



MIT
Working Papers
in Linguistics



MIT Working Papers in Linguistics 82

Edited by Elitzur A. Bar-Asher Siegal

Proceedings of IATL 2015

PROCEEDING OF IATL 31

Israel Association for Theoretical Linguistics

TABLE OF CONTENTS

STRESSED VS. FINAL SYLLABLE IN EARLY SPEECH: WHICH ONE IS STRONGER? <i>Avivit Ben-David and Outi Bat-El</i>	1
THE ACQUISITION OF PRONOUNS IN ENGLISH – IS THERE EVIDENCE FOR A UNIFORM SEMANTICS OF PRONOUNS, INDEXICALS AND THE DEFINITE DETERMINER? <i>Saskia Brockman</i>	15
FUNCTION VARIABLES IN METAPHORIC INTERPRETATION: EVIDENCE FROM ICONICITY IN SIGN LANGUAGES <i>Ariel Cohen and Irit Meir</i>	41
(PSEUDO-)EXTRACTION FROM VP-PROFORMS: THE CASE OF THE COMPARATIVE CLAUSE <i>Vera Hohaus and Andreas Konietzko</i>	53
ITALIAN VERB PARTICLES AND CLAUSAL POSITIONS <i>Stefano Quaglia and Andreas Trotzke</i>	67
GRANULARITY AND APPROXIMATING NUMBER PAIRS <i>Stephanie Solt</i>	83
A PRAGMATICALLY ENRICHED TEMPORAL RELATION <i>Lucia M. Tovená and Xiaoqian Zhang</i>	101
SCALAR PARTICLES AND CONTRASTIVE TOPICS IN ENGLISH AND VIETNAMESE <i>Malte Zimmermann</i>	123

THE ACQUISITION OF PRONOUNS IN ENGLISH – IS THERE EVIDENCE FOR A UNIFORM SEMANTICS OF PRONOUNS, INDEXICALS AND THE DEFINITE DETERMINER?

SASKIA BROCKMAN
University Tübingen

1 Introduction

The present corpus study that looked for the order of production within five longitudinal corpora of monolingual English-speaking children is a first attempt towards combining and testing formal semantic theory with acquisitional data. More specifically, the main aim that this research contributes to is to find evidence for or against two theoretical approaches to the syntax and semantics of pronouns and the proposed uniformity of pronouns and definites in one of the approaches. In recent years, formal semantics has come forward with new suggestions about the relation between pronouns and definites. Acquisitional research already knows a lot about the production and comprehension of both phenomena, but so far, there haven't been many studies that concentrate on their correlation. A number of theories that I subsume under the term 'the classical view' define the definite determiner as a presupposition trigger of uniqueness, while pronouns and demonstratives are formalized as variables that receive their meaning through a pragmatically motivated, contextually available assignment function (Heim and Kratzer 1998). Indexicals in turn are seen as contextual parameters (Kaplan 1989). On the other hand, there are proposals that argue for a uniform treatment of these phenomena (cf. Elbourne 2005, 2013; Postal 1966), saying that they share the semantics of the definite article. In the case of pronouns, the missing NP complement is subject to NP deletion. Indexicals have a slightly more complex semantics, and demonstratives show flexibility when it comes to picking out an overt or covert NP complement. From a purely syntactosemantic point of view, production data of children should demonstrate i) if there is a correlation between these phenomena and ii) if there is an observable, empirically stable order of production. The main hypothesis tested here will mainly be concerned with evidence for or against the uniform view, as the classical view doesn't propose any correlation of the phenomena on purely syntactosemantic grounds. The uniform view suggests the following hypothesis:

(1) **The Uniform View Hypothesis:**

The definite article isn't produced significantly¹ later than pronouns, indexicals or demonstratives.

We will see that the results of this study do not provide clear evidence for or against this hypothesis, which suggests that the pragmatic meaning component of the phenomena is too strong in order to find fine-grained structural differences, at least by means of corpus studies. These results point towards the importance of formal semantic theory when explaining natural occurrences of the phenomena they provide frameworks for. Specifically, in the case of pronouns and definites, it may be the case that the strong pragmatic correlation of picking unique referents out of the discourse context may override semantic differences in natural language data. It is therefore important to see whether the subtle differences between the respective frameworks are even observable in natural language use.

By assuming that children are conservative in their language production, measuring the order of production for related phenomena can give strong insight into inherent learning mechanisms and thus also the systematic build-up of grammar. To be conservative as a learner means to produce only those things that you already understand in their complexity. Taking this assumption as a precondition, two additional logical assumptions about language production follow:

- (2) **Concurring Production:** If the grammatical knowledge (including parameter settings and lexical information) required for construction A, in a given language, is identical to the knowledge required for construction B, then any child learning the language is predicted to acquire A and B at the same time. (Snyder 2007:7)
- (3) **Ordered Production:** If the grammatical knowledge (including parameter settings and lexical information) required for construction A, in a given language, is a proper subset of the knowledge required for construction B, then the age of acquisition for A should always be less than or equal to the age of acquisition for B. (No child should acquire B significantly earlier than A.) (Snyder 2007: 7)

According to (2) and (3), the theoretical claims about syntactic and semantic structures of constructions can be tested using production data: They can be ordered according to their complexity. If a construction A and a construction B share the same syntactic and semantic setup, then we can systematically assume that they will have to be acquired simultaneously. If the syntactic and semantic setup of A is necessary for B, but there are additional things necessary for B, then we can systematically assume that B can't be acquired earlier than A. According to the syntactic-semantic structure proposed for pronouns in the uniform view, the data should demonstrate ordered production. However, this is only the case from a purely syntactosemantic perspective. The results of this study show that the pragmatic dimension of grammatical knowledge must necessarily be a part of the hypotheses regarding ordered production.

¹ Significance is defined here as a statistical significant value as a result of non-parametric statistical analysis conducted with the help of Binominal Tests. A significant p-value is defined here, as is the norm, as being below 0.01.

2 Theoretical background

In the following, I will concentrate on two main accounts of the semantic representations of pronouns, the definite determiner, and indexicals and demonstratives: The classical account is a conglomerate of various analyses of the phenomena, whereas the uniform account represents work initiated by Postal (1966) and taken up, among others, by Elbourne (2005, 2013). In this section, I will give a short and as simple as possible introduction to these analyses and conclude with their predictions for production in first language acquisition.

2.1 The classical view

I will call the analysis put forward in Heim and Kratzer (1998) and Kaplan (1989) the ‘classical view’: In this view, pronouns are interpreted as variables that are assigned a contextual referent through a variable assignment function that maps indices to contextually available individuals:²

- (4) a. John is asleep. He works a lot.
 b. $[[\text{he}_1]]^{g,c} = g(1) = \text{John}$
 c. $[\text{works-a-lot}] (\text{John})$

The pronoun ‘he’ in the second sentence in (4a) is interpreted through the variable assignment g that picks out the index 1 and maps it to the contextually available individual John³. For simplicity’s sake, I will assume that further information provided by the pronoun, such as gender and case, are introduced as presuppositions, see (5), however, I will mostly ignore them:

- (5) $[[\text{he}_1]]^{g,c} = g(1)$ is male. $g(1)$

For the definite determiner, Heim and Kratzer (1998) follow Frege’s original idea in positing an analysis for the definite by treating it as a presupposition trigger:

- (6) a. The Queen of England is a great-grandmother.
 b. $[[\text{the}]]^{g,c} = \lambda f_{\langle e,t \rangle} : \exists! x [f(x)=1]. \iota x [f(x)=1]$
 c. $[\text{is a great-grandmother}] (\text{the unique } x \text{ such that } x \text{ is the Queen of England})$

The definite determiner picks out its NP complement and requires that there be a unique individual in the context for whom the property applies⁴. Thus, for the sentence in (3(6)), the

² There is a further pragmatic restriction on the usage of pronouns: Not only do they have to refer to contextually available referents, but those have to be salient in the context (see Büring (2012) for an overview). This saliency-condition can be implemented as a presupposition to the pronoun, similar to gender and case requirements. I will come back to this issue in the course of the paper.

³ Note that here, and in the rest of the paper, I will concentrate on free and/or anaphoric uses of pronouns, as they are acquired first. I will stay agnostic about other kinds of pronoun uses (e.g. bound pronouns), although both frameworks provide analyses for such cases.

⁴ I will stay agnostic about the state of the familiarity-condition that is often associated with the definite article, even though it could be implemented in this framework. Moreover, I will assume a rather vague definition of what it means to be unique: The referent has to be available within the context of the sentence, without it having been mentioned as an indefinite NP or a DP of another sort in the previous discourse. For a more detailed discussion on uniqueness and familiarity, please see Roberts (2005).

definite determiner requires that there be one unique queen of England for the sentence to be felicitous. According to our current world knowledge of March 2016, that is indeed the case: the current Queen of England is Elizabeth II, and it so happens that she is a great-grandmother.

Indexicals are not discussed in Heim and Kratzer 1998, but I will take Kaplan 1989 as a classical analysis assigned to the phenomenon⁵:

- (7) a. I am hungry.
 b. $[[I]]^{g,c} = c_s$
 c. [hungry] (c_s)

Kaplan argues that ‘I’ and ‘you’, unlike third person pronouns, have fixed values with respect to the context: ‘I’ can only refer to the current speaker, whereas ‘you’ can only refer to the current addressee. Both are thus analyzed as parameters provided by the context. Other parameters of the context are the time, the world and the location of an utterance. In that framework, it is not enough to let the variable assignment do the work, as there wouldn’t necessarily be restrictions on which referent the function will map to the index⁶. Rather, they refer back to individuals that are necessarily part of every conversation, if they are directly referred to or not. Accordingly, the example in (7a) can only be felicitous if the current speaker in the context is hungry. Let’s assume that I myself, Saskia, am uttering this sentence while writing this article at 10.45 am on Thursday, February 18. As this is a reasonable time after having started working, I am indeed already hungry: The sentence is true. However, as soon as the sentence is uttered in another context where someone else is the speaker at a different time of day, it may well be false.

Taking stock of the classical account, we see that all three phenomena are context sensitive but integrate the context in different ways: The definite article comes with a presupposition, so it poses a requirement on the context, while pronouns refer to contextually available referents picked out by an assignment function, and indexicals are an internal part of the context. Thus, strictly speaking, there is no systematicity between the three analyses. However, they all pick out unique and specific referents of the context – a notion that I will come back to at the end of the paper.

2.2 The uniform view

Current theoretical semantic research, e.g. Schwarz (2009) and Elbourne (2005, 2008, 2013), has come back to Postal’s (1966) suggestion to treat pronouns and definites on a par: The reasoning is that, in many cases, the two phenomena are interchangeable and contribute similar truth-conditions. Furthermore, cases have been attested where it seems to be infelicitous to assign a specific referent to a pronoun in the context:

- (8) This year the president is a Republican, but one fine day, he (‘the president’) will be a member of the Green party. ((21c) in Buring (2012))

⁵ I will focus on personal indexicals and assume that for temporal and spatial indexicals, production in FLA will only follow when children have acquired temporality and deixis, respectively. These points in acquisition may overlap, but are beyond the scope of the present enterprise.

⁶ However, there are analyses that assign the first and second person pronoun an additional presupposition that says that the referent picked out by the variable assignment g has to be speaker or addressee in c , respectively: see Buring (2012).

In (8), ‘he’ seems to refer not to the current president, but to a not yet determined individual that will be president in some future, or alternate time. The sentence highlights that treating ‘he’ as a definite would save the semantic composition from resulting in an infelicitous utterance. Taking the semantics for the definite determiner in (3b) as a basis for the uniform view, third person pronouns would be modelled in exactly the same way, ignoring additional presuppositional information about case and gender for now:⁷

$$(9) \llbracket \text{the/he/it} \rrbracket = \lambda f_{\langle e,t \rangle} : \exists ! x [f(x)=1]. \iota x [f(x)=1]$$

The NP complement of pronouns is argued to be subject to NP-deletion:

- (10) a. The cat is sleeping. It snores.
 b. $[\text{IP} [\text{DP it } \text{cat}] [\text{VP snores}]]$
 c. $[\text{snores}]$ (the unique x such that x is a cat)

The second sentence in (10a) would thus be assigned an LF-structure as given in (10b), with the NP ‘cat’ being deleted, and a semantic analysis as given in (10c).

Although recent scholars have focused on the parallels between pronouns and definites, not much has been said regarding a possible uniformity of pronouns/definites and indexicals. Rather, it has been argued that the analysis of indexicals as contextual parameters might not capture the whole range of meanings an indexical can convey. Consider the example below:

- (11) I am traditionally allowed to order whatever I like for my last meal.

Kratzer (2009) and Nunberg (1993) correctly remark that in (11) and similar examples, the first person indexical ‘I’ can’t refer to only the speaker. Rather, it seems that the speaker identifies with a group for which the statement is also true. Both instances of ‘I’ above seem to quantify over individuals, probably triggered by the modifier ‘traditionally’. According to Kaplan, binding indexicals shouldn’t be possible, as they always refer to a fixed individual, the speaker. So how can we account for the sentence in (11)? Grosz and Zobel (2014) propose an analysis along the lines of Elbourne’s (2008) analysis for demonstratives: Indexicals are also treated as definite determiners and receive a complex argument, consisting of a contextually retrievable relation R and an index i :

- (12) a. $[\text{DP I} [\text{NP } R_1 i_c]]$
 b. $[\llbracket R_1 \rrbracket] = \lambda x. \lambda y. \exists P_{\langle e,t \rangle} : P(x) \ \& \ P(y)$
 c. $[\llbracket i_c \rrbracket] = c_s/c_a$

Therefore, the DP has an LF-structure as proposed in **Error! Reference source not found.** The relation R makes a contextually salient property available (similar to pronoun-cases), just that the unique individual referred to by the determiner shares this property with the index. The index can only be the speaker or the addressee, so c_s or c_a . In this way, Kaplan’s analysis is extended rather than refuted. With these elements in place, we can now interpret a simpler alternative version of (11), given below:

⁷ Usually, frameworks within the uniform view quickly extend their analysis to include situation semantics – those are necessary for cases of bound pronouns and donkey pronouns. As we won’t be concerned with these cases, I’ll keep the semantics as simple as possible, although the implementation of situations would be straightforward.

- (13) a. I order my favourite meal.
 b. [IP [DP I [NP R₁ i_c]] [VP order my favourite meal]]
 c. [order my favourite meal] (the unique x such that there is a property P, such that P(x) and P(c_s))
 d. ‘The unique x, such that x and the speaker are condemned prisoners, orders x’s favourite meal.’

Standard cases of indexicals can also receive this analysis, only here, the relation R is the identity-relation:

- (14) a. I am hungry.
 b. [[R₂]] = $\lambda x. \lambda y. x = y$
 c. [hungry] (the unique x such that x is the speaker)

We see that with an analysis of bound indexicals as shown in (13), we get the uniformity of indexicals and definites for free⁸. Thus, definites, pronouns and indexicals share the same basic semantic (and syntax), and differ in the nature of their arguments. In this view, there is a systematic relationship between the phenomena.

2.3 Excursion: demonstratives

Another phenomenon that is treated differently in the two views are demonstrative pronouns. As they can occur with or without overt NPs, children’s production of demonstratives can also be suggestive with regard to the predictions of the classical and the uniform view. Classically, they are derived through contextual parameters, following Kaplan (1989):

- (15) [[this/that]]^{g,c} = the unique proximal/distant object c_s points at in c at c_t (Büring 2012)

Here, the uniqueness condition is mapped to a directly contextually available entity that is identified with the help of Kaplan’s contextual parameters of speaker and time. Even though there is no formalization of this analysis, it seems to be a mixing of a definite with an indexical. Another way of interpreting demonstratives that is in accordance with a variable view treats demonstratives as variables that come with an additional presupposition of proximity or distance (see Büring 2012 for discussion):

- (16) a. [[this₂]]^{g,c} = g(2) is proximal. g(2)
 b. [[that₃]]^{g,c} = g(3) is distant. g(3)

This analysis abandons the indexicality component for demonstratives altogether. Büring (2012) argues against an indexical treatment for demonstratives by highlighting that pronouns as well can be used deictically:

- (17) She/they/you/we (pointing at appropriate group) are better at the game than she/they/you/we (pointing at different appropriate group). (Büring 2012)

⁸ Of course there are various ways and analyses of such cases that go along a variable approach. You can read up on that in Grosz and Zobel (2014), as well as Kratzer (2009).

Examples like these demonstrate that deictic uses should employ a pragmatic inferencing of the referent and therefore, semantic analyses of deictic and anaphoric uses shouldn't differ from one another. That is why I will focus on a variable-type interpretation for demonstratives for the purposes of this paper.

The uniform view can account for demonstratives as well. Following Elbourne (2013), the analysis also infers a proximity/distance presupposition that is then added to a definite determiner meaning:

- (18) a. $[[\text{this}]] = \lambda f_{\langle e, t \rangle} : \exists ! x [f(x)=1 \& x \text{ is proximal}]. \iota x [f(x)=1]$
 b. $[[\text{that}]] = \lambda f_{\langle e, t \rangle} : \exists ! x [f(x)=1 \& x \text{ is distant}]. \iota x [f(x)=1]$

Other versions of Elbourne's theory include a parallel with indexicals, such that they come with a contextually determined index that captures the deictic use. However, Elbourne abandons this analysis in his later work.

Overall, demonstratives are an interesting case that can point towards a systematic relation between pronouns/indexicals and definites regardless of the theoretical differences: Demonstratives allow for overt NP complements, but may be used without them and could thus be seen as a bridge between the phenomena. One possible counterargument is that certain pronouns can collect overt NP arguments as well:

- (19) his car, her house, my bottle, your shoe

However, in both the classical and the uniform view, the possessive has been argued to contribute an individual meaning component that then has to be combined with the usual variable or definite analysis for pronouns. In the case of demonstratives, there is no additional meaning component that can explain their felicity when combined with an overt NP in the same way. I will come back to this at a later point of the paper.

2.4 Predictions for FLA production

As has been previously hinted at, the classical view expects no systematic order of production with regard to the purely syntactosemantic structure of the three phenomena. If we assume that children first learn the syntactic and semantic contribution of phenomena, and that pragmatic reasoning follows, we wouldn't expect any fixed order. But if we consider their pragmatic meaning components, maybe we can establish a correlation. Furthermore, we can't exclude systematicity in production for other reasons of language development that are unrelated to syntax or semantics.

On the other hand, the uniform view predicts systematicity in production of the three phenomena at the syntactosemantic level: If we assume that language is acquired in order of complexity, then it should follow that the definite determiner 'the' with an overt NP complement should be regarded as the simplest of the phenomena. With pronouns, there is an additional mechanism of NP-deletion involved. As for indexicals, the determiner has to combine with two contextually retrievable elements. Therefore, the definite determiner shouldn't be produced significantly later than any of the other phenomena, including demonstratives. The hypothesis is formulated below:

(20) **The Uniform View Hypothesis (former (2)):**

The definite article isn't produced significantly later than pronouns, indexicals or demonstratives.

In other words, as long as we assume that syntactosemantic differences will be observable in acquisitional data, in the classical view, there is no reason why children shouldn't produce demonstratives, indexicals, pronouns and definites in any order available to them. Finding intervariation between different corpora is not surprising under this account. However, it should be surprising for the uniform view if corpora reflect variability regarding the order of the definite article: One corpus demonstrating a significant later production of the definite article would suffice to question the uniform view, given that the complexity of the phenomena is a predictor. Preserving the uniform view in light of intervariation would necessarily mean to abandon the assumption of complexity as introduced in Section 1. The predictions are summarized in the table below:

Ordering Possibilities	Uniform view	Classical View
the \geq {idx, pr, dem}	+	+
{idx, pr, dem} \geq the	-	+

Table 1: Predictions

As is apparent in Table (1), if the data turns out this way, this study will only be able to provide evidence against the uniform view. All other results can only hint towards a preference for the uniform view without falsifying the classical view. As mentioned before, it could well be the case that a correlation in production can be observed that only follows from the pragmatic nature of the three phenomena: Even if pronouns are not definites per se, by virtue of them picking out a unique referent through the variable assignment function, the result will be a unique referent in any case. Thus, pragmatic use could well overwrite semantic differences.

3 Empirical background

Previous studies on both comprehension and production of pronouns, indexicals and definites have mostly looked at the two overall realms of pronouns and definites, with pronominal studies concentrating on the order of acquisition of first versus third person pronouns and studies on definites contrasting acquisition of definites with indefinites. There is also a vast amount of acquisitional studies on pronominal binding.

In the following, I will briefly discuss pronominal studies of comprehension and production first, then studies on the indefinite/definite distinction, and third, studies on binding principles.

The literature on pronouns has observed a production/comprehension asymmetry: Production studies (cf. Cruttenden 1977, Shipley and Shipley 1969, Chiat 1981, Huxley 1970, Halliday 1975) claim that there is an order of production with first person preceding second person, and third person pronouns following both. Felicitous production was observed from 15 months onwards. Some studies report reversal errors of first and second person pronouns, but they are not found consistently across studies.

Comprehension studies (cf. Charney 1980, Loveland 1984) argue for a different order of acquisition: the second pronoun is understood first, followed by first person and followed by third person pronouns. Comparing these results with the production studies above leads to the

suggestion that children use first person pronouns before understanding them. Several arguments have been put forward to explain this paradox that I won't go into here, but they seem to indicate that children have trouble distinguishing addressed and non-addressed speech. Further developing Charney's (1980) study, a recent study on the comprehension of indexicals by Moyer et al. (2014) reveals that two-year-olds are able to understand first and second person pronouns in both addressed and non-addressed speech once the experimental design is sensitive to a natural discourse setting. This study suggests that the acclaimed asymmetry may be due to experimental design and that children's comprehension of pronouns matches up with production after all. Overall, these studies haven't explored the possible correlation of definites and pronouns.

The literature on production and comprehension of definites mainly concentrates on the indefinite-definite distinction, reporting a definite-overuse by children from approx. 3 years onwards (cf. Karmiloff-Smith 1979; Schaeffer and Matthewson 2005; Schafer and de Villiers 2000; van Hout et al 2010; Munn, Miller and Schmitt 2006) when they should produce indefinites. However, comprehension of the definite is mostly adult-like, while comprehension of the indefinite is more flexible. Many studies have argued for a lack of pragmatic reasoning in the definite-overuse (cf. van Hout et al 2010; Munn, Miller and Schmitt 2006) and furthermore demonstrate problems with the experimental setup: Often, children tend to use a definite in paradigms where pictures match target-sentences. In these studies, the existence of a visual referent may well be enough for children to assign it a definite meaning. Overall, the literature has mostly focussed on children older than 3 years. Studies on binding principles reveal that children as young as 30 months are sensitive to binding principles (cf. Lukyanenko, Conroy and Lidz 2013; Lidz 2007; Goodluck 2007). While I'm not interested in bound pronouns in this study, comparing these data with the production of free pronouns suggests that the latter should precede the former in production, given that children have to know the syntax and semantics of free pronouns as well as syntactic binding structure and constraints for bound pronouns.

Song and Fisher (2005) have innovatively tested the pragmatic nature of definites and pronouns inspired by studies on adult use of the phenomena (cf. Hudson-D'Zmura and Tanenhaus 1998; Gordon et al. 1993; Ariel 2001) that revealed that, in contrast to definites, pronouns are restricted to refer to salient discourse referents. Song and Fisher found sensitivity of discourse saliency in the comprehension of pronouns versus definites in three-year-old children. However, this study focused on pragmatic meaning components involved in pronoun and definite interpretation, a component that will only later become relevant for the present study.

All in all, there is evidence of children's adult-like comprehension and production of definites and pronouns from 3 years and 2 years onward, respectively. The present study wants to elaborate on their possible correlation of production by examining longitudinal corpora of spontaneous children's speech. A side effect of this study will be testing the results of previous studies on pronoun production: The present study will reveal that, different from previous results, there is no systematic ordering of first, second and third pronouns, but that most of the time, all three pronoun types are acquired simultaneously.

4 The corpus study

I collected and analyzed five longitudinal corpora⁹ made available through the CHILDES database, choosing corpora that were recorded from an as early as possible age. See Table (2) for the details:¹⁰

Corpus	First Recording	Last Recording	Mean recording rate (per month)
Naima (Providence)	0;11	3;10	2.93
Lily (Providence)	1;1	4;0	2.35
Violet (Providence)	1;2	3;11	1.92
Naomi (Sachs)	1;2	4;9	4.04
Mark (McWhinney)	0;10	5;6	3.96

Table 2: Corpora

4.1 Probability gymnastics

Before coming to the study, let us briefly consider the empirical strength of conducting a study on five children.¹¹ With four groups of phenomena (pronouns, indexicals, definites and demonstratives), there is a total number of 15 possible orderings, given that each of the phenomena can be produced simultaneously. Since only one of the two accounts provides a falsifiable hypothesis on syntactosemantic grounds, we will only consider orderings that go against or according to the uniform hypothesis. As a reminder, evidence in favor of the uniform view should demonstrate that the definite article isn't acquired significantly later than the other phenomena, whereas one ordering such that the definite article follows any of the other phenomena significantly provides evidence against the uniform view. See the types of possible orderings below¹²:

Orderings compatible with Uniform View	Orderings incompatible with Uniform View
{Dem, Pr, Def} < Ix	Ix < {Dem, Pr, Def}
{Def, Ix, Pr} < Dem	Dem < {Ix, Pr, Def}
{Def, Dem, Ix} < Pr	Pr < {Ix, Dem, Def}
{Dem, Def} < {Pr, Ix}	{Pr, Ix} < {Dem, Def}
{Pr, Def} < {Dem, Ix}	{Dem, Ix} < {Pr, Def}
{Def, Ix} < {Dem, Pr}	{Dem, Pr} < {Def, Ix}
Def < {Dem, Pr, Ix}	{Dem, Ix, Pr} < Def
{Def, Dem, Ix, Pr}	

Table 3: Possibilities

⁹ For the presentation at IATL 31, only three of the five corpora were analysed. The Mark and Violet corpora were elicited after the Lily, Naima and Nomi corpus in order to further empirical strength.

¹⁰ The Mark Corpus is a diary study Brian MacWhinney conducted with recording his two sons, Ross and Mark. Even though he recorded them frequently, one should still keep in mind that the selection of recordable data was guided by the researcher's own estimation. Still, the data provided a dense recording of data over a timeframe of 6 years.

¹¹ This probability calculation was conducted as shown in Hohaus et al (2014).

¹² Demonstratives are abbreviated as Dem, Indexicals as Ix, pronouns as Pr and the definite article as Def.

As evidenced by Table (3), 8 out of 15 possibilities go along with the uniform view such that the definite determiner doesn't follow any of the other phenomena. Accordingly, seven ordering possibilities go against the uniform view. As long as all five corpora reveal any of the eight orderings compatible with the uniform view, the uniform view isn't falsified. In the case that the corpora reveal any of the seven orderings not compatible with the uniform view, the uniform view hypothesis is refuted.

(21) a. Probability compatible with uniform view observed across 5 corpora:

$$(8/15)^5 = 0.043^{**}$$

b. Probability incompatible with uniform view observed across 5 corpora:

$$(7/15)^5 = 0.022^{**}$$

However, any of the 15 orderings is compatible with the classical view. Results that demonstrate only orderings among the seven incompatible ones throughout all five corpora would reveal a p-value of 0.022, a significant outcome (see (21b)). On the other hand, if all corpora were to demonstrate only orderings among the eight compatible ones, the p-value would be significant as well (see (21a)). These results would be strong predictors.

4.2 Method

I elicited the data based on Snyder's (2007) suggestion for corpus studies. For all corpora, I searched for all relevant utterances of the children of third person pronouns, indexicals, demonstratives¹³ and the definite article with the search tool CLAN (labels and search commands can be accessed in the appendix). Utterances were labelled according to adult-like and independent usage, excluding imitations, repetitions or routine utterances (cf. Snyder 2007). I also systematically labelled erroneous utterances. The generic reading of 'you', 'that' as a clausal complementizer, as well as the usage of 'it' as a help construction were labelled and excluded from analysis.

4.3 Results

The Age of Acquisition relative to production was determined according to Snyder's (2007) FRU (First of repeated utterances): I isolated those adult-like felicitous utterances that were soon followed by a regular use of the construction within the next two months. Then, these utterances were again checked against a fine-grained detailed analysis considering the utterance context. The table below shows the FRUs:

Constructions	Lily	Naima	Violet	Naomi	Mark
I	1;10	1;5	1;8	1;10	1;5
You	2;1	1;8	1;10	1;11	1;8
Me	2;0	1;8	2;0	1;11	2;6
My	1;11	1;6	1;10	2;0	3;0
Your	2;0	2;5	2;0	2;4	3;0
It	1;11	1;6	1;9	1;10	1;8
He	1;11	1;8	2;0	1;11	2;1

¹³ Demonstratives were elicited after the other three phenomena.

She	2;2	1;7	2;3	2;5	3;0
Him	2;5	1;8	2;0	2;5	3;4
His	1;11	1;6	2;1	2;0	2;10
Her	2;2	1;7	2;0	2;4	3;6
This	1;10	1;5	1;9	1;10	2;4
That	2;0	1;6	1;10	1;10	2;6
this one	2;0	1;7	1;9	2;1	3;0
that one	2;1	1;7	1;9	2;3	3;1
this + NP	1;10	1;7	1;10	1;10	2;4
that + NP	2;2	1;7	1;10	1;10	2;4
The	1;10	1;4	1;9	1;10	2;6

Table 4: First of Repeated Utterances (given in months)

According to the ages shown in Table (4), there is one thing to note first: For all four groups of phenomena, one construction is produced first across corpora: ‘I’ for indexicals, ‘it’ for pronouns, ‘this’ for demonstratives and the definite article ‘the’¹⁴. In the graphs shown in Tables 12-16 in the Appendix, the development of adult-like utterances of the four phenomena is demonstrated on the basis of months. For each corpus, there is a quick escalation of number of utterances per 1000 utterances in total. The immediate beginning of the increase was chosen to be the FRU. For the statistical analysis, I am now able to compare only these first produced constructions and thus comparability across groups can be guaranteed. Also, in most of the corpora, the four phenomena are produced almost within the same month (see Lily, Naomi and Violet). The graphs given in the appendix furthermore demonstrate that the development of utterance increase runs almost parallel for all phenomena across corpora, except in Mark’s case (see Tables 12-16). Thus, we can already conclude that results of previous studies on the order of production cannot be replicated here: There is no systematic ordering of first, second and third pronouns. But there is another result that is already observable when looking at the raw data results from the Mark Corpus: Whereas the other four corpora demonstrate orderings that go along with the uniform view, the Mark Corpus reveals an ordering that goes against it: The definite article is produced at age 2;6, whereas the first person pronoun ‘I’ is already produced at 1;5, and ‘it’ is produced at 1;8. The next step will be to determine if any of the orderings observable in Table 4 is statistically significant.

4.4 Statistical analysis

In accordance with Snyder (2007) and Stromswold (1996), I will employ a nonparametric statistical test that checks significance of one construction being produced before another construction. I will employ the test within corpora, and only for differences in production age on the basis of months. The rationale behind the Binominal Test is to see how unlikely the difference between actual observed age of production and expected age of production of the later construction is, given the relative frequency of both constructions once they are produced on a regular basis. All of the binary tests taken together will then reveal one of the fifteen possible orderings for one corpus. If all five corpora demonstrate the same pattern of ordering on the basis of these individual Binominal Tests, this will in total contribute to the empirical strength illustrated above.

¹⁴ Some of the other constructions demonstrate the same month of FRU, however, for comparability’s sake, I will continue with those constructions that come first across corpora.

On the basis of age of production determined in months, the Naomi Corpus will be excluded from statistical analysis, as all four phenomena are produced within the same month. For the Lily Corpus, there are three orderings that have to be calculated:

Construction 1	Construction 2	# of Earlier Occurences	Relative frequency 1	Relative frequency 2	p-value
I	it	14	0.72	0.28	0.073*
the	it	2	0.58	0.42	0.221
this	it	6	0.37	0.63	0.003***

Table 5: Statistical Analysis Lily-Corpus

The only significant ordering in the Lily Corpus is that the demonstrative ‘this’ significantly precedes ‘it’ (see Table 5). In the Naima Corpus, both ‘the’ and ‘this’ precede ‘it’ significantly. ‘the’ preceding ‘I’ and ‘this’ didn’t turn out to be significant (see Table 5).

Construction 1	Construction 2	# of Earlier Occurences	Relative frequency 1	Relative frequency 2	p-value
The	I	4	0.93	0.07	0.4
The	it	5	0.64	0.36	<0.001***
This	it	10	0.15	0.85	<0.001***
The	this	4	0.91	0.09	0.254

Table 6: Statistical Analysis Naima-Corpus

None of the orderings within the Violet Corpus turned out to be significant (see Table 7).

Construction 1	Construction 2	# of Earlier Occurences	Relative frequency 1	Relative frequency 2	p-value
I	it	4	0.68	0.32	0.145
I	the	4	0.60	0.40	0.078*
I	this	4	0.87	0.13	0.24

Table 7: Statistical Analysis Violet-Corpus

Until now, all four corpora have demonstrated orderings that are compatible with the uniform view: Lily only shows significant orderings regarding ‘it’ following ‘this’ and ‘I’. Also, the Naima Corpus only demonstrates a significant ordering between ‘the’ and ‘this’ and ‘it’. The Violet Corpus didn’t show any significant ordering, neither did the Naomi Corpus, in which all phenomena were produced first within the same month. However, the Mark Corpus reveals evidence against the uniform view, as has been already observable by the raw data.

Construction 1	Construction 2	# of Earlier Occurences	Relative frequency 1	Relative frequency 2	p-value
I	it	3	0.73	0.27	0.295
I	the	44	0.59	0.41	<0.001***
It	the	5	0.21	0.79	<0.001***
This	the	2	0.07	0.93	0.01***
I	this	24	0.95	0.05	0.026*
It	this	3	0.78	0.22	0.168

Table 8: Statistical Analysis Mark-Corpus

Again, all of the results work equally well with the classical view.

4.5 Error analysis

In the following, I want to briefly discuss the errors found across corpora. I conducted an error analysis in analogy to Stromswold (1996) for one relevant error-type. However, as already mentioned by Snyder (2007), if we take conservativity as a predictor for language production of children, errors shouldn't be and in fact aren't a valuable predictor, even if certain erroneous behavior can give additional insights into the acquisitional process. In particular, the two frameworks discussed here make different predictions regarding error emergence:

Error Type	Uniform View	Classical View
Presuppositional Errors (gender- and case-assignment, definite-overuse)	√	√
Structural Errors: combination of pronouns with overt NPs	√	x

Table 9: Error Expectations

Across both theories, one can expect a certain systematicity of errors regarding the presuppositional meaning components of the phenomena, given that pragmatic inferencing is preceded by knowledge of syntactic structure. And indeed, there are errors of both case- and gender-agreement to be found across corpora and also a certain amount of definite-overuse¹⁵. From the point of view of the present investigation, these errors are not very interesting, as both theories predict these errors to come up. Instead, if we found a systematic emergence of children combining pronouns with overt NP complements, this would point towards the uniform view. However, there can only be statistical strength if this error can be observed repeatedly and across corpora over a certain timeframe. Of the five corpora here, only one corpus showed an emergence of this error type. Naomi utters the pronoun 'it' alongside an overt NP. Even though there are no tests available to confirm the complement-relationship, there is at least no break in prosody between the pronoun and the overt NP:

- (22) Naomi: *I throw it ice.
Context: Naomi is playing with an ice-cube at breakfast.
Age: 1;11 (n22.cha)
- (23) Naomi: *I get it egg.
Context: Naomi plays with toy birds and eggs.
Age: 1;10 (n19.cha)
- (24) Naomi: *Need it jacket.
Context: Naomi wants to put her jacket on.
(1;10, n09.cha)

¹⁵ See Wexler (2003), van Hout et al. (2010) and Munn, Miller, Schmitt (2006) for an in-depth discussion on theoretical and empirical reasons that can explain definite-overuse.

There are eight instances of this error type over the course of one and a half months (from 1;10 to 1;11), found in transcripts n09.cha until n28.cha. Calculating an Error Rate according to Stromswold (1996) that compares the number of utterances of the error within the timeframe in which it is uttered with the numbers of all utterances that present opportunities where the child could have made this mistake results in insignificance: The eight instances of the errors only present 5.8% of the overall opportunities Naomi had to use a pronoun or a definite. Thus, even if these utterances can be identified as cases where the NP is a complement of the pronoun, these are not uttered frequently enough by the child in order to be taken as an empirical argument in favor of the uniform view. Furthermore, there is only one additional instance of such an error in one of the other corpora, see (25).

(25) Lily: *And I use it potty, too.
(2;2, lil34.cha)

Thus, the error analysis regarding pronouns combined with an overt NP argument must stay inconclusive regarding the two frameworks.

4.6 Adult input

The pattern of production order could also be explained with an adult input account. A strong hypothesis of this account would be that those constructions that are uttered the most frequently by the parents are produced first. This hypothesis doesn't make any claims about the linguistic structure or complexity of the phenomena. Even though a construction has to be uttered frequently enough in the presence of a child, it would seem arbitrary to explain linguistic competence and its acquisition only on the grounds of such an account. However, we have to provide evidence in order to exclude it as a predictor. In order to do so, all adult utterances of the relevant constructions across all transcripts of the corpora were tested for their frequency, revealing the following ordering of relative frequency per month:

Corpus	Frequency Ordering	Relative Frequency of the phenomena per month
Lily	the < you < I < it < that < this	the = 777; you=733.1; I=329.6; it=270.9; that=169.6; this=57.4
Naima	you < the < I < that < it < this	you=453; the=450.7; I=225.9; that=207.2; it=198.8; this=85.4
Violet	you < the < I < it < that < this	you=181.7; the=178; I=92; it=84; that=75.75; this=72.5
Naomi	you < the < it < that < I < this	you=116.7; the=86.6; it=54.7; that=29.8; I= 27.7; this=17.3
Mark	you < that < the < it < I < this	you=264; that=134.2; the=113; it=82; I=63; this=33.9

Table 10: Adult Frequency

According to Table (10), even though adult frequency reflects that ‘the’, ‘it’ and ‘I’ are among the four constructions that are uttered most frequently, and other constructions like ‘he’ and ‘she’ follow them, the high frequency of ‘you’ is not reflected at all in the order of actual production (see Table 4): In the Lily Corpus, ‘you’ is produced later than the other four constructions. However, frequency would predict it to be produced much earlier. Similarly, in the Naomi Corpus, ‘you’ is produced later than the four other constructions as well, even though frequency would predict it to come first. The same pattern can be observed in the Naima Corpus and in the Violet Corpus. In the Mark Corpus, only ‘I’ is produced earlier than ‘you’, but this result already goes against the adult input frequency, where ‘I’ comes in last compared to the other constructions. Furthermore, frequency reveals that ‘that’ is used more frequently in the adult input than ‘this’, whereas the production data demonstrates the reversed order in production. For ‘this’ and ‘that’, only ‘that’ is among the five most frequent constructions across corpora in the adult input. In the Lily and Mark Corpora, ‘this’ is among the least frequent constructions. One reason for the high frequency of ‘that’ could lie in the fact that with the search tool CLAN, there is no distinction possible between the use of ‘that’ as a determiner and the use of ‘that’ as a relative complementizer. This could be a confounding factor in the input. Nonetheless, we also see for demonstratives that ‘you’ precedes ‘this’ and ‘that’ in frequency but not in the order of production. In order to confirm this discrepancy between input and production, I once more calculated Binominal Tests, comparing the FRU in production of ‘the’, ‘I’ and ‘it’ with ‘you’ within each corpus¹⁶:

Constructions	Lily	Naima	Violet	Naomi	Mark
I < you	< 0.001	< 0.001	0.124	< 0.001	0.204
it < you	< 0.001	< 0.001	0.227	< 0.001	-
the < you	< 0.001	< 0.001	-	< 0.001	-

Table 11: Binominal Tests for 'you'

The p-values of conducting Binominal Tests reveal that for three of the corpora, all constructions, ‘I’, ‘it’ and ‘the’ are produced significantly before ‘you’, going against the frequency hypothesis. In two corpora, the orderings didn’t come out as significant. Still, we would expect to find a constant correlation of the adult ordering with the production data. The adult input however indicates ‘you’ to be the most frequently uttered construction in all corpora except Lily, where it is the second most frequent construction. This pattern is clearly not reflected in the production data.

4.7 Discussion

Results of the corpus study are mixed: On the one hand, we see a systematicity across corpora that reveals that ‘I’, ‘it’, ‘the’ and ‘this’ are those constructions that are produced first within the respective groups of phenomena. For pronouns, this finding goes well with a view that treats ‘it’ as the default pronoun. For the other pronouns, additional presuppositional content has to be learned by the child: ‘he’ can only refer to males, while ‘she’ can only refer to females and for ‘his’, ‘him’ and ‘her’, additional syntactic restrictions and case-marking are present. For indexicals, these results would go together with an account of egocentric perspective of children that argues that they first learn to refer to themselves before referring

¹⁶ Since demonstratives were elicited later, the Binominal Test wasn’t conducted for them at the point of publication. However, this doesn’t change the point: Already with ‘the’, ‘I’ and ‘it’, children demonstrate a use that deviates from the predictions of a purely frequency based account.

to other conversational partners. However, even though these constructions are produced first within the groups, the other constructions soon follow, indicating that there is no significant ordering of these constructions across corpora. The result for demonstratives is not that straightforward: According to the uniform view, the most basic case of the semantic/pragmatic structure should be the combination of ‘this’ with an overt NP complement. However, results do not reveal a significant ordering of ‘this’ and ‘this’ together with an NP-complement, so that it could still be the case that NP deletion doesn’t pose additional acquisitional complications.

Another insight is that the present study demonstrates an overall correlation of the production of the phenomena: except for the Mark Corpus, all phenomena are produced either in the same month or in the course of two to three immediately following months. There hasn’t been observed any statistical ordering of the phenomena that holds across corpora. Thus, the strong claim of previous literature on the ordering of pronominal production cannot be held up here. There is no clear ordering of first, second and third person pronouns. Even if we don’t consider that the phenomena belong to different groups, this study doesn’t reveal consistent orderings of ‘I’, ‘you’ and ‘he’ across corpora (see Table 4). Overall, with a broad perspective on the phenomena, the correlation of production most likely points towards the pragmatic properties of the phenomena that they share: In both the classical and the uniform view, for indexicals, pronouns and definites alike, we assume definiteness to be an underlying core that unites them, even if they are modelled to handle it differently. The uniqueness presupposition of the definite determiner restricts the context in such a way that there has to be a unique referent in the context. The variables approach to pronouns results in the same outcome: For a pronoun to be felicitously uttered, there has to be a unique and salient referent in the discourse the pronoun points to. So, even if in the former case, the semantics itself doesn’t map the determiner and its complement to a specific referent but rather say that the pragmatic machinery has to come up with one, and even if in the latter case, the semantics doesn’t provide any semantics except the features that further restrict the pronoun, but the pragmatic machinery is formalized as a variable assignment function and thus, the meaning of a pronoun can only be the pragmatic referent picked out by the context, both accounts talk about one unique accessible referent. Thus, the present results confirm this overarching correlation of definiteness. It would have been much more surprising if there was a significant ordering of the phenomena across corpora.

However, the study also demonstrates that it might be nearly impossible to tease the classical and the uniform view apart in a more fine-grained fashion, concentrating on the semantic differences. The hypothesis I wanted to test in this study is repeated below:

(26) The Uniform View Hypothesis (former (2) and (10)):

The definite article isn’t produced significantly¹⁷ later than pronouns, indexicals or demonstratives.

We established that only through falsifying this hypothesis, there could be evidence against the uniform view. In all other cases, both theories could still account for the data. A falsification of this claim would have involved the observation that the definite article is produced significantly later than the other phenomena across corpora. Using our probability test, we have seen that this result would have had a p-value of 0.022. However, there is only one corpus that refutes the Uniform View Hypothesis – the other corpora don’t demonstrate the same behavior. Finding one corpus where the uniform view hypothesis is falsified is

¹⁷ Significance is defined here as a statistical significant value as a result of non-parametric statistical analysis conducted with the help of Binominal Tests. A significant p-value is defined here, as is the norm, as being below 0.01.

expected with a probability of 0.47, so almost 50 per cent of the time. This result is far from significant, so there is no strong reason to reject the uniform view. Another reason to question the validity of the Mark Corpus lies in the fact that, even if there had been dense recordings of Mark's spontaneous speech, it was still a diary study and the father himself decided when to record or note down the utterances of his sons. It may be the case that early onset utterances of the definite determiner weren't recorded. However, we see a regular use of the other phenomena early on without significant breaks. Thus, the density of the recordings served to capture the production of the other phenomena.

The errors found in the corpora were analyzed only according to one possible error that could distinguish the two views. However, even though we see the occurrence of the error of combining 'it' with an overt NP complement, the occurrence couldn't be observed across corpora and even within the corpus that demonstrated them, there were only eight utterances. This result isn't significant either.

Finally, adult Frequency as a predictor isn't able to account for the ordering of phenomena. Even though we find that all of the constructions are very frequently used, their order in frequency isn't reflected in the order of production.

5 Conclusions

This study is a first step towards an investigation of theoretical claims that assume a unified semantics for pronouns and definites specifically – a framework that can be extended to indexicals and demonstratives. The way how children perceive and use these constructions is an important source of evidence that can give insights into how and when the constructions are acquired and if there is a correlation of acquisitional ordering of the phenomena is possible. This study makes evident that a rather broad correlation between the phenomena can be observed across corpora: Most of the children produce definites and pronouns as well as demonstratives and indexicals by 23 months of age. However, the issue remains if and how the syntactosemantic properties of both the uniform and the classical view can be tested with acquisitional measures.

Future questions for this enterprise are if comprehension of the phenomena can be observed at an equally early age. Even though early comprehension of indexicals has been attested in 2-year-olds (Moyers et al 2014), most of the studies on the difference between indefinites and definites have tested the difference mostly on 3-year-olds. The goal is therefore to see if 2-year-olds are able to understand third person pronouns and definites as well as indexicals. Once this basis is established, more fine-grained measures are needed to further explore the semantic differences in both adult and child language production, comprehension and processing.

6 Acknowledgements

Many thanks go to Julia Chant who helped me evaluate the data. Also a hearty thank you goes to Sigrid Beck, my advisor, who guided the process of this study, and Vera Hohaus and Sonja Tiemann who gave their valuable time to discuss the results with me. I also wish to extend a special thank you to Valentine Haquard, Jeff Lidz and Alexander Williams for discussing this study and coming up with follow-up questions. I am also thankful for valuable feedback from the IATL 31 conference audience, the Doktorandenkolloquium at the

University of Tübingen and the CNL Lab Meeting Audience at the University of Maryland.
Thank you!

7 Appendix

7.1 Graphs of Production Development

The following graphs demonstrate the increase in utterances of ‘I’, ‘it’, ‘the’ and ‘this’ per month. The numbers of utterances are given in permille rates, giving the mean number of utterances per 100 utterances in a month (cf. Snyder 2007).

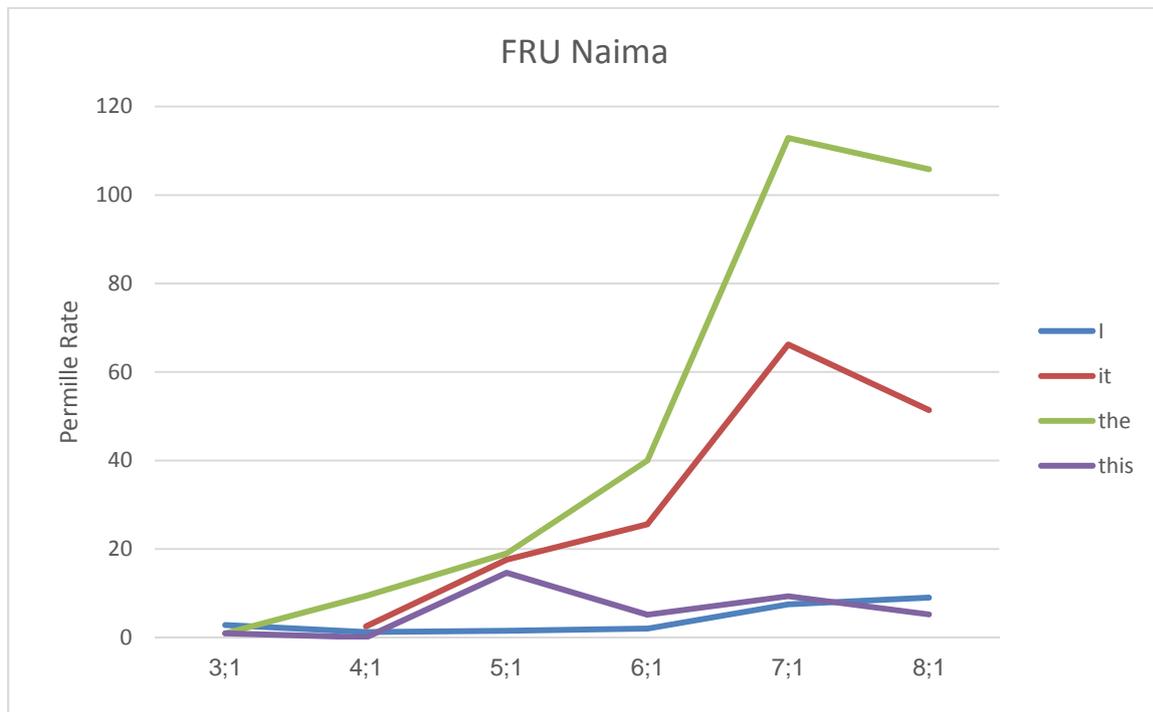


Table 12: Production Development Naima

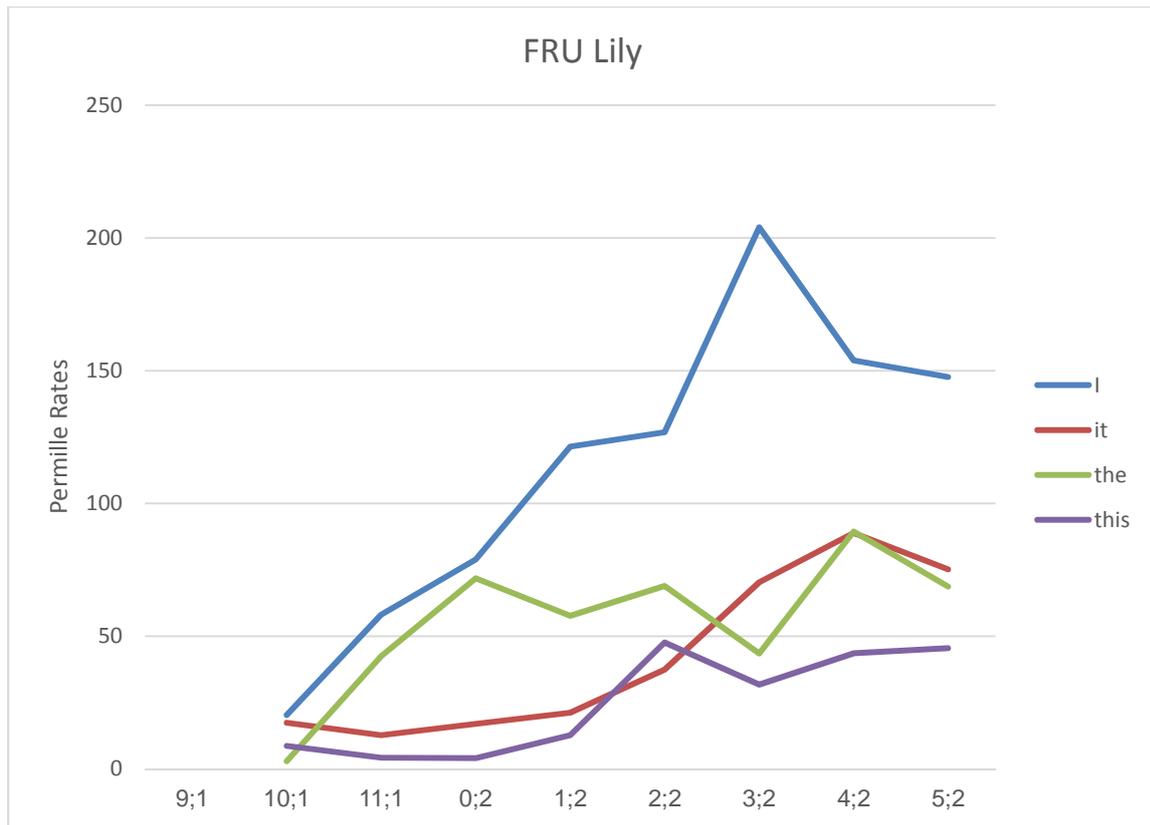


Table 13: Production Development Lily

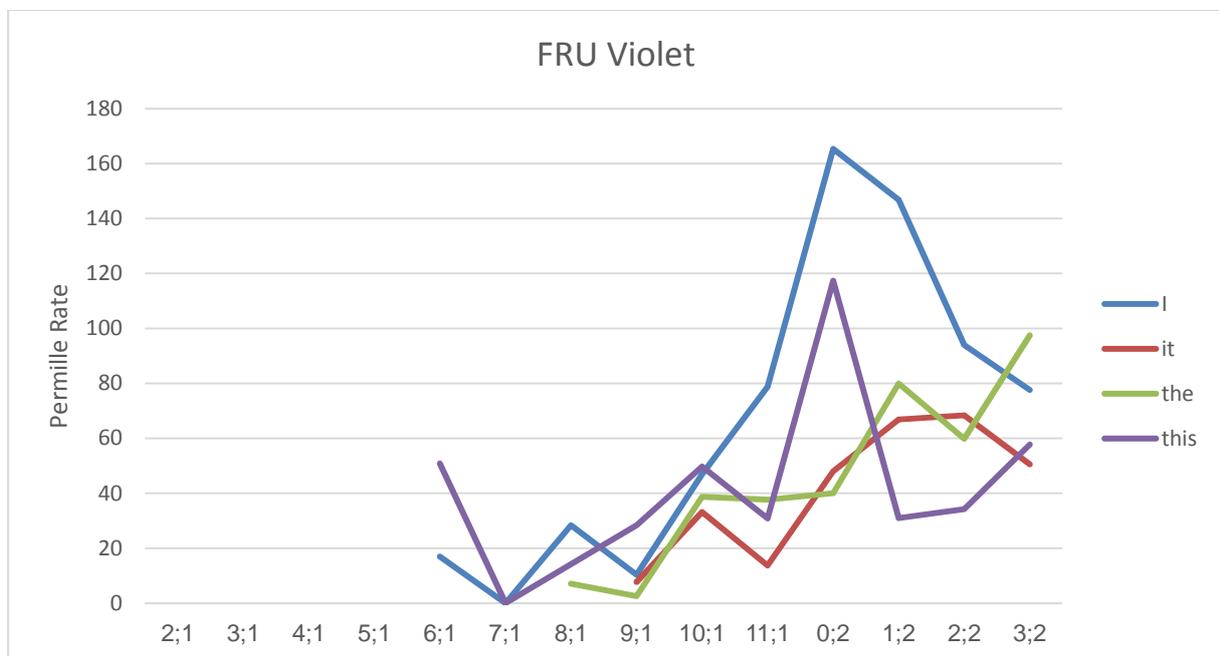


Table 14: Production Development Violet

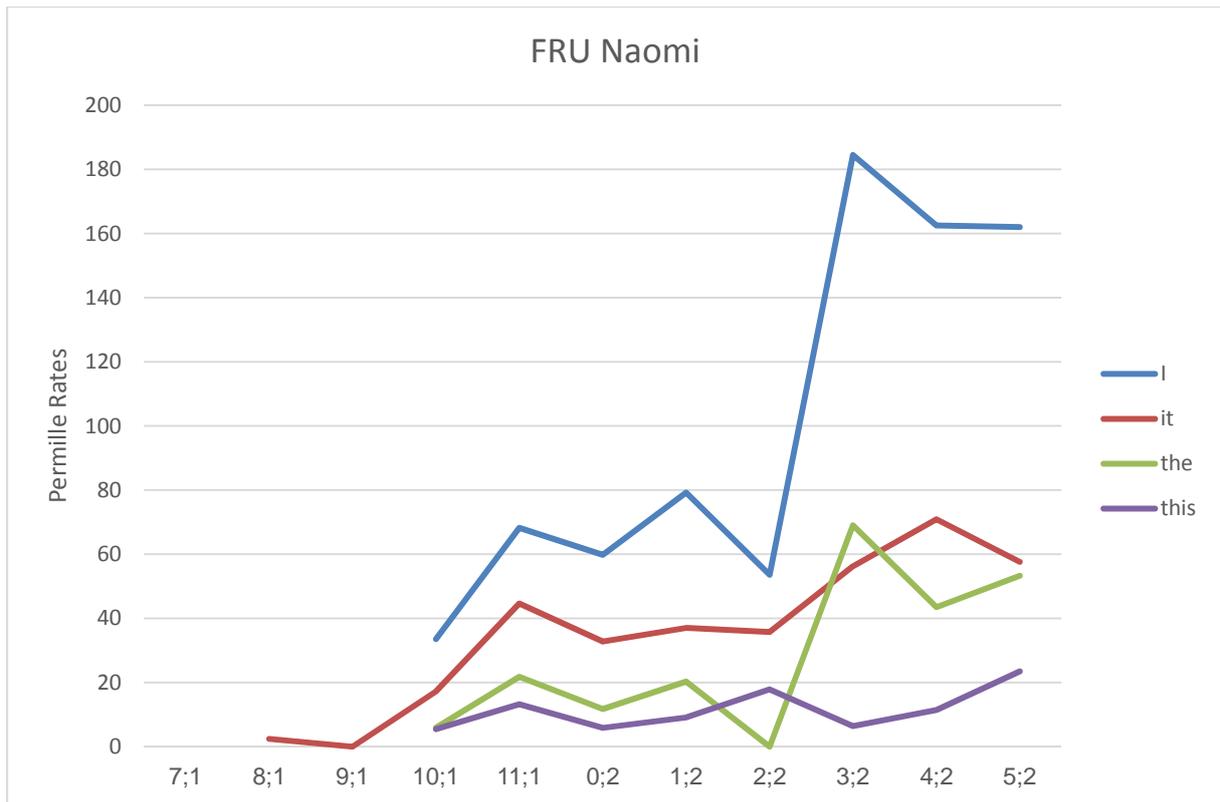


Table 15: Production Development Naomi

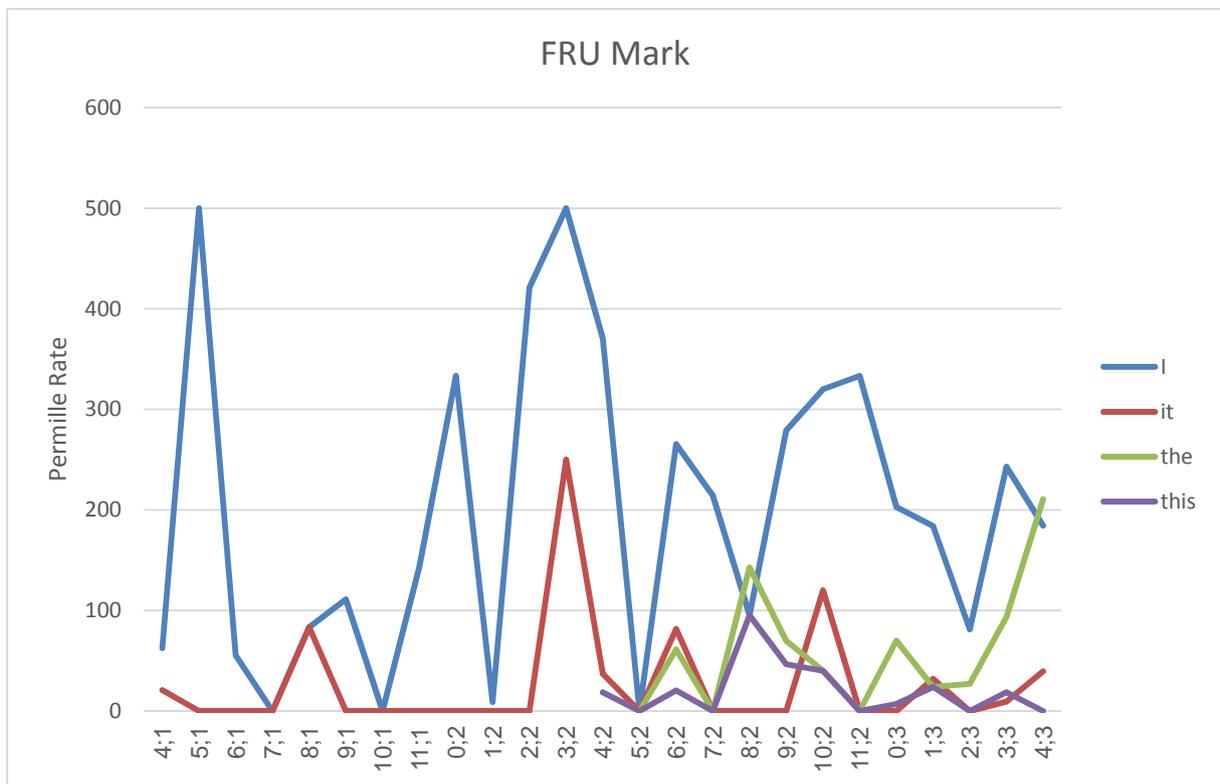


Table 16: Production Development Mark

7.2 List of first (of regular) uses

Pronouns: 'it'

- (27) Naomi: *fix it*. Context: Naomi is playing with a doll and has just fixed her diaper. She refers to the diaper and says that she fixed it, so probably that she put the diaper on the doll. (1;10, n09.cha)
- (28) Naima: *eyes, eyes on it*. Context: Naima is referring to a CD where the eyes of the singer (Patty Larkin) are prominently to be seen. (1;6, nai18.cha)
- (29) Lily: *uh uo it dried off*. Context: Lily is playing with a Polly Pocket doll and a small water pond, but the pond has unfortunately already dried off. (1;11, lil24.cha)
- (30) Violet: *this, it goes where?* Context: Playing puzzles, she talks about a puzzle piece. (1;9, vio16.cha)
- (31) Mark: *I fix it* Context: Mark messed up his father's tape recorder and wants to fix it. (1;8, 43a2.cha)

Indexicals: 'I'

- (32) Naomi: *I get [the blanket]*. Context: Naomi is playing with dolls and pretends to put them to sleep. That's why she needs a blanket for one of the dolls. (1;10, n09.cha)
- (33) Naima: *see, I wanna see*. Context: Naima is talking about spiders. She claims that she wants to see them. (1;5, nai16.cha)
- (34) Lily: *oh I draw right on the floor*. Context: Lily has drawn on the floor with a white crayon. (1;10, lil21.cha)
- (35) Violet: *I choose book*. Context: Mother asks Violet to pick out a book to read. (1;8, vio14.cha)
- (36) Mark: *I help*. Context: Mark wants to help his father. (1;5, 38b2.cha)

The definite determiner 'the':

- (37) Naomi: *baby on the floor*. Context: Naomi is playing with her dolls and one doll is on the floor. (1;10, n14.cha)
- (38) Naima: *I'm on the floor*. Context: Naima is on the floor, playing with crayons, while her parents talk. (1;4, nai11.cha)
- (39) Lily: *in the garbage*. Context: Lily's mother wants her to throw away a dirty shirt. (1;10, lil22.cha)
- (40) Violet: *here is the mouse*. Context: Playing puzzles, she identifies the piece with the mouse on it. (1;9, vio16.cha)
- (41) Mark: *I pushed the button*. Context: Mark reaches the button on the door, probably the doorbell. (2;6, 52a2.cha)

The demonstrative 'this':

- (42) Naomi: *read this*. Context: Naomi wants to read a catalogue that is right next to her. (1;10, n15.cha)
- (43) Naima: *this says...* Context: looks at a book and is wondering about the title. (1;5, nai17.cha)
- (44) Lily: *this is Daddy's tie*. Context: She is playing with her Daddy's tie. (1;10, lil21.cha)

- (45) Violet: *what is this mama?* Context: Looking at a children's book, she wants to know what kind of animal is on the page she's looking at. (1;9, vio15.cha)
(46) Mark: *this is Wompa.* Context: Naming his toys. (2;4, 49a2.cha)

7.3 Error list: selection

In addition to the errors discussed in more detail in Section 4.5, here is a selection of errors related to the presuppositional content of both indexicals and pronouns, e.g. gender and case information or the uniqueness condition of the definite article.

- (47) Naima: *Squirrels eat the rocks. Context: Naima and her mother talk about what squirrels eat in general. (1;6, nai20.cha)
(48) Naima: *Juice me want. (2;1, nai44.cha)
(49) Naomi: *Him have eye. Context: Naomi is talking about a toy dog that has just one eye. (2;0, n38.cha)
(50) Naomi: *He fall down. Context: Naomi talks about a kitty, which is identified through the context as female. (2;0, n38.cha)

In the following, there is a selection of errors demonstrating definite-overuse:

- (51) Naomi: *Where's the Daddy? Context: Naomi talks about her father. (1;10, n14.cha)
(52) Naomi: * Here's the another one. Context: Naomi is outside, she finds two cigarette buds. (2;1, n47.cha)

The data already demonstrates a certain amount of errors where pronouns and indexicals are confused:

- (53) Naomi: *Gonna put it on my floor. Context: She peels an orange and doesn't know where to put the peels. She's in the kitchen. (2;0, n35.cha)
(54) Naomi: *Is it gone? Context: Naomi talks about her father and if he's already gone to work. (1;11, n23.cha)
(55) Lily: *It heart. The operator asks her what is on a picture she shows Lily. (1;11, lil24.cha)

Naima extensively shows a reversed use of 'you' and 'your', as referring to herself:

- (56) Naima: *It's in your hand. Context: Naima herself has a puzzle piece in her hand. (1;8, nai28.cha)

References

- Ariel, M. 2001. Accessibility theory: an overview. In T. Sanders, J. Schilperoord and W. Spooren (eds.), *Text Representation: Linguistic and psycholinguistic aspects*, 29-87. Philadelphia: John Benjamins.

- Büring, Daniel. 2011. Pronouns. In K.v. Heusinger, C. Maienborn and P. Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, Vol. 2: 971-995. Berlin: de Gruyter.
- Charney, R. 1980. Speech roles and the development of personal pronouns. *Journal of Child Language* 7: 509-528.
- Chiat, S. 1981. Context-specificity and generalization in the acquisition of pronominal distinctions. *Journal of Child Language* 8: 75-91.
- Cruttenden, Alan. 1977. The acquisition of personal pronouns and language 'simplification'. *Language and Speech* 20(3): 191-197.
- Elbourne, P. 2005. *Situations and Individuals*. Cambridge, MA: MIT Press.
- Elbourne, P. 2008. Demonstratives as individual concepts. *Linguistics and Philosophy* 31: 409-466.
- Elbourne, P. 2013. *Definite Descriptions*. Oxford: Oxford University Press.
- Goodluck, H. 2007. Formal and Computational Constraints on Language Development. In E. Hoff and M. Shatz (eds.), *Blackwell Handbook of Language Development*, 46-67. Cambridge: Blackwell.
- Gordon P. C., B. J. Grosz and L. A. Gillian. 1993. Pronouns, nouns, and the centering of attention in Discourse. *Cognitive Science* 17: 311-348.
- Grosz, P. and S. Zobel. 2014. A (non-)uniform approach to pronominal semantics. *ESSLLI 2014*, Tübingen. Lecture Notes.
- Halliday, M. and A. Kirkwood. 1975. *Learning how to mean: Explorations in the Development of Language*. London: Edward Arnold.
- Heim, I. and A. Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.
- Hohaus, V., S. Tiemann and S. Beck. 2014. Acquisition of comparison constructions. *Language Acquisition* 21(3): 215-249.
- Hudson-D'Zmura, S. and M.K.Tanenhaus. 1998. Assigning antecedents to ambiguous pronouns: the role of the center of attention as the default assignment. In Walker M. A. Walker, A. K Joshi and E. F. Prince (eds.), *Centering Theory in Discourse*, 199-226. Oxford: Clarendon.
- Huxley, R. 1970. The development of the correct use of subject personal pronouns in two children. In G.B. Flores d'Arcais and W.J.H. Levelt (eds.), *Advances in Psycholinguistics*, 141-165. Amsterdam: North-Holland.
- Kaplan, D. 1989. Demonstratives. In J. Almog and H. Wettstein and J. Perry (eds.), *Themes from Kaplan*, 481-563. Oxford: Oxford University Press..
- Karmiloff-Smith, A. 1979. *A Functional Approach to Child Language: A Study of Determiners and Reference*. Cambridge: CUP.
- Kratzer, A. 2009. Making a pronoun: fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40: 187-237.
- Moyer, M., K. Harrigan, V. Haquard and J. Lidz. 2014. *2-Year-Old's Comprehension of Personal Pronouns*. Manuscript: University of Maryland.
- Munn, A., K. Miller and C. Schmitt. 2006. *Maximality and Plurality in Children's Interpretation of Definites*. Manuscript: MSU.
- Lidz, J. 2007. The abstract nature of syntactic representations: consequences for a theory of learning. In E. Hoff and M. Shatz (eds.), *Blackwell Handbook of Language Development*, 46-67. Cambridge: Blackwell.
- Loveland, K.A. 1984. Learning about points of view: spatial perspective and the acquisition of 'I/you'. *Journal of Child Language* 11(3): 535-556.
- Lukyanenko, C., A. Conroy and J. Lidz. 2013. Is she patting Katie? Constraints on pronominal reference in 30-month-olds. *Language Learning and Development* 10: 328-

- 344.
- Nunberg, G. 1993. Indexicality and deixis. *Linguistics and Philosophy* 16: 1-43.
- Postal, P. 1966. On so-called 'pronouns' in English. In F. Dinnen (ed.), *Report on the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies*, 177-206. Washington, DC: Georgetown UP.
- Roberts, C. 2003. Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26: 287-350.
- Schaeffer, J. and L. Matthewson. 2005. Grammar and pragmatics in the acquisition of article systems. *Natural Language and Linguistic Theory* 23(1): 53-101.
- Schafer, R. and J. de Villiers. 2000. Imagining articles: What a and the can tell us about the emergence of DP. In S. C. Howell, S. A. Fish and T. Keith-Lucas (eds.), *Proceedings of the Boston University Language Development Conference*, 609-620. Somerville: Cascadilla Press.
- Schwarz, F. 2009. *Two Types of Definites in Natural Language*. PhD dissertation, University of Massachusetts Amherst.
- Shipley, E.F. and T. E. Shipley. 1969. Quaker children's use of thee: A relational Analysis. *Journal of Verbal Learning and Verbal Behaviour* 8: 112-117.
- Snyder, W. 2007. *Child Language. The Parametric Approach*. Oxford University Press: OUP.
- Song, H. and C. Fisher. 2005. Who's „she“? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language* 52: 29-57.
- Stromswold, K. 1996. Analyzing children's spontaneous speech. In D. McDaniel, C. McKee and H. Smith Cairns (eds.), *Methods for Assessing Children's Syntax*, 23-54. Cambridge: MIT Press.
- Van Hout, A., K. Harrigan and J. deVilliers. 2010. Asymmetries in the acquisition of definite and indefinite NPs. *Lingua* 120: 1973-1990.
- Wexler, K. 2003. Cues don't explain learning: maximal trouble in the determiner system. In G. Ted and P. Neal (eds.), *Proceedings of the 2003 CUNY Sentence Processing Conference*.