# The development of dependency length minimization in early child language: A case study of the dative alternation

Zoey Liu Department of Linguistics University of Florida liu.ying@ufl.edu

#### Abstract

How does the preference for dependency length minimization (DLM) develop in early child language? This study takes up this question with the dative alternation in English as the test case. We built a large-scale dataset of dative constructions using transcripts of naturalistic child-parent interactions. Across different developmental stages of children, there appears to be a strong tendency for DLM. The tendency emerges between the age range of 12-18 months, slightly decreases until 30-36 months, then becomes more pronounced afterwards and approaches parents' production preferences after 48 months. We further show the extent of DLM depends on how a given dative construction is realized: the tendency for shorter dependencies is much more pronounced in double object structures, whereas the prepositional object structures are associated with longer dependencies.

### 1 Introduction

The principle of Dependency Length Minimization (DLM) (Ferrer-i Cancho, 2004), originally developed based on the framework of Dependency Grammar (Tesnière, 1959), predicts that words or phrases that are syntactically dependent on each other prefer to appear closer in order to minimize the overall dependency distance, thereby reducing its structural complexity.

While research on DLM thus far has been fruitful (Hawkins, 1990; Gildea and Temperley, 2010; Gulordava and Merlo, 2015; Liu, 2020, 2022), one crucial question remains: how does the preference for shorter dependencies develop in early child language? Given that the preference for DLM has been well-documented in the literature, we would expect to see similar preferences in child production as well. That said, it is unclear (1) at what developmental stage the preference for DLM emerges; (2) whether and how the extent of DLM varies along the developmental trajectory; (3) when Stefanie Wulff Department of Linguistics University of Florida UiT The Arctic University of Norway swulff@ufl.edu

children's production of DLM reaches a comparable level to that in parent/adult production.

This study addresses the aforementioned questions using the dative construction in English as the test case. Here (1a) and (1b) are different syntactic variants of the same dative construction: (1a) is a double object construction, (1b) is a prepositional object construction. Within the verb phrase (VP) of (1a), the head verb has two noun phrase (NP) dependents, one as the direct object (*the toy*) and one as the indirect object (*me*); the semantic roles for the two are theme and recipient, respectively. By comparison, in (1b), the direct object dependent of the head verb, *the toy*, is the same as that in (1a), whereas the recipient is realized as a prepositional phrase (PP) dependent instead.

(1) a. give [NP the girl] [NP the lunch box]
b. give [NP the lunch box] [PP to the girl]

Leveraging transcripts of naturalistic childparent interactions and computational techniques, we analyze the developmental patterns of DLM in child production of the dative alternation. We foresee two possible directions regarding the extent of DLM across children's developmental stages. On one hand, at earlier stages, utterances produced by children are comparatively shorter (Brown, 1973); based on evidence from written data that there is a positive correlation between overall dependency length and sentence length (Ferrer-i Cancho et al., 2020; Futrell et al., 2020), this means that during these stages the preference for DLM is potentially weaker, and would gradually increase as utterance lengths increase when children reach later developmental stages. On the other hand, if the primary motivation for DLM is to lessen cognitive load (Gibson et al., 2019; Hawkins, 2007, 2015), then at earlier developmental stages, when children have shorter working memory (Hudson Kam, 2019; Austin et al., 2022), they may have a stronger preference for shorter dependencies than they do in later stages of development.

#### 2 Related work

The dative alternation in English (Levin, 1993), has been studied extensively, specially in first language adult production (Bresnan et al., 2007; Bresnan, 2007; Szmrecsanyi et al., 2017; Engel et al., 2022). In addition, a number of studies have looked into the production patterns of the dative constructions in child (and child-directed) spoken language in English, though from different angles. One line of work probes the generalization (Goldberg et al., 2005; Conwell and Demuth, 2007; Shimpi et al., 2007) and learnability (Gropen et al., 1989; Yang and Montrul, 2017) of the dative structures in children's production. Others attended to the developmental order of the different variants of the dative construction (Campbell and Tomasello, 2001; Snyder and Stromswold, 1997). With syntactic orders in particular, De Marneffe et al. (2012) investigated what structural constraints, such as animacy and pronominality, affect children's syntactic choices.

#### **3** Experiments

#### 3.1 Data and preprocessing

Although prior work has studied the dative alternation in child production, their constructed datasets are not publicly available. In addition, they tended to focus on narrower age ranges of only a handful of children. Therefore we turned to building a dataset of our own. For child (and parent) production data, we resorted to the CHILDES database (MacWhinney, 2000), which contains transcripts of naturalistic child-parent conversational speech. We focused on (monolingual) children with typical development. Child and parent utterances were first taken from the English-NA and the English-UK sections of CHILDES via the childes-db interface (Sanchez et al., 2019). We then automatically assigned part-of-speech (POS) tags as well as syntactic dependencies to each utterance in order to derive morphosyntactic information; the former was performed using Stanza (Qi et al., 2020), a publicly open library for natural language processing; and the latter was achieved using Dia-Parser (Attardi et al., 2021), which has recently been shown to yield good dependency parsing performance for child spoken language in English (Liu and Prud'hommeaux, 2022).

We relied on the classes of dative (N = 336) and benefactive (N = 177) verbs from Levin (1993) as references when extracting utterances that potentially contain a dative structure from the parsed data described above. We searched for VPs where the head verb occurs in either the double object structure (V-NP-NP) or the prepositional object structure (V-NP-PP). (See Appendix A for details on our data extraction process).

Here we used children's age as an index of their developmental stage; therefore we removed utterances where the age information of the corresponding child is not provided. This resulted in an initial dataset of 43,156 utterances. In what follows, we describe our annotation procedures for deciding whether an utterance contains a dative construction. Given the size of the dataset, manually annotating each instance is plausible yet not practical. To remedy that, we also illustrate a simple automatic approach for the identification of dative structures.

#### 3.2 Annotation criteria and process

We determined whether an utterance includes a dative structure or not based on the following two criteria: (1) the verb takes a direct object which is the theme, as well as an indirect object or a prepositional object that serves as either the recipient or the beneficiary; (2) the verb can be understood as expressing some action of transfer from the subject/agent of the sentence to the recipient/beneficiary, even if the action is metaphorical (e.g., (2b)). These restrictions naturally ruled out cases where the head verb takes a verbal complement (which was erroneously parsed as the object by the dependency parser; e.g., 2c); they also excluded cases where the head verb has a PP dependent occurring after the theme, but the semantic role of the PP is purpose (e.g., (2d)) or goal/direction (e.g., (2e)). That said, the annotation criteria were to some extent relaxed for utterances produced by children. For example, while the recipient of the verb is preferred to be animate (Bresnan et al., 2007), if based on preceding context of the utterance, the recipient could be interpreted as being personified (e.g., 2f), we deemed those cases as appropriate dative constructions as well.

(2) a. she **brings** lots of lego to me.  $^{1}$ 

- b. **carry** your dream for you.
- c. \*say thank you to your friend
- d. \*I **took** him for a walk.
- e. \*Daddy **sent** me to **school**.
- f. I made some lunch for my teddy.

<sup>&</sup>lt;sup>1</sup>Examples provided here are adapted from utterances initially extracted from CHILDES; \* marks the types of instances that we did not consider in this study.

Our annotation process for identifying the dative constructions is as follows. From the initial dataset derived from Section 3.1, we constructed three small practice sets for annotators to familiarize themselves with the annotation criteria described above; each practice set contained 50 utterances. Two annotators with advanced training in linguistics independently annotated one practice set first. They were instructed to annotate an instance as yes, if they considered the instance as having a dative structure, no if they considered the opposite, and unsure if they were uncertain about what decision to make. They then cross-checked their own annotations with each other and settled on the unsure cases along with other questions encountered during the annotation process. The annotation procedures for the other two practice sets were the same. Afterwards, the two annotators along with the senior author of this paper each annotated a subset (of different sizes) of the initial dataset, using the same three annotation labels. We ensured that there was overlap between each of these subsets such that a total of 1,000 utterances were independently annotated by two annotators (regardless of which two). Agreement score for these 1,000 utterances, which was measured as the percentage of times when the two annotators agree, was 95.20%. Discrepancies in annotations, including the unsure cases, were eventually resolved through discussions. Given the high agreement score, each annotator continued to independently examine more instances.

In total, we annotated 10,709 utterances taken from the initial dataset (Section 3.1), which we refer to hereafter as the gold-standard. Among these cases, 8,718 have an annotation label of *yes* whereas the remainder have the label *no*.

### **3.3** Automatic identification of the dative constructions

Using the gold-standard utterances, we explored automatic approaches in order to identify which of the remaining utterances in the initial dataset contain dative structures. Specifically, we treated this task as a binary classification task. We randomly split all the gold-standard data into training/test sets at a 4:1 ratio, 3 times. Our classifier was trained with BERT (Devlin et al., 2019) using the default parameters from MaChamp v0.3beta (van der Goot et al., 2021), an open-access multi-task learning toolkit. The input to the classifier was the utterance concatenated with the speaker of the utterance (child or

parent) and the head verb. The performance of the classifier was measured as its prediction accuracy averaged across the three test sets.

Label	Accuracy (%)	
yes	98.43	
no	84.05	

Table 1: Classification accuracy for each label for the gold-standard dative utterances.

Role	Structure	N
Child	double object	5,645
	prepositional object	2,401
Parent	double object	21,865
	prepositional object	8,793

Table 2: Descriptive statistics for the dative constructions in child and parent production.

The classifier was able to perform reasonably well (average accuracy = 94.36%; see also Table 1). Hence we trained the same classifier using all the gold-standard data, then applied it to the remaining instances in the initial dataset. We excluded cases in the original dataset with an annotation label of *no*, whether manually or automatically identified, eventually yielding a dataset of 38,704 utterances (Table 2; see also Figure 3 in Appendix C). Compared to previous work on constituent orderings of the dative alternation in child language development (De Marneffe et al., 2012), our dataset is of much larger scale (including utterances produced by over 900 children from 54 corpora).

#### 3.4 Measures for DLM

Since we used children's age as a proxy of developmental stage, to avoid data sparsity, we set every 6 months as one age bin, then separated all the dative constructions produced by children (and by parents accordingly) into their corresponding age bins. As illustration of our computations for the extent of DLM, consider the examples below. Say the original utterance appears in the double object structure (e.g., (3a)). To check whether DLM is observable in the utterance, we first measured its overall dependency length ( $DL_observed$ ). Then we automatically constructed the syntactic alternative of the utterance (e.g., (3b)).<sup>2</sup> and measured its overall dependency length as well ( $DL_alternative$ ); if the value of  $DL_observed$  is smaller than that of

<sup>&</sup>lt;sup>2</sup>See Appendix B for discussion about using sentences with the heavy NP shift (Wasow, 1997) as syntactic alternatives for the prepositional object structure.



Figure 1: DLM in the dative constructions in child and parent production.



- short\_before\_long - long\_before\_short - equal

Figure 2: DLM in the double object structure (V-NP-NP) and the prepositional object structure (V-NP-PP) in child and parent production.

DL\_alternative, we consider the original utterance to show DLM. For all the utterances produced by children within a certain age bin, we measured the proportion of instances where a preference for shorter dependencies exists, the proportion of cases where the opposite pattern holds, and the proportion of sentences where the overall dependency lengths of the syntacic alternatives are the same. Significance testing was conducted using bootstrapping (Efron and Tibshirani, 1994).



#### **Results** 4

Here we used the production patterns in parent data as benchmarks for analysis of the developmental trajectory of DLM in child data; thus the subplots in each figure often contrast child production patterns with those of parent production. As illustrated in Figure 1, there is consistently a pronounced preference for shorter dependencies across different stages of children's developmental trajectory. This preference seems to emerge in child production between the age range of 12 to 18 months; during this range the proportion of utterances that demonstrate a tendency for DLM is 100%, where all the utterances (N=14) have the double object structures (Figure 3). The overall tendency for DLM is also observable when looking at a few of the most frequent head verbs in the dative dataset, such as give and get (see Appendix D).

When comparing the extent of DLM in child production across different age ranges, it appears that the preference for shorter dependencies gradually gets weaker from 12 to 36 months, then grows noticeably stronger afterwards. In fact, the preference for DLM is the weakest when children are between 30-36 months old; that said, during that age range, the proportion of cases that demonstrate DLM is still 3.09 times that of the utterances that show the opposite observations. When children reach 42-48 months old, their production of DLM becomes more stable and is approaching the production levels in parent data.

When taking a closer look at DLM in the two structural alternatives of the dative constructions, respectively, we see different patterns (Figure 2). In the double object structures, again, there appears to be a strong tendency for shorter dependencies across the developmental trajectory of children. The preference for DLM in child production approximates that in parent production around the age range of 24-30 months.

By contrast, we observe the opposite tendency for the prepositional object structure, that is, across children's age ranges, there seems to be a significant preference against DLM instead. In other words, the observed V-NP-PP utterances produced by children actually have longer dependency length compared to their syntactic alternatives. We conjectured several explanations for this discrepancy and verified them with our data. First, the difference in the overall dependency length between the V-NP-PP instances and their double object alternatives is mostly small. Indeed, in about 63.10% of the prepositional object structures in child production (N=2,401), the overall dependency length difference between them and their structural alternations is equal to one. Second, the direct object/theme of the V-NP-PP utterances is relatively short (De Marneffe et al., 2012); in approximately 67.43% of all these instances, the theme consists of just one word. Third, in 77.14% of cases where the theme is composed of one word, the word is usually pronominal (Bresnan et al., 2007).

The patterns based on the prepositional object structures in turn shed light on the overall developmental trajectory of the preference for DLM in Figure 1: between the age range of 12 to 36 months, the proportion of the V-NP-PP structures in children's production gradually increases (from 20.44% to 39.61%), leading to overall weaker extents of shorter dependencies during this age range; the proportion of the V-NP-PP instances then gradually decreases after 36 months, thereby making the age range of 30-36 months a "turning point" in the development of DLM in child production.

#### 5 Discussion

This study analyzed the developmental trajectory of the preference for DLM in child production using the dative alternation in English as the test case. Our findings illustrated that the tendency for shorter dependencies emerges in child production during the age range of 12-18 months. The extent of the tendency decreases until 30-36 months, then gradually increases and approximates the production level in parent data around 42-48 months.

In this work, we used age as the index of children's developmental stages. For future experiments, we plan to investigate how other alternatives, such as the mean length of utterance, affect observations of children's developmental trajectories of DLM. We would also like to analyze the development of children's syntactic choices via enriching the dataset with annotations for other constraints such as verb semantics. These factors could potentially provide additional explanations for the varying extents of DLM in children's early development. Lastly, given that our dative dataset is much larger than prior ones, we hope that it will be useful to research topics related to acquisition of syntactic alternations more broadly.

#### References

- Jennifer E. Arnold, Anthony Losongco, Thomas Wasow, and Ryan Ginstrom. 2000. Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, 76(1):28– 55.
- Giuseppe Attardi, Daniele Sartiano, and Zhang Yu. 2021. Diaparser attentive dependency parser.
- Alice Austin, Kathryn D. Schuler, S Furlong, and Elissa L. Newport. 2022. Learning a language from inconsistent input: Regularization in child and adult learners. *Language Learning and Development*, 18:249 – 277.
- Joan Bresnan. 2007. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. *Roots: Linguistics in Search of its Evidential Base*, 96:77–96.
- Joan Bresnan, Anna Cueni, Tatiana Nikitina, and R Harald Baayen. 2007. Predicting the dative alternation. In *Cognitive foundations of interpretation*, pages 69– 94. KNAW.
- Roger Brown. 1973. A first language: The early stages. Harvard University Press.

- Aimee L Campbell and Michael Tomasello. 2001. The acquisition of English dative constructions. *Applied psycholinguistics*, 22(2):253–267.
- Erin Conwell and Katherine Demuth. 2007. Early syntactic productivity: Evidence from dative shift. *Cognition*, 103(2):163–179.
- Marie-Catherine De Marneffe, Scott Grimm, Inbal Arnon, Susannah Kirby, and Joan Bresnan. 2012. A statistical model of the grammatical choices in child production of dative sentences. *Language and Cognitive Processes*, 27(1):25–61.
- Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Bradley Efron and Robert J Tibshirani. 1994. An introduction to the bootstrap. CRC press.
- Alexandra Engel, Jason Grafmiller, Laura Rosseel, and Benedikt Szmrecsanyi. 2022. Assessing the complexity of lectal competence: the register-specificity of the dative alternation after give. *Cognitive Linguistics*.
- Ramon Ferrer-i Cancho. 2004. Euclidean distance between syntactically linked words. *Physical Review E*, 70(5):056135.
- Ramon Ferrer-i Cancho, Carlos Gómez-Rodríguez, Juan Luis Esteban, and Lluís Alemany-Puig. 2020. The optimality of syntactic dependency distances. *arXiv preprint arXiv:2007.15342*.
- Richard Futrell, Roger P Levy, and Edward Gibson. 2020. Dependency locality as an explanatory principle for word order. *Language*, 96(2):371–412.
- Edward Gibson, Richard Futrell, Steven Piantadosi, Isabelle Dautriche, Kyle Mahowald, Leon Bergen, and Roger Levy. 2019. How efficiency shapes human language. *Trends in Cognitive Sciences*, 23(5):389–407.
- Daniel Gildea and David Temperley. 2010. Do Grammars Minimize Dependency Length? *Cognitive Science*, 34(2):286–310.
- Adele E Goldberg, Devin M Casenhiser, and Nitya Sethuraman. 2005. The role of prediction in construction-learning. *Journal of Child Language*, 32(2):407–426.
- Jess Gropen, Steven Pinker, Michelle Hollander, Richard Goldberg, and Ronald Wilson. 1989. The learnability and acquisition of the dative alternation in English. *Language*, pages 203–257.

- Kristina Gulordava and Paola Merlo. 2015. Structural and lexical factors in adjective placement in complex noun phrases across Romance languages. In *Proceedings of the Nineteenth Conference on Computational Natural Language Learning*, pages 247–257.
- John A Hawkins. 1990. A Parsing Theory of Word Order Universals. *Linguistic Inquiry*, 21(2):223–261.
- John A. Hawkins. 2007. Processing typology and why psychologists need to know about it. *New Ideas in Psychology*, 25(2):87–107.
- John A. Hawkins. 2015. Typological Variation and Efficient Processing. In *The Handbook of Language Emergence*, chapter 10, pages 215–236. John Wiley Sons, Ltd.
- Carla L. Hudson Kam. 2019. Reconsidering retrieval effects on adult regularization of inconsistent variation in language. *Language Learning and Development*, 15(4):317–337.
- Beth Levin. 1993. English verb classes and alternations: A preliminary investigation. University of Chicago press.
- Zoey Liu. 2020. Mixed evidence for crosslinguistic dependency length minimization. *STUF Language Typology and Universals*, 73(4):605 633.
- Zoey Liu. 2022. A multifactorial approach to crosslinguistic constituent orderings. *Linguistics Vanguard*.
- Zoey Liu and Emily Prud'hommeaux. 2022. Datadriven parsing evaluation for child-parent interactions. arXiv preprint arXiv:2209.13778.
- Brian MacWhinney. 2000. The CHILDES Project: Tools for analyzing talk. transcription format and programs, volume 1. Psychology Press.
- Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D. Manning. 2020. Stanza: A Python Natural Language Processing Toolkit for Many Human Languages. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pages 101–108, Online. Association for Computational Linguistics.
- Alessandro Sanchez, Stephan C Meylan, Mika Braginsky, Kyle E MacDonald, Daniel Yurovsky, and Michael C Frank. 2019. childes-db: A flexible and reproducible interface to the child language data exchange system. *Behavior research methods*, 51(4):1928–1941.
- Priya M Shimpi, Perla B Gámez, Janellen Huttenlocher, and Marina Vasilyeva. 2007. Syntactic priming in 3-and 4-year-old children: evidence for abstract representations of transitive and dative forms. *Developmental psychology*, 43(6):1334.
- William Snyder and Karin Stromswold. 1997. The structure and acquisition of English dative constructions. *Linguistic inquiry*, pages 281–317.

- Lynne M Stallings, Maryellen C MacDonald, and Padraig G O'Seaghdha. 1998. Phrasal ordering constraints in sentence production: Phrase length and verb disposition in heavy-NP shift. *Journal of Memory and Language*, 39(3):392–417.
- Benedikt Szmrecsanyi, Jason Grafmiller, Joan Bresnan, Anette Rosenbach, Sali Tagliamonte, and Simon Todd. 2017. Spoken syntax in a comparative perspective: The dative and genitive alternation in varieties of English. *Glossa: a journal of general linguistics*, 2(1):1–27.
- Lucien Tesnière. 1959. *Eléments de Syntaxe Structurale*. Paris: Klincksieck.
- Rob van der Goot, Ahmet Üstün, Alan Ramponi, Ibrahim Sharaf, and Barbara Plank. 2021. Massive choice, ample tasks (MaChAmp): A toolkit for multitask learning in NLP. In Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations, pages 176–197, Online. Association for Computational Linguistics.
- Thomas Wasow. 1997. End-Weight from the Speaker's Perspective. *Journal of Psycholinguistic Research*, (3):347–361.
- Charles Yang and Silvina Montrul. 2017. Learning datives: The tolerance principle in monolingual and bilingual acquisition. *Second Language Research*, 33(1):119–144.

#### A Notes on data preprocessing

After parsing data from the English sections of the CHILDES database (Section 3.1), we searched for VPs where the head verb takes either a double object structure (V-NP-NP) or a prepositional object structure (V-NP-PP). The part-of-speech (POS) tag of the head verb was VERB, which only includes lexical verbs (as opposed to auxiliaries). For double object structures, we selected VP instances in which the head verb has one direct object and one indirect object, which were identified based on their dependency relations with the head verb (obj and iobj, respectively). For the prepositional object structure, we selected VP instances where the head verb takes one direct object as well as one PP oblique immediately following the direct object; the dependency relation between the PP and the head verb was oblique, and the nominal head of the PP had one of four POS tags: NOUN (lexical noun), NUM (numeral), PRON (pronoun), and PROPN (proper noun). For verbs that only belong to the dative class, the adposition, or the function head of the PP was restricted to to, and for the benefactive verbs, the adposition was for; for verbs

that are included in both classes, the adposition was either *to* or *for*.

Levin class verbs were taken from http://www-personal.umich.edu/ ~jlawler/levin.verbs; there are 23 verbs overlapped in both classes. Note that in the final dative dataset (Section 3.3), there were 67 dative verbs, 52 benefactive verbs, and 15 verbs that belong to both classes.

### **B** Notes on syntactic alternatives for the prepositional object structure

Based on literature related to the heavy NP shift in English (Stallings et al., 1998; Arnold et al., 2000), one might posit that the alternative of an observed prepositional object structure can be constructed another way. For example, if the original sentence is give [NP the bread that she bought at the store yesterday] [PP to her], one grammatical alternative, besides the direct object structure, can also be give [PP to her] [NP the bread that she bought atthe store yesterday]. Nevertheless, structures with (heavy) NP shift as such are rare in child production. We searched for VP instances where the head verb takes one direct object and one prepositional oblique phrase dependent (PP); in addition, the PP has to precede the direct object. This only yielded 128 utterances produced by 56 children. Therefore we left these cases out from our analysis.

## C Descriptive statistics for our dative dataset

Visualizations of the frequency distribution of the double object structure and the prepositional object structure in child and parent speech are presented in Figure 3.

#### **D DLM for specific head verbs**

We present the preferences for DLM in the dative constructions headed by *give* (Figure 4) and *get* (Figure 5) in child and parent production. Of all the head verbs for the dative alternation in our dataset, these two verbs are attested most frequently.



Figure 3: Frequency distribution of the double object structure (V-NP-NP) and the prepositional object structure (V-NP-PP) in child and parent production.



Figure 4: DLM in the dative constructions headed by *give* in child and parent production (Child: N=3,338; Parent: N=12,246).



Figure 5: DLM in the dative constructions headed by *get* in child and parent production (Child: N=1,158; Parent: N=3,414).