Acquisition of the T and C System in Clausal Complements

by

Rebecca L. Norris

B.A. Russian and Linguistics

University of Arizona, 2000

SUBMITTED TO THE DEPARTMENT OF LINGUISTICS AND PHILOSOPHY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE IN LINGUISTICS

AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FEBRUARY 2004

Signature of Author: _____
Department of Linguistics and Philosophy
February 6, 2004

Certified by: _____
David Pesetsky
Professor of Linguistics
Thesis Supervisor

Accepted by: _____
Sabine Iatridou
Professor of Linguistics
Acting Head, Department of Linguistics and Philosophy

Acquisition of the T and C System in Clausal Complements

by

Rebecca L. Norris

Submitted to the Department of Linguistics and Philosophy
on February 6, 2004 in Partial Fulfillment of the
Requirements for the Degree of Master of Science in
Linguistics

ABSTRACT

In order to discover how children acquire the T(ense) and C(omplementizer) system, finite and nonfinite embedded clauses produced by children in the CHILDES database were studied. It was discovered that young children often delete *to* in nonfinite embedded clauses, and that they use *that* in the C position of finite imbedded clauses far less often than adults do.

By adapting Wexler's (1998) theory of optional infinitives to a Pesetsky and Torrego (2001, 2002) framework, I show that the facts about both finite and nonfinite embedded clauses are due to three conflicting constraints: a modified Unique Checking Constraint based on that in Wexler (1998), a conceptual constraint requiring both T and C to appear in every full clause, and a constraint which tells children to use phonological closeness to decide which goal to move when two goals are equally close to a probe. Children cannot avoid violation of at least one constraint, so they are required to violate as few as possible. This results in different possible derivations, each one of which produces results which are seen in child speech.

Thesis Supervisor: David Pesetsky
Title: Professor of Linguistics

1. Introduction*

Language acquisitionists have had a great deal of success in the last decade in their analysis of finite matrix clauses. Detailed accounts of optional infinitives, null subjects, and subject case marking have been fashioned. This paper will expand on this research by looking at the acquisition of tense phrases (TPs) and complementizer phrases (CPs) in embedded clauses.

Children start to use clausal complements very early. As I show in Section 2, there are examples of nonfinite embedded clauses in children's speech as early as one year, three months; children use finite embedded clauses as early as one year, nine months. In nonfinite clauses, children often delete the lexical item *to*, producing sentences like *I want go*. In finite embedded clauses, children do not use the complementizer *that* nearly as often as their parents do.

I will show that these seemingly unrelated facts about the acquisition T and C in finite and nonfinite embedded clauses are due to the same underlying mechanisms. I will be using Pesetsky & Torrego's (2001, 2002) framework, which provides an account of the T and C system in English. In addition, I will use many of the ideas from Wexler's (1998) theory of optional infinitives. Wexler's theory, when adapted to Pesetsky & Torrego's framework, provides the tools necessary to account for the facts about the acquisition of both finite and nonfinite embedded clauses.

* My sincere thanks go out to David Pesetsky for all of his feedback and help on this paper, and for being a wonderful advisor to me during my time at MIT. Thanks also to Ken Wexler for his help on this topic, and to Sabine Iatridou for the support, both personal and professional, she has given me. A big thanks goes to Joey Sabbagh and Aly Adler for their helpful comments when I felt stuck, and to Dónal Coomey for his patience and support at home. All errors are, of course, my own.

The structure of the paper is as follows: Section 2 contains data from the CHILDES database and statistics on children's usage of *to* and *that* in embedded clauses. In Section 3, I provide an overview of Pesetsky & Torrego's theory of T and C. In Section 4, I adapt Wexler's theory of optional infinitives to Pesetsky & Torrego's framework and show how the modified theory accounts for the *to*-deletion facts in children's nonfinite embedded clauses. Section 5 contains an analysis of the *that*-acquisition facts, based on the *to*-deletion analysis from Section 4. Section 6 is the conclusion.

2. The Data

In order to explore children's usage of T and C in finite and non-finite clauses, I searched the CHILDES database (MacWhinney and Snow (1985)) for instances of verbs which take clausal complements. After obtaining search results for these verbs, I saved the sentences of the form *verb* + [full (i.e., non-small) clause].

2.1 Non-Finite Complement Clauses

The verb I searched for to find data on children's use of non-finite complement clauses is *want*. *Want* can be followed by a clause with PRO as its subject (*PRO clause*), or by a clause with a lexical subject (*L.S. clause*).

2.1.1 Clauses with a PRO subject

I searched a total of 947 files, containing the speech of 222 American children. These children used the construction *want* + [PRO clause] starting at the age of 1;3[1]. The oldest children in the files were 6;0.

My search produced 4,613 instances of *want* + [PRO clause]. Since the purpose of this study is partly to study children's acquisition of T, I counted the utterances in which children omit *to* in these embedded infinitives – a process I will call "*to*-deletion". When the verb was transcribed as *wanna* or *wan[t] [t]a*, I analyzed it as *want* + *to*, and therefore did not count it as an example of *to*-deletion[2]. (1a-b) contains examples of *want* + [PRO clause], both with and without *to*[3]:

(1) a. *No* to-*deletion*: adam14 (2;9)
`*CHI: want to write.`

b. to-*deletion*: adam15 (2;10)
`*CHI: want go.`

The results for all children studied are as follows:

---

[1] All ages are specified in the form *years;months;(days)*. All examples from the database are preceded by the name of the file they came from and the age of the child in the file.

[2] Pullum (1997) argues that *wanna* is not syntactic contraction of *want* + *to*, but a headed morphological structure produced by a morpholexical rule. This morphological structure takes a bare VP complement. If this is correct, some (but not all) of the examples I counted as having the structure *want* + [$_{CP}$ [$_{TP}$ … to [$_{VP}$ …]]] should actually have the structure *wanna* + [$_{VP}$ …]. This should not be problematic for my analysis; it would simply mean that the percentage of utterances containing actual deletion of *to* from a clausal complement would be *higher* than what I calculated here.

[3] Is it possible that examples like (1b) are actually a misanalysis of *want* as a verb taking a finite complement clause? Then (1b) could be an OI, not an example of *to*-deletion. (Thanks to Sabine Iatridou and David Pesetsky for pointing out this possibility.) I do not think this analysis is viable, though – out of the thousands of *want* + [clause] examples in the database, there are few or no examples of the clause being an inflected finite clause. Since OIs are, by definition, *optional*, if (1b) is an OI, we would also expect to see some cases of finite clauses in the same environment.

(2) to-*deletion in* want + *[PRO clause]: all children*

Age[4]	*want* + [PRO clause]	~~*to*~~	% ~~*to*~~ of total
1;3 – 2;0	186	68	37%
2;0 – 2;4	708	191	27.0%
2;4 – 2;8	1453	196	13.5%
2;8 – 3;0	899	130	14.5%
3;0 – 3;4	629	51	8.1%
3;4 – 3;8	243	9	3.7%
3;8 – 4;0	170	5	2.9%
4;0 – 4;6	147	0	0%
4;6 – 5;0	124	1	0.8%
5;0 – 6;0	54	1	2%

The trend in Table (2) is clear: by age 4, children have essentially stopped dropping *to*. In fact, most of them have stopped *before* this age, although this is not apparent from the table. One child, Nathaniel, accounts for a large portion of the examples of *to* – deletion after age 3;0. Although only 46 of the 629 examples of *want* + [PRO clause] in the age range 3;0 – 3;4 are from Nathaniel, half of the utterances with *to*-deletion (25 out of 51) are his. Between 3;4 and 3;8, Nathaniel has a total of 15 PRO clauses – but 6 of the 9 instances of *to*-deletion are his. Nathaniel's acquisition of infinitives seems to lag a few months behind that of the other children. Another example of this lagging is his L.S. clauses, which he uses very rarely – he has only 6 utterances of *want* + [L.S. clause], versus 231 *want* + [PRO clause] utterances. His usage of L.S. clauses as a percentage of all of his *want* + [full clause] uterances is 2.5%, while all of the other children between the ages of 2;4 and 3;10 (the age range of Nathaniel's files) use

[4] Ages in tables are abbreviated: *2;0 – 2;4* means *2;0.0 – 2;3.31*.

L.S. clauses 14% of the time. Without Nathaniel's data the table would look as in (3); cells which have changed after the deletion of Nathaniel's data are shaded:[5]

(3) to-*deletion in* want + *[PRO clause]: all children except Nathaniel*

Age	*want* + [PRO clause]	~~*to*~~	% ~~*to*~~ of total
1;3 – 2;0	186	68	37%
2;0 – 2;4	708	191	27.0%
2;4 – 2;8	1308	193	14.8%
2;8 – 3;0	885	130	14.7%
3;0 – 3;4	583	26	4.5%
3;4 – 3;8	228	3	1.3%
3;8 – 4;0	166	3	1.8%
4;0 – 4;6	147	0	0%
4;6 – 5;0	124	1	0.8%
5;0 – 6;0	54	1	2%

The result of *to*-deletion is a bare verb stem which looks remarkably like an Optional Infinitive (OI). I would like to integrate an explanation of *to*-deletion with an explanation of OIs. Table (3) shows that the average child essentially stops deleting *to* in PRO clauses, with few exceptions, around the age of three and a half. Does the OI stage go this late? If the OI stage ends earlier than the *to*-deletion stage, we will have a timing problem. In fact, I found many examples of OIs up to 3;6 and even later; (4) shows some of them:

[5] Strangely, before 3;0 Nathan only deletes *to* 3 times out of 159 *want* + [PRO clause] utterances. I suspect that he sees *wanna / want+to* as a single lexical item at this age and that he has not yet acquired the verb *want*, so *to*-omission is not an issue. If this is the case, then the lexical item *wanna / want+to* in Nathaniel's grammar takes a bare VP complement. By the time Nathaniel turns 3, he has acquired the adult construction *want* + [PRO clause]. I recently discovered that Roberts (1997) makes the same proposal; see fn. 13 as well.

(4) a. nat36>harvey (3;0)
*CHI: he want a go to fishing -.
b. nat36>john (3;0)
*CHI: he want to talk on te:o:n tefelone [= telephone].
c nina35 (2;10.6)
*CHI: she # she want to see the moth.
d. adam32 (3;6.0)
*CHI: and what he want to do with them?
e. adam38 (3;9.0)
*CHI: Paul want to shave.

(4) contains only a few of the many examples I found, simply by glancing through the results of my *want* search, of children using OIs into their late twos and through their threes. The idea that OIs and *to*-deletion can be explained by the same mechanism does not, therefore, seem to suffer from a timing problem. Although children produce far *fewer* OIs as they get older, table (3) shows that their production of *to*-deletion is also decreasing at this age. This is another similarity between the two phenomena.

2.1.2 Clauses with a lexical subject

To find nonfinite clausal complements of *want* with a lexical subject, I used the results of the same *want* search discussed in Section 2.1.1. I counted the number of utterances of the form *want* + [L.S. clause] and, within the results, again counted the occurrences of *to*-deletion (see (5a-b)).

(5) a. *No* to-*deletion*: adam30 (3;5)
*CHI: I want Paul to play with dis.
b. to-*deletion*: adam13 (2;9)
*CHI: want me open?

582 instances of *want* + [L.S. clause] were found in children between the ages 1;3 and 5;0. The distribution of *to*-deletion within these utterances can be found in (6):

(6) to-*deletion in* want + *[LS clause]: all children*

Age	*want* + [L.S. clause]	~~*to*~~	% ~~*to*~~ of total[6]
1;3 – 2;0	10	6	(60%)
2;0 – 2;4	42	18	43%
2;4 – 2;8	100	25	25%
2;8 – 3;0	78	40	51%
3;0 – 3;4	255	166	65.1%
3;4 – 3;8	53	4	8%
3;8 – 4;0	53	4	8%
4;0 – 5;0	41	13	32%

Although there is a dip in the percentage of *to*-deletion between the ages of 3;4 and 4;0, after 4;0 children again are often dropping *to* in L.S. clauses. Compare the percentages of *to*-deletion at various ages in (6) and in (3). There is a noticeable difference – while children gradually stop deleting *to* after the age of 3 in PRO clauses, in L.S. clauses they continue to delete it even into the 4s.

Most of the relevant data on the discrepancy between frequency of *to*-deletion in PRO and L.S. clauses comes from one child, Adam. There are fewer files in CHILDES as ages get higher, and there are few children in the database with as many files as Adam. I had very few children with sufficient data to study over the age of 4. 47 of the 106 L.S. utterances in the age range 3;4 – 4;0 are from Adam, and 37 of the 41 between 4;0 and 5;0 are his. So it is mostly Adam's data that shows the trend that children in their later 3s and 4s can still delete *to* in L.S. clauses, but not in PRO clauses. It is not that the other

[6] When there are not enough data points to produce a statistically significant percentage, the computed percentage is in parentheses.

children's data does *not* show this trend; it is simply that we do not *have* sufficient data from any other child. Adam's data, for both PRO and L.S. clauses, are shown in (7-8)[7]:

(7) to-*deletion in* want + *[PRO clause]: Adam*

Age	*want* + [PRO clause]	~~*to*~~	% ~~*to*~~ of total
2;4 – 2;8	61	1	2%
2;8 – 3;0	67	56	84%
3;0 – 3;4	153	15	9%
3;4 – 3;8	49	2	4%
3;8 – 4;0	63	3	5%
4;0 – 4;6	128	0	0%
4;6 – 5;0	87	1	1%
5;0 – 6;0	41	1	2%

(8) to-*deletion in* want + *[LS clause]: Adam*

Age	*want* + [L.S. clause]	~~*to*~~	% ~~*to*~~ of total
2;4 – 2;8	1	0	(0%)
2;8 – 3;0	41	37	90%
3;0 – 3;4	173	164	95%
3;4 – 3;8	26	2	8%
3;8 – 4;0	21	3	14%
4;0 – 5;0	37	13	35%

Note that Adam, like Nathaniel, has a *lower* frequency of *to*-deletion in PRO clauses at a very young age (2;4 – 2;8); the frequency then spikes right before 3;0 and gradually goes down to near-zero. See fn. 3 for comments on this fact.

[7] Note that Adam seems to stop *to*-deletion in PRO clauses a few months after the average child from Table (3). He also deletes *to* in L.S. clauses at a young age almost 100% of the time. Neither of these facts should affect our analysis: we are only interested in *what happens* during the *to*-deletion stage for PRO and L.S. clauses and in what makes these types of clauses different. The exact percentages of deletion do not matter, and the exact ages of the stages do not matter: all that matters is that children delete *to* some of the time, and they seem to exit the *to*-deletion stage for PRO clauses before they do for L.S. clauses.

In Section 4, I provide an account of *to*-deletion as well as a possible explanation for Adam's asymmetry between PRO and L.S. clauses. If later research shows that this asymmetry exists for other children as well, my proposal will extend to them.

2.2 Finite Complement Clauses

To conduct my search for finite embedded clauses, I chose five verbs which regularly take finite complements. The verbs were *believe, know, say, tell,* and *think.* I searched 1535 files which contained the speech of 338 American children. In the millions of utterances of these files, there were only two examples of children using the construction *believe* + [finite clause]. I did not use data from that verb, so I was left with four verbs in all. The children in these clauses started using these verbs followed by finite clauses at the age of 1;9, and my data goes through age 6;0. I counted the total instances of each utterance of the form *verb* + [finite clause] for each child, and, in my pursuit of data on children's acquisition of C, I counted how many times they used the complementizer[8] *that.* I was careful to only count cases where *that* was clearly introducing a clause, as in (9a'), not acting as a determiner or a pronoun, as in (9a). I found a total of 219 instances of kids using *know* + [finite clause], 190 of *say*, 94 of *tell*, and 936 of *think.*

(9) shows examples of children using all four of these verbs, both with and without *that* (in **boldface**) in C position:

[8] Since *that* ends up in C position, I will call it a complementizer to distinguish it from homophonous *that*s in English. Later, I will be using an analysis that argues that *that* is not actually base-generated in C.

(9) ***know*:**

a. n69 (Naomi) (2;6.14)

```
*CHI: Kimberly didn't know that won't go in .
```

a'. abe159 (4;1.9)

```
*CHI: well # do you think that they would know that it says
their name ?
```

***say*:**

b. eve19 (2;3)

```
*CHI: Papa [//] we said you was a boy .
```

b'. abe159 (4;1.9)

```
*CHI: well # we'll have to say that it's from Abe and Ann
Mommy # I'll do it on the other side ok ?
```

***tell*:**

c. 43-00ursula (3;7)

```
*CHI: she tells him her kitty cat's gone .
```

c'. n89 (Naomi) (3;5.7)

```
*CHI: and he told him that a giant is going to come
outside .
```

***think*:**

d. eve12 (1;11.0)

```
*CHI: I think it going [/] going round now .
```

d'. 38-03susan (3;2.3)

```
*CHI: I think that I can take off this screw here .
```

2.2.1 Usage of that*: totals for all children studied*

Tables (10-13) show, for all children combined, occurrences of *verb* + [finite clause] and occurrences of *that* in C position:

(10) *occurrence of* that *in* think + *[finite clause]: all children*

Age	*think* + [finite clause]	*that*	%*that* of total
1;9 – 2;0	14	0	0
2;0 – 2;4	18	1	6%
2;4 – 2;8	29	2	7%
2;8 – 3;0	63	1	2%
3;0 – 3;4	146	5	3%
3;4 – 3;8	130	0	0
3;8 – 4;0	120	0	0
4;0 – 4;3	96	11	11%
4;3 – 4;6	106	0	0
4;6 – 4;9	96	1	1%
4;9 – 5;0	75	1	1%
5;0 – 6;0	43	0	0%

(11) *occurrence of* that *in* know + *[finite clause]: all children*

Age	*know* + [finite clause]	*that*	%*that* of total
1;9 – 3;0	20	1	5%
3;0 – 3;4	29	3	10%
3;4 – 3;8	30	1	3%
3;8 – 4;0	36	3	8%
4;0 – 4;3	24	5	21%
4;3 – 4;6	15	2	13%
4;6 – 4;9	38	7	18%
4;9 – 5;0	18	1	6%
5;0 – 6;0	8	0	(0)

(12) *occurrence of* that *in* say + *[finite clause]: all children*

Age	*say* + [finite clause]	*that*	%*that* of total
2;0 – 2;6	11	1	9%
2;6 – 3;0	34	1	3%
3;0 – 3;6	46	2	4%
3;6 – 4;0	45	2	4%
4;0 – 4;6	24	3	13%
4;6 – 5;0	25	2	8%
5;0 – 6;0	5	0	(0)

(13) *occurrence of* that *in* tell + *[finite clause]: all children*[9]

Age	*tell* + [finite clause]	*that*	*%that* of total
2;0 – 3;0	21	0	0
3;0 – 3;6	22	2	9%
3;6 – 4;0	22	2	9%
4;0 – 4;6	14	1	7%
4;6 – 5;0	13	1	8%

What the data in (10-13) do not show is that there is considerable discrepancy between different children in their usage of *that*. I also counted parents' instances of *that*, and I discovered that adults' frequencies of use also differ. More telling than a look at all children combined is an examination of individual children and their parents.

2.2.2 *Usage of* that*: totals for individual children*

First, I will show the data for Sarah and her parents. Sarah has a low frequency of *that* with the verbs *think* and *know*; there is not enough data on her usage of *tell* and *say* to calculate statistically significant frequencies with these verbs. (14) shows Sarah's usage of *that* with *think/know* + [finite clause]; (15) shows that of her parents:

(15) *occurrence of* that *with* think *and* know*: Sarah*

Age	*know*	*that*	*%that*	*think*	*that*	*%that*
3;0 – 4;0	6	1	(17%)	25	0	0
4;0 – 6;0	15	0	0	50	0	0

(16) *occurrence of* that *with* think *and* know*: Sarah's parents*

know	*that*	*%that*	*think*	*that*	*%that*
102	5	5%	316	3	0.9%

[9] I have no data beyond 5;0 for *tell* because there are no instances of children past that age using it.

Sarah's parents use *that* in the construction *know* + [finite clause] only 5% of the time, and they use *that* with *think* only 0.9% of the time. Sarah has little input data containing *that* in C position, so it is not surprising that she only uses *that* one time out of 21 utterances of *know*, and does not use it at all with *think*. We see a similar pattern of data with Peter's files: Peter has 47 utterances containing the relevant *verb* + [finite clause] constructions, and none of them contain the complementizer *that*. Peter's parents, likewise, never use *that* in 70 such constructions.

Peter's and Sarah's data therefore tell us little about children's acquisition of the complementizer *that*. We need to look at a child whose parents use it more frequently, and see if that child's usage reflects that of his parents. Abe is such a child. Abe's parents use *that* in C position a large percentage of the time, as the table in (17) shows:

(17) *occurrence of* that *with* think, know, say, *and* tell*: Abe's parents*

know	***that***	***%that***	***think***	***that***	***%that***	***say***	***that***	***%that***	***tell***	***that***	***%that***
69	15	22%	594	55	9.0%	61	20	33%	54	36	67%

With the verbs *know*, *say*, and *tell*, Abe's parents use that a substantial portion of the time – anywhere from 22% to 67%. Abe is therefore receiving a large amount of input data containing *that*, so we might expect him to use it frequently, as his parents do. The tables in (18) show that this is not usually the case:

(18) a. *occurrence of* that *in* know + *[finite clause]: Abe*

Age	***know* + [finite clause]**	***that***	***%that* of total**
2;8 – 3;0	9	1	(11%)
3;0 – 3;6	23	2	9%
3;6 – 4;0	20	0	0
4;0 – 4;6	23	3	13%
4;6 – 5;0	39	9	23%

b. *occurrence of* that *in* think + *[finite clause]: Abe*

Age	*think* + [finite clause]	*that*	*%that* of total
2;8 – 3;0	10	0	0
3;0 – 3;4	67	1	2%
3;4 – 3;8	60	0	0
3;8 – 4;0	72	0	0
4;0 – 4;3	32	5	16%
4;3 – 4;6	16	0	0
4;6 – 4;9	42	0	0
4;9 – 5;0	24	1	4%
5;0 – 6;0	13	0	0

c. *occurrence of* that *in* say + *[finite clause]: Abe*

Age	*say* + [finite clause]	*that*	*%that* of total
2;6 – 3;0	8	1	(13%)
3;0 – 3;6	24	2	8%
3;6 – 4;0	20	0	0
4;0 – 5;0	23	5	22%

d. *occurrence of* that *in* tell + *[finite clause]: Abe*

Age	*tell* + [finite clause]	*that*	*%that* of total
2;6 – 3;0	5	0	(0)
3;0 – 3;6	12	0	0
3;6 – 4;0	14	1	7%
4;0 – 5;0	11	1	9%

With the verb *know*, Abe does not begin to use *that* at the same frequency as his parents until past 4;6. Before 4;0, his frequency is *far* lower than that of his parents. Likewise, with the verbs *think* and *say* his *that*-frequency does not begin to approach that of his parents until 4;0. Most interesting is the verb *tell*. Abe's parents use *that* with *tell* a full two-thirds of the time. Abe, even after age 4, uses it only about a tenth of the time.

Adam is another example of a child who hears his parents use *that* at a high frequency (although not with all verbs). Adam's parents' data are shown in (19):

(19) *occurrence of* that *with* think, know, say, *and* tell*: Adam's parents*

know	*that*	*%that*	*think*	*that*	*%that*	*say*	*that*	*%that*	*tell*	*that*	*%that*
34	9	27%	544	2	0.4%	50	5	10%	19	6	32%

Adam's data can be found in (20); he uses *think* far more than *know*, *say*, and *tell*, but I will show all four verbs:

(20) a. *occurrence of* that *in* know + *[finite clause]: Adam*

Age	*know* + [finite clause]	*that*	*%that* of total
3;0 – 4;0	13	0	0
4;0 – 5;0	12	0	0

b. *occurrence of* that *in* think + *[finite clause]: Adam*

Age	*think* + [finite clause]	*that*	*%that* of total
3;0 – 4;0	15	0	0
4;0 – 4;3	29	0	0
4;3 – 4;6	67	0	0
4;6 – 4;9	38	1	3%
4;9 – 5;0	36	0	0

c. *occurrence of* that *in* say + *[finite clause]: Adam*

Age	*say* + [finite clause]	*that*	*%that* of total
3;0 – 4;0	23	1	4%
4;0 – 5;0	18	0	0

d. *occurrence of* that *in* tell + *[finite clause]: Adam*

Age	*tell* + [finite clause]	*that*	*%that* of total
3;0 – 4;0	8	0	(0)
4;0 – 5;0	11	1	9%

It is not surprising that Adam only uses *that* once out of 186 occurrences of *think*; his parents' frequency of *that* with this verb is less than one percent. With the other verbs, however, Adam's usage of *that* is significantly less frequent than that of his parents. Adam's parents use *that* with *say* one in ten times; Adam uses it one in 40. The disparity between his *that* usage and his parents' is even more striking with *tell* and *know* – his parents introduce complement clauses of these verbs with *that* nearly a third of the time, but he does rarely with *tell* and never with *know*. As with Abe, although Adam hears plenty of examples of the complementizer *that*, his acquisition of the word is delayed with respect to the relevant verbs: he acquires the structure *verb* + [finite clause] before he acquires *that* as a complementizer of that clause.

When we look at these individual children, a trend emerges: not surprisingly, if parents' frequency of *that* in complement clauses is low, their children's frequency is too. However, some parents use *that* very frequently, and we notice two facts about their children. First, at least in some cases children's usage of *that* is delayed in comparison to their usage of *verb* + [finite clause]. Second, even after these children start to use the complementizer *that*, their frequency of use is appreciably lower than that of their parents.

What is keeping children like Abe and Adam, who hear plenty of instances of the complementizer *that*, from using it more often? A child like Abe starts increasing his usage of *that* after the age of 4, although he still does not use it nearly as often as his parents. Before 4;0, his *that*-frequency is far lower – even though he is using finite

complement clauses in his very youngest files (2;6). How do we account for this pattern of acquisition?

In Section 3, I give an overview of Pesetsky & Torrego's (2001, 2002) theory of the nature of T and C. This provides a framework for both the account of *to*-deletion offered in Section 4 and the answers to the above questions, found in Section 5. We will see that both *to*-deletion and the patterns of *that* acquisition can be accounted for by the same mechanisms.

3. The Relation between T and C

All of the questions we need to answer involve the tense node, T, or the complementizer node, C. To review, the questions are as follows:

- Why do young children often fail to pronounce T (*to*) in nonfinite clauses headed by *PRO* or a lexical subject (L.S.?)
- Why is there a period in which children continue to fail to pronounce T in L.S. clauses, but not in PRO clauses?
- Why do some very young children fail to use the complementizer *that* to introduce finite embedded clauses even when their parents use *that* a substantial percentage of the time?
- Why do children who have acquired the complementizer *that* still use it at a frequency below that of their parents?

I argue in Sections 4 and 5 that the answers to all of these questions are related. I will be using the framework developed in Pesetsky & Torrego (2001, 2002) (henceforth *P&T*).

P&T suggest that the "complementizer" *that* is not a simple head base-generated in C position; it is actually the product of movement of T-to-C. This theory could be useful in an explanation of children's patterns of *that*-usage. If *that* were a simple head, it would be difficult to understand why children who hear it used frequently by their parents still have a hard time acquiring it. With a more complicated definition of *that,* the children's data make more sense. In this section, I give an overview of P&T's theory. We will see in later sections that not only does their framework give a basis for a satisfying explanation for the *that* data; it also helps to understand why children delete *to* in nonfinite clauses.

3.1 The Nature of Nominative Case

Linguists have long maintained that the subject agreement found on finite T in languages like English is due to the fact that T has ϕ-features. Since these ϕ-features have no semantic import when attached to Tense, they are uninterpretable. P&T abbreviate uninterpretable features F as *u*F and interpretable ones as *i*F, so we can say that T has $u\phi$. Any DP has $i\phi$, and the subject DP moves out of the verb phrase so that its $i\phi$ can check T's $u\phi$. P&T suggest that this relationship between the subject and T goes both ways: not only does T carry uninterpretable features matched by interpretable ones on the subject DP, but the subject DP also carries an uninterpretable feature, *u*T, matched by *i*T on T. How far does this symmetry reach? $u\phi$ on T is realized as agreement with the subject. Is *u*T on DP realized morphologically as well? P&T say it is:

(20) *The nature of nominative case*
Nominative case is *u*T on D. [(8) in P&T (2001)]

When the proposal in (20) is combined with P&T's hypotheses about the nature of C, a wide variety of facts are accounted for, including the *that*-trace effect, *that*-omission asymmetries, and T-to-C asymmetries in matrix questions. I will briefly recount their analysis of the *that*-trace effect, since it relates to embedded finite clauses, one of the topics of this paper.

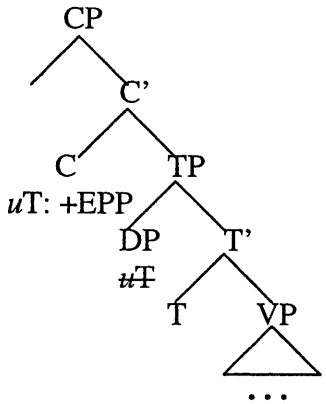
3.2 The *That*-trace Effect

One of the crucial points in P&T's account of the *that*-trace effect is that C has *u*T. (C also has ϕ-features; P&T do not discuss this in-depth but it will be important for our analyses of *to*-deletion and the *that* facts in Sections 4 and 5.) *u*T on C has an *EPP property*. P&T explain EPP effects as follows: when an uninterpretable feature F on X has the EPP property, "F on X requires that an Agree relation with F on Y be followed up with copying of material on Y into the local environment of X" (P&T (2001) p. 359). EPP on C's *u*T means that C probes for the closest instantiation of T to both check *u*T and to move into C's local environment.

A simple embedded clause with these properties, after the subject DP has already raised to Spec-TP[10], would look as follows; only features relevant to the analysis are shown here:

[10] Although *u*T on DP and *u*ϕ on T have been marked for deletion at this point in the derivation, they have not yet actually *been deleted* – this does not happen until after the completion of the phase (more specifically, in P&T (2002) it is argued that checked features are not deleted until after the phase has merged with another head.)

(21)



C probes for a goal to check its *u*T-feature. It finds two possible goals, both of which contain a T-feature: DP and TP. P&T introduce a rule called *Attract Closest Feature* which requires that C choose the closest goal. However, under a hierarchical (rather than linear) definition of closeness like that in (22), DP and TP are equally close to C:

(22) *Closeness*
Y is closer to K than X if K c-commands Y and Y c-commands X.
[(12) in P&T (2001)]

C therefore has a choice: it can Agree with DP, moving DP into its specifier, or it can Agree with TP and adjoin T to C[11]. If the former is chosen, DP moves to Spec-CP and the finished sentence looks as follows:

(23) I think $[_{CP}\ [_{DP}\text{Eustace}]_i\ \text{C}\ [_{TP}\ t_i\ [_{T}\ \text{will}]\ [_{VP}\ \text{smile}]]]$.
"I think Eustace will smile."

If C instead Agrees with TP, T adjoins to C. Why do we not see T moving to C, giving us a sentence like **I think will Eustace smile*? P&T argue that we actually *do* see

[11] TP's head, rather than all of TP, would move because of P&T's (2001) Head Movement Generalization:
(i) Suppose a head H attracts a feature of XP as part of a movement operation.
a. If XP is the complement of H, copy the head of XP onto the local domain of H.
b. Otherwise, copy XP into the local domain of H. [(13) in P&T]

something like this: the "complementizer" *that* is actually the spell-out of T in embedded, finite C[12]. After T-to-C movement, we are left with a sentence like that in (24):

(24) I think[$_{CP}$ [$_{T}$ that]$_j$ + C [$_{TP}$ Eustace will$_j$ [$_{VP}$ smile]]].
"I think that Eustace will smile."

One of several pieces of evidence for this analysis is the *that*-trace effect. When an object *wh*-phrase is extracted out of an embedded finite clause, *that* in C position is optional. When a subject is extracted, however, *that* is ungrammatical:

(25) a. What$_i$ do you think (that) Eustace was wearing t_i?
b. Who$_j$ do you think (*that) t_j was wearing that costume?

P&T explain the *that*-trace effect as follows: In instances such as (25) that involve successive-cyclic movement, embedded C (like matrix C) has an uninterpretable Wh-feature, *u*Wh. As in declarative clauses, it also has *u*T. Both features have the EPP property. Before *wh*-movement, the lower clause in a question like (25a) has a structure like that in (26):

(26) … [$_{CP}$ [C, *u*T, *u*Wh] [$_{TP}$ [$_{DP}$ Eustace, *u*T] was [$_{VP}$ wearing what]]]

Embedded C in (26) must attract and move two different goals: *what*, to check *u*Wh; and either the embedded subject *Eustace* or the embedded T, to check its *u*T. Since *Eustace* and TP are equally close to C, either one can move. If *Eustace* is moved (forming an inner specifier on the embedded CP), (26a) is the final structure; moving T produces (26b):

[12] Of course, it is not as simple as the pronunciation of T (*will* in the example here) changing to *that* – we still see T lower in the clause. P&T compare this to resumptive pronouns left by *wh*-movement: as has been argued for resumptive pronouns, *that* in C is linked by movement to T in its original position.

(26) a. What$_i$ do you think [$_{CP}$ t_i Eustace$_j$ C [$_{TP}$ t_j was [$_{VP}$ wearing t_i]]]
"What do you think Eustace was wearing?"
b. What$_i$ do you think [$_{CP}$ t_i [$_T$ that]$_j$ + C [$_{TP}$ Eustace was$_j$ [$_{VP}$ wearing t_i]]]
"What do you think that Eustace was wearing?"

The optionality of *that* in (26) arises from C's free choice as to which goal to Attract to satisfy *u*T's EPP property.

Subject *wh*-extraction from an embedded clause works differently. Before *wh*-movement, the embedded clause has the structure in (27):

(27) … [$_{CP}$ [C, *u*T, *u*Wh] [$_{TP}$ [$_{DP}$ who, *u*T] was [$_{VP}$ wearing my shirt]]]

In (27), C must attract and move the embedded subject *who* in order to check *u*Wh. It must also check *u*T by moving either *who* or the embedded T. We know, though, because of the ungrammaticality of (25b), that the T-movement option is not available in this case. P&T account for this by proposing an Economy Condition:

(28) *Economy Condition*
A head H triggers the minimum number of operations necessary to satisfy the properties (including EPP) of its uninterpretable features. [(6) in P&T (2001)]

Looking at the derivation again, we see that the choice of one of the possible goals for checking of *u*T on C will violate the condition in (28): if T is the chosen goal, two things will have to move into the local environment of C: *who* (to check *u*Wh), and T (to check *u*T). Using *who* to check *both* of C's uninterpretable features involves only one movement. By the Economy Condition, this is the option that must be followed. This predicts that in cases of subject *wh*-extraction from finite embedded clauses, T will never move to C and *that* will therefore never be spelled out. This prediction, the *that*-trace effect, holds true for Standard English:

(29) *The* that-*trace effect: sentence structures*

a. Who$_i$ do you think [$_{CP}$ t_i C [$_{TP}$ t_i was [$_{VP}$ wearing my shirt]]]
"Who do you think was wearing my shirt?

b. Who$_i$ do you think [$_{CP}$ t_i [$_T$that]$_j$ + C [$_{TP}$ t_i was$_j$ [$_{VP}$ wearing my shirt]]]
* "Who do you think that was wearing my shirt?"

We have now covered the basics of P&T's (2001) theory. In P&T (2002), one of the things they do is extend the theory to account for nonfinite clauses. C in nonfinite clauses has *u*T as well, and in these clauses the spellout of T in C is the "complementizer" *for*:

(30) I would like for John to leave.

Pesetsky & Torrego's account of T and C and the relationship between the two gives us the tools we need to answer the questions that arose in Section 2 regarding *to*-deletion and acquisition of *that*.

4. An Analysis of *to*-Deletion

Why does *to*-deletion occur? The simplest hypothesis, which I will adopt, is that *to* is missing when the T node of a nonfinite clause is omitted.[13] This is reminiscent of Wexler's (1993, 1998) and Schütze & Wexler's (1996) models of optional infinitives (OIs), which state that OIs occur when Tense and/or Agreement is omitted. The similarity of this analysis with the explanation of *to*-deletion as omission of T is attractive because OIs, which contain bare verb stems in English (e.g. *He go home*), look similar to *to*-less nonfinite complement clauses like those in (31b) and (32b), repeated from (1) and (5):

[13] After I had formulated this analysis, it came to my attention that Roberts (1997) was the first to suggest that *to*-deletion is due to T-omission. Roberts uses the presence of *to*-deletion in Adam's and Sarah's speech to argue that the T/INFL-omission explanation of OIs is superior to a clausal truncation analysis.

(31) a. *No* to-*deletion*: adam14 (2;9)
`*CHI: want to write.`
b. to-*deletion*: adam15 (2;10)
`*CHI: want go.`

(32) a. *No* to-*deletion*: adam30 (3;5)
`*CHI: I want Paul to play with dis.`
b. to-*deletion*: adam13 (2;9)
`*CHI: want me open?`

In this section, I provide an overview of Schütze & Wexler's AGR/TNS Omission Model (ATOM) and Wexler's (1998) Unique Checking Constraint (UCC) and show that, by adapting their theories to P&T's framework, *to*-omission can be explained. Section 5 will expand this analysis to account for the facts on the acquisition of *that*.

4.1 ATOM and the UCC

Schütze & Wexler's AGR/TNS Omission Model is summarized in (33). I will be focusing on (33a) for now; I will briefly address (b-c) later. In the following discussion, AGR refers to AGR_S.

(33) a. AGR or TNS (or both) may be deleted.
b. AGR assigns NOM, if no AGR, subject gets default case
c. Default case in English is ACC, in German/Dutch it is NOM
d. Kid knows adult Syntax and Morphology
[simplified from Wexler's (1998) (17)]

Morphological inflection is computed according to principles of Distributed Morphology, as follows:

(34) Features of English Verbal Inflections
/s/ [+3rd, +singular, +present]
/ed/ [+past]
/0/ [] (i.e. no features, the empty inflection is default)
[Wexler's (1998) (18)]

Why would a child decide to omit T or AGR, producing an OI? This question is also important for the question of *to*-deletion, if I am correct in my hypothesis that the mechanism of *to*-deletion is T omission. Another question which must be answered relates to apparent optionality of both of these phenomena. OIs are not found in every declarative clause uttered by a child in the OI stage. Likewise, *to*-deletion does not occur in every nonfinite embedded clause for a child in the *to*-deletion stage. What motivates deletion of T or AGR, and what produces its optionality?

Wexler (1998) expands Schütze & Wexler's theory of OIs to explain why T and AGR are deleted at all, and why children do not always choose to do so. First, he asks why subjects raise out of VP in optional infinitives, even in those where the T node is absent. There is evidence from the placement of negation that the subject does not stay in VP: children place negation after the subject, not before it as they would if the subject remained low in the tree. Wexler's answer to this question is based on Chomsky's (1995: Chapter 4) definition of the EPP:

(35) EPP is a requirement that a D (Determiner) feature be checked.
[Wexler's (1998) (26)]

Wexler assumes that both AGR and T have a D-feature that must be checked by raising a DP. He explains deletion of AGR or T by positing the Unique Checking Constraint (UCC). Children in the OI stage have this constraint in their grammars; as they grow older and move toward adult grammar they lose the constraint.

(36) *Unique Checking Constraint (UCC) (On kids in OI-stage)*
The D-feature of DP can only check against one functional category.
[(35) in Wexler (1998)]

If a child uses a subject DP to delete the D-features on both AGR and T, the UCC will be violated. To avoid this violation, the child deletes either AGR or T. As shown in (34), either deletion produces an OI.

Wexler's next question is: if it is true that kids have this constraint, why is it not the case that they *always* delete AGR or T? This would produce an OI in *every* simple declarative clause[14] – a prediction which the data show to be incorrect. His answer uses the concept, reminiscent of Optimality Theory, of conflicting constraints. Wexler surmises that children in the OI stage *know* that AGR and T are required in simple declarative clauses. They therefore know that omission of either of these heads violates the "interpretive/conceptual property which requires AGRS and TNS" (Wexler (1998), p. 65). Deriving a sentence with both AGR and T violates the UCC, but obeys this interpretive requirement. Each derivation therefore violates one constraint, and Wexler suggests that children have a rule called *Minimize Violations* (*MV*). It is impossible for a child at this stage to construct a simple declarative clause with no violations – either they will violate the UCC, or they will violate the requirement for both AGR and T. The child therefore has free choice between a derivation which results in an OI and one which does not.

4.2 Adapting Wexler's Theory to a P&T Framework

In P&T's framework, there is no need for D-features. EPP is not defined as in (35); rather, it is a property of an uninterpretable feature *u*F on head X which requires that a phrase (or its head; see fn. 9) agreeing with *u*F be moved into X's local environment.

[14] See Wexler (1998) for a discussion of the fact that OIs rarely occur in clauses which are not simple declaratives.

Another difference between Wexler's theories and those of P&T is the AGR node – for P&T, the Infl/T head contains both tense and agreement features, and there is no need for a separate AGR node.

Can Wexler's model be adapted into a P&T-style framework? It would be advantageous to do so, as my explanation for the *that*-acquisition facts will depend heavily on P&T's account. In fact, it is possible to modify Wexler's theories so that they are consistent with those of P&T.

As I mentioned briefly in Section 3.2, C has ϕ-features in P&T's framework. One reason this is necessary is so that CP can be attracted to the subject position of a higher clause, as in the examples in (37):

(37) a. [$_{CP}$That Eustace threw a tantrum] is no surprise.
b. [$_{CP}$For Eustace to throw a tantrum] would be no surprise.
c. [$_{CP}$PRO to lose one's temper] is not encouraged.

Just as occurs with a DP subject, the CP subjects in (37) are attracted by matrix T, which needs a goal to check its *u*ϕ-features. CP can only be the subject of matrix TP if its ϕ-features are interpretable; otherwise T will have no way of valuing its own *u*ϕ-features and there would be no agreement.

The CP-subject sentences in (37) are from Adult English, however. I found no examples in the database of children using CPs as subjects. This leads me to hypothesize that, for small children, ϕ on C is actually *uninterpretable*. Not only does this hypothesis explain children's lack of clausal subjects; it also provides a way to keep the spirit of Wexler's proposal while adapting it to P&T's framework.

My analysis of *to*-deletion will be an extension of my adaptions to Wexler's theory of OIs, so I will first therefore show how the modification of Wexler's theory works. First, every full clause has a C, and every C in child English has both $u\phi$ and uT. In finite declarative clauses, neither $u\phi$ nor uT on C has the EPP property. This is necessary because if these features were +EPP on finite C we would have the option of moving T to C. We do not see this instantiation of T in C, either spelled-out as T (**Is Eustace angry*; ungrammatical if declarative) or as *that* (**That Eustace is angry*; ungrammatical unless *that* is modifying *Eustace*). (38a) shows a declarative sentence structure before the subject raises out of VP and before C has merged with TP; (38b) shows the sentence after the subject has moved to Spec-TP and C has been merged in:

(38) a. $[_{TP}$ [T, $u\phi_{[+EPP]}$] $[_{VP}$ $[_{DP}$ subject, $i\phi$, uT] ....]]]
b. $[_{CP}$ C, $u\phi_{[-EPP]}$, $u\mathrm{T}_{[-EPP]}$ $[_{TP}$ $[_{DP}$ subject, $i\phi$, $\cancel{u\mathrm{T}}]_i$ [T, $\cancel{u\phi}$] $[_{VP}$ t_i ...]]]

In (38b), T's $u\phi$-feature has been checked and had its EPP property satisfied by movement of the subject DP; in addition, the subject's uT feature has been marked for deletion by T. According to P&T's definition of closeness in (22), TP and the raised subject DP are equally close to C, so either one can check $u\phi$ and uT on C.

How does the child decide which goal to choose? I propose that for the child, TP and DP are *not* actually seen as equally viable goals. Children, unlike adults, have to choose the phonologically closest goal, as explained in (39) – this proposal will become very important in Section 5 when we are examining children's difficulty with the complementizer *that*.

(39) *Phonological Closeness*
If a child finds two goals which are equally close to the probe (as closeness is defined in (22), the goal which is *phonologically closest* to the probe (i.e., pronounced first) is chosen.

In the structures in (38), although the child sees two goals which are "closest" by a c-command definition of closeness, only one goal is available according to (39): the raised subject. T is not as phonologically close to C as the raised subject, so it is unavailable as a goal for C.

This notion of closeness is not, of course, enough to explain why children produce OIs. We still need something like Wexler's UCC to motivate the deletion of some head (T for me; T or AGR for Wexler) in OIs. Although I cannot keep Wexler's version of the UCC, which relies on the existence of D-features, I believe the spirit of it is correct – in fact, I believe it should be extended. Wexler (1998) suggests that such an extension of the UCC might be in order:

> "[T]here are a variety of other ways of stating UCC which could also derive ATOM, following the leading idea that it is the *double* checking/movement process which causes a functional category to be omitted, so that convergence takes place. Perhaps in the OI stage, subject DP's can only check once, no matter *which* feature is being checked, and the 2 D-feature checking is only a special case."
>
> [Wexler (1998), p. 62]

(40) contains my modified UCC, which does not rely on D-features:

(40) *Unique Checking Constraint (Modified)*
Each feature F on a DP can only check *u*F on one other head in the derivation.

(40) is very broad, of course, and it may need to be narrowed. For now, I will keep it as it is; it serves our purposes but further research is needed to show if it should be refined.

How does (40) derive the results that Wexler used ATOM and the original UCC to explain?

In finite clauses, which have structures like that in (38b), there are two heads with *u*ϕ: C and T. Due to children's need to pick the phonologically closest goal, the only possible goal for C's uninterpretable features is the subject DP. DP already checked *u*ϕ on T, though; and the modified UCC therefore bars it from checking *u*ϕ on C. In order to avoid violation of the UCC, the child must omit one of the heads containing *u*ϕ: either T or C. If C is deleted, the verb is tensed, and in that respect the sentence looks like adult grammar. If T is deleted, the verb has no tense and an optional infinitive is produced.[15]

One problem for this P&T-inspired analysis of OIs is that it does not easily explain facts about subject case that Wexler's original theory explains. Wexler notes that accusative case is often used on the subject of OI clauses – but it is vanishingly rarely used on the subject of full tensed clauses. He explains this by saying that nominative case is assigned by AGR_S. Recall that in ATOM, OIs occur when either AGR or T is omitted. If T is omitted, AGR is still present and the subject of the OI has nominative case. If AGR is omitted, though, the subject cannot get nominative case. Default case in English is accusative, so subjects of AGR-less OIs receive accusative case.

[15] If T is deleted and there is no AGR-like node, it will be necessary for *u*ϕ and/or *u*T on C to have the EPP property, since the previously mentioned evidence from negation shows that the subject raises out of VP. Possibly the child knows that every clause has to have (at least) one *u*ϕ or *u*T that is [+EPP], so she automatically assigns this property to the feature(s) of C when T is deleted. See fn. 16 for more on this.

I do not have a solid solution to this problem, but I will sketch out one possibility: A crucial point in P&T's proposal is that nominative case is assigned by the agreement of *u*T on a subject DP with the tense feature of finite T. If all OIs occur when T is deleted, as I am suggesting, then the prediction seems to be that all OIs will have accusative case. The feature causing a problem for the child is *u*ϕ – both T and C have this feature, but the UCC says that the subject cannot check the feature on both heads. It might be possible to solve the case problem by suggesting that children do not always delete the entire T head; sometimes they simply delete *u*ϕ on T. If only *u*ϕ is deleted, T can still give nominative case to the subject. One might argue that this produces incorrect surface word order: If the T head is still in the tree but there is no $u\phi_{[+EPP]}$ on T, the subject has no way of escaping its base-generated position in VP. However, as fn. 15 suggests, it may be the case that there is a property of language which says that every full clause must have a [+EPP] ϕ-feature.[16] If this is the case, omission of *u*ϕ on T would force the child to assign the EPP property to C's ϕ-feature. C would now agree with T in order to check its *u*T feature, but it would have to look further for a goal to check *u*ϕ – that goal would be the subject, which would move to Spec-CP. This solution is not perfect, but it shows that it is possible for my modified theory of OIs to account for the case facts.

We now have a theory of optional infinitives that is consistent with P&T's framework. As of yet, we have no evidence for or against Wexler's *Minimize Violations*

[16] I admit that it is rather odd to propose that a *clause* can have a property which requires a *feature* of some head in the clause to have a certain property – in this case, the EPP property (thanks to Sabine Iatridou for pointing this out to me). However, it has been suggested (by Alexiadou and Anagnostopolou, among others) that there is a universal property of language which requires something to move out of VP. This proposal could be a way to achieve this. I will leave the issue open for now.

(*MV*): the child can omit either T or C (or, maybe, simply *u*ϕ on either head); omission of T produces an OI but omission of C does not. We will see in Section 5, though, that MV is still a crucial part of the theory.

Now that I have established a theory of OIs within a P&T framework, I can address another example of T-omission: *to*-deletion.

4.3 The Modified UCC and *to*-Deletion

After the above detailed explanation of OIs, my analysis of *to*-deletion follows easily.

4.3.1 to-*deletion in clauses with a lexical subject*

(41) repeats the data from (5), showing *to*-deletion in nonfinite clauses with a lexical subject (L.S.):

(41) a. *No* to-*deletion*: adam30 (3;5)
`*CHI: I want Paul to play with dis.`
b. to-*deletion*: adam13 (2;9)
`*CHI: want me open?`

In my analysis of nonfinite clauses like those in (41), I assume that T and C have ϕ-features, just as in finite clauses.[17] As was discussed in Section 3, we know that *u*T and/or *u*ϕ on finite embedded C has the EPP property: if no feature on embedded C were [+EPP], T would not move to C and *that* in C position would be impossible. We will assume that one or more of the features on nonfinite embedded C are also [+EPP] (although in Section 5.2 we will have reason to question this assumption).

[17] Nominative case is only assigned to a DP by *finite* T, so there is a mechanism needed for assignment of accusative case to the lexical subject of the infinitive. I will not pursue this issue now; I assume that some sort of external case-marking process is taking place.

(42) shows the structure of a non-finite embedded clause, after both T and C have been merged in and the subject has raised out of VP but before the matrix verb merges with CP:

(42) $[_{CP}\ [C\ uT_{[+EPP]},\ u\phi]\ [_{TP}\ [\text{subject},\ \cancel{uT},\ i\phi]_i\ [\text{to},\cancel{-u\phi}]\ [_{VP}\ t_i \ldots]]]$

As with the finite clause in the examples in Section 4.2, the subject in (42) has already checked $u\phi$ on T. According to the UCC, it cannot check $u\phi$ again. The child can avoid violation of the UCC by omitting either T or C from the derivation. If C is omitted, the clause looks normal, as in (41a). If T is omitted, however, *to* does not appear, and a sentence like (41b) results. The mechanism for *to*-deletion in nonfinite L.S. clauses is therefore exactly the same as the mechanism for OIs in simple declaratives.

4.3.2 to-*deletion in PRO clauses before the age of 4*

Although kids seem to stop deleting *to* in control clauses earlier than they stop deleting it in L.S. clauses, we still need an explanation for what it is they are doing in examples like (43b) (repeated from (1)), from the very early years, when they delete *to* with PRO.

(43) a. *No* to-*deletion*: adam14 (2;9)
`*CHI: want to write.`
b. to-*deletion*: adam15 (2;10)
`*CHI: want go.`

During the stage where children can still delete *to* in PRO clauses, I suggest that they are analyzing PRO exactly as they would analyze an L.S. – basically, that they are analyzing PRO as pro. Null subjects have a far broader distribution in child language than they do in adult language. This suggests that PRO and L.S. in infinitival clauses have the same analysis in the child's mind: the only difference is that one is pronounced and the other is not. PRO, T, and C for the child at this stage, then, are featurally

identical to their counterparts in a child's L.S. clause: PRO has *u*T and *i*ϕ, C has *u*T and *u*ϕ, and T has *u*ϕ. The structure of a PRO clause for a child of this age will therefore look like the structure in (42), and the analysis will be identical to the analysis of *to*-deletion in L.S. clauses.

4.3.3 Lack of to*-deletion in PRO clauses after the age of 4*

We still have not answered one crucial question: Why do children stop deleting *to* in control clauses before they stop in L.S. clauses? So far, the data from *want* have been the only evidence that between about 3;6 and 4;0, most children have stopped deleting *to* when PRO is the subject of the infinitive. In order to check and see if this extends to all PRO clauses, I counted instances of *to* deletion with two other control verbs: the subject control verb *try* and the object control verb *tell*. Examples of children using these verbs can be found in (44-45):

(44) **subject control *try***: adam25 (3;2.0)
```
*CHI: I trying to take out a pencil .
```

(45) **object control *tell***: adam34 (3;7.0)
```
*CHI: Urs(u)la told me to try some (a)nother ones .
```

For the verb *try*, I searched all of Adam's files. I counted the instances of *try* + [PRO clause] and counted how many of these included *to*-deletion. It turned out that children use *tell* as an object control verb far less often than they use subject control *try*, so I searched through all of my *tell* results from Section 2.2, this time counting utterances of the form *try* + *object* + [PRO clause] and finding instances of *to*-deletion within these. In order to verify my observation about children's lack of *to*-deletion in control contexts, it was not necessary for there to be *no* examples of *to*-deletion with these control verbs.

Rather, it was necessary for there to be no (or vanishingly few) examples of *to*-deletion starting at around the age of 3;6 or 4;0.

The results for subject control *try* are as follows: Adam had 76 utterances containing *try* + [PRO clause]. Three of these, listed in (46), contained *to*-deletion.

(46) to-*deletion in the context* try + *[PRO clause]: Adam*

a. adam15 (2;10)
```
*CHI: I'm trying get down # so dere .
```

b. adam40 (3;11)
```
*CHI: are you trying crash it ?
```

c. adam52 (5;2)
```
*CHI: you try look behind my back .
```

(46a-b) are from before the age of 4. Recall from Table (7) that 4;0 is the age at which Adam essentially stops dropping *to* in control clauses with *want*. These examples of *to*-deletion before the age of 4 with the verb *try* are no surprise, then.

(46c) is slightly more problematic because it is after the age of 4. Even with the verb *want*, however, Adam's lack of *to*-deletion after 4;0 is not without exception. Between the ages of 4;0 and 6;0, Adam has two examples of *to*-deletion out of 256 utterances of the form *want* + [PRO clause] – less than one percent. I assume these are simple performance errors. Out of Adam's 76 control-*try* utterances, 25 are between 4;0 and 6;0. (46c) could be a performance error, too. There is one other possible explanation for this "exception", though. As (47) shows, by age 4;5 Adam has acquired the "*try and*" construction:

(47) adam48 (4;5)
```
*CHI: try and stop it .
```

It could be that (46c) is Adam's attempt at a "*try and*" construction, but he either omits the *and* or it is phonologically reduced and the transcriber could not hear it.

The data from *try* therefore offer support to the observation that after the age of 4, Adam does not delete *to* in control clauses.

The data from *tell* include not just Adam, but all 338 children tested for finite embedded clauses. Since object control *tell* is not commonly used by small children, the search turned up only 55 utterances. 7 of these are, on the surface, examples of *to* deletion:

(48) a. 11a (2;3.8) (*child's name unknown*)
```
*CHI: I told horsie go with him .
```
b. gina11 (3;0)
```
*CHI: tell mommy hurry back !
```
c. shem22a (2;8)
```
*SHE: an(d) duh people tell you # get off duh table ?
```
d. shem22b (2;8)
```
*SHE: and this boy tells him get off .
```
e. tre03 (2;0.29)
```
*CHI: tell em go in here .
```
f. tre21 (3;3.4)
```
*CHI: I would tell her give me a new twuck
```
g. abe199 (4;9.24)
```
*CHI: that's what teacher told me do at school say <oh ick>
["] and <oh gross> ["] hey # Momma hey # Mommy .
```

(48a-f) are unproblematic for two reasons. First, the oldest child in any of these utterances is 3;3.4. This is still in the age range within which we expect children to delete *to* in control clauses. Also, in every one of these utterances, it is possible that there is not actually a control clause: each sentence could contain a quote. For example, Trevor in (48f) could be saying, "I would tell her, 'Give me a new truck!'"

(48g), on the other hand, looks like a genuine example of *to*-deletion (if it is not a transcription error). Because *what* is extracted from the clause following *tell*, we know the clause is not a quotation. This therefore looks like a genuine exception to the generalization about children's lack of *to*-deletion in PRO clauses after the age of 4. As

was discussed for *try* and *want*, though, this is not terribly problematic: children do make performance errors, and this error is not surprising because it was made in a sentence with a complex structure. The important fact is that they do not seem to delete *to* in control clauses in a systematic fashion after age 4, as they do in L.S. clauses.

Having established that *to*-deletion is most likely ungrammatical in control clauses for children after the age of 4, how will we account for it? In the previous subsection, I proposed that young children analyze PRO just as they would analyze a lexical subject. If, around 4 years of age, they learn that PRO clauses are featurally different than L.S. clauses, they will now be deriving the PRO clauses in a different way. This new type of derivation could result in ungrammaticality of *to*-deletion in control environments.

Is there a principled reason to believe that PRO and L.S. clauses are fundamentally different in some way? Yes. As observed by Martin (2001) and others, *want* is something of an exceptional verb. Most clauses that take a PRO subject cannot take a lexical subject (49):

(49) a. Eustace tried [PRO to keep his cool.]
b. * Eustace tried [Caspian to keep his cool.]

There must be some difference between the complement clauses of prototypical control verbs and those of non-control verbs. What is the nature of this difference, and do we predict that *to*-deletion in control clauses will be ungrammatical for children who have learned it?

This is a complex question, and the full answer is beyond the scope of this paper. I will sketch out a possible solution, however, and further research will show if it is on the right track.

I propose that both T and PRO in control clauses have a feature that neither non-control T nor lexical subjects have. A prototypical control verb selects a complement clause whose T has this feature, while a non-control verb selects for a T without the feature. PRO and non-control T suffer from feature mismatch, as do L.S. and control T. This mismatch keeps PRO from appearing in non-control clauses and keeps L.S. from appearing in control clauses. Unlike most verbs, English *want* can select for either control or non-control T.

What is this feature found on PRO and control T? It could have something to do with tense; Stowell (1982) and others have observed that control infinitivals have different temporal properties than other types of non-finite clauses, such as raising infinitives. I will remain agnostic on the exact nature of the feature for now, though, and simply call it a PRO feature. T in PRO infinitivals has an uninterpretable PRO feature, and it can only be checked by a PRO subject. Unlike T in an L.S. clause, control T has no ϕ-features[18]. PRO does have $i\phi$, though. Children around the age of 4 have learned the nature of PRO and control T, although they still believe C's ϕ-features to be uninterpretable. After raising of PRO out of VP, and merging of C with TP, a control clause for a 4-year-old child will have the structure in (50):

(50) $[_{CP}$ [C, uT, $u\phi$] $[_{TP}$ $[_{DP}$PRO, ~~uT~~, $i\phi]_i$ [T, ~~uPRO~~] $[_{VP}$ t_i …]]]

[18] Or it may have an impoverished set of ϕ-features; see Norris (2001a,b).

PRO can check both of C's uninterpretable features: it does not have to check $u\phi$ on T because T has no ϕ-features. The UCC is satisfied vacuously, so the child has no reason to omit T. *To*-deletion is no longer grammatical in control clauses.

In L.S. clauses, nothing has changed: the child still knows that $u\phi$ on T and C has to be checked, and needs to delete one of these heads in order to satisfy the UCC. *To*-deletion is still a grammatical option.

This sketch of a theory may or may not turn out to be correct. What it does show, though, is that it is possible to account for the ungrammaticality of *to*-deletion in older children's control clauses by saying that they do not learn the adult analysis of control clauses until around the age of four.

In Section 5, I will show that the mechanisms developed to account for *to*-deletion also predict the facts about acquisition of *that*.

5. Acquisition of *that* in Finite Embedded Clauses

In Section 2.2 we saw that, in general, even when a child's parents use *that* in C position at a high frequency, the child's frequency is far lower.

My proposal about children's restriction on goal selection is repeated in (51):

(51) *Phonological Closeness*
If a child finds two goals which are equally close to the probe (as closeness is defined in (22), the goal which is *phonologically closest* to the probe (i.e., pronounced first) is chosen.

(51) gives us a simple solution to the *that* problem. (52) shows the structure of a child's finite embedded clause after the subject has raised out of VP and C (which, as in adult

English, has the EPP property on at least one of its features but, unlike for adults, has $u\phi$) has merged with TP:

(52) $[_{CP}$ [C, uT$_{[+EPP]}$, $u\phi$] $[_{TP}$ $[_{DP}$ subject, ~~uT~~, $i\phi]_i$ [T, $u\phi$] $[_{VP}$ t_i ...]]]

In (52), the subject is phonologically closer to C than T. The child is therefore unable to move T to C, and *that* is unavailable.

This analysis predicts that the child will *never* use *that* – and this, of course, is not the case. This is not the whole story, though. The child still has the UCC and the conceptual constraint which requires both C and T to be present in any full clause. We can also consider (51) to be a sort of constraint; call it PC. PC states that when a child has two closest goals to choose from, the one which is pronounced earliest is the one that must be chosen.

There are now three constraints: the UCC, the PC, and the T-and-C conceptual constraint (which I will call TC). There is no possible derivation in which the child can obey all three constraints. According to Wexler's (1998) principle of Minimize Violations (MV) if the child cannot obey all constraints she must violate only one, if possible. We now have three possible derivations, each one violating one of the constraints:

1. *Violation of TC*: In order to avoid violation of the UCC, T or C is omitted from the derivation in (52). The UCC is satisfied vacuously because without both T and C in the structure, the subject DP only has to check the ϕ-features of one head. The PC is also satisfied vacuously; if C remains in the derivation then there is no T and DP is C's only possible goal. If T is

deleted, an optional infinitive results. If C is deleted, the clause looks like Adult English. Since *that* is an instantiation of T in C and either T or C is missing, there is no *that*.

2. *Violation of the UCC*: Both T and C are merged into the clause, satisfying TC. PC must be obeyed, so C's only possible goal for checking of its *u*T and *u*ϕ is the subject DP. DP therefore checks *u*ϕ on both T and C, violating the UCC. Since at least one of the uninterpretable features on C is [+EPP], when DP checks these features it must move to Spec-C. T is not in Agreement relation with C, so it stays in situ. Again, *that* does not appear in C.

3. *Violation of PC*: In order to satisfy TC, both T and C are merged in. The subject DP raises to Spec-TP and checks T's *u*ϕ features. After C merges with TP, it looks for a goal to check its own *u*ϕ features. In order to satisfy the UCC, C cannot use DP as its goal, so it chooses T, even though T is phonologically further away than the subject DP. Violating the PC, C Agrees with and moves T. T adjoins to C, checking C's *u*T and *u*ϕ features. *That* is pronounced as the instantiation of T in C.

In only one of the structures resulting from the above three derivations is *that* pronounced. In finite embedded clauses of Adult English, as discussed in Section 3, *that* appears as a result of one in *two* derivations. It therefore makes intuitive sense that children would end up with *that* less frequently in their finite embedded clauses than adults do.

Neither children nor adults choose each possible derivation an equal percentage of the time. Adults do not choose the derivation which produces *that* half the time, and children do not choose it a third of the time. There may be a tendency for *all* speakers, children *or* adults, to choose the linearly closest goal even when the choice of another goal would be equally grammatical. If T-to-C is perceived as taking more effort than the subject raising to C, it is probably not a coincidence that finite complement clauses with *that* are considered more formal than those without *that*. This does not mean that the adult derivation which produces *that* is less grammatical than the non-*that* derivation, or that Derivation 3 above is less grammatical than 1 and 2 for children. All three derivations for children and both derivations for adults are equally grammatical, but both children and adults are likely in most cases to choose the derivation which requires the least effort.

5.1 A Re-Analysis of OIs and *to*-Deletion

Now that children's idea of phonological closeness is a constraint which, under MV, can be violated if other constraints are satisfied, my proposal about OIs and *to*-deletion must be modified. (53) repeats the structure of a either a nonfinite complement clause or a finite matrix declarative, before C's uninterpretable features have been checked:

(53) $[_{CP}$ C, $u\phi$, uT $[_{TP}$ $[_{DP}$ subject, $i\phi$, $\cancel{u}$T$]_i$ [T, $\cancel{u}\phi$] $[_{VP}$ t_i ...]]]

As in finite embedded clauses, there will be three possible derivations, each one violating one constraint:

1. *Violation of TC*: In order to avoid violation of PC and the UCC, either T or C is omitted. If T is omitted, an OI results in a finite clause and *to*-

deletion results in a nonfinite clause. If C is omitted, the verb is inflected normally in a finite clause and *to* appears in a nonfinite clause.

2. *Violation of the UCC*: T and C are both included in the derivation, satisfying the TC. The PC is also satisfied, so the subject is C's only possible goal. The subject checks *u*ϕ on both C and T, violating the UCC. The sentence looks like a sentence from Adult English.

3. *Violation of PC*: Again, both T and C are merged into the clause. Since the subject has already checked T's *u*ϕ, the UCC dictates that it cannot check *u*ϕ on C. C must violate PC and Agree with T to check its uninterpretable features. A finite sentence, in which the features on C do not have the EPP property, will look like adult English. What about a nonfinite clause? If a feature of nonfinite C has the EPP property in child English, as it sometimes does in adult English, T will appear in C. In adult English, this is pronounced as the "nonfinite complementizer" *for*. I found no examples of children using *for* in C position (even when their parents do frequently), so I conclude that the features on C do not have the EPP property in child English.

These three derivations show that we can still explain the "optionality" of OIs and *to*-deletion when factoring in MV and the three relevant constraints. The principles discussed here, based on the theories of Wexler and Pesetsky and Torrego, are therefore able to account for the *that*-acquisition facts as well as the *to*-deletion facts.

6. Conclusions

When examining children's embedded clauses, finite clauses do not look like they have a great deal in common with nonfinite ones. A salient feature of young children's nonfinite clauses is that they often delete *to*, an option which is not available to adults. In finite clauses, one of the most noticeable patterns is that children use *that* in C position far less frequently than adults do. On the surface, these facts do not look related: one has to do with a term generally believed to be merged in the C node, and the other is about a term found in the T node. However, by applying Pesetsky and Torrego's (2001, 2002) theory of T and C, we see that children's acquisition of finite and nonfinite clauses have far more in common than it would appear on the surface.

By adapting Wexler's (1998) theory of optional infinitives to P&T's framework, I have shown that three constraints – the Unique Checking Constraint, Phonological Closeness, and the T-and-C Conceptual Constraint – account for both OIs and *to*-deletion. These three constraints are in conflict, so it is impossible for a child to satisfy all of them at once. The child must therefore Minimize Violations (MV) by violating only one constraint. Violation of the UCC or PC produces no OI or *to*-deletion. Violation of TC results in one of two possibilities: omission of T or omission of C. If T is omitted, an OI or an instance of *to*-deletion results; if C is omitted there is no OI in finite clauses and *to* appears in nonfinite clauses.

The similarity between children's analysis of finite and nonfinite embedded clauses becomes clear when we look at the mechanisms underlying the acquisition of *that* in embedded clauses. The *same three constraints* whose interaction produces *to*-deletion

also provide the basis for my account of *that*-acquisition. Violation of the UCC in the derivation of a finite embedded clause produces a *that*-less, inflected clause. Violation of TC produces a *that*-less clause, as well; if C is deleted the verb is inflected and if T is deleted an OI results. Finally, violation of the PC forces the child to move T to C, with the result that *that* is spelled out. For adults, there are only two possible derivations, one of which produces *that* at spell-out. Since fewer of children's possible derivations result in *that*, it is to be expected that children use *that* less frequently than adults.

Linguistic theory grows when phenomena which seem unrelated are shown to be due to the same underlying principles. This paper has taken two such phenomena, *to*-deletion and the low occurrence of *that* in small children, and argued that they are related at a deep level. This conclusion provides additional evidence for the ideas which form the basis of both Wexler's theory of optional infinitives and Pesetsky and Torrego's theory of T and C.

References

Chomsky, N. 1995. *The minimalist program.* Cambridge, Mass.: MIT Press.

Martin, Roger. 2001. Null case and the distribution of PRO. *Linguistic Inquiry 32*, 141-166.

MacWhinney, B. & C. Snow. 1985. The child language data exchange system. *Journal of child language 12*, 271-296.

Norris, B. 2001. Martin (2001): replies and refinements. Unpublished ms., MIT.

Norris, B. 2001. Case checking and the subjects of infinitives in Belfast and Standard English. Unpublished ms., MIT.

Pesetsky, D. and E. Torrego. 2001. T-to-C movement: causes and consequences. In M. Kenstowicz, ed., *Ken Hale: a life in language*, 355-246. Cambridge, Mass.: MIT Press.

Pesetsky, D. and E. Torrego. 2002. Tense, case, and the nature of syntactic categories. In J. Guéron and J. Lecarme, eds., *The Syntax of Time*. Cambridge, Mass.: MIT Press.

Pullum, G.K. 1997. The morpholexical nature of English *to*-contraction. *Language 73*, 79-102.

Roberts, T. 1997. Evidence for the Optional Tense Hypothesis: tense in subordinate clauses in the acquisition of English. Unpublished ms., MIT.

Schütze, C. and K. Wexler, 1996. Subject case licensing and English root infinitives. In A. Stringfellow, D. Cahan-Amitay, E. Hughes and A. Zukowski, eds., *Proceedings of the 20th Boston University Conference on Language Development*, 670-681. Boston, Mass.: MIT Press.

Stowell, T. 1982. The tense of infinitives. *Linguistic Inquiry 13*, 561-570.

Wexler, K. 1993. Optional infinitives, head movement and the economy of derivations. In Lightfoot, D. and Hornstein, N., eds., *Verb Movement*, 305-350. Cambridge, UK: Cambridge University Press.

Wexler, K. 1998. Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage. In *Lingua 106*, 23-79.