Silent and overt arguments in child language

Ivona Ilić[1] & Artemis Alexiadou[1,2]

SinFonIJa 15, University of Udine, September 22-24, 2022

[1]Humboldt-Universität zu Berlin
[2]ZAS Berlin

Puzzles: PRO vs. RtO & overt PRO

Puzzle: PRO vs. RtO

- In the literature, it has been controversially discussed whether children master PRO structures before or after Raising to Object (RtO) structures.

- While Kirby (2011) argues that RtO (1b), is easier for children, Landau and Thornton (2011) propose that children avoid deficient RtO structures and prefer full complements, (1a), instead.

(1) a. John$_i$ wants [PRO$_i$ to solve the puzzle].

b. John$_i$ wants him$_j$ [t$_j$ to solve the puzzle].

- Our data demonstrates that children have available silent arguments in the form of **PRO** (2-a) and **traces** of the raised argument in RtO (2-b) very early on.

(2) a. I$_i$ want [PRO$_i$ to read a book]. E, 2;00, *M-E-Manchester*

  b. I$_i$ want him$_j$ [t$_j$ to eat]. 4687NH, 2;00, *NewmanRatner*

- However, we observe an apparent **asymmetry** in the acquisition of the verb *want* depending on the subject of the embedded clause.

PRO and RtO differ in four important respects in our data, as PRO:

- **appears first** in child spontaneous production (see also Landau and Thornton 2011) compared to RtO
- has significantly **higher frequency**
- differs in the choice of **Voice** of the embedded verb compared to traces
- differs in realization of infinitival ***to***

Puzzle: Overt PRO in child language

- PRO, a subject argument typically attested in non-finite clauses, has to be **coreferential** with its antecedent (3-a), and **cannot be overtly expressed** (3-b) (for the extensive discussion see also Landau 2000, 2004, Bhatt and Pancheva 2017, McFadden and Sundaresan 2018).

(3)   a.   John$_i$ wants [PRO$_i$ to ride a bike].

      b.   *John$_i$ wants [John$_i$ to ride a bike].

- In contrast to adult grammar, our data provides evidence for **overt PRO** in child language.

- We demonstrate that overt PRO is widely attested across child corpora and comes in **three guises**: (i) overt PRO in the form of the pronoun in nominative case (4-a), (ii) pronoun in accusative case without infinitival *to* (4-b), (iii) pronoun in accusative case with infinitival *to* (4-c).

(4)   a.   I want **I** go.   Olivia, 1;06.00, *Wells*

       b.   I want **me** read the other way.   Nina, 2;02.28, *Suppes*

       c.   I don't want **me** to go into the doctor.   Ross, 4;06, *MW*

Proposal

Meaning First architecture

- **Generator** - a language-independent component consisting of an inventory of logical primitives that combine into complex concepts.

- **Compressor** - a morphological component that radically reduces full conceptual representation (CR) to articulated strings suitable for communication.

- MF adopts the *Late Insertion* of Distributed Morphology.

- However, unlike DM, MF develops *Universal Late Insertion* $\rightarrow$ DM - structure generation is part of grammar, MF - structure generation is *outside of grammar*

The Compressor (Alexiadou 2022)

- **Two types of morphemes (I)**:
  - contributing functional information - tense, number
  - contributing lexical content - roots
- **Two types of morphemes (II)**:
  - realization of a particular unit (Vocabulary items)
  - abstract morphemes that combine to build this unit (Beard 1995)
- Words $\neq$ lexical entries, but are composed on the basis of abstract morphemes in combination with roots, the most deeply embedded morphemes.
- The sequence of functional morphemes is universal.
- Both functional elements and roots are subject to late insertion, i.e. the building blocks of meaning lack phonological content.
- Certain morphemes are simply realized via zero.

PRO as compressed unpronounced subject

- Subjects are always present at the thought level but may stay unpronounced in some languages and environments (Sauerland and Alexiadou 2020).

- This applies to both *pro* and PRO.

- Under MF, a subject corresponding to PRO exists at the thought level (in form of a noun or a pronoun). It is compressed within the Compressor component, resulting in unpronounciation (PRO).

Alexiadou et al. (2021) develop the idea that:

- Individual languages reflect only fragments of thought structures, i.e., conceptual representations (CRs) → big blind spots of cognitive structures that no adult language ever expresses.

- Grammatical sentences in adult language are maximally compressed → the least informative for investigating the underlying universal CRs.

Transparency principle

- Children are biased towards a one-to-one mapping from CRs to language.

- Evidence for one-to-one mapping $\rightarrow$ **errors of commission**
- Often in child grammars overt realizations of material that stays unpronounced in adult grammar can be found.

(5) Das Mädchen sein, das der Opa **das Mädchen**
the girl be that the granddad the girl
umarmt.
hugs

Lit.: 'Be the girl who the granddad hugs the girl.'

<div align="right">Yatsushiro and Sauerland (2018)</div>

- **Errors of commission** serve as a tool for investigating universal conceptual representations (CRs).

Research Question & Hypotheses

Research Question: Why do we find overt PRO in child language?

Hypotheses

- **Unpronounced PRO Hypothesis**: In line with Sauerland and Alexiadou (2020), the subject of the embedded clause is established at the thought level and identical to the subject of the main clause. In the Compressor component, it undergoes the compression mechanism resulting in unpronounced PRO in adult language.

- **Overt child PRO Hypothesis**: Following 1:1 mapping principle between conceptual representations (CRs) and morphology (Alexiadou et al. 2021), children overtly realize subject of an embedded clause that must stay unpronounced in adult language before fully mastering rules of compression.

Corpus study: Methodology & Data

- CHILDES database (MacWhinney 2000) - we extracted all the utterances surfacing with the verb *want*[1] from 49 corpora of typically developing children acquiring UK and US English
- CLAN software
- **UK English**: 210 children, age range: 1;0-7;0
- **US English**: 1112 children, age range: 0;6-8;0

---

[1]We did not collect data including the form *wanna*.

- Total number of utterances containing the target verb: N = 31002

- Utterances in which the verb *want* is followed by an embedded clause: N = 7756 $\rightarrow$ selected for the analysis

- The analysis covers the time span between the age of 1;4 when the first occurrence of control is attested in our corpora, and the age of 6;5 (N = 7563).
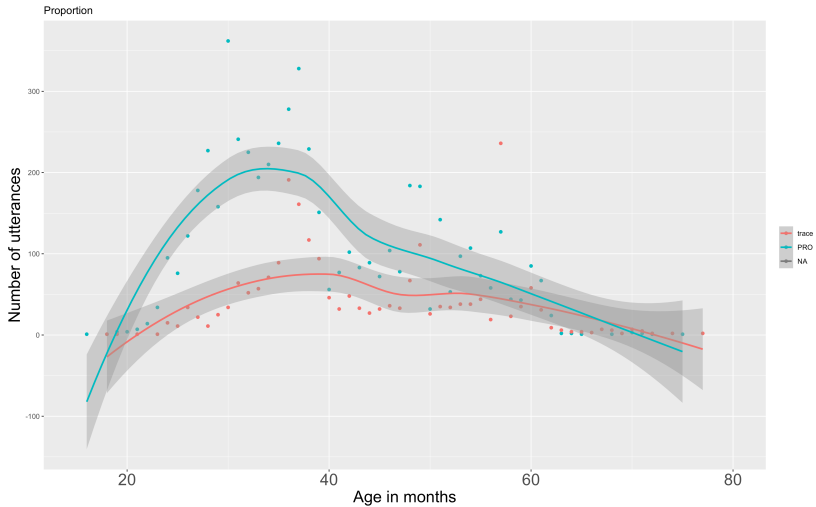
(I) Type of the silent subject

- PRO
- trace

(II) Voice of the verb in the embedded clause

Verb class	Example
non-core transitives (NCTs)	*eat, write*
causatives	*build, open*
unergatives	*walk, run*
stative transitives (STs)	*love, have*
statives	*live, sit*
unaccusatives	*fall, come*
anticausatives	*melt, light up*

Results: PRO vs. RtO

Results: Developmental trajectory

Results: Developmental trajectory

- PRO (6-a) appears first in the child spontaneous speech compared to RtO (7-a), aligning with Landau and Thornton's (2011) results.

- In both cases, *to* is omitted in the first months.

(6)   a.   I wan(t) go boom!²   Ross, 1;04.11, *MacWhinney*

       b.   I want ta have some.   Peter 2;02.13 *Bloom*

(7)   a.   I want Mommy read.   Eve, 1;06.00, *Brown*

       b.   I want him to eat.   4687NH, 2;00, *NewmanRatner*

---

²Examples (6-a) and (7-a) are first occurrences of PRO and RtO, respectively, with omitted *to*; in contrast, (6-b) and (7-b) demonstrate first occurrences involving overt *to*.
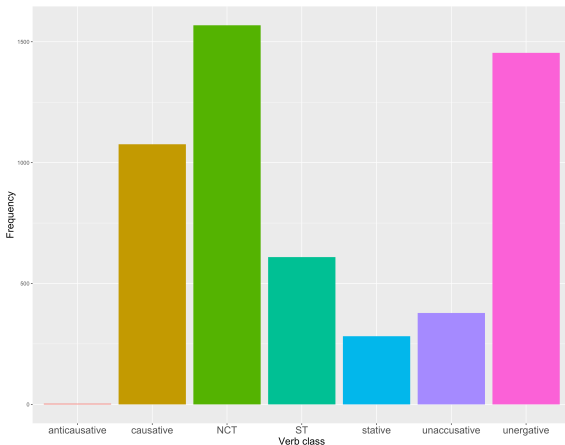
- Embedded clauses surfacing with PRO in the context of the control verb *want* are significantly more frequent.

- Both forms record increase in frequency between 30 and 40 months (age 2;6-3;4).

silent subject	*N*
PRO	5370
RtO	2193

- Strong preference for non-core transitive (NCT), causative, and unergative verbs in the embedded clause.

Different verb classes with PRO subject

NCT verbs

(8)   a.   I want to **read** this paper.   Wanda, 2;04, *Bates*

       b.   Thomas want to **eat** it.   Carl, 2;05.27, *Manchester*

causative verbs

(9)   a.   I want to **make** it wouder[3].   Nanette, 2;02.06, *Gleason*

       b.   He want[4] to **open** the gate.   Carl, 2;02.28, *Manchester*
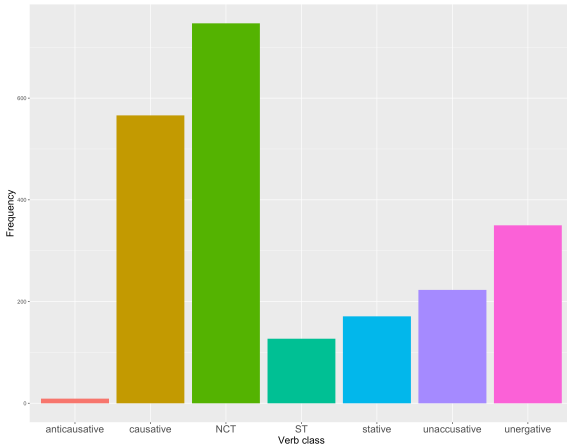
unergative verbs

(10)   a.   She might want to **play** with you.   /, 3;00, *Gelman*

       b.   I want to **go** inside.   Gina. 3;01.08 *MPI-EVA-Manchester*

---

[3][: louder]

[4]In examples (8-b) and (9-b) the 3rd person suffix *-s* is omitted.

- By contrast, RtO construals prefer non-core transitive (NCT) and causative verbs in the embedded clause, while unergative verbs are less frequent.

NCT verbs

(11)   a.   I want mummy to **help** me.   Lara, 2;11.24, *Lara*

      b.   I want you to **read** this.   Naomi, 2;11.18, *Sachs*

causative verbs

(12)   a.   I want you to **put** that on my plate.

                               Fraser, 2;04.02, *MPI-EVA-Manchester*

      b.   I want the farmer to **clean** that side **out**. /, 2;06.12, *Valian*

- The results of the generalized linear regression demonstrate the strong preference for PRO subjects with the following verb classes:

verb class	p-value
causative	0.009224
NCT	0.005886
ST	7.58e-05
unergative	0.000164

- Previous studies converge on the idea that the infinitive *to* is absent both in PRO (13) and RtO (14) construals in the first months after their emergence (Goro 2004, Norris 2004, Landau and Thornton 2011).
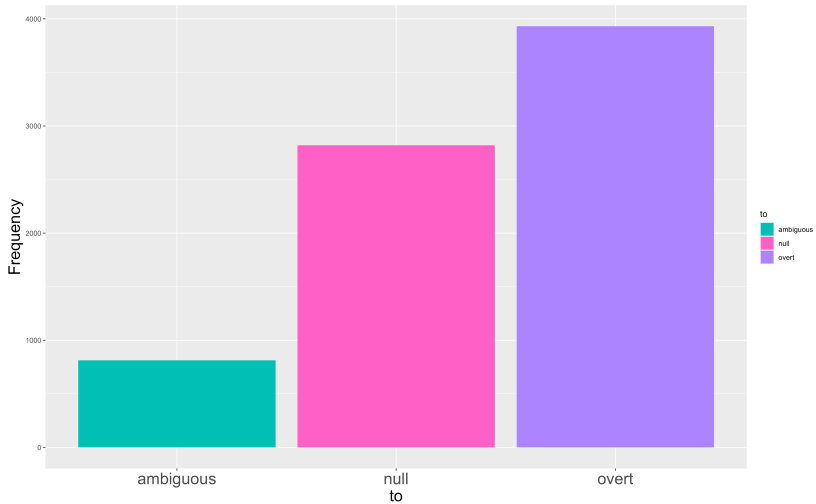
(13) a.   I want go boom!   Ross, 1;04.11, *MacWhinney*

       b.   I want read this.   Naomi, 1;10.18, *Sachs*

(14) a.   I want listen to your heart.

                           Fraser, 2;10.01, *MPI-EVA-Manchester*

       b.   I want Mommy read.   Eve, 1;06.00, *Brown*

- While realization of infinitival *to*, null vs. overt, is identical with PRO, in RtO construals it demonstrates apparent preference for overt *to*.
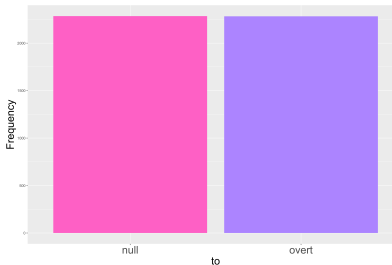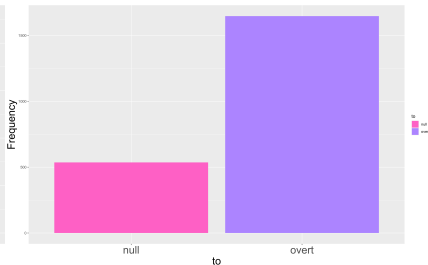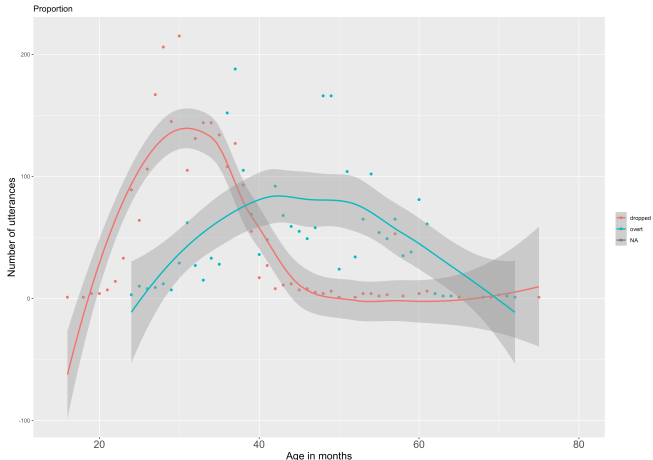

**Figure 1:** *to* & PRO   **Figure 2:** *to* & RtO

- Systematically dropped *to* attested between 20 and 40 months is replaced with overt *to* around the age of 40 months (3;4).
- As the rate of dropped *to* decreases, the rate of overt *to* increases.

- In contrast to PRO, *to* in present at higher rate from the onset of RtO construals.
- Dropped *to* in RtO replicates the pattern observed with PRO and disappears at the same time (3;5).
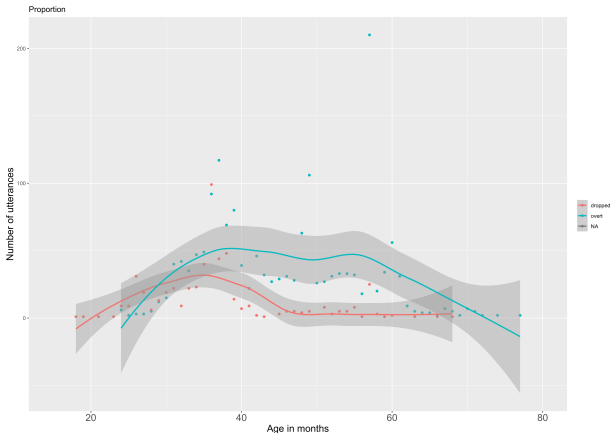
Results: Errors of commission

- In line with Alexiadou et al. (2021), we consider **overt realization of PRO** as an instance of **commissive error**.

- In our data, this pattern is **widely attested** across 11 corpora.

- It covers a **long time span** in acquisition from the age of 1;6, when the first overt PRO in our data is attested, to the age of 4;06.

- Overt PRO in child language comes in **three guises**.

- A child overtly realizes PRO in the form of the pronoun in nominative case.

- This type is attested in production of children around the age of 2 and younger.

(15)   a.   I want **I** go.   Olivia, 1;06.00, *Wells*

        b.   I want **I** put the roof on.   Aran, 2;04.13, *Manchester*

Commissive PRO in child language: Type II

- Instead of adult PRO, children overtly realize the subject argument of the embedded clause in the form of a pronoun in accusative case.
- This pattern is most frequent at the age of 2, when children also use the accusative form *me* to express subjects of matrix clauses (17).
- Accusative has been argued to be the default case in English (Schütze 2001), which might explain its presence in these contexts.

(16) a. I want **me** read the other way.   Nina, 2;02.28, *Suppes*

     b. I want **me** have some.   Abe, 2;10.27, *Kuczaj*

(17) **Me** want see outside.   Laura, 2;09.13, *Wells*

- Children use *me* to mark a subject of the embedded clause, overtly realizing *to*, which is absent in type two.

- This suggests that they analyze the control as an RtO environment, a pattern that is ruled out in adult English.

- This type of errors emerges later in child production compared to previous two, around the age 3-4, when infinitival *to* is already established.

(18) a. Mommy I will want **me** to be right back. Matt, 3;00, *Weist*

b. I don't want **me** to go into the doctor.   Ross, 4;06, *MW*

- Besides commissive PRO, we find commissive prepositions with nouns (19-a) and pronouns (19-b) in RtO construals.

(19) a. I want **for mommy** to do it.   Ross, 3;09.01, *MacWhinney*

      b. Don't you want a zoo game **for me** to play?

                                                      Abe, 3;10.03, *Kuczaj*

- In addition, our data provides evidence for infinitive forms in embedded clauses even with overtly realized *to* (20-b).

(20) a. I want **sitting** very close.   Eve, 2;03.00, *Brown*

      b. I want you to **reading** it.   Geofrrey, 4;11.22, *Wells*

Discussion

- PRO and RtO in child language have been investigated mainly within broader class of infinitival forms.

- The exception is a longitudinal corpus study (1 child, age range: 1;6-2;6) focusing on the acquisition of *want* in Landau and Thornton (2011).

- With respect to the developmental trajectory of PRO and RtO, previous work demonstrates two lines of thought:
  - raising environments are acquired **first** (Becker 2005, Becker 2006, Kirby 2011)
  - raising environments are **delayed** (Hirsch et al. 2007, Hirsch and Wexler 2007, Hirsch et al. 2008, Landau and Thornton 2011, Santos et al. 2016).

PRO vs. RtO asymmetry in child language

- The results of our corpus study support the view that **RtO** construals appear **later** compared to PRO.

- The two forms **significantly differ in frequency** (5370 instances of PRO vs. 2193 examples of RtO), which may provide further argument for the view that PRO is easier for children.

- While **both PRO and RtO** exhibit highest frequency with **NCT** and **causative** verbs, **PRO** displays an **additional strong preference** for **unergative** verbs in the embedded clause.

- **Infinitival *to*** is present at **significantly higher rate from the onset of RtO**, while systematically absent *to* is replaced with systematically present *to* with PRO.

- **Unpronounced PRO Hypothesis**: In line with Sauerland and Alexiadou (2020), the subject of the embedded clause is established at the thought level and identical to the subject of the main clause. In the Compressor component, it undergoes the compression mechanism resulting in unpronounced PRO in adult language.

- Commissive overt realization of the subject argument of an embedded clause attested in our data provides direct evidence for the proposal that argument corresponding to adult PRO is established at the thought level.

- We assume that the subject of the matrix clause and the subject of the embedded clause with the verb in PRO environments are two identical nominals.

- 21-a demonstrates the underlying pre-compressed form at the level of thought structure. 1:1 realization in adult language would result in ungrammatical form.

- However, adults have mastered rules of compression (exponence relation) and map the second occurrence of the identical nominal to $\varnothing$, i.e., PRO.

(21) a. John$_i$ wants [John$_i$ to ride a bike].   *CR-level*

       b. John$_i$ wants [PRO$_i$ to ride a bike].   *compressed*

- As children have not yet acquired exponence relation, following 1:1 mapping principle between CRs and morphology (Alexiadou et al. 2021), they overtly realize the form such as 21-a.

- Our data provides evidence for 1:1 mapping and demonstrates that adult ungrammatical forms such as 21-a are indeed available in child language:

(22)   a.   I want **I** go.   Olivia, 1;06.00, *Wells*

      b.   I want **I** put the roof on.   Aran, 2;04.13, *Manchester*

Acknowledgements

References

Alexiadou, A. (2022). Language and thought. Evidence from language acquisition. Talk given at the LingLab Webinars, Aristotle University of Thessaloniki.

Alexiadou, A., Guasti, M. T., and Sauerland, U. (2021). A Meaning First Approach through a kids first view. Talk given at the University Milano-Bicocca.

Beard, R. (1995). *Lexeme-morpheme base morphology: a general theory of inflection and word formation.* Suny Press.

Becker, M. (2005). Raising, control and the subset principle. In *Proceedings of the 24th West Coast conference on formal linguistics*, volume 24, page 52. Cascadilla Proceedings Project.

Becker, M. (2006). There began to be a learnability puzzle. *Linguistic Inquiry*, 37(3):441–456.

Bhatt, R. and Pancheva, R. (2017). Implicit arguments. *The Wiley Blackwell Companion to Syntax, Second Edition*, pages 1–35.

Goro, T. (2004). On the distribution of to-infinitives in early child English. In *Boston University Conference on Child Language (BUCLD)*, volume 28, pages 199–216.

Hirsch, C., Orfitelli, R., and Wexler, K. (2007). When seem means think: The role of the experiencer-phrase in children's comprehension of raising. In *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, pages 135–146. Cascadilla Press Somerville, MA.

Hirsch, C., Orfitelli, R., and Wexler, K. (2008). The acquisition of raising reconsidered. In *Language acquisition and development: Proceedings of GALA*, pages 253–262.

Hirsch, C. and Wexler, K. (2007). The late development of raising: What children seem to think about seem. In *New horizons in the analysis of control and raising*, pages 35–70. Springer.

Kirby, S. (2011). Move over, control freaks: Syntactic raising as a cognitive default. *BUCLD 35 Online Proceedings Supplement*, pages 1–15.

Landau, I. (2000). *Elements of control: Structure and meaning in infinitival constructions*, volume 51. Springer Science & Business Media.

Landau, I. (2004). The scale of finiteness and the calculus of control. *Natural Language & Linguistic Theory*, 22(4):811–877.

Landau, I. and Thornton, R. (2011). Early child control. *Lingua*, 121(5):920–941.

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk: Volume I: Transcription format and programs, volume II: The database.

McFadden, T. and Sundaresan, S. (2018). Reducing pro and PRO to a single source. *The Linguistic Review*, 35(3):463–518.

Norris, R. L. (2004). *Acquisition of the T and C system in clausal complements*. MA Thesis, Massachusetts Institute of Technology.

Santos, A. L., Gonçalves, A., and Hyams, N. (2016). Aspects of the acquisition of object control and ECM-type verbs in European Portuguese. *Language Acquisition*, 23(3):199–233.

Sauerland, U. and Alexiadou, A. (2020). Generative Grammar: A Meaning First Approach. *Frontiers in Psychology*, 11:3104.

Schütze, C. T. (2001). On the nature of default case. *Syntax*, 4(3):205–238.

Yatsushiro, K. and Sauerland, U. (2018). A filled gap stage in German relative clause acquisition. In Albright, A., Magri, G., Michaels, J., Chan, H., Jacob, H., and Kapia, E., editors, *Proceedings of the 42nd annual Boston university conference on language development*, pages 814–827. Somerville, Mass: Cascadilla.

Appendix

Corpora

- **UK English**: Fletcher, Forrester, Howe, KellyQuingley, Lara, MPI-EVA-Manchester, Manchester, Thomas, Tommerdahl, Wells;

- **US English**: Bates, Bernstein, Bliss, Bloom, Bohannon, Braunwald, Brown, Clark, Demetras, Evans, Feldman, Garvey, Gathercole, Gelman, Gleason, Haggerty, Hall, Higginson, HSLLD, Kuczaj, MacWhinney, McCune, McMillan, Morisset, Nelson, New England, NewmanRatner, Post, Rollins, Sachs, Sawyer, Snow, Sprott, Suppes, Tardif, Valian, Van Houten, Van Kleeck, Warren-Leubecker, Weist