositionality for German particle verbs with *auf*. Of course, German does not have different object-particle orderings, but similar in-depth studies of particle verbs in English (and Dutch, Norwegian, and Swedish, of course) would be expected to find at least some insightful generalizations toward underlying meanings. Semantically, particle verbs therefore seem

to be less problematic than previously thought, and the problem of their structure thus becomes a morphosyntactic one alone.

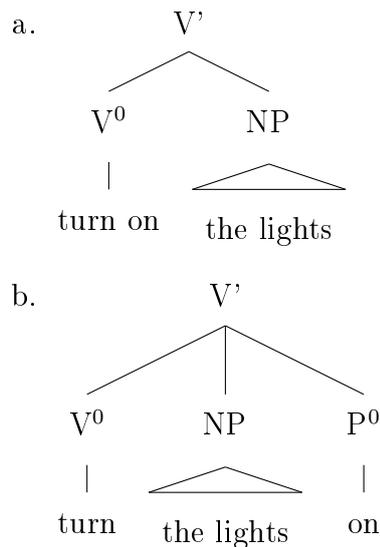
As with the dative alternation, it has been claimed that two very different structures underlie the different object-particle orderings: in the V–Prt–NP order, the particle is supposedly adjoined to the verb, whereas the particle projects a PrtP in the V–NP–Prt order (Kremers 2009:87). This analysis was proposed in order to capture the interesting point that the particle can be modified only in the latter order. However, it does not explain why there should be two different structures to express the same meaning. Kremers therefore turned to phonology for an explanation and found that the particle can create a *phonological* phrase only in the position after the object. This theory avoids the unnecessary complication of syntactic structures (and indeed syntax altogether), but it does not explain why one order is chosen over the other. It could be tested experimentally by studying whether illicit orderings or structures are indeed blocked prosodically, but that would not explain the syntactic structure any further.

Farrell (2005) shows that the V–NP–Prt order has often been assumed to be the basic one, with the other order derived from it in some fashion. The exact opposite has also been proposed: Dehé (2001), for example, argues that V–Prt–NP is the underlying order. Putting that theoretical issue to one side, these transformational approaches are obviously more complex than may be strictly necessary, and a simpler theory is desirable. Farrell attempts to provide one: he assumes that the V–Prt–NP order contains a compound of verb and particle, which then selects an object (see the simplified tree in (12a) below). The other order, however, is less orthodox in this theory: based on the restriction on modifying particles discussed above, Farrell argues that V–NP–Prt can only be explained with ternary branching (see (12b)).

The verb therefore selects two complements, the object and the particle.⁵ However, he also assumes that this restriction is violable if the object NP is heavy (Farrell 2005:101). This is not experimentally supported in any way, and it would not explain why heavy NPs tend to block the V–NP–Prt ordering. It therefore seems that Farrell’s baseless assumptions have backed him into the complicated corner of his analysis. The ungrammaticality of modified particles before objects is instrumental in building other analyses and should not be discarded without strong support.

(12)

(after Farrell 2005:98)



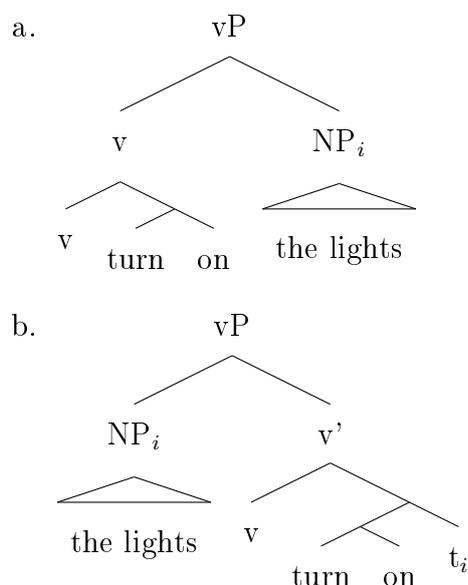
This problem of modifying particles can be connected to problems in object selection by benefactive verbs, as Basilico (2008) argues. In his Distributed Morphology approach, he explains all these problems by merging bare roots with category features at different points. This affects the morphosyntactic behavior of the abstract forms and concrete words thus derived: when

⁵If particle verbs are underlyingly ditransitive, as Oya (2009) assumes, this structure would become even more complicated.

merging a particle with a verb root,⁶ the resulting form must receive the V category feature before it can select its object. In this derivation, the verb and the particle together form one syntactic unit and can therefore not be split up by an intervening object (see (13a) below). The split ordering is derived differently: most particles in particle verbs that allow the V–NP–Prt ordering are homonymous to prepositions, and it has been argued that the particle and the preposition come from the same root (Basilico 2008:732, Oya 2009:288; see Lüdeling 1999 for a line of argument contrary to this view). Basilico deduces from this that the particle has some of the preposition’s inherent features, which means that the abstract particle verb form can select a ‘prepositional’ object before it is assigned its verbal category feature. The object must later move out of this verbal construction to receive case, as shown in (13b), and the bare verb then moves further upwards to create the V–NP–Prt order.

(13)

(after Basilico 2008)



⁶Compare Pause (1999) and Taraldsen (2000) on such incorporation.

As above, this order can be forced by modifying the particle. This theory addresses the problems of incorporation analyses laid out in Lüdeling (1999:123–124), but it fails to explain for the vast majority of cases why one structure is chosen over the other.

Dehé (2002) tries to offer such an explanation. She argues that all factors which influence the object-particle ordering reduce to information structure (except in idiomatic orderings). Experiments on intonation in particle verb constructions support this, but studies based on much more data (see Wasow 2002, summarized in section 3.3 below, and Gries 2002) show that several independent factors indeed affect word order in English. In structural terms, Basilico (2008)’s theory is quite similar to this information structure account in that both assign to particle verbs a word structure that is more complex than derivation or compounding and (somewhat) transparent to syntax (Dehé 2002:239–258).

Of course, there is no reason to assume that general theories of weight effects do not apply to particle verbs. Indeed, these theories have been found to be predictive for object-particle order (Lohse et al. 2004). Lohse and colleagues also found that weight is not the only factor: the medium of language (meaning speech versus writing) and semantic dependencies between particle and verb appear to affect the choice of object-particle ordering. All such factors could probably be included in an iteration of Farrell (2005:125–128)’s Optimality Theoretic account of object-particle order. This, however, would require an established and agreed theory of weight effects, which does not exist yet. As baseless assumptions abound, the next section summarizes and discusses some of the more promising approaches to a theory of weight effects. This discussion should minimize any theory-based bias in the studies below.

3 Theories of syntactic weight

Heavy NP shift, the dative alternation, particle verbs, and other constituent reordering phenomena exhibit weight effects. It is obvious that all of these weight effects will be slightly different: for example, the object of a particle verb can hardly be said to be newer or more accessible in context than the particle. Certain factors, however, appear to affect all phenomena with weight effects. Therefore, a general theory of weight effects is more parsimonious and should be the aim of all investigations into this subject. However, although the phenomenon of weight-licensed reordering has been examined repeatedly, no universally agreed definition of ‘weight’ has emerged. There have been several different theories, but they disagree even on fundamental points (see Zec and Inkelas 1990). The following three sections summarize and evaluate some of these theories. The theories are divided into those that define which feature makes a single constituent heavy, those that proscribe an order of constituents, and finally the multifactorial approach.

3.1 Weight in isolation

Ross (1967:32) discusses heavy NP shift under the name “Complex NP Shift” and finds that the shifted NP always dominates an S node. This structural definition of weight was later expanded to NPs dominating “a phrase or clause” (Wasow 2002:16). Phonological studies have been used to support such definitions: according to Zec and Inkelas (1990)’s Phonological Phrase Algorithm, an NP that dominates a PP contains two phonological phrases and is thus *phonologically* heavy as well. They also note that shifted NPs are marked by other “special intonational effects” (Zec and Inkelas 1990:377).

However, Stallings et al. (1998:411) and Stallings and MacDonald (2011) argue that phonological features are not accessible to the human language faculty at the time when word order decisions are made. This, in turn, has been countered by the argument that heavy NP shift (and presumably other weight-related orderings) is entirely phonological (Shiobara 2008, Lozano and Mendikoetxea 2010). As discussed in section 2, theories which assume this needlessly limit their descriptive and predictive power. While it may be interesting to study if a one-word constituent with a large number of syllables is syntactically heavy, as Akasaka and Tateishi (2001) claim, a purely phonological explanation is therefore unacceptable and cannot be used in any approaches to syntactic weight.

However, support for these structural definitions of weight comes from a more recent corpus study (Wasow 2002:26), which finds that Ross' definition is a relatively good sufficient condition for heavy NP shift, while the expanded structural one might make a better necessary condition. However, it is obvious that none of them is a perfect predictor itself, which explains that further definitions of weight had to be posited.

Another field of linguistics that investigated these essentially syntactic weight effects is pragmatics. It has been proposed repeatedly that “‘given’ material is better when placed early in the sentence and ‘new’ material is better at the end” (Gropen et al. 1989:225). One prominent elaboration of this intuition sees focus as the basis of heavy NP shift and other reorderings: a phrase that occurs in a position towards the end of the clause instead of in its canonical position supposedly has to constitute new information in the context of that clause (Rochemont and Culicover 1990:20–25). While ‘topic’ or ‘new information’ appear to be scalar concepts (Givón 1988, Nomoto 2009:247) and would thus lend themselves to an investigation of gradient phenomena,

they are never rigidly defined (Givón 1988, Hawkins 1994:185) and thus lend themselves to free and imprecise interpretation as well. This may or may not be rectified eventually; but what ‘newness’ by definition cannot explain are the strong ordering preferences found in completely context-less sentences like (14).

- (14) a. Give me that!
b. ? Give that me! (after Ross 1967:36)

Furthermore, there are cases where pragmatic theories like that of Rochemont and Culicover (1990) do make predictions, but speaker judgments clearly show them to be wrong (Hawkins 1994:186).

Of course, relying only on ever-changing speaker data cannot result in perfectly stable theories. There is, however, some evidence that the human language faculty does take previous input (and thus speaker data) into account when considering weight-related reordering. While they fully accept that the length of a constituent is important, the verb disposition hypothesis of Stallings et al. (1998:396) adds the particular verb’s ratio of noncanonical to canonical orderings in previous occurrences into the ordering decision. According to this hypothesis, a verb that is accompanied by the noncanonical ordering more often than other verbs will in future be produced in that ordering more often. The past input thus affects the future output. However, apart from the general problems of frequentist approaches to language cognition, even the a priori assumptions of Stallings et al. (1998) are at odds with speaker data (Wasow 2002:54). Thus, Stallings et al. (1998)’s theory, like those above, is flawed in several regards.

3.2 Relative weight

According to Cooper and Ross (1975), the idea that two elements are ordered according to their relative lengths goes back more than 2000 years. In their discussion of phonological effects on ordering in coordinations, they credit it to Pāṇini. In their formulation, Pāṇini’s law or principle states that “other factors being nearly equal, place 1 elements contain fewer syllables than place 2 elements” (Cooper and Ross 1975:78). Similarly, Scaglione (1979:234) traces weight-related ordering rules to the year 1718. Interestingly, this early rule already refers to the length as well as the modifications and therefore the structure of the elements to be ordered. These early relative ideas of weight were also applied to coordination constructions, and that in turn was used to show how considerations of weight could overrule considerations of emphasis (Cooper and Ross 1975, Scaglione 1979), which Scaglione’s example of Shakespeare’s “Friends, Romans, Countrymen” illustrates well. Although I am not discussing coordination here, these early generalizations and arguments for principles having different ‘strengths’ is important as they surface again in later applications of Pāṇini’s law.

Such applications can be found, for example, in Otto Behaghel’s grammar of German. Like Cooper and Ross (1975), he lists several phonological and semantic criteria for the ordering in coordinations (“Erweiterungsgruppen”). The second of these is length,⁷ and it applies to both words and phrases (Behaghel 1928:367–368). While it has been pointed out that the two phenomena are fundamentally different (Stallings and MacDonald 2011), it is important to note that this very same principle is also applied in Behaghel’s discussion of German adverbials. If there are no strong semantic or pragmatic ties be-

⁷Behaghel apparently does not intend to rank these criteria by strength; however, it surely is no accident that length comes before the vaguer etymological criterion.

tween two adverbials, they are ordered such that “the required ones, that is the most important ones, are at the end, the least required ones . . . at the beginning” (Behaghel 1932:228; own translation). Tempting though it may be to regard this as an early formulation of linguistic gradience, Behaghel never adequately defines the scale of ‘required-ness’—a recurring problem with pragmatic effects in syntax, as will be seen shortly.

Interestingly, Behaghel was also aware that ordering principles could clash: should the above pragmatic principle prefer a different ordering than Pāṇini’s extended principle, which Behaghel (1932:167) calls “Gesetz der wachsenden Glieder” (‘law of increasing constituents’), the long first constituent may split up so that both principles can be satisfied. Furthermore, Behaghel also applies Pāṇini’s extended principle to postverbal constituents: together with pragmatic factors and special accent patterns, the relative lengths of a dative object and an accusative one can force the reversal of the standard (according to Behaghel) order DAT–ACC.

Later theories expanded on this by prescribing an acceptable ordering of sentence constituents according to their syntactic category and function. One example of this is the “Output Condition on Post-Verbal Constituents” (Ross 1967:36), which says that direct object pronouns are to come before indirect object pronouns, which in turn are to come before proper names, and so on. Ross was well aware that a condition making special provision for “*company* in *keep company*” (Ross 1967:36) is rather eclectic. Of course, later work improved on this; but it did also take up some ordering relations of the Output Condition on Post-Verbal Constituents without change.

Two concerns of later ordering rules are motivation and generalization. The most recent example of an explicit ordering rule is perhaps Functional Gram-

mar, where two principles together proscribe a constituent order based on syntactic features: the principle of head proximity says that orderings should “keep the heads of different domains as close together as possible” (Dik 1989:343), and the principle of increasing complexity maintains an order where less complex constituents precede more complex ones. As the functional head of an English phrase is typically the first element of the phrase and complex constituents are typically longer than simple ones, these two principles together produce a strong preference for an ordering where the earlier constituents are the shorter one. Such orderings satisfy both principles. He supports this by the above restriction on modified particles, because the early short constituents minimize the distance to the heads of the later constituents.

From this interaction, Dik deduces the Language-Independent Preferred Order of Constituents (LIPOC). The LIPOC refers to syntactic categories alone, saying that clitics come before pronouns, pronouns before NPs, and so on (Dik 1989:351). However, it does not include particles at all. The complete absence of any grammatical function even makes the LIPOC greatly less predictive than previous ordering rules: there are many possible sentence pairs with strong preferences like those in (14) where the LIPOC can by definition not make a prediction. Of course, establishing a default ordering based on case (as in section 2.2, for example) would rectify that; but surely the desideratum must be one unified theory that combines the categorical power of the LIPOC or default case orderings with the more flexible predictive accuracy of gradient theories. It must therefore be noted that the LIPOC also suffers from some problems. Nevertheless, even though the details are less than perfect, the shift of focus onto connected principles that Functional Grammar exemplifies did lead to further progress.

One major feature of the LIPOC is that it works reasonably well in English and other languages with head-initial phrases. If the LIPOC is really language-independent, however, it fundamentally disagrees with the principle of head proximity in head-final languages: if the heads of all phrases are the last elements, then the principle of head proximity suggests an order of *decreasing* length. This would disagree completely with what the LIPOC proscribes, unless all head-final languages have particles which are longer than pronouns, pronouns which are longer than full NPs, and so on. In other words, length and complexity would have to correlate negatively in all head-final languages, which is obviously false. It will be seen below that the LIPOC provides a reasonably good description of constituent ordering *in English*, but I make no claims about its actual independence here. This problem is arguably due to the development of Functional Grammar and indeed linguistic theory as a whole. In the words of Haider (2009), it is an example of Orwell’s problem of modern linguistics—some languages are more central to the study of all languages than others.

This problem is addressed in the work of John Hawkins. He sees weight effects are as effects of the Early Immediate Constituents (EIC) principle, which says that the human language parser prefers orderings with a high ratio of immediate constituent (IC) heads to all words between the mother node and the last IC head. In effect, this means that the parser tries to keep a mother node and the heads of its ICs as close together as possible (Hawkins 1994:78–79). This is essentially the same as Dik’s principle of head proximity, but Hawkins’ principle also deals with the problem of head-final languages: he simply leaves out the tried and true principle of increasing weight, as it is tried and true for Germanic languages only. The EIC principle thus simply calls for IC heads to be as close together as possible. This works out

to an order of decreasing complexity or length in head-final languages and the well-known order of increasing complexity in head-initial languages, as Hawkins (1990) and Hawkins (1994) demonstrate with examples from various languages. Gibson (1998) and Nomoto (2009) also elaborate on this principle and extend it to other reordering phenomena. Furthermore, Hawkins (2006) addresses problems caused by competing principles and concludes that constructing all syntactic constituents of a node (thus the EIC principle) takes precedence over disambiguating between several possible lexical properties. To summarize, the simple length-based EIC principle captures the order of increasing constituents perfectly and extends the same basic idea to constituent ordering in head-final languages.

While this generalization to one principle is quite insightful, the EIC principle by definition cannot make predictions for cases where both ICs are of equal length, such as (14). Of course, that is only a simple example which could easily be explained away within any theory. However, actual speaker data contain counter-examples to *all* definitions of weight discussed above. Unless a revolutionary monofactorial definition is found, it must be assumed that several factors influence weight (Wasow 2002:20–23). The following section describes how Wasow tries to accomplish such a multifactorial theory.

3.3 The multifactorial approach

As mentioned in section 3.1 above, Wasow (2002) recognizes the power of earlier definitions. He also finds that relative weights correlate slightly (but not notably) better with shifting behavior than categorical definitions of weight do. His experiments and corpus studies do support the definition of weight by relative lengths in words or nodes quite well. Similar support comes from the

recall-production studies of Stallings and MacDonald (2011), who find that a heavy constituent is generally longer in recalled production than the other constituents in the clause. Despite this empirical support, Wasow disagrees with Hawkins' EIC on the grounds that it presumes complete spell-out before production begins—something that is supported neither by speaker data nor by common sense (Wasow 2002:45–46; see also Chang 2009:377). However, there is a readily available solution to this problem of spell-out: Sternberg et al. (1978) found that the length in the speech *plan* rather than in the actual speech can have effects. With this simple amendment, the theoretical foundations of the EIC principle become rather more acceptable.

Apart from syntactic weight, Wasow (2002)'s work shows that a constituent's newness in context can also cause reordering. While he disagrees with Stallings et al. (1998), his investigation of idiomatic expressions and verb bias for or against shifting reveals that both have an effect on such reordering as well. He concludes that his (arguably quite comprehensive) studies address “only small bits of the big picture” (Wasow 2002:109) and that many factors, including length in words and newness of constituents, can influence word and constituent order (see also Stallings and MacDonald 2011:178).

Wasow also uses his results to investigate the reason for weight effects. The most likely one he finds is what he calls “late commitment” (Wasow 2002:49) of the speaker to a certain structure. Supported by corpus data, he argues that the language faculty of speakers generally favors putting in first place those constituents that are found in all subcategorization frames of that verb. Consequently, the constituents whose features are not found in all frames (and are probably unique to just one of them) are placed in later positions. Therefore, the speaker does not commit to one frame early on. Ambiguity avoidance does not appear to be a major influence on word order decisions,

according to Wasow. Information structure does, however, provide provide stimulating facts to this discussion of the causes of weight effects.

Interestingly, one of Wasow's corpus studies suggests that shorter phrases precede longer ones even when that order violates 'given before new', while an experiment shows that given constituents precede new ones even when that order violates 'short before long' (see Arnold et al. 2000). Although these apparently contradictory results are puzzling, they do show that effects of information structure are distinct from weight effects. Further support for this comes from crosslinguistic comparisons: as discussed in section 2.1, heavy phrases seem to go before lighter ones in Korean (Wasow 2002:82) and Japanese (Stallings and MacDonald 2011:185), but new constituents still go after given ones. This, according to Stallings and MacDonald (2011)'s reasoning, is due to the fact that Japanese word order is more free and therefore more defined by information structure. Consequently, constituent length does not determine accessibility and processability in Japanese to the degree that it does in English. Stallings and MacDonald's explanation for weight effects thus tries to account for a variety of languages, and it leads them to posit processing demands as the underlying cause of these weight-licensed optional reorderings. Crosslinguistic evidence not connected to this research supports the independence of semantic effects from syntactic weight effects and therefore the multifactorial approach in general (Rosenbach 2005). Further experimental work on weight effects in Japanese, for example, would certainly provide further insights and advance theories regarding the cause or trigger of weight effects. Indeed, Wasow recommends the effects of information structure on constituent ordering for further research, and I intend to follow his recommendation.

4 Corpus study

Of course, the weight effect phenomena must be acquired somehow. A careful examination of first language acquisition should show how and when weight effects begin to occur, and possibly in what sequence the different factors assert themselves. I chose to do a corpus study for reasons of practicality and coverage: not only is it easier to analyze existing data than to collect new data, it also allows for a far greater number of informants. This, naturally, strengthens any conclusions that are drawn.

4.1 Corpus

The CHILDES project offers child speech corpora from many languages and research backgrounds. Using data from monolingual English-speaking children, I extracted possible constructions with two postverbal constituents using the grammatical role annotations on the *CHI %GRA tier.⁸ Of course, this means I limited myself to corpora where this tier had been annotated. Of the British corpora, these are ‘Belfast’ (Henry 1995, Wilson and Henry 1998), ‘Cruttenden’ (Cruttenden 1978), ‘Fletcher’ (Johnson 1986, Fletcher and Garman 1988), ‘Gathburn’ (Gathercole 1986), ‘Howe’ (Howe 1981), ‘Manchester’ (Theakston et al. 2001), ‘Wells’ (Wells 1981), and ‘Wooten’, a sub-corpus of the Cornell corpus (Hayes 1986). When it emerged that these would yield roughly 5000 items, I selected a sample of the US corpora that would provide roughly as many items (for reasons of comparability). These are ‘Bates’ (Carlson-Luden 1979, Bates et al. 1988), ‘Bernstein’ (Bernstein 1982), ‘Bliss’ (Bliss 1988), ‘Bloom70’ (Bloom 1970, Bloom et al. 1974, 1975), ‘Bloom73’

⁸The actual search string was this:
(*OBJ*+LOC+PRED)~*~!ROOT~!AUX~*~(*OBJ*+LOC+PRED)

(Bloom 1973), ‘Bohannon’ (Bohannon and Marquis 1977, Stine and Bohannon 1983), ‘Brown’ (Brown 1973), ‘Carterette’ (Jones and Carterette 1963, Carterette and Jones 1974), and ‘Clark’ (Clark 1982). As there are some minor differences in annotation practice between all these corpora, the search also returned a number of items without two postverbal constituents. These were removed manually; a method that is inelegant but pays off in the long term (Behrens 2008:xxii).

The remaining items were then annotated for individual factors (age and sex of the children as well as country of recording) and conformance to LIPOC, certain verbs,⁹ length of postverbal constituents in words (see Wasow 2002:32), immediate constituent-to-words ratios (hereafter ‘EIC ratios’) of the observed and the alternative ordering, whether postmodification or coordination was present in either constituent, and whether both constituents were new in context (as far as it could be inferred from the corpus data). I readily concede that the last point is somewhat spurious: not all of the above corpora perfectly recorded the entire context of all utterances; the two ‘newness’ factors were annotated by hand and are therefore subject to myriad biases and inconsistencies; and as discussed in section 3.1, the concept of ‘newness’ is not scientifically rigid. Nevertheless, the theory based on context-new and -old constituents is too important to be entirely left out, and the results below will prove to be interesting in relation to this theory. Finally, the constituent length difference¹⁰ and the EIC ratio difference¹¹ were also calculated for each item.

⁹The particular verbs selected were these: *bring, build, buy, call, get, give, keep, leave, push, put, show, take, tell, and turn*. *Draw, find, move, set, and write* were also annotated, but were dropped from analysis because they each yielded less than 50 items.

¹⁰Length of second constituent – length of first constituent, as it was expected that this would have a positive result for most items.

¹¹Observed EIC ratio – alternative EIC ratio, as it was expected that this would have a positive result for most items.

4.2 Results and discussion

There are some statistically significant ($p < .05$) correlations between variables, but all but two of these had an insignificant correlation coefficient ($-.05 < r < .05$) and will therefore not be discussed. The two strong correlations were between length of the first constituent and the EIC ratio of the observed ordering ($p < .001$, $r < -.99$), and between second constituent length and EIC ratio of the alternative ordering ($p < .001$, $r < -.96$). As the EIC ratios are calculated from the corresponding constituent lengths, these correlations are uninformative too. There were further unconnected correlations among the items with *buy* or *leave*, but as these subsets were quite small ($n = 123$ and $n = 95$, respectively), these correlations are not particularly informative either. For *get* ($n = 880$) and *put* ($n = 3474$), the observed and alternative EIC ratios correlated significantly ($p < .001$, $r \approx 0.13$ for *get*; $p < .001$, $r \approx 0.14$ for *put*). However, this merely indicates that the second constituent tended to be longer than the first in the items containing these two verbs, which is to be expected considering previous work (Wasow 2002:60).

	correct	incorrect	no prediction	sum
LIPOC	7282	45	3257	10584
	68.8%	0.4%	30.8%	
‘old before new’	2709	1058	6817	10584
	25.6%	10.0%	64.4%	
structural weight	586	117	9881	10584
	5.5%	1.1%	93.4%	
EIC	6549	766	3269	10584
	61.9%	7.2%	30.9%	

Table 1: Number of items that are correct, incorrect, and unpredictable according to four ordering theories, with percentages within each theory

Table 1 shows how the LIPOC, pragmatic ordering according to Gropen et al. (1989), the structural definition of heavy constituents as being coordinated or postmodified, and the EIC principle agree with the CHILDES data. For each of these four theories, the corpus contains more items that conform to the theory than items that violate it. However, it is also clear that all four theories fail to make any prediction for a large number of items. In most of these items, the two constituents are of the same type according to the theory in question—for example, both are pronouns, or both are old, or neither is coordinated or postmodified, or both are of the same length.

Of course, the pragmatic ordering theory and the structural definition can only make sense when two constituents are different pragmatically or structurally. As these data show, this is a severe limitation on their predictive power. If the LIPOC and the EIC principle were modified to include concepts like dative and accusative objects and their relative ordering, the data would agree with both of them very well. However, this would make them similar to Ross (1967)'s Output Condition on Post-Verbal Constituents—in other words, eclectic and highly language-specific. If the two theories were to include a last-resort 'anything goes' rule for otherwise equal constituents, the data would agree with them very well too. However, such a rule would be no more than a crutch: admitting that a theory cannot explain certain things does not in itself improve the theory. Nevertheless, even without such a rule, the LIPOC and the EIC principle are much better supported by the data than the other two theories. Thus, the unmodified LIPOC and EIC theories are the best of the theories examined here.

4.3 Statistically modelling the dative alternation

While the above results are quite informative on their own, comparability to other research is a major consideration. Statistical methods have always been common in corpus linguistics, and the same is true for statistical modelling (see Hinrichs and Szmezcany 2007:459 on the benefits). However, the corpus extraction method described above yielded many different types of constructions with two postverbal constituents. As the results so far are not perfectly insightful with regards to the factors that affect constituent ordering, I must for this statistical model abandon the goal of a general description of weight effects and focus on one well-described phenomenon that exhibits weight effects. To achieve meaningful and comparable results, I extracted all dative alternation items from the corpus data described above and fitted a statistical model to that subset. The English dative alternation is limited to a relatively well-defined group of verbs (Whong-Barr and Schwartz 2002:582–583), which makes the dative alternation items easy to extract from any corpus. I took all items with dative alternation verbs from the above data set and manually removed the items where the alternation was not actually possible (see table 2 for a description of these data). Interestingly, no item in this data set violated the LIPOC—in other words, all the prepositional constructions observed the canonical $V-NP_{theme}-PP_{goal}$ order.

Figure 1 shows the percentage of double object dative constructions for verbs with more than 40 occurrences in this dative alternation subset. While they are generally similar to Wasow (2002:87)’s results based on the New York Times Corpus, there are two striking differences: firstly, *give* and *tell* are somewhat more likely to be used in a prepositional construction in my data. It has, however, been claimed that children slightly prefer the prepositional

	<i>bring</i>	<i>build</i>	<i>buy</i>	<i>get</i>	<i>give</i>	<i>pass</i>
prepositional	11	0	1	11	196	2
double object	32	10	78	98	619	1

	<i>read</i>	<i>send</i>	<i>show</i>	<i>take</i>	<i>tell</i>	<i>throw</i>
prepositional	23	8	29	14	3	12
double object	36	2	94	10	72	3

Table 2: Items in the dative alternation data set, by verb

construction (Conwell and Demuth 2007). Therefore, this difference is not surprising. Secondly, my data show a much higher percentage of double object constructions for *bring*. This is due to different a priori assumptions: Wasow excluded all items with pronominal theme arguments, as they absolutely “require the prepositional construction” (2002:87); I did not. This stylistic requirement may be perfectly true for the texts in the New York Times Corpus, but children do not limit their speech in this way: for example, my relatively small dative alternation corpus contains eleven items with the string *bring me it* in them alone. Therefore, these two differences are due to the differences in the corpora and not very informative. In light of these fundamental differences between the data sets, the similarities between Wasow’s and my data for lexical bias in dative alternation are more interesting.

There is an argument that speakers do indeed have information about frequencies of particular verbs in particular constructions, and that this information affects language production (Stallings et al. 1998, Wasow 2002). The results presented here suggest that children have acquired this information and are using it by the time complex sentences with several possible orderings arise in their speech. Of course, their production is not fully adult-like.

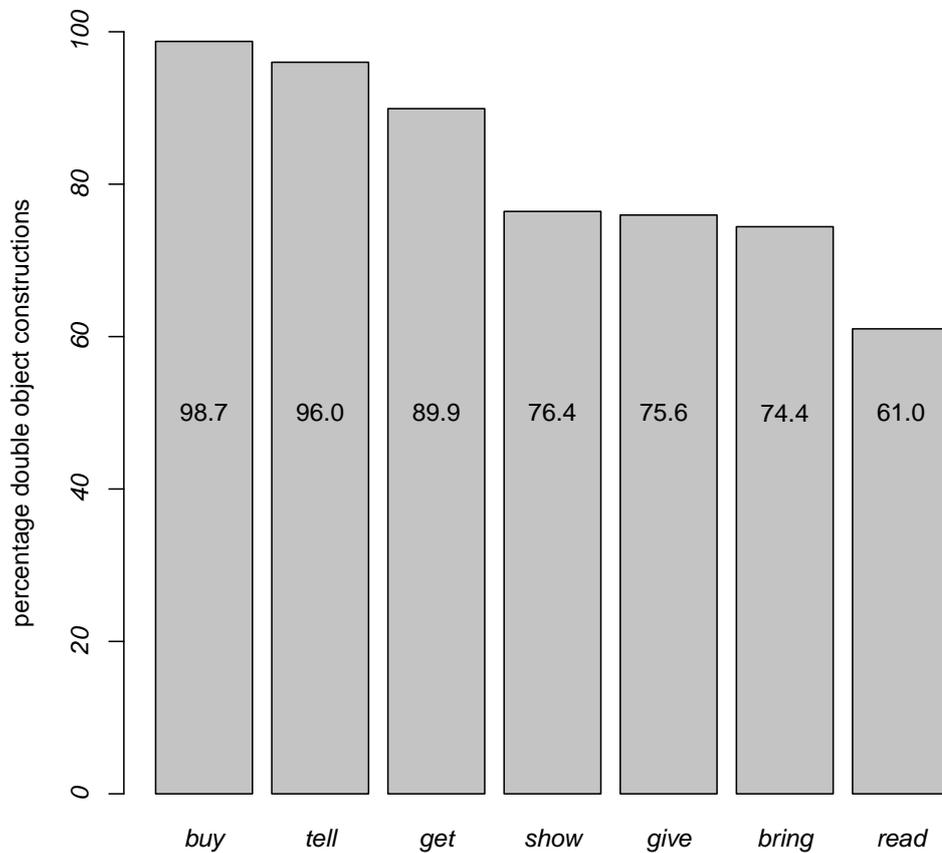


Figure 1: Percentages of double object construction in the dative alternation data, by verb

A study of dative alternation in adults' colloquial spoken English would undoubtedly show key differences not only to this study of children's speech data, but also to studies on more formal newspaper corpora. For the purposes of this study, even this simple comparison shows that child and adult language do not differ greatly in the phenomenon of dative alternation.

4.3.1 The model

For results that are both quantifiable and comparable, I also fitted a regression model to this subset of data. The dative alternation in English has been studied extensively (see for example Conwell and Demuth 2007, Rappaport Hovav and Levin 2008, Gries 2009, Bruening 2010), and statistical models have been devised using dative alternation corpus data from adult speakers (see for example Bresnan et al. 2007). I recoded the ‘newness’ variables so that they referred to theme and goal arguments instead of first and second postverbal constituents, which was straightforward thanks to the strict observance of the canonical constituent orderings. Then, I simply annotated the items in the subset for prepositional or double object construction. The dependent variable I used was ‘prepositionality’, such that a prepositional item received the value 1 and a double object item the value 0. This means that positive correlation coefficients in the following model mean increased likelihood of the prepositional dative construction.¹² There were 1055 double object items and 310 prepositional items in this data set.

A generalized linear model with the logit link function proved to be the most informative with an Akaike information criterion of 1068.524, a value which various other models did not approach. Table 3 shows the relevant coefficients (rounded to two decimal places) and p-values.

The effects of other verbs do not reach significance. *Send* (marked in gray in table 3) is very rare in the data set and therefore adds little predictive power to the model anyway. A model based on a data set where *send* and other rare

¹²Although it has been claimed that successful transfer goes with the double object construction, I did not code whether the intended action was successful. The context would not provide the necessary information in many cases, and (more importantly) the argument against success-based alternation in Levin and Rappaport Hovav (2002) applies.

effect	coefficient	standard error	p	Wald Z
INTERCEPT	-3.58	0.55	< .001	-6.54
second constituent coordinated/postmodified	-4.42	0.95	< .001	-4.66
difference in EIC ratios	6.25	0.95	< .001	6.60
length of first constituent (in words)	2.16	0.28	< .001	7.60
theme new	-1.87	0.21	< .001	-9.10
goal new	2.09	0.32	< .001	6.61
<i>buy</i>	-3.79	1.16	.001	-3.27
<i>throw</i>	2.54	0.85	< .01	2.98
<i>get</i>	-1.61	0.54	< .01	-2.96
<i>tell</i>	-1.89	0.72	< .01	-2.61
<i>send</i>	2.47	1.16	< .05	2.14

Table 3: Coefficients of the best model

(in my dataset) verbs have not been annotated was found to be very similar to this one, but of course more limited and therefore less accurate. Therefore, the above model emerges as the best fit for my data. It is, however, far from perfectly predictive with pseudo- $R^2 = .28$.

4.3.2 Validation

To estimate how well this model generalizes, it was validated using the tenfold cross-validation procedure. The cross-validation estimate of the model's prediction error is 0.1277, the adjusted estimate is 0.1274.¹³

Figure 2 shows diagnostic plots for the model. They give, from left to right and top to bottom, the jackknife deviance residuals against the fitted values, a Q-Q plot of the deviance residuals, the Cook statistic against the leverages,

¹³The estimates of five- and 103-fold cross-validation were also computed, but they do not differ notably from the tenfold estimates: $\Delta_5 = 0.1291$, $\Delta_5^{adjusted} = 0.1284$; $\Delta_{103} = 0.1280$, $\Delta_{103}^{adjusted} = 0.1279$.

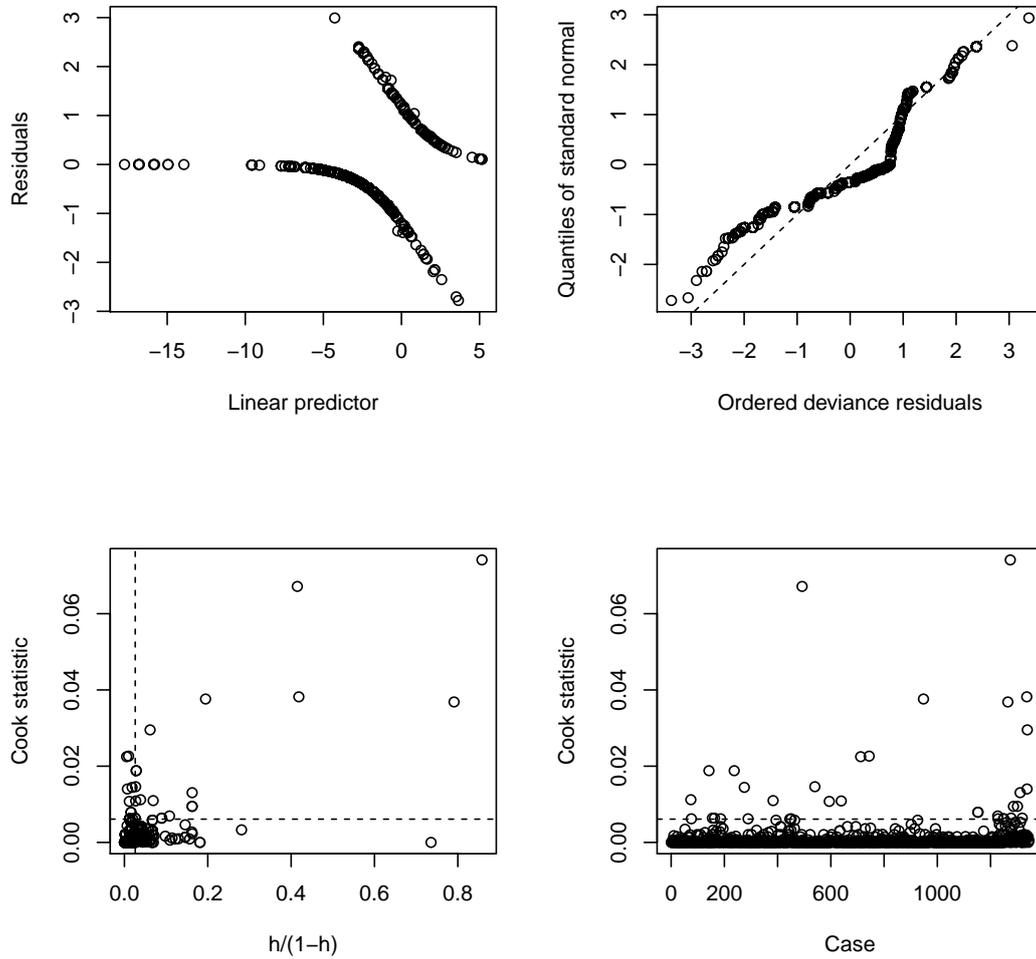


Figure 2: Diagnostics plot for the model discussed here

and the Cook statistic against consecutive item numbers. The bottom right plot shows the items with a possibly unproportionately large effect on the data quite well. Inspection of the 15 most outlying ones of these points shows that some of them have a greatly higher-than-average age or second constituent length, but none of them approaches the maximum values for these variables. 13 of these 15 items are from male children, but the speaker

sex does not enter into the model presented above. Therefore, these supposed outliers are likely artifacts in the statistic.

Table 4 compares the model’s predictions based on the data set to the observed constructions in the same data set. The number of items is slightly lower here than reported above as the statistical model cannot make predictions in cases with missing variables, and there were 21 of those in the data. As the double object construction was coded as 0 and the prepositional construction as 1, I regarded prediction values of .5 and higher as predicting the prepositional construction. As is obvious from table 4, the model’s prediction agrees with the observed data in 81.6% of cases. Of course, this percentage is relatively low; however, based on the variables I used, the present model is simply the best possible model. Future work based on more variables, for example those used in Bresnan et al. (2007), would undoubtedly achieve greater accuracy. In conclusion, the model presented above is acceptable, and the next section discusses the effects in it.

	double object observed	prepositional observed	sum
double object predicted	73.5% (988)	14.4% (193)	1181
prepositional predicted	4.0% (54)	8.1% (109)	163
sum	1042	302	1344

Table 4: Accuracy of the model, with number of items shown in gray

4.3.3 Discussion

If the second constituent is coordinated or postmodified, it is more likely to be in a double object construction; if it is not coordinated or postmodified, it is more likely to be in a prepositional construction. This is hardly surpris-

ing, as the second constituent in the double object structure is an NP and this structural definition of weight was arguably based on and intended for NPs rather than PPs. This means that a correlation between structurally heavy objects and the double object construction is only to be expected in this particular data set. It must also be noted, however, that coordination or postmodification is quite rare in my data (see table 1 on page 36): of the 1365 dative alternation items, only 74 have a coordinated or postmodified second constituent (and a mere two have a coordinated or postmodified first constituent). Therefore, the structural criterion of weight is statistically significant in the model, but only skews the prediction slightly towards a prepositional construction for most of my and arguably any other spoken language data. It is not very informative by itself.

A long first postverbal constituent increases the likelihood of a prepositional construction. While the length of the first constituent correlates with the length of the second constituent in the dative alternation data (Pearson's $r = 0.077$, $p < .01$), the latter is affected somewhat more by outliers and thus naturally does not correlate well with other variables. Interestingly, the two lengths correlate weakly but significantly in the subset of prepositional construction items (Pearson's $r = 0.144$, $p < .01$), but not in the subset of double object items ($p > .1$). Of course, the average PP contains an (average) NP and must therefore be longer than the average NP. 269 of the 310 prepositional construction items (86.8%) do indeed show this pattern. However, it also holds true for 699 of the 1055 double object items (66.3%). It is therefore likely that this effect is indeed caused by the length of the first constituent and not by any hidden interaction, and a Wilcoxon rank-sum test shows that the mean of first constituent lengths in the prepositional construction subset differs from that of the double object subset ($W = 182984$, $p < .0001$). Thus,

the length in words of the first postverbal constituent does indeed appear to affect the choice of construction.

Theme arguments that are new in the context favor the double object construction, while new goal arguments favor the prepositional dative. In the canonical orderings, the theme is the second postverbal constituent in double object constructions, just as the goal is the second constituent in prepositional constructions. While they differ somewhat in their strength, these two effects therefore appear to stem from a common source: new material is ‘heavy’, that is to say it is more often placed near the end of a sentence. While table 3 clearly shows that these effects (or the one more basic effect) are statistically significant, the criticisms of the concept of ‘newness’ or ‘topic’ mentioned in section 3.1 apply here as well.¹⁴ The concept of ‘newness’ has not been defined with ultimately acceptable rigor and probably is too vague inherently to ever be defined. Its effects are therefore accepted as part of the model, but strongly recommended for further experimentation and theorizing. Crucially, table 3 also shows that the correlation coefficients of these effects are not as large as others. The pragmatics of an utterance certainly play a role in shaping it, but this study did by far not prove any effect of context-new arguments on syntactic decisions.

A much more informative factor is the EIC ratio difference. Recall that it was calculated by subtracting the EIC ratio of an item’s possible reordering from the ratio of its observed ordering. A positive ratio difference therefore means that Hawkins’ theory would support the observed ordering over the alternative. The model shows that a positive ratio difference (present in 968 of the 1365 dative alternation items, or 70.9%) means that the item

¹⁴Having another rater annotate the data for ‘newness’ and subsequently calculating measures of inter-rater agreement might be interesting, but is obviously impossible here.

is much more likely to be a prepositional construction. Again, PPs tend to be longer than NPs, and prepositional dative items may therefore tend to have a positive ratio difference. However, recall that there are merely 310 prepositional construction items in the dative alternation data. With its relatively high correlation value and Z statistic, the EIC ratio difference effect therefore must also affect those items that surface as double object constructions.

According to the model, *throw* and *send* go with the prepositional construction and *buy*, *get*, and *tell* go with the double object construction. The Z statistics show that these four effects are comparable in strength. As discussed above (see figure 1), these results are similar to previous research. For example, the model connects *tell* to the double object construction despite children’s preference for the prepositional dative mentioned above. The effects of *throw* and *send* are among the marginally significant ones, and they are based on only a few occurrences of each verb (as mentioned above). Nevertheless, all five of the verb effects are significant. Explaining their effects is straightforward.

Buy, *get*, and *tell* are the three verbs with the highest percentages of double object items in my data, as figure 1 shows. Similarly, 80% of occurrences of *throw* (12 of 15 total) and *send* (8 of 10) are prepositional constructions. A semantic analysis along the lines of Gropen et al. (1989:240–242) explains these effects quite well: the non-theme arguments of *buy*, *get*, and *tell* are generally possessors or benefactives, whereas *throw* and *send* take a goal. As argued in section 2.2, it can be assumed that possession is in the thematic core of the double object construction, while the prepositional construction encodes motion towards a goal. These statistical effects would then connect Wasow’s idea of verbs’ lexical bias with these semantic arguments about

the dative alternation, and the data support both of them. These effects therefore show the underlying ‘lexical bias’ for one construction over the other that certain verbs appear to have.

Finally, some of the factors that are not significant in the model also deserve mention here. As argued above, the simple structural weight of the first postverbal constituent and the raw length of the second are not informative because they are so rare and erratic in the data. Similarly, the child’s sex would not be expected to affect a phenomenon as abstract as the dative alternation.¹⁵ It is however quite interesting that the child’s age does not affect the choice of construction—after all, Gropen et al.(1989) argued that children do not systematically acquire one before the other. Admittedly, there is some overlap between their corpus and mine, but I drew on considerably more (and newer) CHILDES data. Therefore, their conclusion and mine are sufficiently independent from each other and thus both valid.

Bresnan et al. (2007) also investigated the dative alternation by means of statistical models. Their corpus is much larger and arguably also more homogeneous than mine, and they also annotate many more variables. This allows them to address the problem of correlations between variables and to develop a strongly predictive model. Therefore, comparing my model against their model A is a valid method to assess the quality of the former. Interestingly, this comparison shows that both models contain similar effects of ‘newness’: generally speaking, constituents are ordered so that new ones occur last; specifically, new goal arguments correlate with the prepositional construction, which canonically has the goal last, and new themes correlate with the double object construction, which has the theme last. Furthermore,

¹⁵I emphasize that this does not prove the speaker’s sex to be completely irrelevant to dative alternation; this question would certainly need dedicated research.

both models show that longer constituents are more frequently placed after shorter ones, even though the length difference is computed in different ways. Finally, model A in Bresnan et al. (2007) does not address the lexical bias of verbs, but they show the other effects to be significant and strong across different verb meanings. Therefore, my model compares well to theirs: the strongest effects are present in both and can thus be seen as truly informative.

4.4 Summary

In the corpus study described here, I extracted all items with two postverbal constituents from a sizeable part of the CHILDES corpus and annotated them for four proposed measures of weight- or information-related ordering decisions. While it is true for all four that more corpus items fulfil them than violate them, this corpus study also revealed two of the theories, the simple ‘old before new’ and the structural definition of a heavy constituent as being postmodified or coordinated, to be very unproductive: due to the phrasing of these two measures, they are incapable of judging the grammaticality or acceptability of most of the items. The other two measures, the LIPOC and the EIC principle, fared much better. They also make no prediction in many cases, but they both manage to categorize the majority of items. As mentioned, most of these items do indeed conform to these measures. Thus, while they are not perfect, the LIPOC and the EIC principle are much better supported by the data presented here.

As these results are not perfectly comparable to other studies, I fitted a logit-linked regression model to the subset of corpus items that could undergo dative alternation. These items contained both double object and prepositional

constructions, and this study is absolutely uninformative as to how the dative alternation happens syntactically. While age did not affect the choice of construction as such, which previous research has already shown, the double object construction was more frequent in the data set. This may simply be a random effect, however: the study also showed that (some) verbs have a lexical bias for one construction over the other, and more than half of the data contain the verb *give*, which is biased against the prepositional dative. *Give* does not have a significant effect on the choice of construction according to the model, but *buy*, *throw*, *get*, and *tell* do. Their effects are in line with previous work on the lexical bias in dative alternation. The same is true of the effects of the first postverbal constituent's length and the difference in EIC ratios, which are both strong factors in the model. Whether the theme or goal arguments were new also proved to affect the choice of constituent, but the original 'newness' variables were annotated fully by hand and therefore make these the weakest effects to enter into the model. The particular data used here also produces further significant effects in the model, but I have argued that these are statistical artifacts and not true informative effects.

In conclusion, this corpus study achieves two things: firstly, it lends support to the idea of verbs having an inherent bias for certain structures over other other possible ones as well as to the finding that children do not acquire one of the dative alternation constructions before the other. Secondly, it proves the LIPOC and the EIC principle to be the most predictive theories of weight effects on constituent ordering in first language acquisition overall.

5 Questionnaire study

The above corpus study showed Hawkins' EIC principle and Dik's LIPOC to be most predictive for different phenomena with weight effects in the first language acquisition of English. Therefore, I designed a study to test whether these two factors are predictive in second language acquisition as well. As controlled results were desirable, this study uses simple binary grammaticality judgment questionnaire. The following sections will present and discuss the method and results.

5.1 Participants

Participants were recruited from various introduction-level linguistics classes at Konstanz in May and June of 2011. Obviously, introductory classes tend to be larger and thus maximize sample size while minimizing effort. However, this self-imposed limitation also had other benefits: the speakers whose data was analyzed arguably form a homogeneous age group (min = 18, max = 30; mean = 22.27, standard deviation = 2.14), and their experience with English as a second language can be expected to be quite homogeneous as well. The questionnaires were handed out in a randomized order at the beginning of a class session, and students were asked to fill them in as diligently as possible. Students who had filled one in before or were native speakers of English were asked not to participate. This yielded 159 completed questionnaires. To control any possible first language transfer effect, the analysis was limited to native speakers of German ($n = 132$).

5.2 Stimuli

The questionnaire items were declarative English sentences in simple past tense, designed in a 2x2 factorial design of the two factors LIPOC compliance and EIC ratio difference. In LIPOC-compliant sentences, the first postverbal constituent was to the left of the second lower on the LIPOC (for example, a pronoun before a full NP or an NP before a PP), whereas this order was reversed for noncompliant sentences. Sentences with a positive EIC ratio difference had a second postverbal constituent that was two words longer than the first postverbal constituent (resulting in an EIC ratio difference of $0.308\bar{3}$), and sentences with a negative EIC ratio difference had exactly the reverse (with a ratio difference of $-0.308\bar{3}$). The sentences were constructed around the verbs *bring*, *call*, *leave*, *put*, *take (to)*, *turn*, *draw* in the sense of ‘sketch’, and *move*. These verbs have various lexical biases and belong to different verb types, which means that the experiment should show any effects of these facets. All eight verbs are monosyllabic and morphologically simple, so there will be no systematic effects of phonological or morphological complexity. To illustrate, (15) shows the four stimulus sentences for *take*.

- (15) a. Liz brought her dog to the school reunion.
(NEITHER PRINCIPLE VIOLATED)
- b. Liz brought her trusty but old dog to the school.
(EIC VIOLATED)
- c. Liz brought to the school her trusty but old dog.
(LIPOC VIOLATED)
- d. Liz brought to the school reunion her dog. (BOTH VIOLATED)

Four different questionnaires were designed; each contained only one of the four sentences from each of these eight verbs and a total of two sentences for each factor combination. The sentences were interspersed with eight filler items randomly chosen from a list of twelve. The fillers were constructed around the same eight verbs, but crucially had either a clear adverbial (like *by accident*) or an argument of an idiomatic verb construction (like *bring to light*) as one of their two postverbal constituents. Every sentence was immediately followed by two circles, one for ‘grammatical’ and one for ‘ungrammatical’, and participants were instructed to mark the one they felt to be the appropriate judgment for that sentence. The order of items was pseudorandomized for each of the four questionnaires. A sample questionnaire can be found in the appendix.

5.3 Predictions

As the informative value for the four different factor combinations, I will use the percentages of judgments that rate the particular combination as grammatical. Although overt grammaticality judgments are obviously unnatural tasks (Stallings et al. 1998:405, Stefanowitsch 2006:73), they have yielded informative results for decades and, under controlled conditions, are not inherently worse than other method of judgments (Weskott and Fanselow 2009, Bader and Häussler 2010b). As this study relates to the matter of gradience in syntax, it is important to point out that this percentage does not imply a probability of grammaticality—probability may have a place in grammatical knowledge (Bader and Häussler 2010a), this study does in no way examine that question. The percentage of correct answers simply is a single number that captures the sample of binary judgments quite well and lends itself

to comparison between conditions. Therefore, the percentage of grammatical judgments per condition is expected to be a meaningful and informative result.

Obviously, the sentences that conform to both the LIPOC and Hawkins' theory are expected to be judged almost uniformly as grammatical. Similarly, the sentences that conform to neither are expected to get uniform ratings of ungrammaticality. As explained above, the questionnaire sentences are designed to receive a clear prediction from both theories, and section 4.2 showed that the two theories are similarly strong in predicting the ordering in cases where they can make predictions. Therefore, it is expected that all the sentences that violate only one theory will be judged as grammatical less often than those that conform to both, but more often than those that violate both. Due to the similar predictive strengths of the two theories, no difference between the sentences that violate only the LIPOC and those that violate only the EIC principle is expected.

5.4 Results

Figure 3 shows what percentage of participants judged each sentence to be grammatical. Each verb was only used in one set of stimulus sentences, and there were four factor combinations for each verb. Even though only two of the eight sentences that I predicted to be fully grammatical did in fact score a perfect 100% grammatical judgments, the deviation in the other sentences is not significant. The same is true for the sentences that violate only the EIC principle and those that violate both principles tested here. The most striking part of figure 3, however, is the lower left panel. It shows the percentage of grammatical judgments in the condition where only the

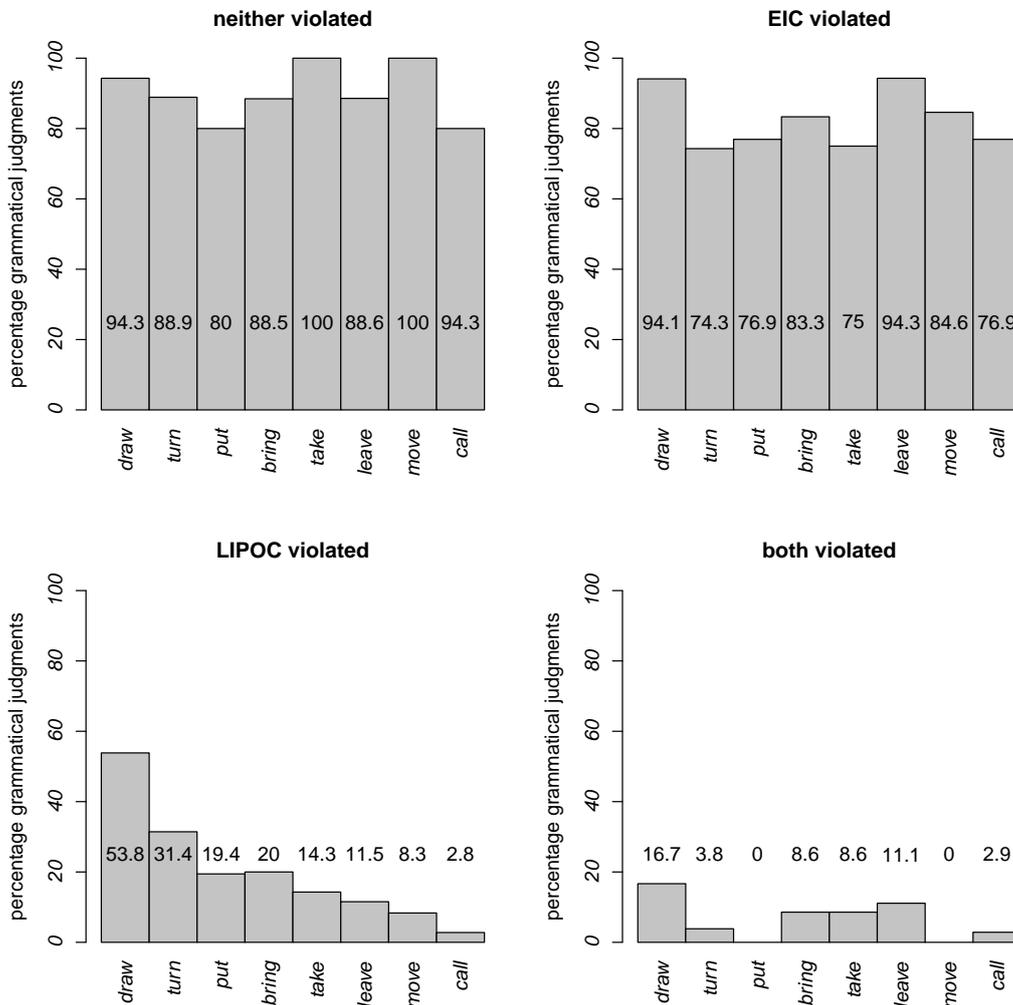


Figure 3: Percentages of “grammatical” judgments by factor combination and by verb

LIPOC was violated, and there appears to be a strong difference: the *draw* sentence was judged as grammatical by more than half of the participants who saw it, while the *call* sentence was judged as ungrammatical by almost all those who saw it. This is puzzling; but as the other sentences in that condition are somewhere between the two, no clear pattern emerges.

However, χ^2 tests revealed no significant differences (at $p < .01$) between the

results for different sentences with the same factor combination. Therefore, the results were compounded by factor combinations. These compounded results are shown in figure 4, alongside the expected values. The comparison immediately shows the relevant differences. The difference between expected and observed numbers for the sentences that conform to or violate both theories is relatively small and does not mean a deviation from the hypotheses. The two conditions where only one theory is violated, however, have surprising results: while those sentences that violate the LIPOC were judged as grammatical less often than the fully compliant ones and more often than the fully violating ones as expected, they are also apparently much more grammatical than the sentences that violate only the EIC principle.

Furthermore, the responses were subjected to an analysis of variance. The interaction between LIPOC and EIC did not reach significance, but the effects of each principle by itself did: $F_{LIPOC} = 1272.48$, $p < .0001$; $F_{EIC} = 22.16$, $p < .0001$.

5.5 Discussion

As the effects of both the LIPOC and the EIC principle are significant, this study proved them to be predictive for constituent ordering decisions in speakers with English as a second language. Of course, either or both could be merely epiphenomena of deeper principles, constraints, or operations. This study, however, is not an exploratory one and is therefore highly unlikely to uncover any such underlying principles serendipitously. Lacking any evidence to the contrary, I must therefore assume that the predictive power lies simply in the two principles.

Interestingly, the effects of the two principles are quite different: violations

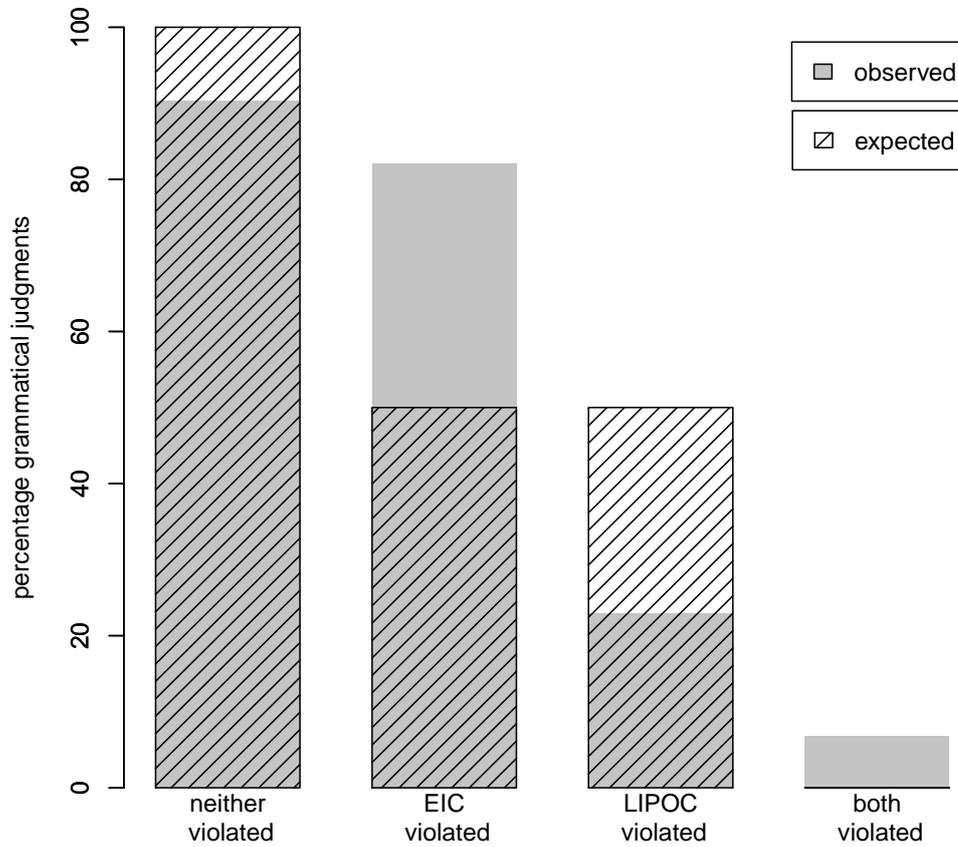


Figure 4: Percentages of “grammatical” judgments per condition, against predicted percentages

of the LIPOC cause a much larger drop in the number of grammatical judgments, which arguably means the LIPOC is stronger and more fundamental than the EIC principle. Table 1 showed that the same is true in children with English as their first language: most of cases where the LIPOC is able to make a prediction (or judgment) in fact conform to the LIPOC, while the EIC principle is violated a little more often. This confirms relatively

common knowledge about the grammar of English: the sentences that violate only the LIPOC have PPs before NPs or full NPs before pronouns; this is categorically ungrammatical according to standard English grammars like Biber (1999) and Huddleston and Pullum (2002). The sentences that violate only the EIC, on the other hand, are merely ‘odd’, and the same grammars make no mention of the reason for this other than general references to the law of increasing constituents. Decades of introspection and the present study therefore both point to this conclusion: the LIPOC captures a categorical rule, while the EIC principle formalizes gradience.

5.6 Summary

The study presented here used a binary grammaticality judgment task to test the relative strengths of Dik (1989)’s LIPOC and Hawkins (1994)’s EIC principle in the constituent ordering decisions made by speakers of English as a second language. The sample only included native speakers of German, but as German syntax is quite different from English syntax, particular in the order of constituents. It is therefore unlikely that speakers of German are subject to any kind of transfer bias, be it positive or negative, when parsing constituent ordering in English. Although further research would of course provide a valuable contribution to the subject, these results and conclusions should therefore generalize well.

The results of this study show that both the LIPOC and the EIC principle affect the constituent order. Both were found to have a significant effect and therefore predictive power, but the effect of the LIPOC was much stronger than that of the EIC principle. This merely reflects the underlying difference between the two theories: the former is a relatively reliable rule of grammar,

whereas the latter is an example of linguistic gradience. It is only to be expected that the grammaticality constraint affects binary grammaticality judgments more than the gradient principle. What this study has shown, then, is that the gradient principle is active even in the English spoken as a second language by speakers who are fully immersed in a first language environment.

6 Conclusion

Researchers from different fields of linguistics have proposed theories about syntactic weight effects in English. While a multifactorial approach is instinctively appealing, not all of these theories are compatible with each other. This thesis adds some data from language acquisition to the discussion. It comprises two studies: a corpus study on first language acquisition with a statistical regression model, and a questionnaire study on second language acquisition. The results support the Language-Independent Preferred Order of Constituents (Dik 1989) and the principle of Early Immediate Constituents (Hawkins 1994). The corpus data also confirm that some verbs have a ‘lexical bias’ for or against certain constructions. Furthermore, it showed constituents that are new to the context to be ‘heavy’. This pragmatic influence has been claimed for all weight effects, but was not found in the larger set of corpus data; this could of course be due to the considerable noise that is to be expected in such a diverse data set. In conclusion, the results of both studies paint a mixed picture: no single theory emerges as the clearly best one, and there may be differences between the weight effects of different phenomena.

As it is recent, substantial, and methodologically sound, Wasow (2002) is the obvious work on weight effects in adult native speakers that these results can be compared against. While there are some minor differences, my results agree with Wasow’s in that they point to a “multiplicity of the influences on ordering” (Wasow 2002:109). Lacking further evidence, the differences can all be explained by the simple fact that a first language acquisition corpus will by nature be different from newspaper and adult speech corpora, or by the limitations of my two studies as discussed above.

The results in empirical linguistics are rarely pure and never simple. While this thesis provides more support to some theories than others, several interesting questions remain unanswered. Is the ‘newness’ of a constituent really a substantive effect in constituent ordering? Can ‘newness’ even be adequately operationalized? How did these syntactic weight effects develop in the history of English?¹⁶ What information is actually available to the human faculty for language at the time it makes constituent ordering decisions? Do the same theories that explain and predict weight effects in English work for weight effects in other languages, and is the LIPOC really “language-independent”? I must leave these questions to future research.

Such research would do well to address one final question that I did not need to deal with here: how conscious are weight effects? Style guides are fond of urging their readers “to keep subjects, verbs, objects, and complements close together for maximum clarity” and to “[m]ove interrupting material to the beginning or end of your new sentence” (Glaser 1999:154)—a good description of heavy NP shift (ignoring the argument to clarity). It is not unreasonable to assume that English-speaking adults may adjust their postverbal constituents consciously. As clear and ‘good’ writing is one priority of university education, such ‘rules’ are probably quite familiar to graduates and students—and they are the participants or subjects in many linguistic experiments and studies. My two studies did not face this problem: the second language learners were not asked to produce sentences, only to rate them; and the children whose production makes up CHILDES will hardly have read any style guides by the time of recording. Production studies on adults, however, should attend to this possible effect, perhaps by collecting data from different registers.

¹⁶I thank Prof. Janet Grijzenhout for pointing out this intriguing approach to me.

References

- Akasaka, Yukiko, and Koichi Tateishi. 2001. Heaviness in interfaces. In *Issues in Japanese phonology and morphology*, ed. Jeroen van de Weijer and Tetsuo Nishihara, 3–46. Berlin: Mouton de Gruyter.
- Arnold, Jennifer, Anthony Losongco, Thomas Wasow, and Ryan Ginstrom. 2000. Heaviness vs. newness: the effects of structural complexity and discourse status on constituent ordering. *Language* 76:28–55.
- Bader, Markus, and Jana Häussler. 2010a. The primacy of grammaticality: a case study on ditransitive verbs in German. Manuscript, University of Konstanz.
- Bader, Markus, and Jana Häussler. 2010b. Toward a model of grammaticality judgments. *Journal of Linguistics* 46:273–300.
- Basilico, David. 2008. Particle verbs and benefactive double objects in English. *Natural Language and Linguistic Theory* 26:731–773.
- Bates, Elizabeth, Inge Bretherton, and Lynn Snyder. 1988. *From first words to grammar*. Cambridge, MA: Cambridge University Press.
- Behaghel, Otto. 1928. *Deutsche Syntax*, volume 3: *Die Satzgebilde*. Heidelberg: Carl Winters.
- Behaghel, Otto. 1932. *Deutsche Syntax*, volume 4: *Wortstellung. Periodenbau*. Heidelberg: Carl Winters.
- Behrens, Heike. 2008. Corpora in language acquisition research. In *Corpora in language acquisition research*, ed. Heike Behrens, xi–xxx. Amsterdam: Benjamins.
- Bernstein, Nan. 1982. Acoustic study of mothers' speech to language-learning children. Doctoral Dissertation, Boston University.
- Biber, Douglas. 1999. *Longman grammar of spoken and written English*. Harlow: Longman.
- Bliss, Lynn. 1988. The development of modals. *The Journal of Applied Developmental Psychology* 9:253–261.
- Bloom, Lois. 1970. *Language development: form and function in emerging grammars*. Cambridge, MA: Massachusetts Institute of Technology Press.

- Bloom, Lois. 1973. *One word at a time*. The Hague: Mouton.
- Bloom, Lois, Lois Hood, and Patsy Lightbown. 1974. Imitation in language development. *Cognitive Psychology* 6:380–420.
- Bloom, Lois, Patsy Lightbown, and Lois Hood. 1975. Structure and variation in child language. Technical Report 40, Monographs of the Society for Research in Child Development.
- Bohannon, John, and Angela Marquis. 1977. Children’s control of adult speech. *Child Development* 48:1002–1008.
- Bresnan, Joan, Anna Cueni, Tatiana Nikitina, and Rolf Baayen. 2007. Predicting the dative alternation. In *Cognitive foundations of interpretation*, ed. Gerlof Bouma, Irene Krämer, and Joost Zwarts, 69–94. Amsterdam: Royal Netherlands Academy of Science.
- Bresnan, Joan, and Tatiana Nikitina. 2003. On the gradience of the dative alternation. Manuscript, Stanford University.
- Brown, Roger. 1973. *A first language: the early stages*. Cambridge, MA: Harvard University Press.
- Bruening, Benjamin. 2010. Double object constructions disguised as prepositional datives. *Linguistic Inquiry* 41:287–305.
- Carlson-Luden, Vicki. 1979. Causal understanding in the 10-month-old. Doctoral Dissertation, University of Colorado at Boulder.
- Carterette, Edward, and Margaret Jones. 1974. *Informal speech*. Berkeley: University of California Press.
- Chang, Franklin. 2009. Learning to order words: a connectionist model of heavy NP shift and accessibility effects in Japanese and English. *Journal of Memory and Language* 61:374–397.
- Clark, Eve. 1982. The young word maker. In *Language acquisition: the state of the art*, ed. Eric Wanner and Lila Gleitman, 390–428. Cambridge, MA: Cambridge University Press.
- Colleman, Timothy, and Bernard de Clerck. 2009. ‘Caused motion’? the semantics of the English to-dative and the Dutch aan-dative. *Cognitive Linguistics* 20:5–42.

- Conwell, Erin, and Katherine Demuth. 2007. Early syntactic productivity. *Cognition* 103:163–179.
- Cooper, William, and John Ross. 1975. World order. In *Papers from the parasession on Functionalism*, ed. Robin Grossman, L. James San, and Timothy Vance, 63–111. Chicago: CLS.
- Cruttenden, Alan. 1978. Assimilation in child language and elsewhere. *Journal of Child Language* 5:373–378.
- Cuervo, María Cristina. 2003. Structural asymmetries but same word order. In *Asymmetry in grammar*, ed. Anna Di Sciullo, volume 1: *Syntax and semantics*, 117–144. Amsterdam: Benjamins.
- de Marneffe, Marie-Catherine, Scott Grimm, Uriel Cohen Priva, Sander Lestrade, Gorkem Ozbek, Tyler Schnoebelen, Susannah Kirby, Misha Becker, Vivienne Fong, and Joan Bresnan. 2007. A statistical model of grammatical choices in children’s productions of dative sentences. Presented at Formal Approaches to Variation in Syntax, University of York, 10–13 May 2007.
- de Wind, Maarten. 1999. Deriving heavy NP-shift in French. In *Formal perspectives of Romance linguistics*, ed. J.-Marc Authier, Barbara Bullock, and Lisa Reed, 89–105. Amsterdam: Benjamins.
- Dehé, Nicole. 2001. Transitive particle verbs in English: the neutral order. In *Structural aspects of semantically complex verbs*, ed. Nicole Dehé and Anja Wanner, 165–189. Frankfurt: Peter Lang.
- Dehé, Nicole. 2002. *Particle verbs in English*. Amsterdam: Benjamins.
- Dik, Simon. 1989. *The theory of Functional Grammar*. Dordrecht: Foris.
- Farrell, Patrick. 2005. English verb-preposition constructions. *Language* 81:96–137.
- Fletcher, Paul, and Michael Garman. 1988. Normal language development and language impairment. *Clinical Linguistics and Phonetics* 2:97–114.
- Gathercole, Virginia. 1986. The acquisition of the present perfect. *Journal of Child Language* 13:537–560.
- Gibson, Edward. 1998. Linguistic complexity: locality of syntactic dependencies. *Cognition* 68:1–76.

- Gilkerson, Jill. 2006. Acquiring English particle verbs. Doctoral Dissertation, University of California, Los Angeles.
- Givón, Talmy. 1988. The pragmatics of word-order. In *Studies in syntactic typology*, ed. Michael Hammond, Edith Moravcsik, and Jessica Wirth, 243–284. Amsterdam: Benjamins.
- Glaser, Joe. 1999. *Understanding style*. Oxford: Oxford University Press.
- Gries, Stefan. 2002. The influence of processing on syntactic variation. In *Verb-particle explorations*, ed. Nicole Dehé, Ray Jackendoff, Andrew McIntyre, and Silke Urban, 269–288. Berlin: Mouton de Gruyter.
- Gries, Stefan. 2009. *Statistics for linguistics with R*. Berlin: de Gruyter.
- Gropen, Jess, Steven Pinker, Michelle Hollander, Richard Goldberg, and Ronald Wilson. 1989. The learnability and acquisition of the dative alternation in English. *Language* 65:203–257.
- Haider, Hubert. 2009. The thin line between facts and fiction. In *The fruits of empirical linguistics*, ed. Sam Featherston and Susanne Winkler, volume 1: *Process*, 75–102. Berlin: Mouton de Gruyter.
- Hawkins, John. 1990. A parsing theory of word order universals. *Linguistic Inquiry* 21:223–261.
- Hawkins, John. 1994. *A performance theory of order and constituency*. Cambridge, Cambs.: Cambridge University Press.
- Hawkins, John. 2006. Gradedness as relative efficiency in the processing of syntax and semantics. In *Gradience in grammar*, ed. Gisbert Fanselow, Caroline Féry, Ralf Vogel, and Matthias Schlesewsky, 207–226. Oxford: Oxford University Press.
- Hayes, Donald. 1986. The Cornell corpus. Technical Report 86-1, Cornell University Department of Sociology, Ithaca, NY.
- Henry, Alison. 1995. *Belfast English and Standard English*. New York: Oxford University Press.
- Hinrichs, Lars, and Benedikt Szmrecsanyi. 2007. Recent changes in the function and frequency of Standard English genitive constructions. *English Language and Linguistics* 11:437–474.

- Howe, Christine. 1981. *Acquiring language in a conversational context*. New York: Academic Press.
- Huddleston, Rodney, and Geoffrey Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge, Cambs.: Cambridge University Press.
- Johnson, M. 1986. A computer-based approach to the analysis of child language data. Manuscript, University of Reading.
- Jones, Margaret, and Edward Carterette. 1963. Redundancy in children's free-reading choices. *Journal of Verbal Learning and Verbal Behavior* 2:489–493.
- Kayne, Richard. 1998. Overt vs. covert movement. *Syntax* 1:128–191.
- Kremers, Joost. 2009. A prosodic approach to particle verbs. *Linguistische Berichte* 217:81–96.
- Lechler, Andrea, and Antje Roßdeutscher. 2009. German particle verbs with auf. reconstructing their composition in a DRT-based framework. *Linguistische Berichte* 220:439–478.
- Levin, Beth, and Malka Rappaport Hovav. 2002. What alternates in the dative alternation? Presented at the Conference on Role and Reference Grammar, Universidad de La Rioja, Logrono, Spain.
- Liu, Feng-hsi. 2006. Dative constructions in Chinese. *Language and Linguistics* 7:863–904.
- Lohse, Barbara, John Hawkins, and Thomas Wasow. 2004. Domain minimization in English verb-particle constructions. *Language* 80:238–261.
- Lozano, Cristóbal, and Amaya Mendikoetxea. 2010. Interface conditions on postverbal subjects. *Bilingualism: Language and Cognition* 13:475–497.
- Lüdeling, Anke. 1999. On particle verbs and similar constructions in German. Doctoral Dissertation, Universität Tübingen.
- MacWhinney, Brian. 2000. *The CHILDES project: tools for analyzing talk*. Mahwah: Lawrence Erlbaum Associates, third edition.
- Nomoto, Hiroki. 2009. On indefinite subjects of pivot verbs in Malay. *International Journal of Corpus Linguistics* 14:221–254.
- Ogawa, Yoshiki. 2008. The dative alternation as A-movement out of a small clause CP. *English Linguistics* 25:93–126.

- Ormazabal, Javier, and Juan Romero. 2010. The derivation of dative alternation. In *Argument structure and syntactic relations*, ed. Maia Duguine, Susana Huidobro, and Nerea Madariaga, 203–232. Amsterdam: Benjamins.
- Oya, Toshiaki. 2009. Ground arguments in German particle verbs. *Journal of Germanic Linguistics* 21:257–296.
- Oyharçabal, Beñat. 2010. Basque ditransitives. In *Argument structure and syntactic relations*, ed. Maia Duguine, Susana Huidobro, and Nerea Madariaga, 233–260. Amsterdam: Benjamins.
- Pause, Peter. 1999. Locative alternation in German particle verbs and their French equivalents. *Studi Italiani di Linguistica Teorica e Applicata* 28:373–393.
- R Development Core Team. 2011. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna. URL <http://www.R-project.org>, ISBN 3-900051-07-0.
- Rappaport Hovav, Malka, and Beth Levin. 2008. The English dative alternation: the case for verb sensitivity. *Journal of Linguistics* 44:129–167.
- Rizzi, Luigi. 2006. Satisfying the subject criterion by a non subject. In *Phases of interpretation*, ed. Mara Frascarelli, 341–361. Berlin: Mouton de Gruyter.
- Rochemont, Michael, and Peter Culicover. 1990. *English focus constructions and the theory of grammar*. Cambridge, Cambs.: Cambridge University Press.
- Rosenbach, Anette. 2005. Animacy versus weight as determinants of grammatical variation in English. *Language* 81:613–644.
- Ross, John. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.
- Saito, Mamoru, and Naoki Fuku. 1998. Order in phrase structure and movement. *Linguistic Inquiry* 29:439–474.
- Scaglione, Aldo. 1979. Beauzée, Behaghel, and J. R. Ross on heavy-NP shift. *Historiographia linguistica* 6:231–245.
- Shiobara, Kayono. 2008. Prosodic phase and left-to-right structure building. *The Canadian Journal of Linguistics* 53:253–274.

- Sigurd, Bengt. 1995. Analysis of particle verbs for automatic translation. *Nordic Journal of Linguistics* 18:55–65.
- Snyder, William, and Karin Stromswold. 1997. The structure and acquisition of English dative constructions. *Linguistic Inquiry* 28:281–317.
- Stallings, Lynne, and Maryellen MacDonald. 2011. It’s not just the “heavy NP”: relative phrase length modulates the production of heavy-NP shift. *Journal of Psycholinguistic Research* 40:177–187.
- Stallings, Lynne, Maryellen MacDonald, and Padraig O’Seaghdha. 1998. Phrasal ordering constraints in sentence production. *Journal of Memory and Language* 39:392–417.
- Stefanowitsch, Anatol. 2006. Negative evidence and the raw frequency fallacy. *Corpus Linguistics and Linguistic Theory* 2:61–77.
- Sternberg, Saul, Stephen Monsell, Ronald Knoll, and Charles Wright. 1978. The latency and duration of rapid movement sequences. In *Information processing in motor control and learning*, ed. G. E. Stelmach, 117–151. New York: Academic Press.
- Stine, Elizabeth, and John Bohannon. 1983. Imitations, interactions, and language acquisition. *Journal of Child Language* 10:589–603.
- Takano, Yuji. 2003. How antisymmetric is syntax? *Linguistic Inquiry* 34:516–526.
- Taraldsen, Knut. 2000. V-movement and VP-movement in derivations leading to VO order. In *The derivation of VO and OV*, ed. Peter Svenonius, 97–122. Amsterdam: Benjamins.
- Theakston, Anna, Elena Lieven, Julian Pine, and Caroline Rowland. 2001. The role of performance limitations in the acquisition of verb-argument structure. *Journal of Child Language* 28:127–152.
- Wasow, Thomas. 2002. *Postverbal behavior*. Stanford: CSLI.
- Wells, Gordon. 1981. *Learning through interaction*. Cambridge, Cambs.: Cambridge University Press.
- Weskott, Thomas, and Gisbert Fanselow. 2009. Scaling issues in the measurement of linguistic acceptability. In *The fruits of empirical linguistics*, ed. Sam Featherston and Susanne Winkler, volume 1: *Process*, 231–245. Berlin: Mouton de Gruyter.

- Whong-Barr, Melinda, and Bonnie Schwartz. 2002. Morphological and syntactic transfer in child L2 acquisition of the English dative alternation. *Studies in Second Language Acquisition* 24:579–616.
- Wilson, John, and Alison Henry. 1998. Parameter setting within a socially realistic linguistics. *Language in Society* 27:1–21.
- Zec, Draga, and Sharon Inkelas. 1990. Prosodically constrained syntax. In *The phonology-syntax connection*, ed. Sharon Inkelas and Draga Zec, 365–378. Chicago: University of Chicago Press.
- Zovko Dinković, Irena. 2007. Dative alternation in Croatian. *Suvremena lingvistika* 63:65–83.

A Sample questionnaire of the second language study

The following is one of the four questionnaires used in the study in section 5. The odd-numbered items are the fillers.

Alter: _____ Geschlecht: _____ Muttersprache(n): _____

Beurteilen Sie bitte jeden dieser Sätze danach, ob er grammatisch korrekt ist oder nicht. Bitte lassen Sie dabei möglichst keinen Satz aus!

1. Jennifer moved the hearts of everyone that heard her short speech.
 grammatisch ungrammatisch
2. Cathy put her nice and naive fiancé in this position.
 grammatisch ungrammatisch
3. Jewel called her boss “evil maniac who is lazy”.
 grammatisch ungrammatisch
4. Karen moved the very heavy wooden boxes to the library.
 grammatisch ungrammatisch
5. Adam left his car keys by accident.
 grammatisch ungrammatisch
6. Peter turned toward the locked door the chair.
 grammatisch ungrammatisch
7. Simon put rules that were very complicated in place.
 grammatisch ungrammatisch
8. Matt drew on the paper some strange and wavy lines.
 grammatisch ungrammatisch
9. Kate turned the tables on her baffled enemies.
 grammatisch ungrammatisch
10. Liz brought her dog to the school reunion.
 grammatisch ungrammatisch
11. Gina put to rest these dangerous speculations.
 grammatisch ungrammatisch
12. Billie left at home her new improved cell phone.
 grammatisch ungrammatisch
13. Nathan brought to her attention this strange but interesting topic.
 grammatisch ungrammatisch
14. Chris took his uncle to that upscale restaurant.
 grammatisch ungrammatisch
15. Tim brought to light some very surprising evidence.
 grammatisch ungrammatisch
16. David called the surprised and confused policeman “alien that smells”.
 grammatisch ungrammatisch

A