

# A statistical model of grammatical choices in children's productions of dative sentences

Marie-Catherine de Marneffe, Scott Grimm, Uriel Cohen Priva,  
Sander Lestrade, Gorkem Ozbek, Tyler Schnoebelen, Susannah Kirby,  
Misha Becker, Vivienne Fong and Joan Bresnan

## 1 Introduction

One of the most well-studied questions in child language acquisition is what type of knowledge children possess to guide their syntactic production.<sup>1</sup> Some posit that children are able to construct abstract representations to facilitate the acquisition of specific items, whereas others take the specific items as primary, claiming children learn each construction individually as anchored to a specific lexical item. A less frequently pursued question is what is the range of influential factors which weigh upon a child's production choice. Understanding which factors potentially motivate children's production choices will broaden the context in which one can pursue questions about children's acquisition of syntactic production capacity. For instance, it is becoming increasingly clear that adult production is sensitive to multiple factors, including both discourse and grammatical factors (see representative studies by Wasow (2002), Szmrecsanyi (2005), Jaeger (2006), Hinrichs & Szmrecsanyi (2007) and references therein). It is thus important to test whether such factors play a role in children's syntax, so as to gain a better understanding of what factors are at stake in the acquisition process.

This study focuses on the English dative alternation, such as shown in 1a and 1b.

- 1a. The child gave the drawings to the teacher.
- 1b. The child gave the teacher the drawings.

This construction has received much attention both in the literature primarily based on adult production and from the perspective of language acquisition. Corpus studies of adult English have found that grammatical and discourse properties of the recipient and theme have a quantitative influence on dative syntax (Thompson 1990, Collins 1995, Snyder 2003, Gries 2003, *inter alia*). In this paper, we build on the work of Bresnan et al. (2007), who have shown that the effects of discourse accessibility, animacy, definiteness, nominal expression type, and syntactic weight are significant factors influencing adult dative construction choice. We will apply a similar methodology to the question of children's production of dative sentences.

This approach has an added benefit in that examining the construction in terms of responses to a set of factors for both adults and children enables a comparison between the two. Accordingly, we will discuss how conjoining two models of child and adult production allows us to make convincing comparisons of the similarities and differences between the two populations in a quantitative fashion.

---

<sup>1</sup>The research reported here was supported in part by NSF Award No. IS-0624345 to Stanford University.

## 2 Modeling multiple factors

### 2.1 Logistic regression

Logistic regression modeling techniques have increasingly been employed to analyze and evaluate the multiple factors present in language production (Baayen 2006). A logistic regression model is a function of a set of factors which predicts a binary outcome. Thus it is particularly suited to studies of alternations between two choices (e.g., genitive alternation (Rosenbach 2003, Hinrichs & Szmeccanyi), dative alternation (Bresnan et al. 2007), presence/absence of complementizer (Roland et al. 2006)).

This method permits simultaneous evaluation of all the factors in the model, assessing the strength of each factor relative to all the others. For example, one can evaluate whether “givenness” is more important to a construction than “heaviness” in terms of how the two factors perform relative to one another within a data set. Further this allows one to determine which factors best explain the variance in the data set.

One of the pitfalls of analysis that logistic regression avoids is confounded factors. For instance, “givenness” might follow from apparent underlying effects of weight (or conversely), since given referents tend to be described by shorter phrases. Logistic regression can assess whether all the factors introduced in the model make independent contributions.

This technique delivers a statistical model of the data, which in turn allows quantitative generalizations to emerge from an often noisy data set. One striking instance of this is found in Bresnan et al.’s 2007 study of the dative alternation where the factors form a clear pattern. The values of the factors associated with a higher level of salience (for instance, “discourse given”, “animate” or “definite”) were all predictive of the argument occupying the first complement position in the dative construction whereas those associated with a lower level of salience (“not discourse given”, “inanimate”, “indefinite”) were predictive of the argument occupying the second complement position. Thus, the values indicative of higher prominence aligned with the more prominent syntactic position, and the converse. This pattern corresponds to *harmonic alignment*, a well-noted phenomenon in the functional and typological literature. With the aid of the logistic regression model, this qualitative pattern is quantitatively derived from the data.

### 2.2 Modeling the dative alternation in adult production (Bresnan et al. 2007)

The Bresnan et al. model mentioned above focused on the adult production of dative sentences, and was the first to effectively show the broad range of factors that independently contribute to this construction.

The study is founded on spoken speech, culling 2360 dative observations from the 3 million word Switchboard collection of recorded telephone conversations (Godfrey, Holliman, and McDaniel 1992). Each sentence was annotated for the following factors: “animacy”, “accessibility”, “concreteness”, “definiteness”, “nominal expression type”, “length” (number of words), “person”, “number”, “verb” and “verb semantic class”, as well as “persistence”, a measure of parallelism or structural repetition.<sup>2</sup> All the factors cited significantly contributed to explain the variance in the data, and the model correctly predicted 94% of the production choices of the dative sentences.

---

<sup>2</sup>Their data set is publicly available for download as part of the languageR package.

### 3 Modeling the dative alternation in child production

To assess which factors are pertinent to child production, we began with the hypothesis that factors pertinent to adult production are good candidates. To parallel the study of Bresnan et al. (2007), we also present a logistic regression model of the data for child production. This approach will facilitate testing whether children are sensitive to the same factors as adults with respect to the dative alternation. Below we describe the factors hypothesized as driving the variation in the dative alternation.

#### 3.1 Data and factors

The data of children's speech come from the Child Language Data Exchange System (CHILDES, MacWhinney 2000). CHILDES is a publicly available database of children's speech that is produced in an ecologically natural environment. We focused on the following seven children: Abe, Adam, Naomi, Nina, Sarah, Shem, and Trevor. These children were selected based on the amount of data they had available compared to other children: both their total number of utterances and the number of utterances containing a dative construction. The utterances were taken from children's production between the ages of 2 to 5 years old. The data yielded a sufficient number of utterances to investigate in depth two verbs, *give* and *show*.

We only selected dative constructions following the “verb NP NP” (double object construction) or “verb NP PP” (prepositional dative) patterns.<sup>3</sup> We allowed constructions where the child had elided the preposition but had put the arguments in the NP PP order (theme, recipient), as in “I wanna show it Daddy” [Sarah, 4;5.14], “give dat Ursula” [Adam, 2;6.17]. We ended up with 530 utterances.

The different factors taken into consideration when building the model for children's production are essentially the same as the ones used in the adult model of Bresnan et al. (2007):

**Animacy of themes and recipients** Syntactic structures can be sensitive to animacy: a number of studies have found it to be an important factor in the choice of genitive constructions in English, for example Rosenbach (2003). Animacy has also been identified as an important factor in the dative alternation of German-speaking children (Drenhaus & Féry, to appear), and also in earlier corpus studies of English (e.g., Thompson 1990).

Children in the 2-5 year-old age group distinguish animate from inanimate NPs in a largely adult-like manner, both in linguistic tasks (Becker 2007) and in non-linguistic, conceptual tasks (Massey & Gelman 1988). In order to verify this, we also coded for whether a particular theme/recipient was a toy, just in case toys had any particular properties (e.g., being treated more like animates than inanimates). However toys did not differ significantly from inanimates in their effect on construction choice; therefore, the

---

<sup>3</sup>We removed the data points where the theme and the recipient did not occur postverbally, i.e., in instances of topicalization, question formation or passivization. We also removed data which did not have both a theme and a recipient:

- There were 221 utterances that did not have a theme, e.g., “I give you win no tickling” [Abe, 4;3.11]. Only one of these suffered from a partially unformed theme: “I'm gonna show you what +...” [Adam, 4;7.0].
- There were 150 utterances that had a theme but did not have a recipient, e.g., “You give nice lollipops” [Naomi, 2;5.8]. Only one of these had a partially-formed recipient (“I going(g) show it to my +...” [Adam, 4;2.17]), all the others we eliminated simply did not have any recipient at all.

animacy variable only takes into account the opposition between true animates and inanimates in our investigations.

**Length of themes and recipients** Length has long been noted as an important factor in adult speech, for example, heavy NP shift places a longer constituent at the end of the clause (Behagel 1909, Wasow 2002, Bresnan et al. 2007). We measured this factor in terms of the number of words.<sup>4</sup>

**Nominal Expression Type** The choice of a pronoun over a full NP has been known to affect the acceptability of and the preference for the different dative constructions (Green 1971, Collins 1995). We coded for the nominal expression type of themes and recipients in the following way. Pronouns include:

- definite pronouns (*it*, “show it to Mike” [Abe, 2;8.6])
- demonstratives (*that*, “I I gave Bruno that for that to sleep with” [Nina, 3;2.12])
- personal pronouns (e.g., *me*, *them*, “yeah an(d) den after our truck will [?] give dem back to Marianne” [Shem, 3;0.13], including pronouns followed by a lexical NP, such as in “she gave them all her children a spanking” [Naomi, 3;3.27])
- reflexive pronouns (e.g., *myself*, “I give the bag to myself” [Adam, 3;7.7])

Names, sentential complements, verb phrases, adverbs, indefinite pronouns (*something*, *any*, “I if if I gave you some, you I will gwab [:grab] it away” [Trevor, 2;8.10]) and wh-words (e.g., “I’ll show you where” [Abe, 3;10.7]) were categorized as non-pronouns.

**Givenness** A number of authors have shown the importance of information structure in dative constructions – given information typically comes before new information (Collins 1995, Arnold et al. 2000, Wasow 2002, Snyder 2003, Bresnan et al. 2005). We therefore coded whether a theme or a recipient had been mentioned in the previous 10 speaker turns.

**Persistence** Researchers have long known that repetition and parallelism play a role in how people choose a construction: speakers reuse what they have just heard or just used. Syntactic priming effects have also been reported in young children (see Savage et al. 2003, Huttenlocher et al. 2004, Conwell and Demuth 2007, and the references therein). Most recently, Szmrecsanyi (2005) studied structural persistence from a corpus-based, variationist perspective. He found that persistence plays a significant role in linguistic choice for three different English alternations: analytic vs. synthetic comparatives, particle placement, and future marker choice.

We encoded the persistence factor in the following way: we examined the 10 previous speaker turns and when a dative construction was found, we marked the speaker of that construction (adult vs. child) and counted its distance from the current dative construction by the number of clauses.

---

<sup>4</sup>We also considered the possibility that phonological length would be a more appropriate measure for children’s speech, in part since children use fewer words in their utterances. We approximated phonological length by counting the number of syllables. However, this measure did not perform significantly differently from the standard measure in word length. Therefore, we used word length as the unit of measurement, which also facilitated the comparison with the adults.

$$\text{Probability \{Response = NP PP\}} = \frac{1}{1+e^{-X\beta}} \text{ where:}$$

$$\begin{aligned}
X\beta &= 0.1916 && + \\
&-0.7092 \cdot \{\text{the number of words in the theme}\} && + \\
&2.3717 \cdot \{\text{nominal expression type of the theme = pronoun}\} && + \\
&-1.5391 \cdot \{\text{nominal expression type of the recipient = pronoun}\} && + \\
&-1.4154 \cdot \{\text{previous NP NP construction in the last ten turns = yes}\} && + \\
&1.8178 \cdot \{\text{previous NP PP construction in the last ten turns = yes}\} && +
\end{aligned}$$

Table 1: The Model Formula

**Age and MLU** We consider it likely that some of our measures (like length of theme/recipient) could be confounded with developmental advances allowing for longer utterances overall. Since there is considerable variation among children, age is not a sufficient measure of developmental progress. One of the standard metrics used since Brown (1973) is the mean length of utterance (MLU), which attempts to capture the syntactic complexity of children’s utterances. The CLAN program, which is linked to the CHILDES database, makes it fairly straightforward to compute the MLU for each recording session in CHILDES. We added this information to the data. However, consistent with recent research in language acquisition (Legendre 2006), none of these measures proved to be significant.

### 3.2 Resulting model

The final logistic regression model, computed with the R statistical modeling software, is summarized in the formula in Table 1. Four factors turned out to be significant ( $p < .05$ ): length in words of the theme, nominal expression type of theme and recipient, and structural persistence. The other factors lacked predictive value and were eliminated from the final model shown in Table 1.

The model predicts the likelihood of the prepositional construction, stating the baseline, the intercept value, and quantifying the influence of each factor, viz. the coefficients in the formula (see Table 1). Any positive value for a coefficient increases the likelihood of the prepositional construction. Thus the values of the coefficient of the previous NP NP construction and the length of the theme are negative as they decrease the odds of realizing a NP PP construction. The length of the recipient and the nominal expression type of the theme have positive coefficient values, as they increase the odds of the NP PP construction being used. The relative odds of each factor can be seen in Table 2, as well as the detailed p-values. This model predicts the choice of the dative construction with a high accuracy (90.7%). Also the model captures 56.7% of the variance (Nagelkerke  $R^2$ ).<sup>5</sup>

### 3.3 Discussion

The model delivers not only information about which factors are predictive, but also the strength of their predictive power measured in terms of log odds. The model predictions are shown in Figures 1-4. Figure 1 shows the effect of the previous dative construction, which tends to persist. If the previous construction was a prepositional dative (PP), the current construction is more likely to be a prepositional dative than if

<sup>5</sup>This value was validated against overfitting employing the bootstrap validation technique, under which the variance captured remained 56.2%.

Factor	Odds	P-Value
theme type=pronoun	10.72	0.0000
recipient type=pronoun	0.21	0.0000
theme length	0.49	0.0073
previous dative=NP	0.24	0.0003
previous dative=PP	6.16	0.0000

Table 2: Odds and P-Values of the Significant Factors

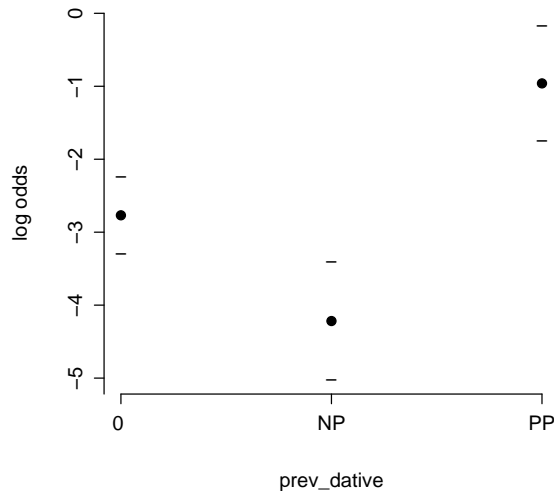


Figure 1: Probability of Prepositional Dative Given Persistence

there was no previous dative construction. Conversely, if the previous construction was a double object (NP), the current construction is less likely to be a prepositional dative than if there was no previous dative construction. The influence of the length of the theme on the probability of the occurrence of the prepositional dative becomes more feeble as it increases, as the linear relationship shows in Figure 2. If the theme is realized as a pronoun, the probability of the prepositional dative is greater than if the theme is realized as a lexical NP (Figure 3). Figure 4 shows the inverse: if the recipient is realized as a pronoun, the probability of the prepositional dative is less than if the recipient is realized as a lexical NP.

The model illuminates the relationship of harmonic alignment within the children’s production data, parallel to the findings in Bresnan et al. (2007). Thus far, we have discussed the factors in terms of predicting the likelihood of the prepositional dative construction, but these results can also be viewed from the perspective of predicting which syntactic position the argument will fill. For instance, referring to Figure 4, if the recipient is pronominal, it favors the double object construction (i.e., it disfavors the prepositional dative construction), in which the recipient appears in first position. Similarly, if the theme is pronominal, it favors the prepositional dative construction, whereby the theme will appear in first position. Clearly then, there is a correspondence between the values of the factors and syntactic

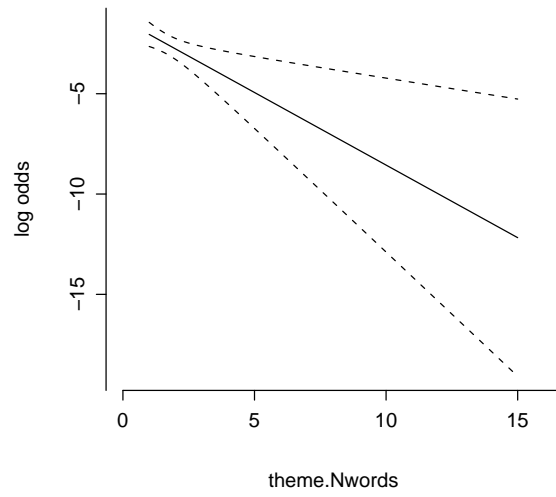


Figure 2: Probability of Prepositional Dative as a Function of the Length of the Theme

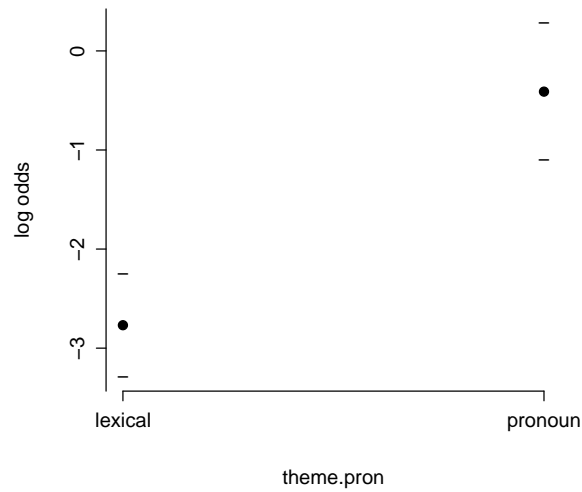


Figure 3: Probability of Prepositional Dative as a Function of the Nominal Expression Type of the Theme

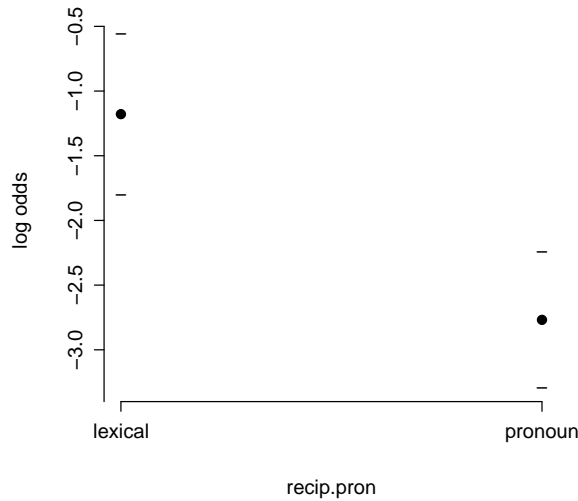


Figure 4: Probability of Prepositional Dative as a Function of the Nominal Expression Type of the Recipient

given	>	<b>not given</b>
shorter	>	<b>longer</b>
pronoun	>	<b>non-pronoun</b>
V NP NP		V recipient <b>theme</b>
V NP PP		V theme <b>recipient</b>

Figure 5: Harmonic Alignment Effects Observed in Child Dative Constructions

position: shorter and pronominal NPs (more prominent) align with the first syntactic position of the dative construction, while longer and non-pronominal NPs align with the second position, as schematically represented in Figure 5.

### 3.4 Mixed-effect model

The results above assumed that the data is homogenous in terms of speaker variation. However this is not necessarily a safe assumption since children are known to vary widely in their individual developmental trajectories (Clark 2003).

We can examine the legitimacy of this assumption by introducing individual children as random effects in the model, a technique known as “mixed effects modeling” (Pinheiro and Bates 2000). Taking each child as a random effect will show whether there is significant variation among the production patterns of the children. In essence, the mixed-effect model makes an adjustment for each child representing that child’s individual bias towards the prepositional dative construction.

Surprisingly, the global trends reported above hold locally for each child, both in terms of direction and magnitude of response. That is, using each child as a random effect did not lead to a significant



Factor	Fixed effect model coefficients	Mixed effect model coefficients
(Intercept)	+ 0.19	+ 0.21
theme type=pronoun	+ 2.37	+ 2.36
recipient type=pronoun	- 1.54	- 1.56
theme length	- 0.71	- 0.71
previous dative=NP	- 1.42	- 1.43
previous dative=PP	+ 1.82	+ 1.81

Table 3: Intercept and Coefficients of the Fixed-effect and Mixed-effect Models

Child	Intercept
abe	+ 0.028
adam	- 0.059
naomi	- 0.040
nina	+ 0.119
sarah	- 0.071
shem	+ 0.067
trevor	- 0.042

Table 4: Intercept Adjustments for each Child in the Mixed-effect Model

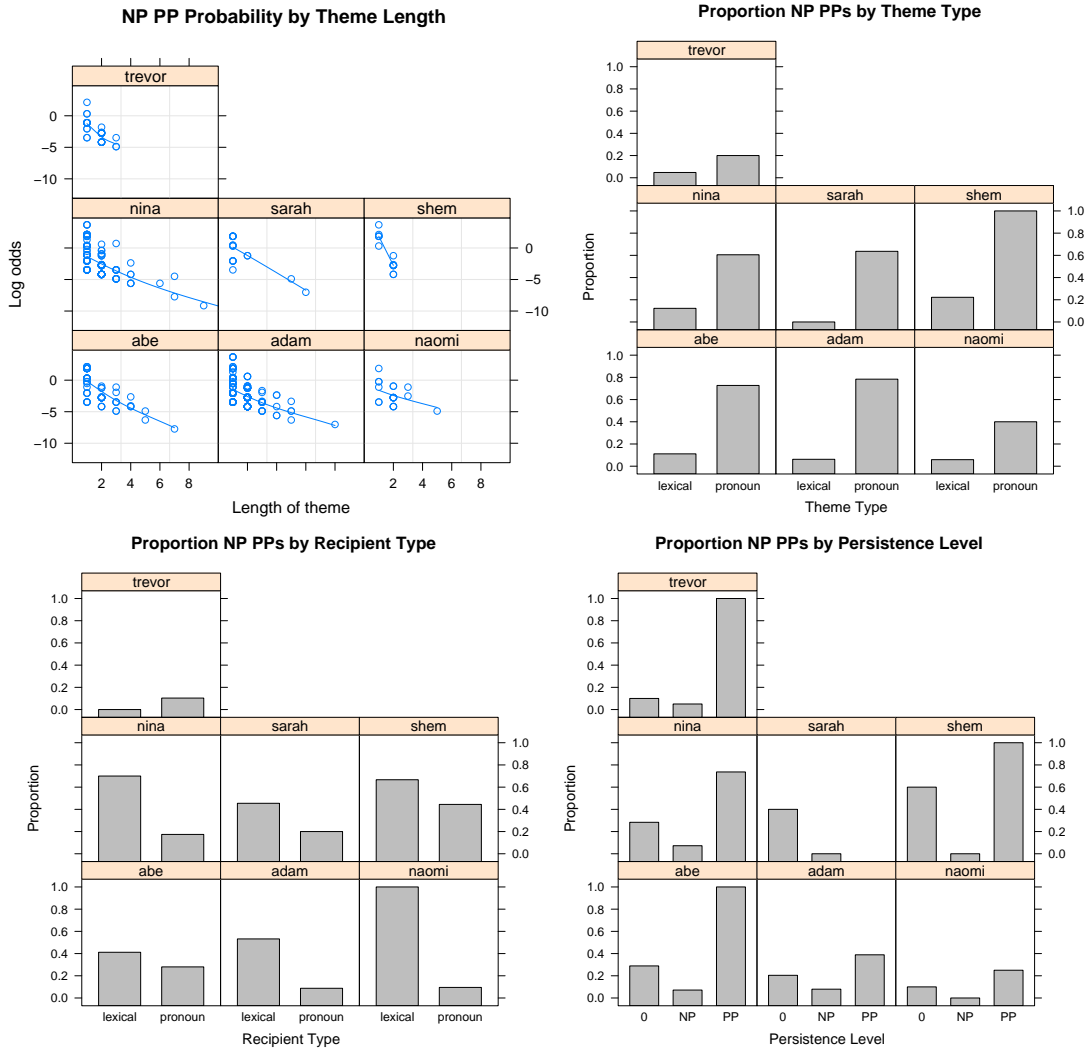


Figure 6: Effects of the Factors by Child

difference. The magnitude of the responses visibly varies by child, as can be seen in Figure 6. However the model informs us that this variation is not significant: Table 3 displays the coefficients for the fixed- and mixed-effect models, respectively, and Table 4 gives the intercept adjustments by child. Moreover, as the graphs show (Figure 6), the direction of the response is constant by child. The sole exception is Trevor’s greater production of prepositional datives with pronominal recipients (lower left corner). Yet, this exception is only apparent and not significant.<sup>6</sup>

### 3.5 Interim conclusion

The findings show that the overall picture of child production of dative sentences is much the same as what Bresnan et al. (2007) found for adults. First, construction choice is governed by multiple factors, in particular a subset of the factors found in the adult model of Bresnan et al. (2007). Second, the significant factors in the child model align harmonically. Finally, these findings are robust and found

<sup>6</sup>Fisher’s exact test,  $p = 1$ .

across the entire sample from CHILDES.

## 4 Comparison with adults

While the model of child production of the previous section follows the trend one would expect from Bresnan et al. (2007), there are some salient differences between the child and adult models. In particular, the models differ in the number of factors that are significant, with the child model having fewer factors. Also, contrary to our expectations, animacy was not found to be a significant factor in the child model.

Given that we have a quantitative model of both child and adult production, we can compare the responses of both populations, assessing the similarities and differences between the populations in a quantitative fashion.

### 4.1 Conjoined model

One cannot draw conclusions about the similarities and differences among different populations from isolated models of different datasets, because there is no way to determine whether the model differences are significant. To properly compare child and adult production of dative sentences, we constructed a conjoined model pooling the data together from both studies, and examined how the group variable (children vs. adults) interacted with the other predictors.

To make the adult data comparable to that of the children, we limited the adult data to the verbs “give” and “show”, yielding 611 data points, comparable to the 530 occurrences for the child data. Also, we re-coded persistence in the adult data to approximate the 10 speaker-turn range operationalized in the child data, making use of “persistence” as annotated in the spoken dative database developed at Stanford in the summer of 2006 (Recchia 2006). To ensure that restricting the adult data to only two verbs does not compromise the findings of Bresnan et al. (2007), we re-ran their model on this restricted data set. Apart from the verb type, we found no differences in the main effects, e.g., animacy remained significant. The verb type ceased to be significant simply because there was no longer enough variation. Assured of the integrity of the model, we were then able to pool the child and restricted adult data together.

The conjoined model obtained a high accuracy (95.7%), and captures 70.3% of the variance (Nagelkerke  $R^2$ ).<sup>7</sup> Table 5 shows the conjoined model, in terms of odds, as well as listing the p-values. The table also lists five main effects as well as two interactions between population type and nominal expression type of the theme and recipient.

### 4.2 Discussion

The conjoined model again confirms the direction of the responses across both populations. As can be seen in Figure 7, the persistence of previous dative construction in the conjoined model has the same configuration as that of the previous model.

Similarly, the effects of length of the recipient and the theme maintain the same quantitative relations noted in the previous model. An increase of recipient length favors the prepositional dative, while a decrease in theme length favors the prepositional dative (Figure 8).

---

<sup>7</sup>The bootstrap validation technique confirmed the Nagelkerke  $R^2$  value (69.2%).

Factor	Odds	P-Value
intercept	0.284	0.0824
recipient type=pronoun	0.021	0.0000
theme type=pronoun	1536.0	0.0000
recipient length	2.6	0.0021
theme length	0.646	0.0008
previous dative=NP	0.240	0.0000
previous dative=PP	5.5	0.0000
<b>group=child * recipient type=pronoun</b>	11.0	0.0073
<b>group=child * theme type=pronoun</b>	0.008	0.0000

Table 5: Odds and P-Values of Factors in the Conjoined Model

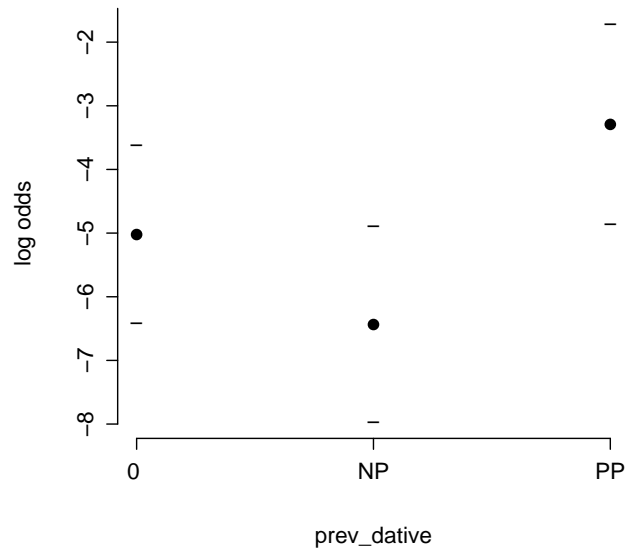


Figure 7: Persistence in the Conjoined Model

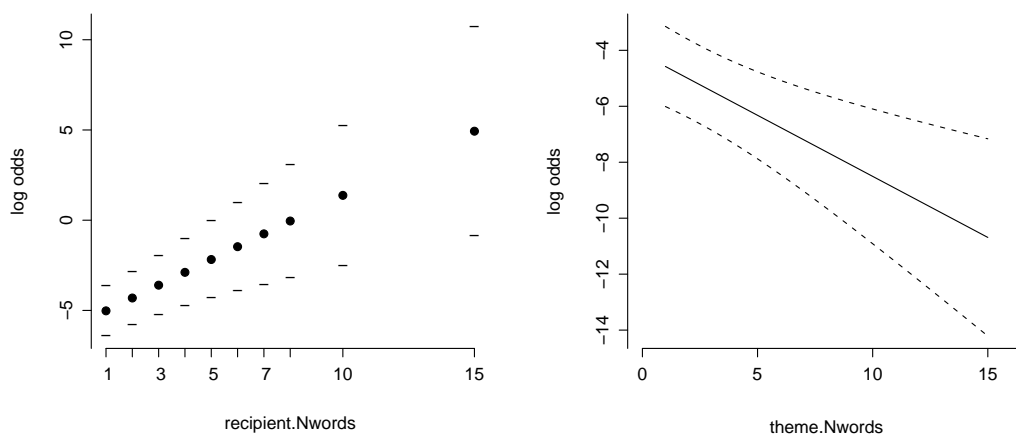


Figure 8: Effect of Length of Recipient and Theme

The effect of the nominal expression type of the recipient and theme also conformed to the pattern of the previous model: a lexical recipient favors the prepositional dative construction as does a pronominal theme (Figure 9).

While the overall patterns of the previous model are echoed in the conjoined model, interaction effects show that the two populations differ in their sensitivity to the shared factors. Figure 10 shows that the directions of the effects are the same, but the children and adults differ in the degree to which the factor influences their choice. In particular, the nominal expression type of the recipient and theme has a greater influence on the adults' production choice.

## 5 Conclusion

The models of children's production described in this paper suggest that children's construction choices for dative are founded, at least in part, in factors to which adults are also sensitive. The production choices made by children and adults are neither identical nor radically different, rather a core set of factors are shared. Moreover, the conjoined model indicates that, for the dative construction, child speech only differs from adult speech in degree, not in kind.

These findings resonate with much current work in language acquisition which contends that there is a continuity between the grammars young children use and those used by adults, namely that "the child's developing grammar is hypothesized to be always a grammar that is a possible adult grammar" (Goodluck 2007, p. 48). Further, the models shown here demonstrate that children's production patterns echo the probabilities of adult production patterns, which is unexpected if children are assumed to go through a period in which they maximalize to only one of the alternation's variants. Some studies have shown evidence that children fare worse on probability matching tasks than adults (Hudson Kam & Newport 2005; see discussion in Ramsar & Gitcho 2007); however, the naturally-occurring data considered here manifests an apparent sensitivity on the part of the children to production probabilities.

This study has demonstrated the feasibility of comparing child and adult speech, and shown that

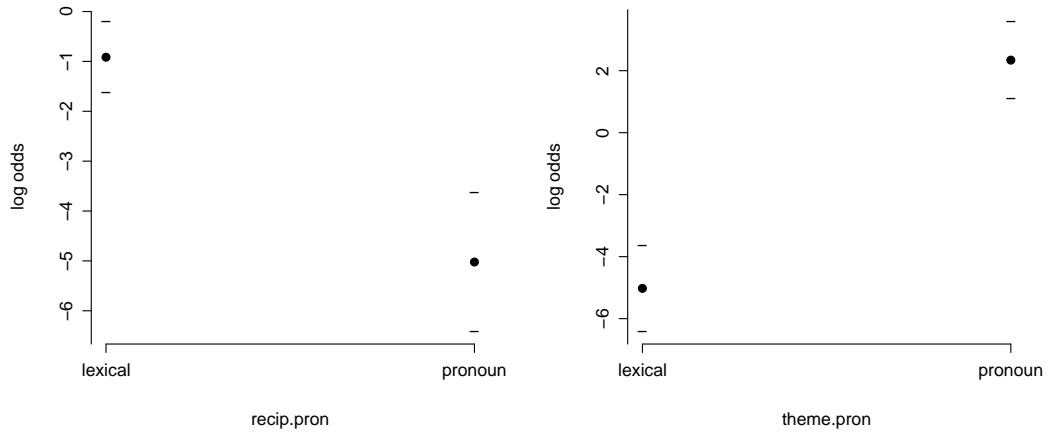


Figure 9: Effect of Nominal Expression Type of Recipient and Theme

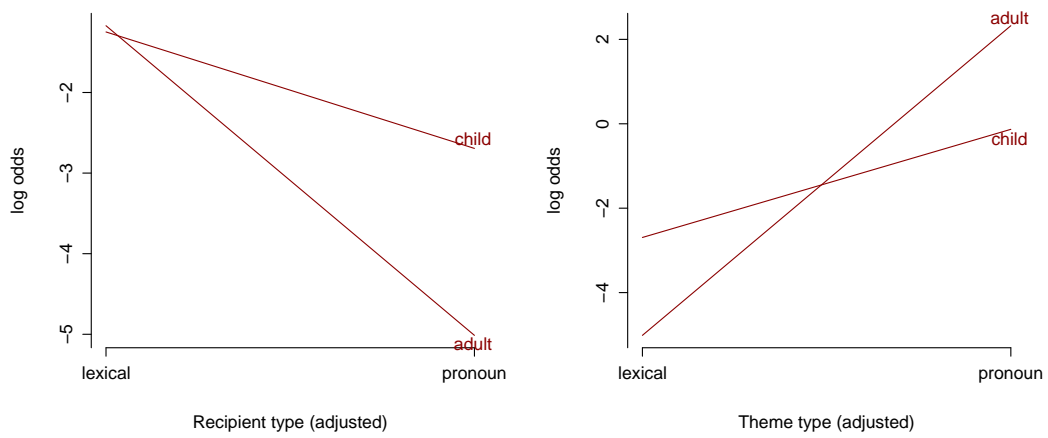


Figure 10: Interaction Effects

statistical modeling techniques can yield insight into the factors at play in children's speech production. Further research may shed light upon why the differences between these patterns of production were observed, for instance by exploring interactions with the learning environment, i.e., child-directed speech, and with processing capacities, such as resource limitations. Given the size of the corpus, our results are promising rather than definitive, yet already indicate that new evidence can be brought to bear on the acquisition of alternations using quantitative modeling methods.

## References

- ARNOLD, J., WASOW, T., LOSONGCO, A., & GINSTROM, R. 2000 Heaviness vs. newness: The effects of complexity and information structure on constituent ordering. *Language*, **76**, 28–55.
- BAAYEN, H. 2006 *Practical Data Analysis for the Language Sciences with R*. (forthcoming).
- BECKER, M. 2007 Animacy, Expletives, and the Learning of the Raising-Control Distinction. In BELIKOVA, A., MERONI, L., & UMEDA, M. (editors), *Generative Approaches to Language Acquisition North America 2*, 12–20.
- BEHAGEL, O. 1909 Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen*, **25/110**.
- BRESNAN, J., CUENI, A., NIKITINA, T., & BAAYEN, H. 2007 Predicting the Dative Alternation. In BOUME, G., KRAEMER, I., & ZWARTS, J. (editors), *Cognitive Foundations of Interpretation*, Amsterdam: Royal Netherlands Academy of Science.
- BROWN, R. 1973 *A First Language*. Harvard University Press, Cambridge, MA.
- CLARK, E. 2003 *First Language Acquisition*. Cambridge University Press.
- COLLINS, P. 1995 The indirect object construction in English: an informational approach. *Linguistics*, **33**, 35–49.
- CONWELL, E. & DEMUTH, K. 2007 Early syntactic productivity: Evidence from dative shift. *Cognition*, **103**, 63–179.
- DRENHAUS, H. & FÉRY, C. 2007 Animacy and Child Grammar: an OT account. In LAMERS, M., LESTRADE, S., & DE SWAR, P. (editors), *Lingua special issue on animacy*, to appear.
- GODFREY, J., HOLLIMAN, E., & MCDANIEL, J. 1992 SWITCHBOARD: Telephone speech corpus for research and development. *Proceedings of ICASSP-92, San Francisco*, 517–20.
- GOODLUCK, H. 2007 Formal and computational constraints on language development. In HOFF, E. & SHATZ, M. (editors), *Blackwell Handbook of Language Development*, 46–67, Blackwell Publishing.
- GREEN, G. 1971 Some implications of an interaction among constraints. In *Papers from the Seventh Regional Meeting, Chicago Linguistic Society*, 85–100, The University of Chicago.

- GRIES, S. 2003 Towards a corpus-based identification of prototypical instances of constructions. *Annual Review of Cognitive Linguistics*, **1**, 1–28.
- HINRICHS, L. & SZMRECSANYI, B. 2007 Recent changes in the function and frequency of Standard English genitive constructions: a multivariate analysis of tagged corpora. *English Language and Linguistics*, **11**(3).
- HUDSON KAM, C. L. & NEWPORT, E. L. 2005 Regularizing Unpredictable Variation: The Roles of Adult and Child Learners in Language Formation and Change. *Language Learning and Development*, **1**(2), 151–195.
- HUTTENLOCHER, J., VASILYEVA, M., & SHIMPI, P. 2004 Syntactic priming in young children. *Journal of Memory and Language*, **50**, 182–195.
- JAEGER, T. F. 2006 *Redundancy and Syntactic Reduction in Spontaneous Speech*. PhD thesis, Stanford University Department of Linguistics.
- LEGENDRE, G. 2006 Early Child Grammars: Qualitative and Quantitative Analysis of Morphosyntactic Production. *Cognitive Science*, **30**:5, 803–835.
- MACWHINNEY, B. 2000 *The CHILDES Project: Tools for analyzing talk. 3rd Edition. Vol. 2: The Database*. Lawrence Erlbaum Associates.
- MASSEY, C. & ROCHEL, G. 1988 Preschooler's ability to decide whether a photographed unfamiliar object can move itself. *Developmental Psychology*, **24**, 307–317.
- PINHEIRO, J. C. & BATES, D. M. 2000 *Mixed-Effects Models in S and S-PLUS*. Springer-Verlag, New York.
- R programming language 2004 R: A language and environment for statistical computing, Development Core Team. <http://www.R-project.org>.
- RAMSCAR, M. & GITCHO, N. 2007 Developmental change and the nature of learning in childhood. *TRENDS in Cognitive Science*, **11**:7, 274–279.
- RECCHIA, G. 2006 The Spoken Syntax Lab. Presented at the Stanford Linguistics Department Colloquium, Stanford University, California. 20 Oct 2006.
- ROLAND, D., ELMAN, J. L., & FERREIRA, V. S. 2006 Why is that? Structural prediction and ambiguity resolution in a very large corpus of English sentences. *Cognition*, **98**, 245–272.
- ROSENBACH, A. 2003 Iconicity and economy in the choice between the 's genitive and the of-genitive in English. In ROHDENBURG, G. & MONDORF, B. (editors), *Determinants of Grammatical Variation in English*, Berlin/New York: Mouton de Gruyter.
- SAVAGE, C., LIEVEN, E., THEAKSTON, A., & TOMASELLO, M. 2003 Testing the abstractness of young children's linguistic representations: Lexical and structural priming of syntactic constructions? *Developmental Science*, **6**-5, 557–567.



- SNYDER, K. 2003 *The relationship between form and function in ditransitive constructions*. PhD thesis, Philadelphia: University of Pennsylvania Department of Linguistics.
- SZMRECSANYI, B. 2005 Language users as creatures of habit: A corpus-based analysis of persistence in spoken English. *Corpus Linguistics and Linguistics Theory*, **1**, 113–150.
- THOMPSON, S. 1990 Information flow and dative shift in English discourse. In EDMONDSON, J. A., CRAWFORD, F., & MUHLHAUSLER, P. (editors), *Development and Diversity, Language Variation Across Space and Time*, 239–253, Summer Institute of Linguistics, Dallas, Texas.
- WASOW, T. 2002 *Postverbal Behavior*. Stanford: CSLI.