Language Acquisition and Development

Proceedings of GALA 2015

Edited by
Jiyoung Choi, Hamida Demirdache, Oana Lungu and Laurence Voeltzel
# Table of Contents

**Introduction**  
The editors

**Comprehension of Wh-Questions in Child Romanian: A Case About Case and Lexical Restriction**  
Anamaria Bențea  
1

**Free Relative Clauses and the Over-generalization of Wh-Movement**  
Michael Clauss  
20

**Is Agree Also Delayed as well as Movement?**  
Yoshiki Fujiwara  
36

**School-Age Sequential Mandarin-Italian Bilingual Children's Comprehension of Relative Clauses**  
Maria Teresa Guasti & Shenai Hu  
55

**Absolute Implications in Children’s Understanding of Gradable Expressions**  
Koji Kawahara  
70

**On the Acquisition of Distributive Numerals in Serbian**  
Nataša Knežević & Hamida Demirdache  
86

**The First Steps in the Acquisition of Syllables and Geminates in Tashlhiyt Berber**  
Mohamed Lahrouchi  
111

**L2-English Definites and Non-Maximality**  
Bert Le Bruyn & Xiaoli Dong  
134
Factors Affecting Performance in Child Heritage Portuguese in Germany ......................................................... 150
Tatjana Lein, Cornelia Hamann, Monika Rothweiler & Solveig Chilla

Acquiring the Denotation of Object Denoting Nouns .......................................................... 182
Suzi Lima, Peggy Li & Jesse Snedeker

Resolving Pronoun Ambiguity in European Portuguese: Children vs. Adults ........................................... 204
Maria Lobo & Carolina Silva

Phonological Categories of Voice Onset Time and Phonetic Compromise in L2 Acquisition .................. 224
Kuniya Nasukawa, Phillip Backley & Hitomi Onuma

Actionality Speaks Louder than Felicity: English-Speaking Children’s Comprehension of Passives ........ 240
Emma Nguyen, Diane Lillo-Martin & William Snyder

Heritage Signers: Bimodal Bilingual Children from Deaf Families ....................................................... 255
Wanette Reynolds, Jeffrey Levi Palmer, Deborah Chen Pichler, Ronice Müller de Quadros, Viola Laura Kozak & Diane Lillo-Martin

Vowel Epenthesis in Children’s Oral and Written Productions of Consonants Clusters ............................ 278
Rita Santos

The Acquisition of Recursive Locative PPs and Relative Clauses in Child English ................................. 296
Anca Sevcenco, Thomas Roeper & Barbara Pearson

The Lack of Agent-Orientation for Zibun and the EPP Parameter in Child Japanese .............................. 311
Hiroyuki Shimada
The Role of Executive Functions in the Acquisition of Reference: The Production of Demonstrative Pronouns by German Monolingual Children ................................. 325
Jacopo Torregrossa

List of contributors
Introduction

The conference *Generative Approaches to Language Acquisition* is the biannual meeting of researchers working on language development from a generative perspective. The conference provides a forum for discussion of the most recent advances on first and second language acquisition, heritage language and bilingual acquisition, language pathology, the acquisition of sign language and neurolinguistics.

The 2015 edition was organized by the *Laboratoire de Linguistique de Nantes* (LLING, – UMR 6310 CNRS/Université de Nantes) and was held in Nantes from September 10-12, 2015. The high quality of the research presented at the conference is reflected in the important number of submitted papers (137), from all over the world (21 countries), of which only 35% were accepted for oral presentation.

This 12th edition of GALA included, alongside the general session, a special session on *Heritage Language Acquisition: Native vs. Heritage vs. Second language acquisition*, co-organized with Janet Grijzenhout (BSL, University of Konstanz), and supported by the AThEME (Advancing the European Multilingual Experience) collaborative research project studying multilingualism in Europe. Heritage Language users are unbalanced/passive bilinguals who have acquired simultaneously or successively a Heritage Language (HL)—the language of the parents (minority or immigrant) learned from birth in the home environment—and an ambient language spoken outside the home, the Dominant Language of their society (DL). As such, heritage speakers do not fit into either the native vs. non-native or L1 vs. L2 dichotomy, since unlike L2 (but like native) speakers, heritage speakers are exposed to the target language during the critical period, but just like L2 speakers, heritage speakers fail to converge on the target language, exhibiting variability in ultimate attainment.

GALA 12 also included three other workshops. The workshop on *The Role of Prosody in Early Speech Perception* (co-organized with Judit Gervain, LPP – UMR 8242, CNRS/Paris V) focused on recent experimental work on the role of prosody in child language. The workshop on the *Acquisition of Causation: Culmination Entailments and Agency* (co-organized with Fabienne Martin, University of Stuttgart) targeted experimental research probing children’s understanding of causation and
agenthood. The workshop on *Segments and Interactions in Phonological Acquisition* focused on (a-)typical acquisition of segments and interactions between segments in production and perception. The invited speakers of GALA 2015 were Maria Polinsky (Harvard University), René Kager (University of Utrecht), Angeliek van Hout (University of Groningen), Nina Kazanina (University of Bristol), and Eirini Sanoudaki (Bangor University).

The present volume gathers 21 papers presented at GALA 2015 and reflects the current research in the field of language acquisition at the interface with syntax, semantics and phonology. The papers on the acquisition of syntax and semantics cover (a-)typical first and second language acquisition, heritage language and bilingual acquisition including bimodal bilingualism with heritage signers. They address a wide variety of topics at the syntax semantics interface: wh-movement (comprehension of *who* vs. *which* questions, *wh*-expressions in free relatives, relative clauses in bilingual children), passivisation (actional vs. non-actional (long) passives, delayed acquisition of verbal passives), the acquisition of distributivity (in languages marking the distributive share), the acquisition of reference (count nouns in languages with no count/mass grammatical distinction, (non) maximal readings of definites), comparative and positive uses of gradable expressions, anaphora resolution (with null vs. overt pronouns, d-pronouns vs. personal pronouns, long-distance anaphors), the acquisition of indirect recursion and locality, the role of input factors in heritage language acquisition, and heritage language characteristics in bimodal bilingual children (in the domains of phonology, syntax, and discourse). The papers on phonological acquisition focus on the production of consonant clusters, the acquisition of Voice Onset Time, the development of the speech system in infants, and cover both first and second language acquisition.

Given the quality of the papers and the diversity of the subject matter, the Proceedings of GALA 2015 will make, without any doubt, a significant contribution to the field, imposing itself, together with the previous Proceedings volumes, as a valuable reference guide for the researchers working in the domain of language acquisition and language development from a generative perspective.

We are very grateful to the contributors to the volume, to the reviewers, and to the following sponsors for their support in the organization of the conference and the preparation of the volume: the large-scale integrating
project AThEME (Advancing the European Multilingual Experience) funded by the European Seventh Framework Programme for research, technological development and demonstration under grant agreement no 613465, the Laboratoire de Linguistique de Nantes (LLING), the Laboratoire Psychologie de la Perception – UMR 8242, CNRS/Université de Paris V, the project Grasping Meaning across Languages and Learners (GraMALL) funded by the Netherlands Organisation for Scientific Research (NWO), the project Incremental Specification in Context (SFB 732) funded by the National German Science Foundation (DFG).

Nantes, October 11, 2016

Jiyoung Choi
Hamida Demirdache
Oana Lungu
Laurence Voeltzel
COMPREHENSION OF WH-QUESTIONS IN CHILD ROMANIAN: A CASE ABOUT CASE AND LEXICAL RESTRICTION

ANAMARIA BENTEÀ

1. Introduction

Cross-linguistic studies on children’s comprehension of wh-questions have revealed a twofold asymmetry: (i) subject which-questions (1) are easier for children to comprehend than object which-questions (2) and (ii) object who-questions (3) yield better accuracy than object which-questions (De Vincenzi et al. 1999; Avrutin 2000; Philip et al. 2001; Friedmann et al. 2009).

(1) Which giraffe tickled the turtle?
(2) Which giraffe did the turtle tickle?
(3) Who did the turtle tickle?

Friedmann et al. (2009) propose that the source of children’s difficulties with questions like (2) lies in the intervention of the subject (the turtle) in the interpretive chain formed by the A’-object (which giraffe) with its canonical position in the structure (i.e. object of the verb tickle). The gist of this approach is that children encounter difficulties with movement structures in which one element containing a lexical restriction (i.e. a [+NP] feature) intervenes in the movement of another [+NP] element, as shown in (4). These intervention effects are traced back to the general locality principle of Relativized Minimality (RM), which imposes constraints on the syntactic relations that can hold between a displaced element and its original position in the sentence.

\[
\begin{align*}
\text{X} & \quad \text{Z} & \quad \text{Y} \\
+\text{NP} & \quad +\text{NP} & \quad +\text{NP}
\end{align*}
\]

(4) \([\text{DP}_{\text{Object}} \quad [\text{DP}_{\text{Subject}} \ [\text{V} <\text{DP}_{\text{Object}}> ]]]\)

In other words, children only struggle with those structures in which (i) the A’-chain linking the object to its base-generated position crosses an intervening subject and (ii) the intervening element shares a partial featural specification with the moved constituent. This results in an intervention configuration in which-object questions since the subject DP represents a potential competitor in establishing the correct grammatical dependency between the object and its original position as argument of the verb. Children’s improved performance with object who-questions shows that they can easily establish a movement dependency once the object and the intervening subject do not share the feature [+NP].

Wh-questions in Romanian are particularly relevant in this context as which-questions can appear without an overt NP. Although such which-questions do not carry an expressed NP, they still have lexical specificity, which limits the set of possible referents for the wh-expression to those existing in the discourse context. This offers a good testing ground to determine whether children’s grammar is sensitive to the presence of a lexical restriction or NP feature on both the moved element and the intervener, even when the lexical NP restriction is not overtly expressed in the wh-expression.

Moreover, object questions in Romanian are disambiguated from subject questions through the presence of the accusative marker pe. The form of the interrogative pronoun thus overtly signals whether the structure should be interpreted as a subject or object question. Testing the comprehension of wh-questions can therefore prove insightful into the role that pe plays in the processing of A’-dependencies in Romanian.

The rest of this paper is organized as follows: section 2 briefly describes the syntax of wh-questions in Romanian; section 3 outlines the experimental study and the results, which are discussed in section 4; section 5 concludes the paper.

2. Wh-questions in Romanian

Two wh-elements introduce subject and object questions referring to [+animate] entities: cine (“who”) for non-lexically-restricted or bare wh-questions (see (5-6)) and care (“which”) for lexically-restricted questions, as in (7) and (8). The noun phrase in parentheses in these latter examples illustrates the option of having care-questions without an overtly expressed lexical NP. For the sake of clarity, we will refer to such questions as which –NP (vs. which +NP questions, i.e. those questions containing an overt lexical restriction):
(5) Cine l-a interviievat pe student? 
who him,ACC-has interviewed pe.ACC student, i 
Who interviewed the student?

(6) Pe cine a interviievat profesorul? 
pe.ACC who has interviewed professor.the.M.SG 
Who did the professor interview?

(7) Care (profesor) l-a interviievat pe student? 
which professor him,ACC-has interviewed pe.ACC student, i 
Which professor interviewed the student?

(8) Pe care (student) l-a interviievat 
pe.ACC which student, i him,ACC-has interviewed 
profesorul? 
professor.the.M.SG 
Which student did the professor interview?

Note that in object which-questions (8), contrary to object who-questions (6), a co-indexed clitic pronoun l (“him”) doubles the direct object pe profesor\textsuperscript{1}. The contrast between (6) and (8) shows that clitic pronouns are illicit in non-lexically-restricted interrogatives, but are obligatory with lexically-restricted wh-phrases, as illustrated by the ungrammaticality of examples (9) and (10):

(9) *Pe cine; l-a interviievat profesorul? 
pe.ACC who; him,-has interviewed professor.the.M.SG 
Who did the professor interview?

(10) *Pe care (student) a interviievat profesorul? 
pe.ACC which student has interviewed professor.the.M.SG 
Which student did the professor interview?

Given the parallelism between which +NP and which –NP structures with respect to the obligatory presence of a clitic pronoun corresponding to the moved argument, it follows that they also share the same structural properties and that which –NP is only an apparent instantiation of a bare wh-element\textsuperscript{2}. Investigating whether Romanian children are sensitive to the differences between which-questions, both +NP and –NP, and who-questions can prove very insightful for the study of children’s syntactic development.
3. Experimental study

In light of the above considerations, the current study explores Romanian children’s sensitivity to morphosyntactic information in the interpretation of bare and lexically-restricted wh-questions and investigates to what extent children’s selective difficulties with movement dependencies can be modulated by language-specific properties, such as case-marking. To date, this is the first paper that investigates the comprehension of who- and which-questions by Romanian children, and thus aims to fill the gap in the acquisition literature on the comprehension of wh-questions in Romanian.

3.1. Participants

We report the results of 44 typically developing monolingual Romanian-speaking children aged 3;8 to 7;2. They were recruited at a preschool and a school in Bistrița, Romania and divided across two age groups, as illustrated in Table 1. In addition, 10 adult native speakers of Romanian were included as a control group.

Table 1: Participant data per age group (number per group, age range, mean age and standard deviation)

<table>
<thead>
<tr>
<th>Age group</th>
<th>No. of participants</th>
<th>Age range</th>
<th>Mean Age (S.D.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 y.o</td>
<td>21</td>
<td>3;8 - 5;2</td>
<td>4;4 (0.60)</td>
</tr>
<tr>
<td>6 y.o.</td>
<td>23</td>
<td>5;8 - 7;2</td>
<td>6;5 (0.62)</td>
</tr>
<tr>
<td>Adults</td>
<td>10</td>
<td>18;0 - 40</td>
<td>31 (6.73)</td>
</tr>
</tbody>
</table>

3.2. Materials

The experiment investigated the comprehension of subject and object questions introduced by various types of wh-elements. The study used a 2 X 3 design with type of structure (subject vs object) and type of wh-element (who vs which +NP vs which –NP) as within-participant factors. 18 sets of 6 sentences like those exemplified in (11) were constructed. The which +NP and which –NP notation distinguishes between care (“which”) questions with and without an overt full lexical noun. In using this notation, we adopt the terminology of previous studies (Friedmann et al.
2009, Belletti et al. 2012) where the +NP feature characterized cases involving the presence of a lexical noun and, accordingly, we initially classify wh-questions headed only by care and pe care as –NP.

(11) a. **Subject cine (“who”) question**
   
   Cine a  gâdilat țestoasa?  
   who has tickled turtle.the.F.SG
   Who tickled the turtle?

   b. **Object cine (“who”) question**
   
   Pe  cine a  gâdilat țestoasa?  
   pe.ACC who has tickled turtle.the.F.SG
   Who did the turtle tickle?

   c. **Subject care +NP (“which +NP”) question**
   
   Care girafă a  gâdilat țestoasa?  
   which giraffe has tickled turtle.the.F.SG
   Which giraffe tickled the turtle?

   d. **Object care +NP (“which +NP”) question**
   
   Pe  care girafă a  gâdilat-o  
   pe.ACC which giraffe has tickled-her.ACC țestoasa?
   turtle.the.F.SG
   Which giraffe did the turtle tickle?

   e. **Subject care –NP (“which –NP”) question**
   
   Care a  gâdilat țestoasa?  
   which has tickled turtle.the.F.SG
   Which one tickled the turtle?

   f. **Object care –NP (“which –NP”) question**
   
   Pe  care  a  gâdilat-o  țestoasa?  
   pe.ACC which has tickled-her.ACC turtle.the.F.SG
   Which one did the turtle tickle?

   In the test items, (i) pe was the only element that signaled an object interpretation vs. a subject interpretation for cine (“who”) questions in (11a-b); (ii) subject and object care NP “which +NP” and care –NP (“which –NP”) questions in (11c-d) and (11e-f) were disambiguated through the use of pe and the clitic pronoun3; (iii) object cine (“who”) and object care (“which –NP”) questions were only distinguished, at a surface level, by the use of the clitic pronoun in the latter structures. All target sentences contained transitive verbs and nouns that matched in gender and number. The characters were always either two animals or two humans. A post-verbal subject was used in all object questions.
Comprehension of Wh-questions in Child Romanian

Children saw each item twice. The 6 sentences of each set were divided into three lists such that each list included 6 items for each of the 6 conditions. Items were presented in a randomized order within each list and were mixed with 18 fillers. Additionally, there were 2 practice trials at the beginning of each list. Consequently, each list included 2 practice trials, 36 target sentences and 18 fillers, summing up to a total of 56 trials per list. Each experimental session lasted approximately 35 minutes, with a short intermediary break. Children were also told that they could stop the experiment and have a break at any time.

The task used was a character-selection task adapted from Friedmann, et al.’s (2009) study on wh-questions in Hebrew and Adani’s (2011) study on relative clauses (RCs) in Italian. The visual stimuli were cartoon movies realized with an animation software. Each cartoon displayed three characters (e.g. two giraffes and a turtle) involved in the same action and was accompanied by a pre-recorded description of the scenes. A sample description and an example of test question are illustrated in (12). Figure 1 shows the sequence of events in the visual display.

(12) **Lead-in:** Uite două girafe și o țestoasă!

*Look! There are two giraffes and a turtle.*

1st *action:* Să vedem ce se întâmplă! O girafă știindă țestoasa.

*Let’s see what happens next! A giraffe is tickling the turtle.*

2nd *action:* Acum uite! Țestoasa știindă o altă girafă.

*Now look! The turtle is tickling another giraffe.*

**Question:** Care girafă a știindă țestoasa?

*Which giraffe tickled the turtle?*

In order to create a pragmatically felicitous context for the use of care –NP (“which –NP”) questions Care a știindă țestoasa? (“Which one tickled the turtle?”), the target question was always preceded by an introductory question (e.g. Ai văzut girafele? *Did you see the giraffes?*).
After each question, a puppet—the reindeer Rudolf—asked the child a wh-question, thus prompting her to point to the correct character or reply by describing the color of the character (e.g. “the pink giraffe”). When the children heard the target question, the characters appeared in their original position. The adult participants saw the same experimental stimuli. The procedure used with them was similar to that used with the children, but the experimenter would ask the question directly, without using the puppet. The experiment with the adults took approximately 15 minutes.

3.3. Predictions

Several predictions hold for the study. First, as shown in section 2, the syntactic behavior of care “which –NP” phrases patterns with that of which-expressions with an overtly expressed noun and is distinct from that of bare wh-phrase like who. In addition, as far as their meaning is concerned, both types of which-expressions presuppose the availability of a given set of referents in the discourse context. Contrary to questions introduced by cine “who”, questions introduced by care “which” cannot be asked out-of-the-blue since they imply a choice from “a set of individuals previously introduced into the discourse, or . . . part of the ‘common ground’ shared by speaker and hearer” (Pesetsky 2000, 16). Despite the similarity in surface form between cine “who” and care “which –NP”, only the first wh-element can be truly considered as –NP, whereas the latter only apparently lacks a lexical restriction. The apparently bare care is in fact analogous to the English which one. It could thus be assigned the form [\wh \NP \N e]], along the lines of Dobrovie-Sorin (1994), where the N has been deleted, contrary to care NP cases (e.g. care “which one” vs care girafă “which giraffe”), where an NP is present. As such, the instances of care “which” are not followed by an overt full lexical noun still contain a [+NP] specification.
This has interesting consequences for an account of children’s comprehension difficulties with object A’-dependencies in terms of intervention effects driven by a similar featural specification on the A’-moved element and the intervening subject. An important observation of the intervention account is that the features that enter the computation of locality are the features that function as attractors for movement. Independent evidence in favour of considering lexical restriction as an attractor in different constructions comes from Northeastern Italian Dialects. Munaro (1999, cited in Poletto & Pollock 2000) showed that bare wh-words and lexically-restricted wh-elements in Bellunese occur in different positions at the left periphery of the clause, as lexically-restricted wh-phrases target a sentence-initial position, while bare wh-phrases appear sentence-finally:

(13) a. Ha-tu magnà che?
    have-you eaten what?
    *What did you eat?
b. *Che ha-tu magnà?
    what have-you eaten?
c. Che vestito à-la comprà?
    what dress has-she bought
    *Which dress did she buy?
d. *Ha-la comprà che vestito?
    has-she bought what dress?

The difference in syntactic behavior between lexically-restricted and bare wh-phrases points to the presence of different attractors for the movement of the two types of wh-elements. Romanian also supports this conclusion based on data from multiple wh-questions in which lexically-restricted wh-elements target a higher position than bare wh-words. Thus, the data from Bellunese and Romanian related to the behaviour of wh-elements with or without a lexical restriction clearly show that the NP feature acts as an attractor for movement of which-elements to a higher position than that of bare wh-phrases.

The NP feature thus forms part of the array of morphosyntactic features that trigger movement and, as such, should enter into the computation of intervention. Consequently, if NP enters the computation of locality, then the child’s grammar system should be sensitive to this feature although it is not overtly expressed through the presence of a noun from the contentive lexicon. If the child’s grammar system only paid attention to the surface form of the lexical elements and if this were enough to overcome comprehension difficulties, then performance for
bare who-questions should be on a par with performance for which-questions that do not have a full nominal element. Moreover, children should comprehend these questions better than which +NP ones.

The second prediction is linked to the effect that case-marking on the wh-pronoun has on the comprehension of questions in child Romanian. If the presence of pe at the very onset of the wh-question is an informative enough cue to signal that the wh-phrase should be assigned a patient theta-role and interpreted as the object of the verb and if children are able to draw upon this information, as well as the mismatch in case features between the A′-object and the intervening subject, then the presence of pe should greatly reduce any subject-object asymmetry attested in the comprehension of wh-questions. In addition, the fact that case-marking is present on both bare (cine) and lexically-restricted (care) wh-questions should equally facilitate processing of the two structures and modulate the difference between questions with (which) or without (who) a lexically restricted +NP feature attested cross-linguistically.

3.4. Results

Response accuracy was the dependent variable in the experiment: an answer was coded as correct when the child pointed to or described the corresponding character targeted by the wh-question. While the adults performed at ceiling for all the experimental conditions, children’s comprehension scores within the two age groups revealed that the 6-year-old group gave overall more accurate responses than the 4-year-old group. However, the same performance patterns could be observed within the groups for each of the experimental conditions (see Figure 2).
As illustrated above, both age groups performed equally well for subject and object who-questions. The 4-year-olds performed on a par for object which +NP and which −NP questions, whereas they were more accurate with both subject which +NP and which −NP conditions. The comprehension scores for the 6-year-old group revealed similar performance for all three types of subject wh-questions. In contrast, when it comes to object questions, children comprehended who-questions better than both which −NP and which +NP. Like for the 4-year-old children, the accuracy scores of the 6 year-olds show that they struggled most with object which +NP questions and this led to a sharper subject-object contrast in the comprehension of which +NP questions.

The data obtained were fit to a mixed logit model with type of structure (i.e. Structure Type), type of wh-element (i.e. Wh-word Type) and age group as fixed predictors. The reference level was the response accuracy mean for Object Which −NP against which the means for each of the other levels of the variables Structure Type, Wh-word Type and Age Group were compared. The maximal random effect structure justified by the data included intercepts for subjects and items, as well as by-subject random slopes for Structure Type. The analysis revealed that, while the interaction between Structure Type and Wh-word Type with Age Group did not add significance to the model (χ²(5) = 4.01, p = 0.54), the effect of Age Group was significant (χ²(1) = 5.79, p < .01), as was the interaction between Structure Type and Wh-word Type (χ²(2) = 11.06, p < .001).

The results showed that children overall comprehended subject wh-
questions better than object questions ($\beta = 0.79$, $SE = 0.39$, $z = 2.02$, $p < .05$). Children also performed better with object who questions than with object which –NP questions ($\beta = 0.54$, $SE = 0.26$, $z = 2.07$, $p < .05$), whereas there was no significant difference in comprehension between object which +NP and object which –NP questions ($\beta = 0.20$, $SE = 0.23$, $z = 0.85$, $p < .05$). The 6-year-old group gave significantly more accurate responses ($\beta = 0.84$, $SE = 0.35$, $z = 2.43$, $p < .05$), illustrating that comprehension of wh-dependencies improves with age. The interaction between Structure Type and Wh-word Type was also significant and was due to Wh-word Type having a different effect on subject and object questions: while comprehension of subject who and subject which –NP questions was on a par, children performed better with subject which +NP questions. The type of wh-word had a different effect on object questions: whereas there was no difference in performance between object which –NP and object which +NP questions, children were more accurate with object who questions.

4. Discussion

Given that the adult data showed no asymmetries in comprehension, as adults performed at ceiling for all the experimental conditions, we will focus on the child data in the discussion.

We will first examine the role that an overt or covert lexical restriction or +NP feature plays in the comprehension of wh-questions in Romanian. Such a comparison is particularly relevant given the findings for who-questions, i.e. [–NP] questions that pose no difficulties for comprehension. The results revealed a subject-object asymmetry for which-questions: children comprehended subject which-questions better than object which-questions, not only when the lexical NP overtly followed the wh-expression, but also when the lexical restriction was not expressed on the wh-phrase. No difference in performance emerged between subject who- and object who-questions. These findings are consistent with the results of a number of studies on a variety of languages (De Vincenzi et al. 1999; Avrutin 2000; Friedmann et al. 2009 etc.) illustrating that children struggle more with object which-questions and reinforce the idea that not all types of object movement configurations are problematic for acquisition, but only those that give rise to an inclusion relation between the set of features on the moved wh-object and the intervening subject due to the presence of a +NP lexical restriction on both elements, as represented in (14a):
Comprehension of Wh-questions in Child Romanian

\[+Q+NP \quad +NP\]

(14) a. Pe care (girafă) a gâdilat-o țestoasa?  
pe.ACC which giraffe has tickled-her turtle.the  
Which giraffe did the turtle tickle?  

b. Pe cine a gâdilat țestoasa?  
pe.ACC who has tickled turtle.the  
Who did the turtle tickle?

The computation of the inclusion relation in (14a) is more problematic for children than the computation of a disjunction relation (14b) (Friedmann et al. 2009; Belletti et al. 2012), because the intervening subject in (14a) hinders the realization of the correct dependency between the moved element and its base-position as it becomes a potential candidate for this relation. This holds for both care +NP and care –NP questions. In contrast, who-questions, in which the fronted constituent is –NP, do not give rise to such intervention effects, since the featural sets of the subject and the object are in a disjunction relation (14b). Children’s performance with which-questions further underscores the role that even the covert presence of a +NP element plays in the processing of wh-dependencies.

The second finding of this experimental study concerns the role of the preposition pe in children’s parsing of wh-questions. The use of pe in the experimental material should straightforwardly disambiguate between a subject and an object interpretation, irrespective of the type of wh-word, and should help children identify the thematic role of the wh-argument in relation with the verb. The data revealed that case marking on the wh-pronoun in Romanian considerably improved comprehension of object questions, but did not eliminate the subject-object difference in which-questions, nor the asymmetry between who and which-object interrogatives (both +NP and –NP).

Note however that the subject-object asymmetry surfacing in the comprehension of which-questions in Romanian is greatly reduced with respect to earlier studies with other languages. The children tested in this study comprehended object which +NP questions 77% of the time, compared with earlier findings of 58% for Hebrew (Friedmann et al. 2009) and 50% for Italian (De Vincenzi et al. 1999). The improved results for the comprehension of which-questions seem likely due to properties of wh-questions in Romanian that facilitate comprehension of object questions to a greater extent. More specifically, the improved performance of Romanian-speaking children compared to their Hebrew- and Italian-
speaking peers could be explained through the additive effect of the case marker *pe* and the post-verbal subject. The latter, coupled with the presence of *pe* at the very onset of the wh-question, which signals that the phrase should receive an object interpretation, seems to boost children’s comprehension of object wh-questions. These two properties distinguish Romanian questions from both Hebrew and Italian questions⁵.

The idea we would like to put forth is that the presence of a post-verbal subject allows to by-pass the strong intervention effects found in structures in which the subject is pre-verbal, such as RCs. This can be achieved by postulating an analysis of *care* structures that builds on Cecchetto’s (2000) analysis of clitic left-dislocated (CLLD-ed) objects. According to Cecchetto (2000), the CLLD-ed object originates in the argument position of the verb as a Big DP headed by the clitic (*la* in examples (15) and (16)) and they obligatorily reconstruct in an IP-internal position outside the vP, located higher than the one occupied by a post-verbal subject, but lower than the one occupied by a pre-verbal subject.

(15) *L’opera prima di uno scrittore*, lui la scrive sempre (volentieri).
A writer’s first work, he always writes it with pleasure.

Thus, the co-indexation between the DP contained in the CLLD-ed object and the preverbal pronominal subject gives rise to a Principle C violation because the referential expression is reconstructed within the c-command domain of the pronoun. Conversely, the DP *uno scrittore* can be coindexed with the post-verbal subject in (16) without violating Principle C since the referential expression is reconstructed outside the c-command domain of the post-verbal subject.

Such an account has interesting consequences for an analysis of *which*-object questions in Romanian. Contrary to *cine*-questions, *which*-questions do not give rise to weak cross-over effects (WCO) (see also endnote ii), as the wh-element *pe* care băiat in (16a) can felicitously bind the pronoun inside the post-verbal subject DP:

A writer’s first work is always written by the writer himself.
Comprehension of Wh-questions in Child Romanian

(17) a. *Pe cine i-a certat mama lui? pe.ACC who has scolded mother.the his 

*Who did his mother scold?
b. Pe care băiat i-a certat mama lui? pe.ACC which boy him-has scolded mother.the his 

Which boy was scolded by his own mother?

Along the lines of Cecchetto’s (2000) analysis of CLLD-ed objects, we explain the absence of WCO effects with *care-structures by postulating that the wh-element and the clitic are initially generated as a Big DP in the internal argument position of the verb and together undergo a first movement to an intermediate position within the IP domain, higher than the post-verbal subject. Once in this intermediate position, the wh-element undergoes A’-movement to the left periphery and it is reconstructed above the c-command domain of the pronoun lui. The co-indexation between the referential expression băiat and the pronoun lui can thus take place without triggering WCO effects. In the case of cine-questions, since no clitic is present in the structure, the wh-element does not reconstruct in an intermediate position within the IP, thus giving rise to the prototypical WCO configuration where the pronoun inside the post-verbal subject is c-commanded by the operator, but not by the variable.

According to this analysis, the only instance in which the lexically restricted wh-phrase crosses over the subject is when the Big DP moves to the intermediate IP-internal position preceding the post-verbal subject. Since what is moved across the intervening subject is a larger constituent, namely the Big DP containing both the clitic and the +NP wh-double, this could alleviate the intervention effects found in object which-questions (much like in the case of the smuggling analysis proposed by Belletti 2014; Contemori & Belletti 2014 for circumventing intervention effects in passive object relatives). In the case of cine object questions, no intervention effects hold since the features of the moved A’-object are in a disjunction relation with the features of the intervening post-verbal subject.

If this analysis is on the right track, then the presence of a pre-verbal subject in care-structures would lead to stronger intervention effects, since the pre-verbal subject will appear in a position between the moved wh-phrase and its intermediate reconstruction site inside the IP. Moreover, the presence of a pre-verbal subject would also give rise to a typical case of inclusion relation when the wh-element containing the +NP feature moves out of the intermediate site of the Big DP to a position at the left periphery. This second movement inevitably crosses over the pre-verbal
subject, thus creating an inclusion configuration with the feature of the embedded subject that also contains a +NP specification. Though it is impossible to test this prediction with wh-questions, since wh-elements must be adjacent to the verb, RCs in Romanian offer a good test case, as they can have both a pre-verbal and a post-verbal subject. The analysis outlined above predicts that, in Romanian, object RCs with a post-verbal subject should be easier to comprehend than object RCs with a pre-verbal subject. Bentea (2012) reports low comprehension scores for object RCs with pre-verbal subjects, while the results of Sevcenco & Avram (2012) do not show a clear distinction between the comprehension of RCs with a pre- or a post-verbal subject. The question whether the presence of a pre- or a post-verbal subject can alleviate intervention effects therefore requires further investigation. The preliminary results of an ongoing study investigating the comprehension of RCs with a pre- or a post-verbal subject and using the same methodology as in the experiment reported in this paper seem to bring converging evidence for the analysis outlined above, however more children need to be tested and the data analyzed in more detail.

Let us now turn to the comparison between the results obtained for Romanian wh-questions and those reported for Hebrew (Friedmann et al. 2009) and Italian (De Vincenzi et al. 1999; Belletti & Guasti 2015). The comparison with the results for Hebrew becomes more straightforward: while in both languages which-elements are preceded by an external case-marker (et in Hebrew, pe in Romanian), only Romanian wh-questions have both a post-verbal subject and a clitic. The wh-questions tested in Hebrew (Friedmann et al. 2009) all contained pre-verbal subjects that, according to the analysis outlined above, lead to strong intervention effects. As for the difference found between the comprehension of object which-questions in Romanian and Italian, despite the presence of a post-verbal subject in both languages, we can account for it by linking it to the absence of clitic doubling in Italian questions. It follows that which-phrases in Italian cannot undergo reconstruction in an IP internal position and therefore cross the post-verbal subject on their way to the left periphery, thus giving rise to an inclusion configuration.

This brings us to the role that case-marking has on the comprehension of wh-questions in Romanian. The findings of the present study suggest that case per se does not affect comprehension of wh-questions in Romanian. More specifically, we see that despite the presence of a case-marker in both care and cine questions, we still find an asymmetry in comprehension between the two types of questions, even when care questions do not contain an overt lexical restriction and are similar in
surface form to *cine* questions. Although case-marking appears very early in the structure, its effect is not unselective, but depends on the particular syntactic configuration in which it is found. If children used a general strategy drawing on the presence of a case-marker on the wh-word to determine the object and hence the patient of the verb, we would expect no difference between subject and object questions. In addition, we would expect no difference between structures with a pre- and a post-verbal subject. Studying the comprehension of Romanian RCs will thus provide crucial insight into whether the effect of case-marking on A’-dependency comprehension is linked to a specific syntactic configuration. More cross-linguistic data on the effect of case-marking on wh-comprehension could also help to shed light on the role it plays in circumventing or reducing intervention effects in acquisition.

5. Conclusions

This study aimed at investigating whether a subject-object asymmetry also surfaces in the comprehension of *who* and *which*-questions in Romanian, as well as determine whether the [+NP] feature impacts the processing of wh-dependencies even in the absence of an overt lexical restriction or lexical noun on the moved wh-object.

Whereas no difference in comprehension emerged from the adult data, children’s results revealed two asymmetries: (i) a subject-object asymmetry in the comprehension of *which*-questions, but no difference in the comprehension of subject and object *who*-questions; (ii) an asymmetry in performance with object *who*-dependencies and object *which*-dependencies, independent of the absence of a full lexical noun after the wh-phrase. That children comprehended *care* “which –NP” questions on a par with *care NP* “which +NP” object questions suggests that the child’s grammar system also takes into consideration features that do not have an overt manifestation, but which play a role in the computation of locality. Importantly, the findings for Romanian children’s comprehension of *care* “which” object questions show that the subject-object asymmetry found cross-linguistically is considerably reduced, contrary to what has been reported so far for the comprehension of *which*-object questions in Hebrew-speaking and Italian-speaking children, for example.

We account for such differences by postulating an analysis of *care* “which” object questions building on the idea that intervention effects can be alleviated in structures with a post-verbal subject due to a two-step movement: (i) movement of a Big DP (containing the clitic and the wh-
double) over the post-verbal subject to an intermediate position internal to the IP, but higher than the subject in the specifier of vP, and (ii) movement of the wh-element to the left-periphery out of this intermediate reconstruction site.

Notes

1 This is an instantiation of the “clitic doubling” phenomenon present in languages like Romanian and Spanish, whereby an accusative or dative cliticpronoun appears together with a co-referential full lexical noun phrase.

2 The presence of a clitic pronoun in care–structures determines additional differences with respect to cine–structures as to the presence or absence of weak crossover (WCO) effects and whether they license or not parasitic gaps. See Dobrovie-Sorin (1994) for a more detailed discussion and analysis of clitic doubling in care– vs. cine-questions.

3 The clitic pronoun can appear either post-verbally (in the case of the feminine clitic) or pre-verbally (in the case of the masculine clitic)

4 Final model: Response Accuracy ~ Structure Type + Wh-word Type + Age Group + Structure Type * Wh-word Type + (1 + Structure Type | Participant) + (1 | Item); N = 1584, AIC = 1186.4, BIC = 1245.4, log-likelihood = -582.20, $\chi^2 = 5.27.$

5 In Hebrew, the wh-object is also preceded by an accusative marker – et – similar to pe in Romanian, but the subject appears in a pre-verbal position. In Italian, on the other hand, the subject appears in a post-verbal position in object wh-questions, but there is no case marking on the wh-phrase. This leads to the same surface order in both subject and object questions in Italian, rendering the WH Verb NP order ambiguous between a subject and object interpretation. Agreement with the verb is therefore crucial to disambiguate the structure.

6 Following Belletti (1999 and subsequent work), we consider that the doubled lexical argument appears in the PP complement of the clitic head in the Big DP structure, and not in its specifier position (Uriagereka 1995).

7 Cecchetto (2000) calls this position FP, however the fine details of this intermediate landing site require further investigation. See Coene & Avram (2012) for an analysis of the final landing site of the clitic.

References


Coene, Martine, and Larisa Avram (2012), “(A)symmetries in the acquisition of Romanian clitics,” in: Pedro Guijarro-Fuentes and Maria Pilar Larranaga (Eds.), *Pronouns and clitics in early language*, 131-173, Berlin: Mouton de Gruyter.
Philip, William, Peter Coopmans, Wouter van Atteveldt and Matthijs van der Meer (2001), “Subject-object asymmetry in child comprehension of WH-questions,” in: Anna Do, Laura Dominguez and Aimee Johansen (Eds.), *Proceedings of the 25th annual Boston University


Free Relative Clauses and the Over-Generalization of Wh-Movement

Michael Clauss

1. Free Relative Clauses and Wh-Movement

Wh-movement has, in much classic syntactic work, been assumed to be essentially a single generalized operation used to derive a number of constructions by moving a Wh operator to the specifier of a clause (Chomsky 1977, inter alia). This syntactic assumption is thought to hold for a variety of constructions, such as Wh question formation, Relativization (1), as well as other operations like Tough Movement and Clefting.

(1) Wh-Movement
   a. [CP What did John see (t)]
   b. The thing [CP which John saw (t)]

However, different Wh-movement constructions also are subject to different constraints on output form. One example of this is the availability of complex (i.e. phrasal) Wh expressions in Wh questions compared with Free Relative Clauses (FRCs), discussed in some detail recently by Donati (2006) and Cecchetto & Donati (2011, 2015).

(2) Wh-NP sequences
   a. What (gift) did Ben bring?
   b. I wonder what (gift) Ben brought
   c. I opened what (*gift) Ben brought

This effect, which I will refer to as the *Wh-NP rule, will be the focus here, though other such rules exist which distinguish FRCs from other Wh constructions. One byproduct of this rule is that certain potentially ambiguous sentences which could be interpreted as either embedded questions or FRCs (for example, sentences containing verbs which select both nominal and clausal arguments) are disambiguated in the presence of
Wh-NP sequences. We will see in the following discussion that children do not make use of disambiguation by *Wh-NP.

The discussion will proceed as follows: Section 2 will discuss cases where Wh movement sentences might be ambiguous between different sorts of constructions; Section 3 will discuss children's knowledge of the semantics and syntax of Free Relative Clauses, and in particular failure to use the *Wh-NP rule to disambiguate between FRCs and questions; Section 4 will discuss two possible ways in which children might be making over-generalizations of Wh Movement syntax which would result in this performance. Section 5 concludes with general remarks about the role of over-generalization in acquisition paths.

2. The semantics of CP/DP ambiguities

As discussed above, the *Wh-NP rule is a way in which the syntax of FRCs and embedded Wh questions can be distinguished. Generally, it's the case, as Donati (2006) points out, that categorial selection distinguishes where these constraints apply; FRCs are nominals (DPs), and so in a nominal position one cannot have a Wh-NP sequence (etc.); Wh-questions are clauses (CPs), so in a clausal position one can.

(3) CP positions
   a. Ben asked {if Molly had arrived, what gift Molly brought}
   b. {That Molly came, What gift Molly brought} surprised Ben

(4) DP positions
   a. Ben opened {the box, what (*gift) Molly brought}
   b. {The box, what (*gift) Molly brought} lay open on the table

Cases like (3-4), where the selecting predicate takes exclusively clausal or nominal arguments, make this distinction clearest. A subtler effect, however, comes up in contexts where the syntactic context allows for either clausal or nominal arguments. Such an environment is the complement of *see (5). This gives rise to categorial ambiguity between DPs and CPs with strings like what Molly brought, which can be either an embedded question or an FRC. But, this ambiguity disappears where the *Wh-NP rule would be violated (6).

(5) Complement of *see
   a. Ben saw [DP the book]
b. Ben saw \([_{CP} \text{that Molly wrote the book}]\)

(6) a. Ben saw \([_{CP/DP} \text{what Molly brought}]\)
b. Ben saw \([_{CP/*DP} \text{what gift Molly brought}]\)

Just as the *Wh-NP rule serves to rule out strings like \(\text{what gift Molly brought}\) in (4a-b) above, it rules out the same string being labeled a DP in (6b). This variable categorial ambiguity is not semantically vacuous; there are more possible semantic values for (6a) than for (6b).

What are the two readings for (6a)? FRCs are definite descriptions (or more particularly, definite descriptions derived from questions; see Caponigro 2000, 2003) of entities with the property denoted by the gapped clause, while questions are propositions or sets of propositions. Basically, a DP reading is True as long as the subject sees something with the property of Molly having brought it; a clausal interpretation requires that the seer's belief state come to include knowledge of the answer to the question, \(\text{What did Molly bring?}\).

Certain conditions make this distinction clear: if one sees something with the property of Molly having brought it, without any knowledge about it (as in the situation in 8), then only the DP reading is true, while the CP reading is false. And crucially, in this case, sentences which only include \(\text{see what}\) can be evaluated as true (8a), while sentences including \(\text{see what NP}\) must be evaluated as false.

(8) Situation: Ben sees several books laying open on a table. Among them is the book that Molly brought, though he cannot tell which is which.
   a. Ben saw what Molly brought = True
   b. Ben saw what book Molly brought = False

This is a case where the *Wh-NP rule doesn't rule out possible strings, but it does limit the cases in which a particular string (\(\text{see what book Molly brought}\)) is true. Thus, complete acquisition of English FRCs includes not only failing to produce or accept strings like \(\text{write what book Molly read}\)—applying the *Wh-NP constraint, among other constraints on form, to surface string—but also correctly discerning that constraints on form of FRCs constrain the available readings of licit surface strings.
3. Children's knowledge of FRCs

Some previous findings in acquisition suggest that from a very early age children have adultlike knowledge of the representation of Free Relative Clauses. The first point of note is that children acquiring English begin producing adultlike FRCs as early as 3;6, their first form of relativization in production (Flynn & Lust 1980: 9). And while instances can be found in the CHILDES database (MacWhinney 2000) of children producing Wh-NP sequences in embedded questions from around the same age (10), there are no instances to be found of illicit Wh-NP sequences.

(9) Naturally produced FRC, 3;6
   Cookie monster eats what pushes Big Bird

(10) Adult-like Wh-NP production: Julia, 3;5
   I don't know which day we have these

Evidence that children also know the semantic distinction between embedded Wh and FRCs can be seen in several ways. One comes from comparing studies of children's knowledge of exhaustivity in different constructions. Many studies have shown that children up to around age 6 or 7 have difficulty interpreting definite descriptions with the as exhaustive (Maratsos 1974, inter alia). More recent studies show similar performance for FRCs, specifically that the age of onset for exhaustive FRCs is roughly the same as that for prototypical definite DPs (Modyanova & Wexler 2007; Caponigro et al 2012).

Conversely, exhaustivity in Wh questions does not seem so delayed, and children begin giving exhaustive responses to Wh questions closer to 3 or 4 (de Villiers & Roeper 2011). Thus, children seem to know that the exhaustivity requirement of FRCs tracks not with the Wh word itself but with the semantics of definite descriptions, and that FRCs are the latter.

A second piece of evidence comes from observations of children's answers to syntactically complex questions. While children often erroneously treat embedded Wh questions as matrix questions (de Villiers et al 2008; de Villiers et al 2012: 11), they rarely if ever treat Wh clauses in nominal positions as matrix questions (Clauss 2013: 12).

(11) Medial Wh answers
   How did Ben say [where he had gone]?
   • Adult response: By showing it on a map
Free Relative Clauses and the Over-generalization of Wh-movement

- Child response: To the store

(12) No medial Wh answers
   Was [what Molly saw] scary?
- Adult and child response: Yes
- Never: A wolf

All of these results suggest quite clearly that children do very well at distinguishing, based on factors like selecting predicates, the semantic content of FRCs from that of embedded Wh questions. But, do they do similarly well using the internal syntax of the clauses to disambiguate?

As shown above, adult-like knowledge of English FRCs includes use of this fact in determining where sentences are and are not ambiguous between FRC and question readings. To assure that this is in fact the case for adult speakers of English, I performed a Truth Value Judgment Task with adult speakers of English using sentences of the relevant forms.

The design of the experiment was two-by-two, with the factors being syntax (see what [S] vs. see what NP [S]) and story type (whether or not the question (Q) reading of the target sentence was true). A sample item is shown in (13), with both versions of the story and the prompt questions, as well as their expected truth values.

(13) An experimental item: *Goat is going to Cow's birthday party*

**Q reading False:**
Goat brings Cow a present, but does not see Cow before he puts the gift down among all the other gifts. Cow opens all the gifts, but doesn't know who brought what.
(a) Cow saw what Goat brought (True)
(b) Cow saw what gift Goat brought (False)

**Q reading True:**
Goat brings Cow a present. He hands it to her as he gets to the party. She opens it along with all her other gifts, and sees that Goat brought her a hat.
(c) Cow saw what Goat brought (True)
(d) Cow saw what gift Goat brought (True)

The condition in (b) is critical, as it is the condition where the syntax of the sentence to be evaluated (containing Wh-NP) isn't allowed in the semantic environment (the Question reading is false).

This study was first performed online, with 16 adults, all native English speakers, of various levels of education. There were four items of
the type above, counterbalanced, and four filler items of various types, presented in random order in text form. Participants were not given a time limit to choose an answer. The results are presented in the Table 1, which shows, by condition, rates of “True” responses and standard errors. A two-way ANOVA shows main effects of both Story type ($F = 6.58, p < .05$) and syntax ($F = 20.14, p < .001$), with a marginally significant interaction ($F = 3.7, p = .059$). In the Q False condition, simple Wh sentences are significantly less likely to be evaluated as true ($t = 2.67, p = .01$).

Table 1: Adult results

<table>
<thead>
<tr>
<th></th>
<th>Q-False</th>
<th>Q-True</th>
</tr>
</thead>
<tbody>
<tr>
<td>See Wh</td>
<td>68.75% (12.0)</td>
<td>93.75 (6.3)</td>
</tr>
<tr>
<td>See Wh-NP</td>
<td>25% (11.2)</td>
<td>87.5% (8.5)</td>
</tr>
</tbody>
</table>

The child version of the experiment used similar materials, but with slight edits to the text, and presented with recorded stories and key sentences, accompanied by a series of pictures. Control items included other Wh embedding predicates and both full relatives and embedded questions. Since most of the experimental items have a target adult-like response of True, four out of five control items had a target answer of False, to control for possible “Yes bias” in responses, and included items to control for problems with false belief tasks. Four children were excluded from the data based on failures on one of these two measures.

The participants who were not excluded from the data were 16 monolingual English-speaking children from Western Massachusetts, ages 5;3 to 6;10 (mean of 6;4). The results are summarized in Table 2, again with rates of True responses and standard errors.

Table 2: Child results

<table>
<thead>
<tr>
<th></th>
<th>Q-False</th>
<th>Q-True</th>
</tr>
</thead>
<tbody>
<tr>
<td>See Wh</td>
<td>62.5% (12.5)</td>
<td>81.25% (10.1)</td>
</tr>
<tr>
<td>See Wh-NP</td>
<td>68.75% (12.0)</td>
<td>81.25% (10.1)</td>
</tr>
</tbody>
</table>

A two-way ANOVA shows no main effect of syntax ($F = .08, p = .78$) or story ($F = 1.94, p = .17$), and no interaction ($F = .08, p = .78$). Children
are equally likely to judge sentences as True regardless of syntax for both conditions, though the data shows a trend towards more True responses in the Q True condition in both syntactic conditions. Children were significantly more likely than adults to give True responses in the Wh-NP condition with Q-False stories (Condition b; $t = 2.67, p = .01$).

There are several possible explanations of children's non-adultlike performance: that children are not making the same conceptual assumptions about the sentences as adults, that children are garden-pathed by *see what* into the FRC reading and fail to revise, and that they have different syntactic representations of FRCs than adults. I suggest that the best explanation is grammatical.

A challenge to a conceptual account is that children seem to understand the relationship between seeing and knowing by 4 years, a full year younger than any participants in this study (Schmidt & Pyers 2011; O'Neill, Astington, & Flavell 1992). So these 5-6 year olds know perfectly well that someone can see something by knowing it, as well as simply just seeing it. This suggestion would imply that children should be at ceiling in both of the Q-False conditions; that children behave wholly adultlike with the condition (a) suggests that they know that the Q reading is available, and does not necessarily fit with the situation they are presented.

An explanation based on children's parsing of the sentences is confounded by the fact that children typically hear *see Wh* in Question environments. The Adam corpus on CHILDES shows no examples of *see Wh* in an unambiguous FRC environment, and overwhelming evidence that children should expect Question semantics given this string. 57 instances of *see Wh* in the input (excluding cases where the Wh word is not the complement of *see*, in situ/echo questions) have question semantics. (14) shows relevant examples, including one (14b) with a possible, but unlikely, FRC reading.

(14) a. Adam, 2;8
   Investigator: Do you want to see what it says?
   Investigator: Ask your Mommy what it says
b. Adam, 3;10
   Mother: Go backwards and see what happens.

There is thus no reason to expect that children would be garden-pathed into an FRC interpretation by *see what*; if anything such an explanation would predict that children trend toward not accepting the sentences where the Q reading is false.
Given these problems with conceptual and processing-based explanations, the most likely explanation is that children are giving the strings non-adultlike representations which allow *What NP sequences with FRC semantics.

4. The analysis of FRCs for adults and children

If we assume that children's performance on this task diverges from adults' due to a difference in representation of FRCs, we may then ask what sort of representation would allow FRC and Wh semantics to both allow Wh-NP. First we'll want to consider what the target representation is: why do adults bar Wh-NP in FRCs but not questions?

The syntax assumed here is based on two essential ideas: (i) that in the type of FRCs under discussion here, the Wh word occupies the D₀ of the entire DP within which it appears (following Donati 2006; Cecchetto & Donati 2011, 2015); and that (ii) FRCs are syntactically (as well as semantically) derived from Wh questions, and consist of a +Wh CP dominated by a DP with a definite operator (following Caponigro 2000, 2003; Chierchia & Caponigro 2013).

Cecchetto & Donati (2015) argue that *Wh-NP effects occur because, for a Wh structure to be in a nominal position, the Wh word must be in a position where it can “relabel” the entire substructure as a DP; in the structure in (15), either the C head or the D head can label the structure, resulting in this string having the distribution of both DPs and CPs (as discussed above). In (16), a phrasal Wh item is moved, and an adjoined XP cannot relabel a structure in which it moves; so, only the C can project, and the associated string only has the distribution of a CP.

(15) Ambiguous syntax

```
         DP/CP
           |   D
what   C   TP
        └── Goat brought t
```
A concern for this sort of structure might be the question of how the resulting structure is interpreted. But the inclusion of a definiteness operator in the FRC structure (after Caponigro 2000, *et seq*) can address this issue, while still embracing the idea that the Wh word in an FRC is the head of the entire DP. The crucial idea here is that the Wh word must move to fill the definiteness operator (17). *Wh-NP effects thus arise in English because movement of *what out of a Wh-NP sequence would be an illicit type of left-branch extraction (18).

*Wh-NP effects are thus a type of conspiracy between two general facts about English: the need to fill definite D heads (in this case by
movement of a Wh head) and the Left-Branch condition. We will see below that the analogy to the need for overt definite heads can be helpful in understanding children's performance.

A number of possible grammars would allow for children's mostly adultlike knowledge of the semantics of FRCs as well as their lack of complete adultlike knowledge of *Wh-NP. I will consider two here.

The first possibility is that children have an adultlike knowledge of the semantic pieces of FRCs: they involve the construction of a Wh question clause and a type-shifting definiteness operator. What they lack is the knowledge that the Wh word must head-move to the definiteness operator. I will call this the “Partial Derivation” analysis. The tree in (19) shows what the representation of an FRC would be on this analysis.

(19) FRC in Partial Derivation analysis

```
  DP
   \_________\     \_____\_____\_____
     |        |     C'         CP
     |        |     \__________\_____
     |        |                      DP
     |        |                      \_____
     |        |                    D_def
     |        |                    \_____\_____
                     \_____
                      \_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\_____\____
(21) FRCs in Pseudo-Free Relative analysis.

```
DP
  D
  what
  N
  gift
  CP
  goat brought t
```

While both of these analyses capture the idea that children's representation of FRCs allows for Wh-NP sequences which are illicit for adults, there are several facts about children's acquisition of nominals in general, and relative clauses in particular, that suggest that the Partial Derivation analysis is preferable to the Pseudo-Free Relative analysis.

The first fact involves the connection made above between *Wh-NP and the requirement of filled D heads in English definites. Children acquiring English and typologically similar languages have difficulty consistently requiring that bare nominals (nominals without an overt determiner) have indefinite readings (Pérez-Leroux et al 2004; Gavarró et al 2006), particularly in “NP/DP” languages like English, where there is a correlation between definiteness and overt D heads (Pérez-Leroux, Gavarró & Roeper 2011); that is, for children learning English in particular, DPs without articles such as (22a) are over-assigned definite semantics expected for something like (22b). Thus, while something like (22a) can only include a silent existential/indefinite D head for adults, children allow a silent definiteness operator here.

(22) a. The girl needs shoes
   Adult: indefinite
   Child: indefinite or definite
   b. The girl needs the shoes
   Adult and child: definite

In view of the Partial Derivation analysis of children's performance on the FRC task, this is akin to saying children at preschool ages acquiring English do not consistently apply the rule that definite semantics must be accompanied by a phonologically overt D head. Pérez-Leroux, Gavarró, & Roeper (2011) suggest that this effect with bare NPs correlates with the parameter which allows mixed NP/DP phenomena in English. If the form of English FRCs is a type of extension of this definiteness/D correlation,
then a delay in its acquisition is predicted as long as a delay in adultlike interpretation of bare NPs is seen, or longer.

Other studies of the acquisition of relative clauses also show some issues for the Pseudo-Free Relative analysis. As cited above, is that FRCs appear in production earlier than other relatives (Flynn & Lust 1980). Thus this analysis would involve children learning FRCs first as if they were headed relatives, but still not producing headed relatives early on.2

Of course, the children in the experiment described above are 5-6, an age where children have begun producing headed relatives, so it is possible that at this age, children have one unified, over-generalized representation for relatives of all kinds. This would involve a sort of U-shaped acquisition path for FRCs, where their first representation is apparently adultlike, then they learn a generalized form for all relatives (allowing Wh-NP), and then they eventually learn the adultlike form (disallowing Wh-NP).

While this sort of acquisition path is possible, the acquisition path predicted by the Partial Derivation analysis seems more likely: this would involve making a generalization about all Wh movement (that Wh expressions move to a specifier above C), and later learning a particular feature of a single Wh construction (that FRCs involve Wh head movement to $D_{def}$).3

Choosing between these two analyses is a non-trivial task. But, slightly different predictions are made by the two analyses. In particular, if children are analyzing FRCs as equivalent to Pseudo-FRCs, similar disambiguation tasks using other differences between pseudo- and Standard FRCs should give similar results. One might consider *It-clefts in particular, which are available in *Wh-ever Relatives but in no other relatives, including standard FRCs (23). Thus we would expect that if children are making use of a Pseudo Free Relative analysis for all FRCs they would erroneously permit ambiguity of sentences involving see and an *It-cleft (24)

(23) a. Ben ate whatever it was that Molly brought  
    b. *Ben ate what it was that Molly brought  
    c. *Ben ate the thing it was that Molly brought

(24) a. Ben saw what Molly brought → Ambiguous  
    b. Ben saw what it was that Molly brought → Ambiguous?

Both analyses are viable, and further research to choose between the two must make use of subtle facts about FRCs and related constructions.
5. Over-generalization as a part of the acquisition path

The previous section proposed two possible analyses to account for the behavior of FRCs. An important feature of both of them is that they involve an early representation of FRCs that generalizes features from another construction: the Partial Derivation analyses has children begin by analogizing FRCs to other embedded Wh movement, as involving movement of a +Wh DP to the specifier of a CP, and no further movement (as well as analogizing them with other definite DPs, involving a definiteness operator which they eventually learn must be phonologically realized); the Pseudo-Free Relative analysis has them begin by analogizing FRCs with other relatives, involving a DP which contains a CP and potentially an overt NP restrictor.

This sort of over-extension of one more common pattern in the target language, which is eventually replaced by the more idiosyncratic properties of a particular pattern, is seen in other domains of language acquisition, most familiarly in the acquisition of exceptions to regular morphology \((\text{goed} \rightarrow \text{went})\). Finding subtle ways in which similar paths can be seen in syntactic acquisition allows for a unified understanding of acquisition paths in general, a generally desirable outcome.

Notes

1 The responses for condition b, Q-False with Wh-NP, are higher than expected due to an item effect; 3/4 of the relevant responses were for the sentence in (i).
   
   (i) Bill sees what articles Mary writes
   The use of the imperfective and a plural NP makes this sentence admissible as a Free Relative, with a sort of “silent-ever” reading (roughly, Bill sees such articles as Mary writes). In an unambiguous FR context (such as complement of \(\text{read}\)), a non-plural NP is ungrammatical.
   
   (ii) Bill reads what article-(s) Mary writes.
   With this taken into account, adults' “true” answers to (i) are licit, and so adults' performance on this condition is in fact 98% as expected. Thus we can take it that adults are in fact wholly reliable at distinguishing where FRC readings of \(\text{see Wh}\) are licit. See Section 4 for a discussion on types of FRCs which allow Wh-NP.

2 This assumes that children entertain one grammatical representation for a construction at a time, eventually changing them to a different one if given sufficient input; a different model would allow both analyses to be “competitors” (in the sense of Yang 2003) for the best representation of FRCs until the input convinces them to allow one over the other. Under this assumption the acquisition path involving the Pseudo-Free Relative analysis becomes more possible.
A similar line of reasoning is pursued in Guasti & Shlonsky (1995) in order to explain the early emergence of FRCs compared to other relatives: FRCs allow for obvious, overt Wh movement, and do not require the more complex sort of null operator (or, equivalently, copy-deletion movement) syntax associated with headed relatives, and thus are acquirable early.

Acknowledgement

This work owes a great deal to helpful feedback from Tom Roeper, Jeremy Hartman, Barbara Pearson, Luiz Amaral, Jill de Villiers, Rajesh Bhatt, Adriana Belletti, Nick LaCara, Megan Somerday, and audiences at BUCLD, LSA, UConn, and UM ass.

References

de Villiers, Jill, and Tom Roeper (2011), "The acquisition path for Wh-questions," in: Jill de Villiers and Tom Roeper (Eds.), The Handbook
Free Relative Clauses and the Over-generalization of Wh-movement

of Generative Approaches to Language Acquisition, 189-246, Dodrecht: Springer.
O'Neill, Daniela, Janet Wilde Astington, and John Flavell (1992), "Young children's understanding of the role that sensory experiences play in knowledge acquisition," Child Development 63, 474-490.

Yang, Charles (2003), Knowledge and Learning in Natural Language, Oxford University Press.
IS AGREE ALSO DELAYED AS WELL AS MOVEMENT?

YOSHIKI FUJIWARA

1. Introduction

Children’s difficulty in comprehending verbal passives has been reported in a wide variety of languages: for example, Dutch (Verrips 1996), English (Maratsos et al. 1985), Greek (Terzi and Wexler 2002), Hebrew (Berman 1985), and Japanese (Sugisaki 1999; Minai 2000; Sano 2000). To account for this from the perspective of grammatical maturation, Wexler (2004) proposes that children lack the ability to create defective/weak phases. This maturational hypothesis is called the Universal Phase Requirement (UPR) as illustrated in (1). According to him, the delayed comprehension of verbal passives is due to their violation of the Phase Impenetrability Condition (PIC; Chomsky 2000) in (2).

(1) Universal Phase Requirement:
Children take all vP and CP to define phases.

(2) Phase Impenetrability Condition:
In phase $\alpha$ with head H, the domain of H is not accessible to operations outside $\alpha$, only H and its edge are accessible to such operations.

The essence of the PIC is that the complement domain of a phase head H is not accessible to operations from outside of the phase HP because it is spelled out to PF/LF interfaces when the phase HP is constructed. Keeping this in mind, consider how to explain children’s difficulty with verbal passives under the UPR. The structure of the English passive sentence (3) is roughly as shown in (4). Suppose the $v$ in passives is a defective phase head for adults (Chomsky 2000).
(3) Mary was kissed.

For adults, the internal argument DP moves to Spec TP after undergoing Agree with T. This does not violate the PIC because the v in passives does not constitute a phase for adults. In contrast for children, since even a defective v defines a phase under the UPR, the movement as well as the Agree relation induces their violation of the PIC. This is why children have trouble with verbal passives under the UPR.\(^2\)

The UPR has been examined with constructions that involve movement; for example, passive, raising, and unaccusative constructions. However, given that Agree is a prerequisite for movement, we cannot identify which one causes the delayed acquisition of these constructions, Agree or movement. To address this issue, we have to test constructions that involve only Agree. Nominative Object Construction (NOC) in Japanese is one test case. The result of my experiment with NOCs shows that NOCs are not delayed in acquisition, contrary to the prediction of the UPR. It suggests that children’s observed difficulty with passives and other constructions is related to their inability to create a certain type of movement, rather than the Agree relation between T and objects.

This paper is organized as follows. In Section 2, I briefly introduce the syntactic analysis of NOCs, and then verify the prediction of the UPR for the acquisition of NOCs in Section 3. Section 4 reviews previous studies. Section 5 reports the results of my experiment, which suggests that children around age five have already acquired NOCs. Section 6 concludes the paper.
2. Nominative Object Constructions in Japanese

In Japanese, the nominative case marker –ga is canonically attached to a subject. On the other hand, an object is canonically marked with accusative case marker –o.

(5) John-ga nihongo-o hanas-u.  
John-NOM Japanese-ACC speak-PRES  
*John speaks Japanese.

However, objects can be marked with the nominative marker –ga in certain stative constructions as shown in (6) (Kuno 1973). This is the so-called Nominative Object Construction (NOC). In some NOCs, subjects may be marked with the dative case marker –ni.

John-NOM/DAT Japanese-NOM speak-can-PRES  
*John can speak Japanese.

In the literature it has been argued that nominative objects do not qualify as subjects even though they are marked with nominative case (Shibatani 1977; Ura 1999; Kishimoto 2004; Koizumi 2008; Takano 2003). The subject-oriented anaphor zibun is often used as a diagnostic for subjecthood.

According to Kuroda (1965), zibun takes only a subject as its antecedent. The example in (7) shows that only the subject Taro, not the object Hanako, qualifies as a possible antecedent of zibun.

(7) Taro-ga Hanako-j-ni zibun-j-no e-o mise-ta.  
Taro-NOM Hanako-DAT self-GEN picture-ACC show-PAST.  
(lit.) Taro showed Hanako self’s picture.

I follow Katada (1991) for a theoretical explanation of the subject-orientation property of zibun. She attempts to explain it in terms of Binding Condition A. According to her proposal, zibun is an anaphoric operator and it adjoins to VP at LF as shown in (8). Hence, zibun can be bound only by a subject.

(8) [TP Taro-ga [zibun [VP Hanako-ni t-no e-o mise]-ta]]
Example (9) shows that nominative objects do not qualify as antecedents of *zibun* (Shibatani 1977). Note that in the passive sentence in (10), the internal argument *Mary*, which is also marked with the nominative case, can be an antecedent of *zibun*.

(9) John$_i$-ga/ni Mary$_j$-ga zibun$_{ij}$-no heya-kara mi-e-ta.  
    John-NOM/DAT Mary-NOM self-GEN room-from see-can-PAST  
    (lit.) John$_i$ can see Mary$_j$ from self$_{ij}$’s room.

(10) Mary$_i$-ga John$_j$-ni zibun$_{ij}$-no imouto-to kurabe-rare-ta.  
    Mary-NOM John-DAT self-GEN sister-with compare-PASS-PAST  
    (lit.) Mary$_i$ was compared by John$_j$ with self$_{ij}$’s sister.

This suggests that nominative objects do not undergo A-movement to TP Spec, unlike internal arguments of passives (Takano 2003; cf. Saito 2010).

I assume Takahashi’s (2010) proposal for the NOCs. He proposes that *v* becomes a phase head if and only if it values accusative Case. Therefore, *v* in the NOCs does not constitute a phase head because it does not value accusative Case. Accordingly, nominative objects are case-valued by T via Agree without any PIC violation, as shown in (11).

(11) $[_{TP} \text{Subj} [_{T'} [_{vP} t [_{v'} [_{VP} \text{Obj}_{NOM} V] v] ] T ] ]$

### 3. UPR’s prediction for NOC

If the structure of the NOC in (11) is subject to the UPR, nominative Case of the object is not licensed as in (12) because the PIC blocks AGREE between T and the nominative object.\(^3\)
Thus, the UPR predicts that children should have trouble in generating nominative objects because uCase of the object remains unvalued.

Note that Japanese children have difficulty in comprehending passives until around age 6 (Minai 2000; Sano 2000; Sano, Endo, and Yamakoshi 2001; cf. Okabe and Sano 2002). Given this, it is predicted under UPR that nominative objects are difficult for children until around age 6.

However, this prediction seems wrong since Japanese children even at age 2 are able to produce nominative case marker –ga attached to objects in stative predicates as shown in (13) (Matsuoka 1998).

\[(13)\]
\[
a. \text{Aki-chan are-ga hoshi-i yo.} \quad (\text{Aki, 2;10;7})
\]
\[
\text{Aki} \quad \text{that-NOM want-PRES SFC}
\]
\[
\text{Aki (= I) wants that thing.}
\]

\[
b. \text{Kore-ga d eki-na-i.} \quad (\text{Kan, 2;4;25})
\]
\[
\text{this-NOM can-NEG-PRES}
\]
\[
(I) \text{ can’t do this.}
\]

\[
c. \text{hiru-wa omanju-ga tabe-ta-i.} \quad (\text{Sumi, 2;7})
\]
\[
\text{noon-TOP sweet.bun-NOM eat-want-PRES}
\]
\[
I \text{ want to eat a sweet bun for lunch/afternoon snack.}
\]

Although this data clearly shows the acquisition of NOCs, one might think of a possible way for the UPR to account for it by assuming children apply a different structure from the adult’s to NOCs. A similar idea is proposed by Babyonyshev et al. (2001). They propose that children give an unergative structure to unaccusatives in order to explain children’s early production of unaccusative verbs, which are expected to be delayed.
under the UPR. For example, the English unaccusative sentence in (14) has a structure like (15), which is identical to an unergative’s.

(14) The mail arrived.

(15) \([TP \ [T \ [vP \ t_i \ v \ [vP \ arrived]]]] \) (no object trace)

In (15), the “internal” argument the mail is base-generated in the external position and then moves to TP Spec as the subjects of unergative verbs do. This is an example of what I call misanalysis. According to Babyonyshev et al. (2001), misanalysis takes place when there is a syntactic-homophones defined in (16). Such a misanalysis approach is also proposed for Japanese unaccusatives by Machida, Miyagawa, and Wexler (2004).

(16) A phrase \( \alpha \) is a syntactic homophone (s-homophone) of \( \beta \) if \( \alpha \) and \( \beta \) have distinct structure but common pronunciation.

With the idea of misanalysis, the UPR is able to account for children’s production of NOCs as reported by Matsuoka (1998). Given s-homophones, children may misanalyse nominative objects as nominative marked phrases other than nominative objects, that is, (nominative marked) subjects. If so, the structure of NOCs for children is not (11) but rather (17).

(17)

\[
\begin{array}{c}
TP \\
\text{Subj} \\
"\text{Obj}\"_{\text{NOM}} \\
\text{vP} \\
\text{t}_{\text{Subj}} \\
\text{t}_{\text{Obj}} \\
\text{VP} \\
\text{V}
\end{array}
\]

Note that the nominative “object” behaves like a usual subject in (17); it is base-generated in the external argument position. In this structure, even if children’s vP becomes a phase, uCase of the “object” can be
Is Agree Also Delayed as well as Movement?

valued by T at vP Spec without violating the PIC.\(^7\) Therefore, there are two subjects in (17); the original subject and the misanalysed “subject”. It seems to be likely that Japanese children generate two “subjects” in a sentence, given that Japanese allows multiple subjects as in (18) even in adult grammar.

\[(18) \text{John}_i\text{-ga okusan}_j\text{-ga zibun}_i/j\text{-no heya-de nemutte } i\text{-ru.} \]
\[
\begin{align*}
\text{John-NOM} & \quad \text{wife-NOM} & \quad \text{self-GEN} & \quad \text{room-in} & \quad \text{sleep} & \quad \text{be-PRES} \\
\text{(lit.) As for John}_i, \text{his wife}_j \text{is sleeping in self}_i/j\text{'s room.}
\end{align*}
\]

In (18), the subject-oriented anaphor zibun can refer to the first NP John or the second NP wife.

Therefore, one might be able to account for the early production of NOCs under the UPR by assuming this type of misanalysis. However, as a prediction, if children misanalysed nominative objects as subjects in NOCs, then their nominative “objects” should become a possible antecedent of zibun, unlike adults’, because they are base-generated in vP Spec and can bind the zibun-anaphor in the VP. To test the prediction, I conducted an experiment with NOCs including zibun.

4. Previous studies

Before going to the experiment section, I review the previous studies of the acquisition of zibun and nominative objects. Otsu (1999) reported that Japanese children at age 3-5 have already acquired knowledge of zibun. One of his test items is shown in (19), repeated from (8). In his experiment, almost all the participants knew that zibun refers to the subject, not the indirect object.

\[(19) \text{Taro}_i\text{-ga Hanako}_j\text{-ni zibun}_i/*j\text{-no e-o mise-ta.} \]
\[
\begin{align*}
\text{Taro-NOM} & \quad \text{Hanako-DAT} & \quad \text{self-GEN} & \quad \text{picture-ACC} & \quad \text{show-PAST.} \\
\text{(lit.) Taro}_i \text{showed Hanako}_j \text{self}_i/*_j\text{'s picture.}
\end{align*}
\]

Based on his observation, I assume that young children at age 3 to 5 already have knowledge of the subject-orientation of zibun.

There are some studies investigating whether children’s nominative object is truly an object. In Fujiwara (2013), I examined NOCs such as (20).
The task for children in this experiment is choosing between two pictures, which differ as to whether Baikinman piggybacks the rabbit in Baikinman’s garden or the rabbit’s garden. The antecedent of zibun in (20) is Baikinman, not rabbit. The result is that children chose the matching (Baikinman’s garden) picture at the rate of 45.5% (10/22). One possible problem with the experiment is that, since there is no context in the picture selection task, the potential sentence in (20) is not used in a felicitous way. Another possible problem is that the sentence yields a structural ambiguity as shown in (21).

(21) a. [TP SubjTOPi [VP ObjNOMj [PP zibun\_\(i/(*)j\) …] V] T]
   b. [ ObjTOPi [TP SubjNOMj [VP [PP zibun\_\(i/(*)j\) …] tOBJ V] T]

(21a) is canonical regarding the subject-object order, but (21b) is not, in that the first phrase in (21b) is a topicalized object. Thus, (21b) displays the object-subject order. Therefore, the antecedent of zibun in (20) can vary depending on which structure is applied to the sentence.

However, this ambiguity can be avoided by using the dative subject in NOCs. Sano, Shimada, and Fujiwara (2014) used exactly such a sentence in an experiment with the Truth Value Judgment Task (TVJT; Crain and Thornton 1998). We tested sentences such as (23) in the situation illustrated in (22): A dog, an elephant, a pig and a monkey are in the dog’s class (on the left side). A pig, a squirrel, an elephant and a cat are in the pig’s class (on the right side). The dog is about to give a medal to an animal who he finds the most interesting in each class. He considers the elephant the most interesting in his class, while in the pig’s class, he is interested the most in the pig.
Is Agree Also Delayed as well as Movement?

In (23), *zibun*'s antecedent is the dative phrase *dog* but not the nominative phrase *pig*. Note that the sentence in (23) does not yield a structural ambiguity because a dative-marker can be attached only to the subject in this construction. In this experiment, children around age 5 correctly rejected (23) for the situation in (22). The result that children around age 5 disallowed nominative objects as *zibun*'s antecedent is unexpected under the UPR. However, it is unclear whether the acquisition of sentences such as (23) is predicted to be delayed under the UPR because its predicate is adjectival and its structure is not clearly defined. To fix this problem, I used verbal predicates in NOCs in my experiment.

In addition, Sano et al.'s experiment itself may have a problem. Consider the possibility that children correctly rejected the target items in the situation in (22) without interpreting *zibun*. If children have ignored *zibun* in (23), the same interpretation should arise in (24).

(24) Inu-ni-wa buta-ga kumi-no-naka-de ichiban
    dog-DAT-TOP pig-NOM class-in the most
    omosiro-i interesting-PRES

*The dog is interested the most in the pig in a class.*

Contrary to the target items such as (23), the sentence in (24) is true of the situation in (22). As a prediction, if children ignored *zibun* in interpreting the target sentences, they would also reject the example in (24) as they rejected the target items in the situation in (22). Although this
test was not conducted with children, out of the 18 adult participants, 14 adults wrongly rejected Example (24) in the situation in (22). Most of them rejected it for the reason that they paid attention only to the dog’s class because the dog who gave the medals belonged to that class. Thus, whether zibun is used in a test sentence or not, adults tend to answer on the basis of the dog’s class in the scenario of (22). Given that, it can be reasonably speculated that whether zibun is used in a test sentence or not, the children in our experiment would also have given their answers only paying attention to the dog’s class as the adults did in the test with (24).

5. Experiment

The UPR predicts that NOCs are difficult for children until around age 6 like passives. Recall that in order to account for children’s production of NOCs reported by Matsuoka (1998), under the UPR one may need to assume that nominative objects are misanalyzed as subjects. If such a misanalysis took place, however, children’s nominative objects should be zibun’s antecedent. The purpose of the experiment is to see whether this prediction is borne out.

5.1. Method

15 Japanese children, aged from 4;9 to 5;11 (mean 5;4) participated in this experiment with TVJT. Example (25) is one of the target items in this experiment. In (25), the dative marker –ni is attached to the subject, instead of the nominative marker –ga, in order to avoid a structural ambiguity.

(25) Buta_i-ni-wa inuj-ga zibun_i/*j-no heya-kara tatak-e-ta.
    pig-DAT-TOP dog-NOM self-GEN room-from hit-can-PAST 
    (lit.) The pig_i was able to hit the dog_j from self_i/*j’s room.

The topic marker –wa is added to the dative subject to make the sentence sound natural (Shibatani 2001). Two types of verbs – tatak “hit” and tutuku “poke” – were used in the target items, attached to the potential suffix –e “can”. Thus, the predicates are stative as a whole. Zibun’s antecedent can be the dative subject pig but not the nominative object dog.
Four target items were given to children in matching and mismatching scenarios. A sample mismatching scenario for the target items is the following: There are two rooms near a tree. One is a dog’s, the other a pig’s. The dog and pig play tag with a stick. The pig, as tagger, looks for the dog and goes to the PIG’s room. He finds the dog hiding behind the tree and tries to hit the dog from the PIG’s room, but he fails because the tree is in the way. Then, he goes to the DOG’s room, and succeeds in hitting the dog from the DOG’s room (For a matching scenario, capitalized PIG replaces DOG and vice versa). The final scenes of the mismatching and matching situation are shown in (26) and (27), respectively.

(26) Situation: the pig hits the dog from the dog’s room.

(27) Situation: the pig hits the dog from the pig’s room.

Sentence (25) is false for the mismatching situation in (26), but if children regarded the nominative object dog as a subject, then they would accept it in the story. In the matching scenario in (27), the sentence in (25) is true. If children know the dative-marked phrase pig is the subject, children should accept it in the matching situation.
The result of the target items is as follows. In the mismatching scenario in (26), children correctly rejected the target items such as (25) at the rate of 96.7% (29/30). On the other hand, in the matching scenario in (27), they correctly accepted them at the rate of 86.7% (26/30). It means that children did not take the nominative object as zibun’s antecedent but instead took the dative subject as the antecedent. It seems that the nominative-marked DPs in sentences like (25) do not behave as subjects in child Japanese, contrary to the prediction discussed above. However, it is possible to suspect that the children gave correct answers in this task for other reasons.

First, consider the possibility that children gave correct answers to the target items without interpreting the sentences. Recall that in the matching scenario in (27), the pig is in his room in the final scene of the story, while he is not in his room but in the dog’s room in the last scene of the mismatching scenario (26). Therefore, children may have accepted/rejected the target stimulus sentences, depending on whether or not the pig is in his room in the last scene. To exclude this possibility, control sentences such as (28) were included just before the target items.

(28) Buta-ni-wa inu-ga inu-no heya-kara tatak-e-ta.
   pig-DAT-TOP dog-NOM dog-GEN room-from hit-can-PAST
   The pig was able to hit the dog from the dog’s room.

A phrase inu-no heya-kara “from the dog’s room” is used in (28), instead of zibun-no heya-kara “from self’s room” in the target item (25). The control items are presented in the same scenarios used in the target items. Contrary to the target items, this control items become true for the situation in (26) and false for the situation in (27). If children responded “true” or “false” depending on whether the pig is in his room in the final scene of the story, then they would give wrong answers to the control items like (28) because it is true in the situation where the pig is in the dog’s room.

There were two trials of this control item for each child: one for Situation (26) and the other for Situation (27). All of the 30 trials with (28) except one response in the situation in (26) were correct answers, which means that children did not rely on such a strategy in giving correct answers to the target items.

There was another control item. In the target item (25), since pig is not only the subject but also the first phrase of the sentence, it is possible to suspect that children took it as zibun’s antecedent because it is the first phrase. This strategy was tested by means of control sentences like (29).
Here, I used the verb *nage* “throw,” which takes an indirect object. Since (29) does not have the potential suffix –e, the predicate is not a stative, which means nominative objects and also dative subjects cannot appear in (29). Note, however, that this control item is the same as the target item in (25) regarding the NP-ni-wa NP-ga order. The difference is that the first phrase *pig* is a topicalized indirect object in (29). The subject of (29) is the nominative marked phrase *dog*. Therefore, the antecedent of *zibun* is the second phrase *dog* but not the first phrase *pig*. If children took the first phrase in a sentence as *zibun*’s antecedent, they would incorrectly interpret the control item in (29).

There were two trials with control items like (29): one with a matching story and the other with a mismatching story. The matching story for (29) is as follows: a dog and a pig play with a ball. The dog tries to throw the ball to the pig from the PIG’s room, but he cannot because there is a tree between the dog and the pig. Then, he moves to the DOG’s room and throws the pig the ball from there (For the mismatching story, capitalized PIG replaces DOG and vice versa).

Out of 30 trials with this control, only one response in the matching scenario was a wrong answer. The correct answer rate is at 96.7% (29/30). This result suggests that children do not simply take the first phrase in a sentence as *zibun*’s antecedent.

5.2. Discussion

Recall that it is predicted under the UPR that immature children cannot construct the Agree relation between objects and T and move objects to TP Spec because both operations are blocked by phases created by the UPR. Accordingly, NOCs and passives should both be delayed in acquisition under the UPR because the former includes the Agree relation and the latter both operations.

The result of the experiment demonstrates that 5-year-old children know that nominative objects cannot be the antecedent of *zibun*. Considering that they easily identify *zibun*’s antecedent in NOCs, it clearly shows that they do not “misanalyse” nominative objects as
subjects. This suggests that the Agree relation between the nominative object and T is not problematic for children at this age, in contrast to the prediction of the UPR.

Note that the age of success with NOCs is before the age of success with the Japanese passive. Given this, the observed delay in passives is not relevant to the Agree relation but to movement of the object. Thus, it seems that the UPR is too strong to be maintained.

6. Conclusion

In this paper, I investigate what causes children’s delayed acquisition of verbal passives. Given the Universal Phase Requirement proposed by Wexler (2004) and the Phase Impenetrability Condition, one possible answer to this issue is the movement of the internal argument to Spec TP. Another possible answer is the Agree relation between the internal argument and T. If such an Agree relation is problematic for children, Japanese nominative object constructions should also be delayed in acquisition.

The result of my experiment demonstrates that Japanese children around age 5 do not have trouble with nominative objects. Acquisition of Japanese nominative objects suggests that children allow the object to be a goal for the T probe in contrast to the prediction of UPR. It suggests that the delay of passives is not caused by Agree between T and an object. Rather, it would be related to movement of the object to TP Spec. In this sense, UPR is too strong to be maintained. Thus, the maturational hypothesis should be related to movement as stated in the A-Chain Deficit Hypothesis (Borer and Wexler 1987), the Universal Freezing Hypothesis (Hyams and Snyder 2006; cf. Snyder and Hyams 2015) or the Argument Intervention Hypothesis (Orfitelli 2012). These hypotheses all claim that children’s grammar is unable to create a certain type of movement. In order to determine what type of movement is delayed, we need to wait for further studies.

Notes

1 The PIC introduced in Chomsky (2000) is different from the one given in (Chomsky 2001) in terms of the condition of Spell-Out. Under the former, Spell-Out is applied once a phase is completed, while it must wait until the next higher phase head is introduced under the latter. Wexler (2004) seems to adopt the former
version. It appears that the UPR cannot attribute children’s delay of passives to violation of the latter version of PIC. In this paper, I assume the version of the PIC introduced in Chomsky (2000).

2 The UPR seems to assume that children cannot take an option of moving the internal argument to the target position, TP Spec, through the edge of vP. If this option were available for children, the UPR would lose the ability to explain children’s delayed passives because such a movement does not violate the PIC.

3 Here, I assume that the nominative object cannot undergo movement or scrambling to the edge of vP in order to be Case-valued by T without the PIC violation. See also f.n.2.

4 Similarly, it is reported by Hirsch and Wexler (2007) that 6-year-old children also have difficulty with verbal passives in English.

5 As to theta roles in (15), Babyonyshev et al. (2001) mention that children fully know Baker’s (1988) Uniformity Theta Assignment Hypothesis (UTAH) but they can violate it. Thus, although the “internal” argument the mail is base-generated in the external position, that is, the specifier of vP, it possesses the correct theta role theme in (15).

6 This is similar to the structures proposed by Saito and Hoshi (1998), Shibatani (2001), and Takano (2003) in that nominative objects are base-generated outside of the predicate. See Takahashi (2010) for arguments against supposing such a structure in NOCs.

7 Although I suppose that the nominative “object” moves to Spec TP to satisfy EPP in (17) as usual subjects do, it is possible to consider that the nominative “object” stays in the base-generated position Spec,vP. The choice of the movement does not affect the discussion.


9 One may suspect that children just selected the dative subject as zibun’s antecedent in the target sentence in (25) although they regarded nominative objects as subjects. It means children have a tendency to choose a distant and dative-marked subject between two possible antecedents of zibun. Okabe’s (2008) experiments might be helpful to consider this point although constructions used in her experiments and my experiment are different. Studying a construction that has two subjects, she found that children from 4 to 6 prefer to take the nominative-marked subject as zibun’s antecedent in sentences like (i), in which the embedded subject is marked with dative Case. And if both subjects are nominative marked as in (ii), children prefer the one that is nearer to zibun.

(Bold shows children’s preference for zibun’s antecedent)

(i) **Buta**-**ga** [kuma-ni zibun-no bousi-o kabur]-ase-ta.
pig-NOM bear-DAT self-GEN hat-ACC put.on-cause-PAST
(lit.) The pig put the bear’s hat on zibun’s.

(ii) **Buta**-**wa** [kuma-ni keeki-o tabe-ta] no-o mi-ta.
pig-TOP bear-NOM cake-ACC eat-PAST COMP-ACC see-PAST
(lit.) The pig saw the bear ate zibun’s cake.

Given this, if children consider that there are two subjects in (25), they should choose the nominative object as zibun’s antecedent because it is nominative-
marked, and it is the closest potential antecedent to zibun. From this, I speculate that children’s performance on the target items in my experiment reflects their knowledge of grammar, not their preference in choosing zibun’s antecedent.

Acknowledgements

I am grateful to Takuya Goro, Ken Hiraiwa, Miwa Isobe, Diane Lillo-Martin, Reiko Okabe, Hiroyuki Shimada, William Snyder, and especially to Tetsuya Sano for valuable comments. I would also like to thank the children and the stuff at their day-care center for offering me the opportunity to conduct the experiments.

References


1. Introduction

In most languages with head-initial relative clauses (RCs), English and Italian among them, children comprehended subject RCs (e.g. *the cat that hits the dog*) at ceiling level from three years of age on, whereas object RCs (e.g. *the cat that the dog hits*) are challenging for typically developing school-age children (e.g. Adani 2011; Corrêa 1995; Costa Lobo and Silva 2011; Guasti and Cardinaletti 2003; Labelle 1990, 1996; Utzeri 2007). This subject advantage in RCs with animate lexical NPs has also been observed in atypical language acquisition (e.g. Friedmann and Novogrodsky 2004; Håkansson and Hansson 2000; Zukowski 2009). On the other hand, in languages with head-final RCs, such as Mandarin and Cantonese, both the subject advantage and the object advantage have been reported (e.g. Chan, Matthews and Yip 2011; Chen and Shirai 2015; Hu et al. 2016b; Lee 1992).

Only a relatively small number of studies have examined the acquisition of relative clauses in bilingual children. The current study investigated school-age sequential Mandarin-Italian bilingual children’s comprehension of Mandarin and Italian RCs. Our results will provide evidence for the subject advantage in the acquisition of Mandarin RCs. We organize the present paper as follows. First, we present the properties of Mandarin and Italian RCs and the prediction of the current study. Then, we report Experiment 1 on the comprehension of Mandarin RCs, followed by Experiment 2 on the comprehension of Italian RCs. Finally, we offer a general discussion.
2. Mandarin and Italian relative clauses

Mandarin RCs are head-final, namely, the RC linearly precedes the relative marker *de* and the relative head, as exemplified in (1a) for subject RCs and in (1b) for object RCs.

(1) a. da xiaogou de xiaomao
    hit dog DE cat
    *the cat that hits the dog*

   b. xiaogou da de xiaomao
    dog hit DE cat
    *the cat that the dog hits*

By contrast, Italian RCs are head-initial, namely, the relative head precedes the relative marker and the RC, as illustrated in (2a) for subject RCs and (2b) for object RCs.

(2) a. i gatti che colpiscono il cane
    the cat-pl that hit-pl the dog-sg
    *the cats that hit the dog*

   b. i gatti che il cane colpisce
    the cat-pl that the dog-sg hit-sg
    *the cats that the dog hits*

Although Mandarin RCs are superficially different from Italian RCs, their hierarchical structure is quite similar to that of Italian RCs. We illustrate the hierarchical structure of Mandarin RCs in (3) and that of Italian RCs in (4).
To properly interpret RCs, a dependency between the relative head and its copy should be established. In Mandarin object RCs (3b), this dependency crosses the embedded subject *xiaogou* “dog”, namely, the subject structurally intervenes in this dependency, as it is the first e-
commanded argument found by the relative head. In contrast, there is no structural intervener between the relative head and its copy in Mandarin subject RCs (3a). Similarly, in Italian object RCs (4b), the embedded subject *il cane* “the dog” structurally intervenes in the dependency between the relative head *i gatti* “the cats” and its copy, whereas in Italian subject RCs (4a), there is no structural intervener between them. To sum up, in both languages, the dependency between the relative head and its copy in subject RCs does not cross any argument, but it does in object RCs.

Mandarin and Italian RCs have been studied in first language acquisition. It is well-established that Italian monolingual children comprehended subject RCs much better than object RCs (Adani 2011). The same asymmetry has been revealed by Mandarin studies (Hu et al. 2016b), but the contrasting result has also been reported (Chen and Shirai 2015). The subject/object asymmetry has been given various explanations based on the active filler hypothesis (Frazier and Flores D’Arcais 1989), the minimal chain principle (De Vincenzi 1991) and the Dependency Locality theory (Gibson 1998, 2000). In this paper, we assume a recent proposal, the featural Relativized Minimality (fRM; Rizzi 1990, 2004; Starke 2001) approach sketched in Friedmann, Belletti and Rizzi (2009). The fRM approach claims that the nature of the relative head and of the embedded NP modulates the difficulty in the acquisition of object RCs. The central idea of fRM is that the local relation between X and Y is blocked when a Z, structurally similar to X (and Y), intervenes. See (5), taken from (Rizzi 2004: 225).

(5) a. X … Z … Y
   b. Z intervenes between X and Y iff Z c-commands Y and Z does not c-command X

Under a raising analysis (e.g. Donati and Cecchetto 2011; Vergnaud 1974), the relative head is attracted by a complex attractor [R, NP], where R and NP represent the relative feature and the lexical restriction feature, respectively. The embedded subject and object both have a [NP] feature and the relative copy additionally has a [R] feature. As shown in (6a) for subject RCs, there is no structural intervener between the relative head and its copy. By contrast, as illustrated in (6b) for object RCs, the embedded subject intervenes in the dependency between the relative head and its copy, and its features are a subset of the feature of the relative head.
According to Friedmann, Belletti and Rizzi (2009), to establish the dependency in object RCs, the subset relation has to be computed to decide whether the intervening subject (Z) is similar to X and Y. Children are assumed to have more limited computation resources than adults and may fail to compute this subset relation, ending up treating the relative head and the embedded subject as featurally similar. Accordingly, this would yield an fRM violation (for relevant discussion on agrammatic comprehension, see Garraffa and Grillo 2008; for more discussion on child acquisition, see Adani 2011; Belletti and Rizzi 2013; Guasti, Stavrakaki and Arosio 2012; Hu, Gavarró and Guasti 2016). Based on this account and on previous findings, we expect that Mandarin–Italian bilingual children will have problems in comprehending object RCs, since in both languages these structures display structural intervention by the embedded subject.

We also know that language-specific properties influence bilingual children’s acquisition of RCs. Results from studies on Cantonese-English bilingual children show that surface structure overlap is one key factor claimed to account for influence (Yip and Matthews 2007; Kidd, Chan and Chiu 2015). As in Mandarin RCs, Cantonese RCs are head-final. Kidd, Chan and Chiu (2015) reported that Cantonese-English bilingual children (N = 20, aged 4;10–11;11, M = 8;11) comprehended subject RCs significantly better than object RCs in both languages, while monolingual Cantonese children (N = 20, aged 5;2–9;2, M = 6;4) did not show a preference. The authors argued that bilingual children’s difficulty in Cantonese object RCs is related to the fact that Cantonese object RCs have a canonical SVO word order, which competes with the canonical SVO word order in Cantonese and crucially in English. Based on the canonical order, children assign the Agent thematic role to the first NP in Cantonese object RCs. While monolingual children reanalyze this assignment, the bilinguals do not and thus fail to comprehend object RCs in Cantonese. We should be cautious with this conclusion due to the large age range of the children tested in the study, namely, from age 4 to age 11. Averaging across different ages may conceal the subject/object asymmetry. Bilingual children may be less advanced and be at a stage in which the asymmetry is still evident (for a review of inconsistent results of previous studies on acquisition of Cantonese RCs, see Chan, Matthews and Yip 2011). If our explanation of Kidd et al.’s findings is correct, we expect that the subject/object asymmetry is still evident in older Mandarin–Italian
bilingual children, while it is no longer evident in older monolingual Mandarin-speaking children (although it will be evident in younger monolingual children).

To test these predictions we investigated for the first time sequential Mandarin–Italian bilingual children’s comprehension of RCs. Control data were also collected from age-matched monolingual children.

### 3. Experiment 1: Comprehension of Mandarin relative clauses

#### 3.1. Methods

48 Mandarin–Italian bilingual children and 64 Mandarin-speaking children participated in this study. The children were further divided into a younger (6;0–7;11) and an older (8;0–9;11) group. Details for the participants are provided in Table 1.

<table>
<thead>
<tr>
<th>Group</th>
<th>#</th>
<th>Age range</th>
<th>Mean age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bilingual younger children</td>
<td>24</td>
<td>6;0–7;11</td>
<td>6;11</td>
</tr>
<tr>
<td>Bilingual older children</td>
<td>24</td>
<td>8;0–9;11</td>
<td>8;10</td>
</tr>
<tr>
<td>Mandarin younger children</td>
<td>40</td>
<td>6;1–7;11</td>
<td>6;11</td>
</tr>
<tr>
<td>Mandarin older children</td>
<td>24</td>
<td>8;0–9;11</td>
<td>8;8</td>
</tr>
</tbody>
</table>

The bilingual children were recruited from the Chinese community in Milan (Italy). After an initial screening done by the first author together with parents and teachers, all the parents of suitable bilingual children completed a questionnaire on the linguistic input children received since birth, and on children’s ability to use both languages. The criteria for the selection of the bilingual children were the following: no history of language impairment or hearing loss; regular exposure to Mandarin Chinese from birth; systematic exposure to Italian started at the kindergarten by age 4; use of both languages on a daily basis. The
Mandarin control children were recruited from primary schools in Zhejiang (China). They had regular exposure to Mandarin from birth and no significant exposure to other languages.

The Mandarin materials consisted of 8 subject RCs like (7a) and 8 object RCs like (7b), the same as in Hu et al. (2016b).

(7) a. Na yi-ge shi da xiaogou de xiaomao?
   which one-CL is hit dog DE cat
   Which one is the cat that hits the dog?

b. Na yi-ge shi qingwa da de laoshu?
   which one-CL is frog hit DE mouse
   Which one is the mouse that the frog hits?

A character-sentence matching task was used. All the sentences were matched with 16 sets of pictures. Each set of pictures consisted of two black and white drawings, and in each set of pictures there were four characters, as exemplified in Figure 1 (Hu et al. 2016b). The characters participated in the same event, but with reversed thematic roles. Children were asked to choose one character (out of four), because the function of RCs is to single out the referent described by the RC. In addition, there were 8 filler sentences involving intransitive verbs (e.g. sit) or actional irreversible verbs (e.g. wear). In total, there were 24 items.

When participants did not choose the correct character, we coded the response as Error. Due to limitation of space, we do not discuss the detail in this paper.

Figure 1: A set of experimental pictures used in Experiment 1

3.2. Results

Figure 2 reports the percentages of correct responses by Mandarin–Italian bilingual children and their Mandarin monolingual peers. It is
evident that bilingual children perform less accurately than their monolingual peers. Improvement is observed both in bilingual and in monolingual children. For all the groups, but the older monolingual children, subject RCs are easier to comprehend than object RCs. These observations are confirmed by the statistical analysis.

Figure 2: Percentages of correct responses in Experiment 1

The data were analyzed using the lme4 package (Bates et al. 2015) in the R environment (R Core Team 2015). The best-fitting model included Sentence Type (subject RCs vs. object RCs), Age (younger vs. older) and Group (bilingual group vs. control group) as fixed factors. Reference categories were object RCs for Sentence Type, older for Age and bilingual group for Group. As random effects, we included a by-subjects and a by-items intercept. The analysis revealed an effect of Sentence Type (Wald $Z = 6.80, p < .001$), Age (Wald $Z = -7.45, p < .001$) and Group (Wald $Z = 5.45, p < .001$) and a significant interaction of Sentence Type and Group (Wald $Z = 3.43, p < .001$).

To unpack this interaction, we further analyzed each group separately. Crucially, the main effect of Sentence Type was found in the Mandarin control group ($\chi^2(1) = 41.50, p < .001$; Wald $Z = 11.55, p < .001$) and in the bilingual group ($\chi^2(1) = 18.10, p < .001$; Wald $Z = 5.75, p < .001$), showing that subject RCs were performed significantly better than object RCs in each group.

Additionally, a new model was run to compare the younger Mandarin control group, the older Mandarin control group, the bilingual younger age group and the bilingual older age group. We found that the older Mandarin children’s performance was at ceiling, and their performance was significantly better than their age-matched bilingual peers ($\chi^2(3) = 99.76,$
These older bilingual children did not differ from the younger Mandarin control children (Wald $Z = -0.41$, $p = .68$).

4. Experiment 2: Comprehension of Italian relative clauses

4.1. Methods

The same bilingual children in Experiment 1 took part in Experiment 2. In addition, 41 Italian-speaking children served as controls, and we divided them as a younger Italian control group ($N = 14$, aged 6;0–7;10, $M = 7;1$) and an older Italian control group ($N = 27$, aged 8;1–9;9, $M = 8;10$). The Italian controls were recruited from primary schools in Milan and Como (Italy). The criteria for the selection of the Italian control children were the same as that for the Mandarin control group in Experiment 1.

The Italian materials consisted of 8 subject RCs as in (8a) and 8 object RCs as in (8b), which was used by Adani (2011). All the experimental sentences were unambiguous sentences as the relative head and the embedded noun always had different number features.

(8)  

(a) *Fammi vedere le tartarughe che rincorrono la gallina.*  
*Show me the turtles that chase the hen.*

(b) *Fammi vedere i leoni che il cavallo rincorre.*  
*Show me the lions that the horse chases.*

We used the character-sentence matching task, with only one black and white picture for each experimental sentence (Adani 2011). All the pictures had the same structure (e.g. a pair of animals X on the left, animal Y in the middle and a pair of animals X on the right), as exemplified in Figure 3. In addition, there were 8 filler sentences. Similar to the Mandarin filler sentences, the Italian fillers involved either intransitive verbs or actional irreversible verbs. In total, there were 24 items.
Figure 3: An experimental picture used in Experiment 2

**4.2. Results**

Figure 4 reports the percentages of correct responses by bilingual children and their Italian monolingual peers. It is evident that bilingual children obtain lower scores than monolingual children and subject RCs are easier to comprehend than object RCs. Improvement with object RCs is observed in both monolingual and bilingual children.

![Figure 4: Percentages of correct responses in Experiment 2](image)

We first analyzed the Italian controls and bilingual children’s responses. The best-fitting model included Sentence Type (subject RCs vs. object RCs), Age (younger vs. older) and Group (bilingual group vs. control group) as fixed factors. The analysis revealed an effect of Sentence Type (Wald $Z = 4.75$, $p < .001$), Age (Wald $Z = -.529$, $p < .001$) and Group (Wald $Z = 3.51$, $p < .001$), a significant interaction of Sentence Type and Age (Wald $Z = 3.43$, $p < .001$) and a significant interaction of Sentence Type and Group (Wald $Z = -3.20$, $p < .01$).
To explore these interactions, we further analyzed each group separately. Crucially, the main effect of Sentence Type was found in the Italian control group ($\chi^2(1) = 8.96, p < .01; \text{ Wald } Z = 3.32, p < .001$) and in the bilingual group ($\chi^2(1) = 14.79, p < .001; \text{ Wald } Z = -4.02, p < .001$), showing the fact that subject RCs are more accurate than object RCs in each group. The main effect of Age was observed in the Italian control group ($\chi^2(1) = 5.95, p < .05; \text{ Wald } Z = -2.54, p < .01$) and in the bilingual group ($\chi^2(1) = 36.04, p < .001; \text{ Wald } Z = 10.21, p < .001$). Importantly, the significant interaction of Sentence Type and Age was found in the control group (Wald $Z = 3.58, p < .001$) and in the bilingual group (Wald $Z = 3.39, p < .001$), revealing the fact that the older age groups are more accurate than that of the younger age groups in object RCs, but not in subject RCs.

Moreover, in additional models comparing the younger Italian control group, the older Italian control group, the bilingual younger age group and the bilingual older age group, we found the bilingual older children performed significantly worse than their age-matched monolingual peers ($\chi^2(3) = 44.76, p < .001; \text{ Wald } Z = 2.88, p < .01$).

5. Discussion

In this study, we examined school-age sequential Mandarin–Italian bilingual children’s comprehension of RCs, as compared to their monolingual peers’ performance. This study confirmed with bilingual children a subject/object asymmetry both in Mandarin and Italian RCs. Moreover, as we predicted, the asymmetry was still evident in the older bilingual children, but not so in older monolingual children. This shows that bilingual children are weaker in the comprehension of RCs than the monolinguals. In the following, we discuss the findings.

Our results confirmed the subject/object asymmetry in both languages: children comprehended subject RCs quite well, but object RCs were found to be much less accurate. We conjecture that the asymmetry is due to the structural intervention of the subject in object RCs. In Mandarin and Italian object RCs, locality has disruptive effects in the dependency between the relative head and its copy. According to the fRM approach, one has to compute the subset relation in order to distinguish the relative head from the embedded subject (see references in section 2). Children are assumed to have limited computation resources and have trouble computing the subset relation. As a consequence, they treat the relative
head and the embedded subject as similar and an fRM violation incurs. Therefore, they may fail in understanding object RCs.

Bilingual children were less advanced than their age-matched monolingual peers. With regard to Mandarin data, Mandarin monolingual children in the older age group were at ceiling level in their comprehension of both subject and object RCs, while their age-matched bilingual peers were still far from the ceiling performance (80% for subject RCs and 55% for object RCs). With respect to Italian data, Italian monolingual children in the older age group were almost at ceiling on object RCs, while their age-matched bilingual peers comprehend 69% of the RCs correctly. These contrasting results suggest that the acquisition of two typological distinct languages may slow down for complex structures (especially for object RCs). Taken together with the same asymmetry observed in bilingual children and in the two control groups, we propose that the bilingual children’s parsing mechanism is basically the same as that of monolingual children, but their ability of integrating multiple source of information to form an interpretation is less efficient than that of monolinguals’ (see also Clahsen and Felser (2006) for adult L2 learners and Serratrice (2007) for simultaneous bilingual children).

To sum up, we found that bilingual children and monolingual younger children, regardless of whether RCs are head-final (Mandarin) or head-initial (Italian), showed more difficulty in comprehending object RCs than in subject RCs. Such a subject/object asymmetry is consistent with the prediction of the fRM approach along the line of Friedmann et al. (2009). We also showed that the competence of RC comprehension requires several more years of exposure for bilingual children. All in all, our study provides new evidence that learning two languages may temporally slow down the acquisition of each language for complex structures such as relative clauses, but it does not affect the qualitative aspects of development.

Note

In Italian, there is another type of object RCs, in which the embedded subject is a postverbal subject as in (i). See Adani (2011) for the relevant discussion.

(i) i gatti che colpisce il cane
the cat-pl that hit-sg the dog-sg
the cats that the dog hits
Acknowledgements

We sincerely thank all the participants, their parents and teachers. We are grateful to Fabrizio Arosio and Mirta Vernice who helped us out with the data collection, to HU Shuangling, HU Guanshao, YU Weichang and ZHANG Meilian for their help in obtaining permission to work in schools, and to the reviewers for their valuable comments. The research leading to these results has received funding from the European Union’s Seventh Framework Programme for research, technological development and demonstration under grant agreement no.613465, and from the PRIN (Programmi di Ricerca scientifica di rilevante Interesse Nazionale) project no.20128YAFKB (“Theory, Experimentation, Applications: Long distance dependencies in forms of linguistic diversity”).

References


Guasti, Maria Teresa, and Anna Cardinaletti (2003), “Relative Clause Formation in Romance Child’s Production,” *Probus* 15, 47-89.

Guasti, Maria Teresa, Stavroula Stavrukaki, and Fabrizio Arosio (2012), “Cross-linguistic Differences and Similarities in the Acquisition of Relative Clauses: Evidence from Greek and Italian,” *Lingua* 122, 700-713.


1. Introduction

Comparing objects according to some standard is one of the central components in human cognitive ability. Every human language encodes a way to construct comparisons. Comparisons are typically expressed by gradable adjectives that are not simply treated as a predicate but somehow encode degrees that are ordered along an ordered set of abstract dimensions or scales (Seuren 1973; Cresswell 1977; von Stechow 1984; Heim 1985, 2000; Kennedy 1999; Kennedy & McNally 2005; Kennedy 2007b). Comparisons are made up of a target of comparison, a standard of comparison and a standard of abstract measurement. European languages such as English manifest the core relations linguistically: a target of comparison is realized as a main subject, a standard of comparison is introduced by than and a scale is associated with a gradable adjective that makes an ordering relation explicit by a degree morphology. The relevant examples are listed in (1):

(1)  a. David is taller than Nancy.
     b. That is more interesting than watching England play football.

Truth conditions of a sentence that contains a gradable adjective are determined by the comparison between two degrees; one is derived by applying the gradable adjective meaning to its sentential subject and the other by a standard expression hosted by than. The meaning of (1a) can be true if and only if the degree of height David possesses exceeds the degree of height Nancy possesses. Assuming that gradable adjectives are measure functions that take an object to return a degree and the comparative morpheme -er combines with the measure function to denote a property of individuals, the meaning of the complex predicate in (1a) can be illustrated as follows à la Kennedy (1999):
(2)\[
\begin{array}{c}
\text{DegP}\\
\lambda x.\text{tall}(x) > \text{tall}(\text{Nancy})
\end{array}
\]

Even if a standard of comparison is not linguistically explicit or the gradable adjective is used in its positive or unmarked form as in (3a), a degree relation must be somehow hypothesised. One way to express the relation is to assume the abstract degree morpheme \textit{pos} following Kennedy (2007b). The meaning of the positive degree morpheme is shown in (3b), where \textit{s} is a context sensitive function that derives a standard of comparison in order that the objects stand out in the context of utterance. Since Chris in (3a) is tall compared with other individuals, the proposition that Chris is tall is true in an absolute sense; the statement such as “Chris is short” cannot follow as shown in (3c).

(3) a. Chris is tall.
   \[ [[[\text{Deg } \textit{pos}]]] = \lambda g \lambda x. g(x) \geq \textit{s}(g) \]
   b. Chris is tall, but he is short for an adult male.

By contrast, the ‘absolute’ implication is not relevant to the comparative counterpart in (1a). It is possible that David is short while (1a) is true: the context in which both David and Nancy are short, but David is at least taller than Nancy. The distinction between the comparative form of gradable adjectives and the unmarked form is manifested by the existence of the degree morpheme; while the comparative morpheme makes the ordering relation clear and does not take the contextual information into account, the absence of the degree morpheme requires the postulation of the abstract degree morpheme that encodes an ordering relation with respect to a context of utterance. A question to be asked is whether the distinction between the comparative meaning of gradable adjectives and the non-comparative meaning is upheld universally; is it really the case that they are in complementary
distribution? Logically speaking, it is possible that their core meanings are somehow combined. We can hypothesize that the ordering relation between a target of comparison and a standard of comparison is made explicit linguistically and that a target holds degrees that stand out with respect to a contextual standard. Roughly put, the semantics like (4) would be possible, contrary to fact:

\[(4) \quad [[\text{pos}]]( [[\text{taller than Nancy}}]]) = \lambda x. \text{tall}(x) > \text{tall}(\text{Nancy}) \land \\
\lambda x. \text{tall}(x) \geq s( [[\text{tall}}]])\]

Japanese is a good candidate to investigate the availability of the ‘non-neutral’ comparison that is a comparative construction with contextual information, because this language does not distinguish the comparative form from the non-comparative form because of the lack of the degree morpheme. In Japanese comparative meanings are derived by the existence of the standard marker yori ‘than’ or sometimes the directional nominal hoo ‘towards’ that is attached to a target of comparison as shown in (5a) (Kubota and Matsui, 2012). The example in (5b) shows that hoo can be omitted rather freely. The existence of hoo in (5c) leads to a neutral interpretation; David can be either tall or not tall depending on a context. In the absence of hoo or yori a non-neutral or positive interpretation is required as in (5d).

(5)  
\begin{align*}
\text{a. Nancy-yori(mo) & David-no & hoo-ga} \\
& \text{Nancy-than & David-GEN & toward-NOM} \\
& \text{se-ga & takai.} \\
& \text{height-NOM & tall(er)} \\
& \text{‘David is taller than Nancy.’}
\end{align*}

\begin{align*}
\text{b. Nancy-yori(mo) & David-ga & se-ga & takai.} \\
& \text{Nancy-than & David-NOM & height-NOM & tall(er)} \\
& \text{‘David is taller than Nancy.’}
\end{align*}

\begin{align*}
\text{c. Dotira-mo (seizing-tositewa) & ookiku-nai kedo,} & \\
& \text{Which-FOC (adult-for) & big-NEG though} \\
& \text{David-no & hoo-ga & se-ga & takai.} \\
& \text{David-GEN & toward-NOM & height-NOM & tall(er)} \\
& \text{‘Both of them are not large (for adults), but Davis is taller.’}
\end{align*}

\begin{align*}
\text{d. David-ga & se-ga & takai.} \\
& \text{David-NOM & height-NOM & tall(er)} \\
& \text{‘Davis is tall.’}
\end{align*}
The purpose of this paper is twofold; one is to investigate whether young Japanese children are sensitive to the distinction between comparative and positive meanings of gradable expressions based on absolute implications and the other is to explore whether the existence of degree morphemes has an influence on the availability of non-neutral comparisons. Based on three experiments, I show that child Japanese is sensitive to the distinction between comparative and positive meanings and non-neutral comparisons are not observed. The central conclusion of this paper is to support the view that Japanese comparisons are based on degrees after all (Sawada, 2013; Shimoyama, 2014; Sudo, 2015). If the analysis is on the right track, it follows that the availability of degree morphology is not directly related to the availability of degree comparisons. The organization of this paper is as follows. Section 2 is a starting point of the discussion to achieve the goal. I discuss my past experiment to show that Japanese young children are sensitive to the distinction between the comparative and the positive uses of gradable expressions. In section 3 I analyze the distinction between the two uses based on absolute implications. I first show that an absolute implication is not relevant in processing comparisons. I then point out that there are some variations in establishing a contextual standard in the positive uses of gradable expressions among young children. Finally, in section 4 I summarize the discussion and consider the availability of non-neutral comparisons cross-linguistically.

2. Experiment I

2.1. Research Questions

In order to explore whether Japanese-speaking preschool children are sensitive to the distinction between the comparative and non-comparative uses of gradable adjectives, it is necessary to set up a context, in which these two uses are differentiated. In addition, for the positive or non-comparative use of gradable adjectives to be interpreted in the same way as the adult grammar, a general conception that is associated with an adjective must be established. The purpose of the present section is to investigate this. The experiment has contained both the comparative and the non-comparative uses of gradable adjectives in order to check whether each child has a solid knowledge about comparisons and whether each child has a general conception about the size of bears even after seeing the smallest bear. We assume that the bear is large in an absolute sense and the
assumption is supported by the control group; all the university students think that the bear is large in general.

2.2. Procedure

We have carried out an experiment with twenty-five children, ranging in age from 3;10 to 6;10 (mean age 4;23) and with sixteen university students as control\(^1\). The experiment is based on a modified version of the Truth-Value Judgement Task by Crain and Thornton (1998). In this task, each child was asked about the size of bears by an experimenter. The test sentences are shown in (7) and the experiment used three bears (15cm, 35cm, 45cm). At the end of each story, the experimenter verbally asked each child about the size of bears. The task for the child was to provide a correct answer to the puppet, Mickey Mouse.

(6) *Sample Story:*
Mickey Mouse has come to Sapporo for the first time. Since he has been in the dream land all his life, he has no idea about the size of animals on earth.

(7) (After introducing many animals (bears, foxes, dogs and mice), a child was asked to tell him the size of animals by an examiner.

a. Docchi-ga ookii?
   Which-NOM large
   ‘which is bigger?’

b. Kumasan-wa ookii-no? tiisai-no?
   bear-TOP large-Q small-Q
   ‘Is the bear large or small?’

2.3. Results

The result of the responses is summarized in Table 1, where 1 stands for a correct answer, while 0 an incorrect answer and the result is expressed by a percentage. It is assumed that the bear is large for the question 3. All the children gave a positive answer for (7b) even after seeing the smallest one. The experiment shows that the small bear (15 cm) in front of them does not have an influence on the judgement about the size of bears.
Table 1: Results of the Experiment I: children’s vs. adults’ responses

<table>
<thead>
<tr>
<th>Question</th>
<th>children</th>
<th>adults</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Which is bigger? (45 cm vs. 35 cm)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2 Which is bigger? (35 cm vs. 15 cm)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>3 Is the bear big or small?</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

The experiment conducted here shows that preschool children can compare the size of animals and they have a general conception about the size of bears. All the examinees including university students live in Hokkaido, which is famous for brown bears and there are plenty of information about bears. The original aim of this experiment was to investigate whether young children have acquired the knowledge of gradable adjectives irrespective of scalar structural differences (Syrett et al., 2009). New experiments to be reported in section 3 are a further attempt to explore whether young children are sensitive to the distinction between the comparative and non-comparative uses of gradable expressions based on absolute implications.

3. Experiment II

3.1. Research Questions

Having shown that Japanese preschool children have knowledge of comparisons, a new experiment has been conducted to investigate whether young children are sensitive to the distinction between the comparative and the non-comparative uses of gradable expressions. The purpose of the experiment is to investigate whether children distinguish the meanings of fast and slowly in comparative uses from those of non-comparative uses. The aim of this experiment is to investigate whether the comparison based on the scale of speed is possible even if the ‘positive’ meaning of the gradable expression hayai ‘fast’ is used in a very slow competition that would count as slowly without an explicit comparison and whether the comparison based on the ‘negative’ meaning of the gradable expression osoi ‘slowly’ is possible (Kennedy 2001).
3.2. Procedure

Twenty-two Japanese-speaking children have participated in this experiment, ranging in age from 4;11 to 5;11 (mean age, 5;09). Twelve university students have also participated in the experiment as a control group. As in the experiment reported in section 2, we have made use of a modified version of the Truth-Value Judgement Task by Crain and Thornton (1998). In this task, each child watched a story on a tablet computer with a figure called Zibanyan. At the end of each story, a puppet asked the child and the figure about the content of the story. The task for the child was to judge whether the figure’s description of the story was true or false, by giving him either a lovely chocolate bar called Chocobo (This is the figure’s favorite food and every child knows about this.) as a prize for the correct answer or a very spicy red chili as punishment for the wrong answer. The story and the test sentences uttered by the figure are listed in (8) and (9).

(8) Sample Story:
Today, we have an athletic festival. The hero, Anpanman and the enemy Baikinman are competing with each other in running and jumping. Unfortunately, it has begun to rain and the two characters have got wet. Since his face’s got wet, Anpanman has been feeling sick. Anpanman said, “Oh no, I cannot do my best, because my face has got wet”. (Anpanman’s weak point is so famous that every Japanese child knows about this problem.) The enemy, Baikinman, felt easy and relaxed because he knew Anpanman was in trouble. Baikinman said, “Wow, it’s my day, today. It is super easy to beat Anpanman!”. They both ran slowly and Anpanman almost collapsed. Baikinman ran slowly but at least he ran faster than Anpanman.

(9) a. Baikinman-wa Anpanman-yori hayaku
    Baikinman-TOP Anpanman-than fast(er)
    hasit-ta nya².
    run-PAST meow
    ‘Baikinman ran faster than Anpanman.’

b. Anpanman-wa Baikinman-yori osoku
    Anpanman-TOP Baikinman-than slow(er)
    hasit-ta nya.
    run-PAST meow
‘Anpanman ran more slowly than Baikinman.’

### 3.3. Results

The results of the experiment are listed in Table 2 and Table 3 and the percentage is basically the same between children and adults ($p = 0.5195$. Hence, no significant difference between them).

Table 2: The percentage of children’s responses

<table>
<thead>
<tr>
<th></th>
<th>(9a)</th>
<th>(9b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a chocolate</td>
<td>95</td>
<td>68</td>
</tr>
<tr>
<td>a red chilli</td>
<td>0</td>
<td>27</td>
</tr>
<tr>
<td>no response</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 3: The percentage of students’ responses

<table>
<thead>
<tr>
<th></th>
<th>(9a)</th>
<th>(9b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a chocolate</td>
<td>92</td>
<td>67</td>
</tr>
<tr>
<td>a red chilli</td>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td>no response</td>
<td>8</td>
<td>0</td>
</tr>
</tbody>
</table>

Variable responses are observed for (9b). Six children thought that the (correct) statement for (9b) is actually wrong. This is reasonable for native speakers of Japanese, because the comparison based on negative gradable expressions is somehow difficult to process. The reason may be due to the fact that Japanese does not have the counterpart of *less than* comparatives and we often make use of negation for comparisons based on negative gradable expressions. Actually four university students thought that something is wrong about (9b); they speculate that there was some special reason for Anpanman to ran slowly and thought about many different scenarios for this situation in order to understand why the comparison is based on slowness.

### 3.4. An Additional Experiment

We have conducted another experiment with the same participants. The story is the continuation of the experiment above. At the athletic festival,
the two characters competed with each other in jumping and Anpanman was still sick, because his face was still wet. Baikinman felt easy and relaxed and jumped not high but at least he jumped higher than Anpanman. Since both of these characters jumped not high, it is expected that children think that even the winner Baikinman do not jump high. In the animation broadcasted on TV they can fly freely like Superman or Batman. The story and the test sentences are presented in (10) and (11).

(10) Sample Story:
Unfortunately, Anpanman was lost in running. This time they compete with each other in jumping but it was still raining. Anpanman was weakened because his face was still wet. Anpanman jumped first but the height was very low, only a few centimeters, although he can fly usually. Anpanman said, “Oh no, I’m still sick with my face wet”. The enemy, Baikinman, smiled and was happy. Baikinman jumped not high, but at least higher than Anpanman. Baikinman said, “Oh, how happy I am today. I’m so lucky!”.

    Baikinman-TOP    high    jump-PAST    meow
    ‘Baikinman jumped high.’

b. Anpanman-wa    hikuku    ton-da    nya.
    Anpanman-TOP    low    jump-PAST    meow
    ‘Anpanman jumped low.’

The results of the experiment are summarized in Table 4 and 5. Since both these characters can fly rather freely in the cartoon, our expectation is that children will give a red chili to Zibanyan for (11a), contrary to fact. In fact, some children said takaai ‘high’ or boku-yori takai ‘higher than me” during the experiment. From their reaction, I suppose that what counts as high varies among young children, because establishing a contextual standard needs some or much experience. Five university students also gave a chocolate; they thought that the height is not good enough to count as high and thus they reinterpret (11a) as a comparison between the height of Baikinman and Anpanman. After all, they thought that Baikinman did not jump high in an absolute sense. The p-value for (11a) is relatively small between children and adults ($p = 0.0749$) but it is still higher than 0.05. Hence, it might be safe to say that there is no significant difference.

As for (11b) all the children gave a chocolate, because the gradable expression based on a negative scale hikui ‘low’ is acceptable for them.
Hence, it is concluded that gradable expressions with a negative meaning are usable in their non-comparative uses, while they are hard to process in comparisons in child Japanese. Four university students gave a red chili for (11b), because they thought that the height of Anpanman is not good enough to count as jump; according to their judgement, some reasonable amount of height is necessary if you want to say “X jumped.” The $p$-value for (11b) is rather low ($p = 0.00394$) and thus there is significant difference between children and adults. It is possible to say that the verb tobu ‘tobu’ is not compatible with a ‘negative’ modifier for some adults, but I leave for future research with respect to how and why that kind of semantic restriction is observed.

Table 4: The percentage of children’s responses

<table>
<thead>
<tr>
<th></th>
<th>(11a)</th>
<th>(11b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a chocolate</td>
<td>73</td>
<td>100</td>
</tr>
<tr>
<td>a red chilli</td>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>no response</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5: The percentage of students’ responses

<table>
<thead>
<tr>
<th></th>
<th>(11a)</th>
<th>(11b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a chocolate</td>
<td>42</td>
<td>67</td>
</tr>
<tr>
<td>a red chilli</td>
<td>58</td>
<td>33</td>
</tr>
<tr>
<td>no response</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

All the results of the experiment II is summarized in the following graph, where c-(9a) expresses the number of responses by children and a-(9a) the number of responses by adults and *ejusdem generis*. The graph shows that child Japanese and adult Japanese are basically the same except that some verbs (e.g. tobu ‘tobu’) are not compatible with a ‘negative’ gradable expression that is reflected in the responses for (11b).
4. Summary and Discussion

This paper has used the absolute implications observed in gradable expressions as an empirical probe on the distinction between the comparative and the positive uses, showing that Japanese young children are aware of the distinction between these two in spite of the fact that there is no degree morpheme on gradable expressions. First, I have shown that by 6 years of age, children can compare the size of bears and even after seeing a small bear they have a solid conception about the size of bears; the bear is large in an absolute sense. Second, we have investigated whether a gradable expression in comparisons holds a degree that is evaluated by a contextual standard. The experiments conducted in section 3 show that by 5 years of age, Japanese young children are aware that contextual information does not need to be calculated in processing comparisons. In one story, two characters ran slowly and the comparison based on hayai ‘fast’ that is a ‘positive’ part of the scale of speed is possible for both children and university students. In the same situation, the comparison based on the negative scale derived by osoi ‘slowly’ seems to be hard to process for both of them. The variability observed here might indicate that there is no difference about the knowledge of comparisons between children and adults. The other experiment tries to investigate how the non-comparative use of gradable expressions receives an interpretation.
in the similar situation. The experiment shows that processing the positive form of the negative gradable expression *hikui* ‘low’ is relatively easy to comprehend, while there is variability about the judgement of *takai* ‘high’ in its positive use. We have assumed that jumping a few centimeters does not count high for the two characters, because they can fly rather freely in the cartoon. According to the children’s reactions, however, it seems that they compare the height based on their own experiences, leading to the variability. We have estimated that establishing a contextual standard needs some experience but a detailed analysis of how to establish a standard will be left for future research.

It is important to point out that it is not necessarily true that non-neutral comparisons are not an option available by Universal Grammar. Bogal-Allbritten (2013) discusses the comparisons in Navaho and argues that the clitic =*go* marks an adverbial clause, or an adverbializer. The bracketed expression can be moved to the left edge and sometimes can be omitted. If deleted, the sentence will be interpreted as ‘my little sister is pretty’. Since the statement holds true, the comparison can include contextual information.

(12) Shideezhí [shi-lááh ‘át’ée=go] 
    my.little.sister 1-SG.OBJ-beyond 3-SUBJ.BE=SUB 
nízhóní. 
    3-SUBJ.pretty.AA 
    ‘My little sister is prettier than me.’

In English, it is known that the complement of *than* can be a variety of expressions (Corver 2006). The examples in (13a) and (13b) below show that the clausal comparative, where the complement of *than* is hypothesized to compose a clausal structure, contains an absolute implication for some speakers, while the phrasal counterpart in (13c) is insensitive to the contextual factors.

(13) a. John plays the piano better than Mary plays the piano. 
    b. John plays the piano better than Mary does. 
    c. John plays the piano better than Mary.

If the semantic difference between the clausal and the phrasal comparatives is the real one, it is necessary to hypothesize a different semantic entry for the comparative morpheme. This is because if the phrasal comparative is analyzed as a type of clausal structure at LF and the comparative morpheme does not distinguish the phrasal comparative from
the clausal one, there would not be any difference between the clausal and the phrasal comparatives (Bhatt & Takahashi 2011). If two semantic entries for the comparative morpheme are required, it follows that comparisons are not necessarily incompatible with contextual information that would be derived from the standard expression hosted by (clausal) than.

Comparative constructions are not the only option in which degree expressions are combined with gradable expressions. Assuming that measure phrases denote degrees, measure phrase modification provides another type of complex degree constructions. Watanabe (2013) points out that measure phrase modification with an absolute implication is available in Japanese if the degree nominal can be treated as a type of gradable adjectives. The difference between the two examples below is subtle. The copular verb aru in (14b) leads to a non-neutral interpretation, while the complex copular verb dearu in (14a) is neutral with respect to contextual information. Hence, the statement It is very short is compatible with (14a), but it is not with (14b).

(14) a. Kono biru-wa takasa 10-meeteru dearu.  
this building-TOP height 10-meter is 
Totemo hikui.  
very short  
‘This building is 10 meters high. It is very short.’

b. Kono biru-wa takasa 10-meeteru aru.  
this building-TOP height 10-meter is  
#Totemo hikui.  
very short  
‘This building is 10 meters high. It is very short.’

If the non-neutral interpretation of measure phrase modification is an option available, a question to be asked is how and why Japanese preschool children are sensitive to the distinction between the comparative and the non-comparative uses of gradable expressions. Degree morphemes cannot trigger the difference, simply because they are not realized in Japanese. Therefore, at least visibility guideline in the sense of Fukui and Sakai (2003) for adjectives and adverbs does not play an important role in distinguishing the comparatives from non-comparatives based on an absolute implication. The experiments conducted in this paper suggest that Japanese comparisons are based on semantic degrees as it is the case in English comparisons. Finally, it is possible that comparisons in some languages can lack a degree morpheme and they are not dependent
on degrees. Bochnak (2015) points out that comparisons in Washo (isolte/Hokan) do not adopt degrees and they make reference to contextual information. In Washo, comparisons are built by a conjunctive strategy like (15).

(15) \text{t'è:liwhu delkáykayiP k'ePí daPmóPmoP} \\
\text{t'e:liwhu de-Pil-kaykay-iP k'-eP-i daPmoPmoP} \\
\text{man NMLZ-ATTR-tall-ATTR 3-COP-IPFV woman} \\
\text{delkáykayiPé:s k'áPaš.} \\
\text{de-Pil-kaykay-iP-e:s k'-eP-aP-š} \\
\text{NMLZ-ATTR-tall-ATTR-NEG 3-COP-AOR-SR} \\
\text{‘The man is taller than the woman} \\
\text{(Lit: ‘The man is tall, the woman is not tall.’).’}

Bochnak (2015: 4)

If Bochnak (2015)’s analysis and description are on the right track, the availability of degree morphemes can be a trigger for the semantic parameter in degree constructions. Japanese data presented here at least indicate that comparisons in Japanese are based on degrees and there is no essential difference between English and Japanese (Sawada 2013; Shimoyama 2014; Sudo 2015).

Notes

1 Part of this experiment was presented at Kawahara (2013).
2 Nya is onomatopoeia or the characteristic crying sound of cats in Japan. Since he is a ghost cat, Zibanyan uses this word at the end of his utterances to indicate what kind of role he plays.
3 See also the discussion about the typology of degree comparisons by Kennedy (2007a) where it is claimed that Japanese and English build different comparisons.

Acknowledgement

This research is partially based upon work supported by JSPS Grant-in-Aid (Kakenhi) for Young Scientists B under Grant Number 26770173. I am grateful to the audience of GALA 12 for valuable comments on, and discussion of the material in this paper. Of course, all the remaining errors or inconsistencies are my responsibility.
References


ON THE ACQUISITION OF DISTRIBUTIVE NUMERALS IN SERBIAN

NATAŠA KNEŽEVIĆ
AND HAMIDA DEMIRDACHE

1. Introduction

This paper investigates child and adult comprehension of distributed numerals – that is, bare singular and numeral NPs with and without the overt distributive marker po – in Serbian. So-called distributive anti-quantifiers (in the sense of Choe 1987, Zimmermann 2002) or distributive-share quantifiers (Gil 1995), such as po, differ crucially from distributive (universal) quantifiers, such as svaki (every) in Serbian or every/each in English, in that they select as argument the NP serving as the distributed share, while the latter select the NP serving as the distributive/sorting key. The basis of this inquiry is Knežević’s (2015) claim that the Serbian distributive marker po is indeed not a universal quantifier, quantifying (distributing) universally over individuals, but rather a marker of event plurality, enforcing distribution of the nominal it combines with over events. Consequently, the truth conditions of sentences with po-numerals in Serbian, will be more vague, looser than those of sentences with universally quantified numerals and, as such, compatible with a variety of distributive contexts (namely, participant, as well as spatial and/or temporal distributive contexts).

Our experimental investigation of Serbian children and adult comprehension of sentences with bare singular/numeral NPs with and without the distributive marker po validates the analysis of po as enforcing distribution of the numeral NP it combines with over spatiotemporally separate (sub)events. Our findings suggest, moreover, two developmental stages in the acquisition of po: a first stage where children do not know the meaning of po and also allow distributive readings of numerals without an (overt) distributor (unlike adults), and a second stage where children know that po marks the distributive share and distributes over events and also start showing the adult preference for marking distributivity overtly with po.
2. Theoretical background

In the literature, distributive markers such as Serbian *po* are known as “anti-quantifiers” (Choe 1987), “distributive-share quantifiers” (Gil 1995) or “distance-distributivity markers” (Zimmermann 2002), as opposed to universal (determiner-like) quantifiers such as *every* and *each* in English or *svaki* (*every*) in Serbian.

The core empirical difference between universal distributive quantifiers such as *svaki* (*every*) and anti-quantifiers/distributive share markers such as *po* is that, while the former select as argument the NP serving as the *sorting key* (the argument over which the distribution takes place), the latter select as argument the NP serving as the *distributive share* (the argument that is distributed over the sorting key) (cf. Choe 1987, Zimmermann 2002). Thus, in both (1) and (2), the distributive share is the NP *boat* and the sorting key is the NP *girl(s)*: boats are distributed over a set of girls. Crucially, while in (1), the universal quantifier *svaki* or its English counterpart *every* attaches to the subject/sorting key (*girl*), in (2), the distributor *po* attaches to the object/distributive share (*boat*).

(1) a. *Svaki*-sentence  
[Svaka [sorting key devojka]] pere brod.  

b. *Every*-sentence  
[Every [sorting key girl]] is washing a boat.

c. The girls are each washing a boat.  
(participant-distributive reading)  
#The girls are washing a boat at different times/places  
(event-distributive reading)  
?All the girls are washing a boat together.  
(collective reading)  

(2) a. *Po*-sentence  
Devojke peru [po [distributive share brod]].  
girl-nom.pl  wash-3.pl.pres.  DIST  boat-acc.sg.

b. The girls are each washing a boat.  
(participant-distributive reading)
c. The girls are washing a boat at different times/ places.
   (event-distributive reading)
   #The girls are washing a boat together.
   (collective reading)

The reader should bear in mind, throughout our discussion of the readings available in (1) vs. (2) that Serbian is a language with determiner-less NPs that can be interpreted as either definite or (non) specific indefinites. Importantly, since po numerals and more generally, distributive shares, must be construed as non-specific (cf. Choe 1987), only the non-specific reading of the po-nominal is licensed in (2) (see Knežević 2015 for extensive discussion).

The sentences in (1) and (2) all yield a so-called “participant-distributive” reading, where the girls each wash a boat separately – that is, where boats (the distributive share) appear to be distributed over girls (the sorting key) – that is, over the agent participants in the described event. The difference between po and svaki is that, besides this participant-distributive reading, po-sentences also yield a so-called “event-distributive” reading where the girls are washing boats at different times and/or places – that is, where boats (the distributive share) are distributed over spatially and/or temporally separated washing events, each involving the girls acting together as an agent. This reading is not available with svaki in Serbian or every in English.

In sum, there are two correlated differences between universal quantifiers such as svaki and anti-quantifiers such as po: the former syntactically combine with the NP serving as the sorting key and yield only participant-distributive readings, while the latter syntactically combine with the NP serving as a distributive share and yield two types of distributive readings (participant-distributive and event-distributive), as the glosses provided for (1) vs. (2) illustrate.

Choe (1987) and Zimmermann (2002) propose a unified account for quantifiers and anti-quantifiers, under the assumption that both involve a distributive operator (each). When the distributor combines with the sorting key, the distributive operator is ranging over individuals. When the distributor combines with the distributive share, the distributive operator can range over either individuals or over events. In contrast, Knežević (2015), building on Cable’s (2014) analysis of Tlingit distributive numerals, argues that po does not universally quantify/distribute over individuals, but rather distributes over spatiotemporal locations (more precisely, over spatially and/or temporally separated (sub)events and that
participant-distributive readings can be derived as instance of spatiotemporal distribution).

On Knežević’s (2015) proposal, *po* is a marker of event plurality enforcing rather weak truth conditions – that is, merely that the situation described by the sentence containing a *po*-numeral NP involves at least two events that are spatially or temporally separated, each of which must involve *nNP*(s) (where *n* stands for the numeral with which *po* combines). The resulting sentence will be true under various scenarios, as long as the described events involving *nNP*(s) have different running times and/or running spaces and/or participants. Essentially, *po* contributes to the meaning of a sentence by conveying that there is a plurality of events, each involving *nNP*(s). For illustration, the *po* sentence in (3a), will be assigned the semantics in (3b) where ε stands for a type of event involving *nNP*(s) (here, one boat) and * for a cumulative denotation of the predicate (Link 1983).

(3) a. Dve devojke peru *po* (jedan) brod.
   two girl-nom.pl. wash-3.pl.pres DIST one boat-acc.sg

   b. λe. e ∈ *ε₁boat & e ∉ ε₁boat & *wash(e) & \exists x *boat(x)
   & Theme-sum(e)(x) & \exists y*girl(y) & | At(y) | = 2
   & Agent-sum(e)(y)

   c. There is an event constructed out of (sub)events each involving one boat, and this is an event of two girls cumulatively washing one boat.

We can informally read (3b) as in (3c). On this proposal, the sentence in (3a) describes an event constructed out of (at least) two subevents involving two girls cumulatively (that is, together or individually) washing a boat. Since here *po* attaches to the singular NP “(one) boat” (the distributive share), *po* requires that each subevent involve exactly one boat. Now, suppose that the set of two girls is itself partitioned atomically and each girl atom individually washes a different boat. Then a participant-distributive reading arises since the ensuing reading is tantamount to distributing boats over individuals (girl) atoms. A context that would make (3a) true under this so-called participant-distributive reading is illustrated in (4a).
Distributive contexts

a. Participant-distributive (PartD)

b. (Spatial)Event-distributive (EventD)

The truth conditions of (3a) require that there be an event constructed out of (at least two) subevents involving one boat each (since po in (3a) marks ‘(one) boat’ as the distributive share). Suppose, however, that the group (of two) girls is not partitioned into atoms and that the respective agent of each subevent is a group of two girls washing (together) a different boat (per subevent). A context that would make (3a) true under this scenario is illustrated in (4b). The ensuing reading is tantamount to distributing events (of two girls washing a boat) over spatial locations, since the two washing subevents in (4b) are spatially but not temporally separated (that is, have different running spaces, but not running times).

But (3a) would just as well be true on a scenario where the washing events are distributed over different temporal locations—that is, where the subevents are temporally separated, have different running times (whether or not they also have different running spaces). This would be the case on a scenario where say the two girls (either collectively or individually) wash a (different) boat (at least) twice during the course of the week.

Consider now the sentence in (5), where po combines with the numeral NP “two girls”.

(5) a. Po dve devojke peru (jedan) brod.
    DIST two girl-nom.pl. wash-3.pl.pres. one  boat-acc.sg
b. \(\lambda e. e \in *_{\text{2girls}} & e \notin *_{\text{2girls}} & \ast_{\text{wash}}(e) & \exists x *_{\text{girl}}(x)
\& \text{Agent-sum}(e)(x) & \exists y *_{\text{boat}}(y) & | \text{At}(y) | = 1
\& \text{Theme-sum}(e)(y)\)

c. There is an event constructed out of (sub)events each involving two girls, and this is an event of girls cumulatively washing a boat.

In (5), since \(po\) combines with the NP “two girls”, there must be at least two subevents each involving a different set of two girls washing a boat. As a consequence, (5a) will be true under the spatial distributive context in (4b), but crucially not under the participant-distributive context in (4a) since only the former satisfies the requirement enforced by \(po\) – namely, that there be two girls per subevent.

(5a), however, would be true be it on a spatial or a temporal distributive reading. On the spatial-distributive reading, there are (at least two) simultaneous subevents each involving two girls washing a boat, while on the temporal-distributive reading there are (at least two) temporally separated subevents each involving the same/a different boat. This would be the case on a scenario where say groups of two girls (either collectively or individually) wash a different/the same boat, one group after another.¹

Finally, note that \(po\) only combines with numerals, weak quantifiers (e.g. \(mnogo\) (many)) or singular bare nouns. In the latter case, since the NP serving as a distributive share must always be a quantity denoting expression, Knežević (2015) analyses the NP combining with \(po\), which appears on the surface to be bare, as combining with the silent numeral “one”—e.g. \(po\) \(\text{brod}\) (in (2)/(3)) is analyzed as \([\text{distributive share} \ po \ [\ ONE \ \text{boat}]\].

To conclude, \(po\) does not universally quantify over individuals, unlike the universal quantifier \(svaki\) (every). Instead, the semantics of \(po\) is that it merely requires a plurality of events involving \(po\)-numeral NP(s). On this proposal, \(po\) is neither a quantifier, nor an anti-quantifier, but rather a marker of event plurality enforcing distribution of the \(n\)NP(s) it combines with over different (sub)events. Both so-called participant and event-distributive readings arises via the plurality of events (involving \(po\)-numeral NP(s)) that \(po\) enforces.

Our goal here is to contribute to the large literature on the acquisition of distributivity which exclusively bears on the acquisition of distributive quantifiers, such as \(every\) and \(each\), which, unlike \(po\), do not yield event-
distributive readings, by experimentally investigating the acquisition of the distributive-share marker *po*.²

### 3. Experiment

The goal of the experiment reported here is to explore the interpretations of numerals with and without distributive *po* in Serbian. In particular, we seek to answer the following questions:

(6) Research questions

**Adult and Child Grammar**

(a) Is distributive marking on the numeral obligatory?

**Child Grammar**

(b) Do children know that the distributive marker *po* enforces distributing the numeral NP to which it attaches over (sub)events? That is,

i. Do children know that that *po* attaches to the distributive share (i.e. that the position of *po* determines what is being distributed)?

ii. Do children know that *po* yields both participant-distributive and event-distributive readings?

To this effect, three sentence types were tested: *po*-less, *po*-object and *po*-subject sentences, as illustrated in (7), (8) and (9) respectively.

(7) *po*-less sentences

Dve devojke peru brod.
two girl-nom.pl. wash-3.pl.pres. boat-acc.sg.
‘Two girls are washing a boat.’

(8) *po*-object sentences

Dve devojke peru po brod.
two girl-nom.pl. wash-3.pl.pres. DIST boat-acc.sg.
‘Two girls are each/separately washing a boat.’

‘Two girls are washing a boat at different times/places.’

(9) *po*-subject sentences

Po dve devojke peru brod.
DIST two girl-nom.pl. wash-3.pl.pres. boat-acc.sg.
‘Groups of two girls are each/separately washing a boat.’
‘Two girls are washing a boat at different times/places.’

These sentences were proposed under the following three experimental conditions.

(10) Experimental conditions
   a. Collective (COLL)
   b. Participant-distributive (PartD)
   c. (Spatial)Event-distributive (EventD)

Note that, for simplicity of the experimental design, we chose to test po-sentences in spatial distributive (event) contexts, but not temporal-distributive (event) contexts.

Notice also that the test sentences in (7), (8) and (9) are almost identical, differing in: the absence vs. presence of po ((7) vs. (8)/(9)), and whether po attaches to the singular object “boat” ((8)), or the plural subject “two girls” ((9)).

Recall from our discussion above of the truth conditions of (3a) (= (8)) vs. (5a) (= (9)), that the po-object sentence in (3a)/(8) will be true under
both distributive contexts, since in both (10b-c) there is one boat per subevent (whether the two girls wash a boat individually or together). In contrast, the po-subject sentence (5a)/(9) will only be true under the distributive context in (10c) since only the latter satisfies the requirement enforced by po – that there be two girls per subevent. This contrast in the truth conditions of (8) vs. (9) yields the following additional question:

(11) Do children know that, with a plural subject and a singular object, po on the object yields both participant-distributive and event-distributive reading, while po on the subject only yields event-distributive readings?

3.1. Participants

Adults.
The group consisted of 21 native speakers (f=14, m=7, MA=18;7, SD=11.7). Most of them (N=16) were from the northern region of Serbia, within or around the city of Novi Sad. There were also four participants from Niš, a southern city of Serbia. One participant was from Sarajevo, Bosnia and Herzegovina.

Children.
We tested three age groups of monolingual Serbian-speaking children: between the ages 4;1 and 5;7 (N=24, MA=4;6, SD=0.7, 9 girls, 15 boys); between the ages 7;3 and 8;2 (N=25, MA=7;2, SD=0.4, 12 girls, 13 boys); and between the ages 9;4 and 10;6 years (N=25, MA=9;1, SD=0.3, 12 girls, 13 boys). We henceforth distinguish these three groups on the basis of (roughly) their mean age. The youngest children (5-years-olds) were recruited in the kindergarten “Maslačak” in Niš, Serbia. The older children (7- and 9-year-olds) were from the primary school “Dura Daničić” in Novi Sad, Serbia.

3.2. Design and materials

Girls played short scenes that were filmed and then compiled with a recorded female voice describing the scenes as they were going on. A program in Python was created to screen the scenes in a random order and register the answers.

The three types of sentences, illustrated in (7), (8) and (9), were
matched with the three experimental conditions—collective (COLL), participant-distributive (PartD), and event-distributive (EventD), given in (10). This 3×3 design was supplied by six different predicates (different verb-object combinations). The total number of test items was 54. The experiment also contained 12 distracters and 12 control items. The distracters were of the type “NP-pl V-intransitive”, such as “Devojke plešu” (“The girls are dancing”), six of which matched the situation played out (while the other six did not). The control items were of the type “Dve/tri devojke plešu” (“Two/three girls are dancing”) and their role was to ensure that children understood the meaning of the numeral tested. Again, six out of twelve of the control items (containing the numeral three girls) did not properly describe the situations, since all control movies involved two girls. The other six (with the numeral two girls) truthfully described the given situations. This brought the total number of items to 78.

The full experiment consisted of three blocks, each containing only one sentence type, po-less (sentence type (7)), po-object (sentence type (8)), or po-subject (sentence type (9)), and all three experimental conditions, collective (COLL) participant-distributive (PartD) and event-distributive (EventD). The control items and distracters were equally distributed over the three blocks and interspersed among the experimental items.

The total of 78 items was thus divided into three blocks of 26 items each.

3.3. Procedure

A truth-value judgment task was used. The participants were individually presented with the short scenes filmed, compiled with the oral descriptions of each scene as it was screening. The children were instructed to say whether they agreed or not with a given description of a scene. The experimenter entered 5-years-olds’ answers. 7- and 9-year-olds did the experiment autonomously by clicking (on their own) on the yes/no button. The experimenter was present but did not intervene. The children were also asked to justify their responses. These justifications were recorded.

Each child was tested on one experiment block per day, one day after the other. The order in which the blocks were presented was balanced across subjects. Each block lasted around 10-15 minutes. Short warm-up eliminatory pretests were run with the 5-year-olds to ensure that they
understood the task.

Adult participants did all three blocks of the experiment in the same day, one after the other. The order in which the three blocks were presented was balanced across adult subjects as well. The experimenter was present but did not intervene and was not able to see the movies or the adults’ answers during the sessions. The participants were encouraged to give comments on their responses at the end of the three sessions. The experimenter collected these comments manually.

3.4. Hypotheses and predictions

(12) H1: *Po*-less sentences yield both collective and participant-distributive readings.
H2: *Po* yields both participant-distributive and event-distributive but not collective readings.
H3: *Po*-object sentences allow both participant- and event-distributive readings.
H4: *Po*-subject sentences allow only event-distributive readings.

These hypotheses make the following predictions.

Starting with the collective (COLL) condition, we expect *po*-sentences to be rejected under COLL, since distributors block collective readings.

Turning to the participant distributive (PartD) condition, we expect *po*-object sentences to be accepted under this condition, since the situation depicted can be construed as two (spatially) separate washing subevents, each of which involves one boat (with the group of two girls partitioned atomically, and each girl individually washing a boat, as discussed in section 1). *Po*-less sentences should also be accepted under this condition since sentences with a plural subject and a distributive predicate are expected to allow participant-distributive readings (alongside collective readings) in the absence of *po*. In contrast, *po*-subject sentences are not expected to be accepted under this condition since each subevent depicted does not involve two girls. Finally, turning to the event-distributive (EventD) condition. We expect both *po*-object and *po*-subject sentences to be accepted under this scenario, since the situation depicted can be construed as two (spatially) separate washing events, each of which involves one boat (satisfying the truth conditions of the *po*-object sentence) and two girls (satisfying the truth conditions of the *po*-subject sentence). These predictions are summarized below.
Table 1: Predictions

<table>
<thead>
<tr>
<th></th>
<th>COLL</th>
<th>PARTD</th>
<th>EventD</th>
</tr>
</thead>
<tbody>
<tr>
<td>po-less</td>
<td>YES</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>po-object</td>
<td>NO</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>po-subject</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

3.5. Adult Results

The results for our 21 adult speakers are given below.

Table 2: Percentages of “yes” responses for adults

<table>
<thead>
<tr>
<th></th>
<th>COLL</th>
<th>PARTD</th>
<th>EventD</th>
</tr>
</thead>
<tbody>
<tr>
<td>po-less</td>
<td>99.2</td>
<td>37.3</td>
<td>27.0</td>
</tr>
<tr>
<td>po-object</td>
<td>2.4</td>
<td>93.7</td>
<td>26.2</td>
</tr>
<tr>
<td>po-subject</td>
<td>3.2</td>
<td>0.8</td>
<td>60.3</td>
</tr>
</tbody>
</table>

Performance is almost at ceiling level for po-less sentences under the collective scenario (99.2% acceptance rate). PartD and EventD are not completely rejected with po-less sentences (37.3% and 27%, respectively). The difference between the rates of acceptance on the COLL vs. PartD/EventD conditions is, however, significant ($\chi^2(2) = 155.05, p < .001$).

Turning to po-sentences. We see that they are not accepted under the collective scenario as expected, regardless of the position of po (2.4% and 3.2%). Po-object sentences are accepted at high rates (93.7%) on the PartD condition, but only at 26.2 % on the EventD condition. A Log Linear Analysis revealed that the percentage of “yes” responses for PartD differs significantly according to the sentence type, po-less vs. po-object, $\chi^2(2) = 223.91, p < .001$.

As for po-subject sentences, we see that they are (almost) totally rejected under the PartD condition, as expected. They are, moreover,
accepted under the EventD condition, though only at 60.3%. The EventD reading is also, to a certain extent, accepted with the two other sentence types, *po*-less (27%) and *po*-object (26.2%) Nevertheless, the acceptance rates with *po*-less/*po*-object sentences were significantly lower than the acceptance rates with *po*-subject sentences under EventD, $\chi^2(2) = 40.65, p < .001$.

### 3.6. Children results

Table 3: Percentages of ‘yes’ responses for 5-year-olds

<table>
<thead>
<tr>
<th></th>
<th>COLL</th>
<th>PartD</th>
<th>EventD</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>po</em>-less</td>
<td>93.8</td>
<td>95.8</td>
<td>73.6</td>
</tr>
<tr>
<td><em>po</em>-object</td>
<td>86.2</td>
<td>97.8</td>
<td>75.4</td>
</tr>
<tr>
<td><em>po</em>-subject</td>
<td>88.0</td>
<td>94.0</td>
<td>72.7</td>
</tr>
</tbody>
</table>

The 5-year-olds showed high acceptance rates on all conditions, regardless of sentence type.

With *po*-less sentences, the difference in acceptance rates was significant between EventD and COLL (and therefore between EventD and PartD), $\chi^2(1) = 19.93, p < .001$.

With *po*-object sentences, no significant difference is found between EventD and COLL (and therefore PartD), $\chi^2(1) = 4.58, p = .03$. A significant difference was found between PartD and COLL, $\chi^2(1) = 11.11, p < 0.001$. With *po*-subject sentences, the rates for EventD and COLL (and therefore PartD,) were significantly different, $\chi^2(1) = 10.21, p < .001$. The difference between PartD and COLL was not significant, $\chi^2(1) = 2.6, p = .11$. 
Table 4: Percentages of ‘yes’ responses for 7-year-olds

<table>
<thead>
<tr>
<th></th>
<th>COLL</th>
<th>PartD</th>
<th>EventD</th>
</tr>
</thead>
<tbody>
<tr>
<td>po-less</td>
<td>98.7</td>
<td>64.7</td>
<td>40.0</td>
</tr>
<tr>
<td>po-object</td>
<td>48.6</td>
<td>73.3</td>
<td>48.7</td>
</tr>
<tr>
<td>po-subject</td>
<td>64.7</td>
<td>42.7</td>
<td>68.7</td>
</tr>
</tbody>
</table>

Although 7-year olds also show relatively high acceptance for all conditions, an adult-like pattern emerges.

First, *po*-less sentences are preferred under COLL over distributive readings, especially over EventD. There was a significant difference between PartD (64.7%) and COLL (98.7%), $\chi^2(1) = 55, p < 0.001$, and also between EventD (40%) and COLL (98.7%), $\chi^2(1) = 118.66, p < .001$. Likewise, we found a significant difference between PartD (64.7%) and EventD (40%), $\chi^2(1) = 17.3, p < .001$.

Second, with *po*-object sentences, PartD (73.3%) is accepted significantly more often than EventD (48.7%) and COLL (48.6%), $\chi^2(2) = 24.81, p < .001$. Finally, with *po*-subject sentences, children were not sensitive to the distinction between EventD and COLL, $\chi^2(1) = 0.37, p = .54$. The rates of acceptance for PartD (42.7%) were significantly lower than for EventD (68.7%) and COLL (64.7%), $\chi^2(2) = 24.25, p < .001$.

Table 5: Percentages of ‘yes’ responses for 9-year olds

<table>
<thead>
<tr>
<th></th>
<th>COLL</th>
<th>PartD</th>
<th>EventD</th>
</tr>
</thead>
<tbody>
<tr>
<td>po-less</td>
<td>95.3</td>
<td>33.3</td>
<td>34.0</td>
</tr>
<tr>
<td>po-object</td>
<td>44.7</td>
<td>58.7</td>
<td>64.0</td>
</tr>
<tr>
<td>po-subject</td>
<td>66.7</td>
<td>21.3</td>
<td>71.3</td>
</tr>
</tbody>
</table>

With *po*-less sentences, the COLL (95.3%) reading is significantly preferred to PartD (33.3%) and EventD (34%), $\chi^2(2) = 153.22, p < .001$.

With *po*-object sentences, no significant difference is found between PartD (58.7%) and EventD (64%), $\chi^2(1) = 0.69, p = 0.4$. The difference
between the COLL reading (44.7%) and PartD reading (58.7%) was not found to be significant, ($\chi^2(1) = 5.33, p = .02$). There was a significant difference between COLL (44.7%) and EventD (64%), $\chi^2(1) = 10.53, p < .001$. With $po$-subject sentences, no significant difference is found between EventD (71.3%) and COLL (66.7%), ($\chi^2(1) = 0.56, p = .4$). PartD (21.3%) is significantly disliked compared to the two other readings, ($\chi^2(2) = 91.89, p < .001$).

4. Discussion

4.1. Discussion of the adult results

Recall from (12) above, repeated below, the theoretical generalizations concerning the adult grammar of $po$ we sought to experimentally probe:

(12) H1: $Po$-less sentences yield both collective and participant-distributive readings.
H2: $Po$ yields both participant and event-distributive readings, but not collective readings.
H3: $Po$-object sentences allow both participant and event-distributive readings.
H4: $Po$-subject sentences allow only event-distributive readings.

The findings reported in Table 2 validate the four hypotheses under scrutiny.

Concerning H1, we see that $po$-less sentences readily allow collective readings, but are also accepted with distributive readings, though the latter are clearly dispreferred since the difference between the rates of acceptance on the COLL vs. PartD/EventD was significant. The adults’ marginal acceptance of bare numerals in distributive scenarios corroborates the experimental findings reported in Knežević (2012) that overt distributive marking is enforced in adult Serbian. Interestingly, the acceptance rate for $po$-less sentences on PartD is actually higher here than in Knežević (2012) which tested exclusively the readings (distributive vs. collective) of $po$-less sentences: 37.3% vs. 15%, respectively. Note, however, that the material was slightly different since in Knežević (2012), static pictures were used, while in the present experiment, videos were used. On the basis of these two experiments, we conclude that $po$-less numerals in Serbian do allow participant-distributive readings, although
these readings are strongly dispreferred. This is arguably because the option of overtly marking distributivity (via po) is available in the grammar. Indeed, adults spontaneously volunteered po-sentences to describe distributive scenarios, thus corroborating the claim that adults prefer to explicitly mark distributivity with po.

Concerning H2, we see that po-sentences do not yield collective readings and are accepted under both the PartD and the EventD conditions. Moreover, depending on the position of po (i.e. whether it attaches to the subject or to the object), po-sentences give rise to different acceptance patterns, as expected. That is, po-subject sentences allow only event-distributive readings (H4), since they are rejected on participant-distributive readings (predicted) and accepted on event-distributive readings (predicted). In contrast, po-object sentences are accepted on both participant-distributive and event-distributive readings (H3). The surprising finding, however, is that, while po-object sentences are strongly accepted under participant-distributive readings (predicted), they are disliked under event-distributive readings (not predicted). Indeed, more generally, we find that although all po-sentences are predicted to be true under the event-distributive scenarios, the scores for event-distributive readings with po-sentences are not as high as expected. With po-objects, EventD readings were rejected over 70% of the time. With po-subjects, the acceptance rate was also lower than expected (60.3%).

Let us first tackle the second issue: with po-subject sentences, why are the results for the EventD condition not as high as expected, although significantly higher than with po-object sentences? We offer the following conjecture, based on the comments spontaneously volunteered by adult speakers. Recall that Serbian is a determiner-less language and that a bare noun such as brod (“boat”) in (9) can thus be construed as either a definite or (non) specific indefinite, depending on context. However, as their volunteered comments reveal, some speakers preferred to interpret the singular noun brod as referring to a specific boat, while the scenario depicted in (10c) involved two boats. By hypothesis, po enforces distribution (of the nNP to which it attaches) over different spatially and/or temporally separated subevents. Since po in (9) is on the subject “two girls”, what is being distributed is a set of two girls. Crucially, distribution over different (temporally simultaneous) spatial locations, as is the case in (10c), requires a different boat at each location. We could then impute the lower acceptance rate of the po-subject sentence in (9) on the EventD condition in (10c) to the conflict between the preference that speakers show for interpreting the bare noun “boat” in (9) as specific and the spatial event-distributive context in (10c) which requires interpreting
the bare noun “boat” in (9) as nonspecific. On this account, we crucially expect, (9), however, to be accepted on a temporal event-distributive context, such as the one suggested below.\(^5\)

(13) Temporal EventD context for the po-subject sentence (9)

We are at a sailing class during the month of August attended by eight different girls. Every Monday, two different girls must wash the same sailing boat, the beautiful ocean blue one.

As illustrated with the context in (13), distribution over different temporal locations (but at the same spatial location) is compatible with the same boat being washed by different sets of two girls. We would thus expect performance for the po-subject sentence in (9) to be at ceiling level in a temporal event-distributive context such as the one proposed in (13).

Turning to the other unexpected finding. Why were po-object sentences significantly preferred under the PartD condition (93.7\% acceptance rate) over the spatial EventD condition (26.2\% acceptance rate)? Now, the po-object test sentence in (8) is acceptable under both distributive readings, and indeed was accepted on both readings. Speakers, however, appear to strongly favor one of its readings. The question then is what is the source of this adult preference and, in particular, why is the preference for the PartD over the EventD context, and not conversely for the EventD over the PartD? Moreover, the acceptance rate under the spatial EventD context is even worse for po-object sentences (26.2\%), than for po-subject sentences (60.3\%). Again why? We suggest that the explanation just offered for the lower acceptance rate of po-subject sentences under the EventD condition can be extended to po-object sentences.

Once again, distribution (of the nNP po attaches to) is over different spatially and/or temporally separated subevents. Since po in (8) is on the object “boat”, each subevent must involve a different boat. The girls could in principle be washing a boat either together or individually, in each (temporally/spatially separated) subevent. Distribution, however, over different (temporally simultaneous) spatial locations requires different sets of two girls at each location. We could then impute the lower acceptance rate of the po-object sentences in (8) on the EventD condition in (10c) to the conflict between the preference that speakers show for interpreting the determiner-less NP “two girls” in (8) as specific and the spatial event-distributive context in (10c) which requires interpreting “two girls” in (8) as nonspecific. On this account, we expect (8) to be
strongly accepted (indeed at ceiling level) on a temporal event-distributive context, such as the one suggested below:

(14) Temporal EventD context for the po-object sentence (8)

We are at a sailing class during the month of August attended by eight children, two of which are girls.
Every Monday in August, the only two girls must wash a different sailing boat.

We have offered a uniform explanation for the lower acceptance rate of po-sentences under the Spatial EventD condition: both subject and object arguments must be interpreted as nonspecific in this context, while speakers have a preference for interpreting the po-less argument (the subject in (8), object in (9)) as specific. But this explanation still begs the question of why the acceptance rate under the Spatial EventD context is even worse for po-object sentences (26.2%), then for po-subject sentences (60.3%)? We offer the following conjecture. Serbian is a language displaying relatively free word order, where information structure plays an important role. The subject position of the sentence is canonically associated with topichood and topics are preferably interpreted as specific and as the sorting keys (see Choe (1987)). We could then impute the very low acceptance rate of the po-object sentence in (8) on the EventD condition in (10c) to the conflicting requirements that (8) must satisfy on this reading. On the one hand, the subject “two girls” is preferably interpreted as a topic/sorting key and as specific. On the other hand, the spatial event-distributive context depicted in (10c) requires interpreting the spatiotemporal argument of the verb as the topic/sorting key and the subject “two girls” as nonspecific. In other words, the acceptance rate is worse with po-object sentences than with po-subject sentences, because the former violate the preference for interpreting the subject as a topic/sorting key, as well as the preference for construing a bare NP as specific, while po-subject sentences only violate the latter preference.

We thus might expect the percentage of acceptance for the po-object sentence in (8) to go up, even on the spatial EventD condition, if the word order is altered so as to overcome the preference for interpreting the subject as a topic/sorting key, as illustrated with (15) which would have the structure in (16a), paralleling in terms of information structure, the po-subject sentence in (9), as shown in (16b).
4.2. Discussion of children results

We take the results presented in section 2 above to provide evidence for two development stages in the acquisition of distributive po.

4.2.1. Stage 1: children do not know the meaning of po and allow distributive readings without an (overt) distributor.

Recall from Table 3 that, overall, 5-year-old children accept all three sentence types across all three conditions at very high rates (ranging between roughly 72% and 90%). We conclude that 5-years-olds do not display any sensitivity to either the distributive force of po or to its syntactic position. They accept collective and distributive readings at very high rates (regardless of the presence or the absence of po), showing however a slight dispreference for EventD context (irrespective of sentence type). We assume that 5-year-olds simply do not know the meaning of po.

Notice, however, that the 5-year-olds show two target-like tendencies. They prefer po-less sentences in collective (COLL) (see Figure 1) and participant-distributive (PartD) over event-distributive (EventD) contexts (see Table 3), and they prefer po-object sentences in PartD over COLL contexts (see Table 3). Crucially, this is the expected pattern since po-less sentences with a plural subject and a distributive predicate should indeed be acceptable under either a collective or a participant-distributive reading. The only difference with adults is that the latter require distributivity to be overtly marked with po, while children do not. We thus take these two tendencies to show that children at this stage allow distributive readings but, unlike adults, do not require an overt distributor to license such readings.
In sum, at this stage, children ignore po and take numerals to yield both collective and distributive readings. In this respect, they pattern like English children, as reported in Musolino (2009), who tested the interpretation of sentences with two numerals (in subject and object positions). The data revealed that English 5-year-olds have access to both (participant) distributive and collective readings of numerals (just like English adults).

4.2.2. Stage 2: children know that po marks the distributive share and distributes over events, and prefer marking distributivity overtly with po

The graph below shows the acceptance rates across age groups, including adults (for comparison) for po-subject sentences under the two distributive conditions: PartD vs. EventD.
We see that *po*-subject sentences are accepted roughly 70 % of the time on the EventD condition by 7- and 9-year-old children. Indeed, both 7- and 9-year-old children do better than adults on this condition (the difference between the 9-year-olds and the adults was found to be significant: $\chi^2(3) = 5.68, p < .001$, though not the one between the 7-year-olds and the adults). More importantly, we also see that *po*-subject sentences are rejected on the PartD condition roughly 80% of the time by 9-year-olds and 60% of the time by 7-year-olds (the difference between 7- and 9-year-olds was found significant, $(\chi^2(1) = 14.72, p < .001)$, as was the difference between 9- (and thus 7-) year-olds and adults $(\chi^2(1) = 25.53, p < .001)$.

Recall that the semantics of *po* merely requires a plurality of events involving *po*-numeral NP. In (9), since *po* combines with the *nNP* “two girls”, there must be at least two subevents each involving two girls washing a boat. As a consequence, (9) will be true under the spatial event-distributive context in (10c), but crucially not under the participant-distributive context in (10b) since only the former satisfies the requirement enforced by *po* - that there be two girls per subevent. That 7- to 9-year-old children reject (9) in the PartD context in (10b) at significant rates, while accepting it under the EventD context in (10c), tells us that they know that *po* enforces distribution over events and that the argument to which *po* attaches serves as the distributive share.
Why, however, do children accept (9) under the EventD context significantly more than adults do? Recall the explanation we offered for this unexpected finding with adults: under the Spatial EventD condition both subject and object arguments must be interpreted as nonspecific, but adult speakers have a preference for interpreting the \textit{po}-less argument (the bare singular noun ‘boat’ in (9)) as specific. This preference is not as strong with children and this is why they accept (9) more than adults under the Spatial EventD condition.

Let us now turn to \textit{po}-object sentences. The graph below shows the acceptance rates across age groups (including adults) for this sentence type under the two distributive conditions: PartD vs. EventD.

![Figure 3. % of “yes” responses across age groups with \textit{po}-object sentence under PartD & EventD](image)

Recall that our prediction for \textit{po}-object sentences was that they should be accepted under both distributive conditions, but adults, however, significantly preferred \textit{po}-object sentences under the PartD condition over the spatial EventD condition. In contrast, 7- and 9-year-olds score significantly higher than adults on both readings. In particular, under the PartD condition, while no significant difference was found between 7- and 9-year-olds, ($\chi^2(1) = 6.55$, $p = .01$), the difference between 9-year-olds and adults was significant ($\chi^2(1) = 42.45$, $p < .001$). Likewise, on the EventD condition, no significant difference was found between 7- and 9-
year-olds, although both groups scored significantly higher than adults \( (\chi^2(1) = 37.82, p < .001 \) for 9-year olds; \( (\chi^2(1) = 13.69, p < .001 \) for 7-year-olds).

We imputed the adult’s very low scores on the EventD condition to the conflicting requirements that \textit{po}-object sentences must satisfy on this reading: under the spatial EventD condition both subject and object arguments must be interpreted as nonspecific, but this violates the adult preference for interpreting the \((\text{po-less})\) subject NP as a specific topic. Again, we conclude that this preference is not strong with 7- and 9-year-olds and this is why they accept (8) more than adults do under the spatial EventD condition.

Putting all these results together, two important conclusions emerge. First, the children’s performance on \textit{po}-subject sentences shows that they know that \textit{po} enforces distribution of the \( n \) NP with which it combines over events, since they reject \textit{po}-subject sentences under the PartD context. Second, the older children accept significantly more often than adults both \textit{po}-object and \textit{po}-subject sentences under the EventD condition. We take this to show that they have not yet acquired the preference for interpreting the \textit{po}-less NP as specific.

Finally, turning to \textit{po-less} sentences, comparing the percentage of \textit{yes}-responses for this sentence type under all three conditions and across all age groups, we see that while 5-year-olds accept \textit{po-less} sentences in the two distributive contexts over 70\% of the time, 7-year-olds start showing an adult pattern since they reject \textit{po-less} sentences under the EventD condition 60\% of the time (though significantly less under the PartD condition). By the age of 9, children have the adult preference for overtly marking distributivity with \textit{po} since they reject \textit{po-less} sentences under both distributive conditions over 65\% of the time (roughly like adults).

5. Conclusion

The goal of our experimental investigation of \textit{po}-numerals was twofold. As regards the adult grammar, we sought to experimentally validate Knežević’s (2015) analysis of the distributive-share marker \textit{po} as a marker of event plurality enforcing distribution of the \( n \) NP it combines with over (at least two) different (sub)events. Both so-called participant- and event-distributive readings arise via the plurality of events (involving \textit{po}-numeral NPs) that \textit{po} enforces. The findings reported here experimentally
validate this analysis, since adults consistently accepted po-sentences only in contexts where the po-nominal NP was appropriately distributed over (sub)events. That is, adults rejected almost at ceiling level (99.2% of the time) po-subject sentences under the PartD condition in (10b) where the subject NP was not distributed.

What was surprising, however, is that although po-sentences (be it po-object or po-subject) were accepted only in contexts where the po-argument was distributed, they were not accept across the board, in all contexts where the po-nominal NP was distributed, since the scores for spatial event-distributive readings with po-sentences were not as high as expected. We contend that this surprising pattern is due to a conflict between an adult preference for interpreting the po-less argument as specific and the spatial EventD context, which requires interpreting both arguments of the verb as non-specific. This conjecture makes strong predictions that can be falsified – namely, that both sentence types in (8) and (9) should be accepted at ceiling level under the temporal EventD contexts proposed in (13) and (14), respectively.

As regards the child grammar, we identified two developmental stages in the acquisition of po. Children in the youngest age group (5-year-olds) do not know the meaning of po and allow distributive readings of numerals without an overt distributor. Older children know that po marks the distributive share and distributes over events and prefer marking distributivity overtly with po.

Notes

1 Recall that bare po-less nominals allow both definite and indefinite construals, but that a po-nominal (the distributive share) can only be construed as non-specific. Consequently, since po in (5a) attaches to the subject “two girls”, then different (non-specific) groups of two girls must be distributed over at least two boat washing events. In contrast, the po-less object “boat” in (5a) could, in principle, be construed as either definite or indefinite. When distribution is over temporally separated events, the boat involved in each washing subevent could thus be the same or a different. However, when distribution is over spatially separated simultaneous events, as is the case on the scenario depicted in (4b), the boats involved have to be different. This constraint, as the reader shall see, will impact our experimental findings.

2 For further ongoing experimental research on Serbian po see Bosnić, Knežević, Spenader, Demirdache & Hollebrandse (2016) and Knežević, Bosnić, Spenader, Demirdache & Hollebrandse (2016).
This is an important qualification, as should be clear from the previous discussion in section 1. If the po-less object were a plural term, a so-called participant-distributive reading would be also available for (5)—e.g. “Po-two girls wash two boats” would be true under a PartD scenario involving two boats (the theme participants in the described event), where each boat is being washed by a different set of two girls (the distributive share).

For this type of study, the usual confidence level for statistic tests (such as Chi-Square) is 95%, and the result is considered as significant if $p < 0.05$. However, given that we have performed multiple comparisons, we applied Bonferroni correction. To do so, we divide the risk level (0.05) by the number of possible comparisons across treatments (or conditions) and age groups. In our experiment, this is equivalent to taking a risk level of 0.001. Therefore, results are considered significant when $p < 0.001$.

See Bosnić et al. (2016) for further interesting experimental findings and discussion of spatial distributive readings with po-numerals.

The astute reader will have noticed that the preference for interpreting the subject/topic position as mapping the sorting key should not be relevant under Kneževič’s (2015) proposal where distribution is always over events. We leave this question open pending further investigation of po-numerals in sentences with alternative word orders/information structure, such as (15)/(16). Competition between po-subject and po-object sentences on the EventD condition could also be playing a role here since the former are only acceptable under the EventD condition, while the latter are acceptable under both distributive contexts.

É. Kiss, Gerőcs and Zétényi (2013) report that Hungarian 6-year-olds also allow both distributive and collective readings with numerals in sentences with the distributive particle is, which blocks collective readings in the adult grammar just like po. Cross-linguistic data suggest that distributive interpretations are early but not mandatory in child language.

Acknowledgement

We are very thankful to all children and adults that voluntarily participated in the experiment, as well as the schoolteachers. We are indebted to Ana Bosnić, as well as Boban Arsenijević, Ken Drozd, Bart Hollebrandse, Patricia Cabredo Hofherr, Angeliek van Hout, Jennifer Spenader, Brenda Laca, Orin Percus, and Irina Sekerina for discussion and precious comments.
References


THE FIRST STEPS IN THE ACQUISITION OF SYLLABLES AND GEMINATES IN TASHLHIYT BERBER

MOHAMED LAHROUCHI

1. Introduction

In this paper, we present preliminary results of a longitudinal study that focuses on the developmental trajectory of speech production capacities in two Berber children acquiring Tashlhiyt from the babbling period to the emergence of early grammar. Very few studies, if any, are devoted to the Berber language acquisition, and Tashlhiyt, the variety spoken in South-west Morocco, presents very interesting phonetic and phonological characteristics to study in a developmental perspective. Probably the most salient characteristic is the use of complex consonants clusters, resulting in a highly marked syllable structure where any segment, even a voiceless obstruent, may occur in the nucleus position (cf. Dell & Elmedlaoui 1985, 2002). Another feature, which deserves to be studied relates to consonant length. Tashlhiyt Berber contrasts singleton and geminated consonants in various contexts, including word-initial, medial and final positions. Geminates raise many issues with regard to their syllabic representation, especially when they occur in contexts where ambisylllabicity is required.

Traditionally speaking, an ambisyllabic consonant is a single intervocalic consonant which belongs to two syllables at the same time. It is often used by phonologists to account for allophonic variation (“flapping”) of coronal alveolar plosives in English (see Kahn 1976; Gussenhoven 1986) and the distribution of short and long vowels in German (see Becker 1998; Vennemann 1990). The example in (1) illustrates the ambisyllabic character of flapped /t/ in English atom.
Using a standard syllable structure, intervocalic /t/ occupies only one x-slot but attaches simultaneously to the coda of the first syllable and to the onset of the second syllable. Interestingly, the timing tier (also called skeleton) allows distinguishing an ambisyllabic consonant from a geminate. The latter requires two x-slots, one dominated by coda and the other by onset. According to Dell & Elmedlaoui (2002, 2011), this is exactly the configuration that an intervocalic geminate displays in Tashlhiyt Berber (e.g. *ittu* “he forgot”). Their syllabic model further allows a geminate to form a complex coda (e.g. *a33* “let”, *uff* “inflate”) or onset (e.g. *ff* “pour”, *ddu* “go”). It can also branch into nucleus + onset positions (e.g. *islla* “he heard”, *ignna* “sky”), but never into onset + nucleus, relying on the idea that onsets are weightless (cf. Hayes 1989).

However, it is not clear how geminates are syllabified when they occur in isolation such as in *kk* “cross” and *ff* “eat”. According to Dell & Elmedlaoui’s syllable algorithm, such forms would be normally analysed as closed syllables since they eschew prohibited branching into onset + nucleus positions. Why don’t they simply form a light syllable? What kind of evidence can be marshalled on each possibilities? These questions in relation with the ambisyllabicity issue are given special attention in this paper.

The paper is structured as follows. Section 1 provides an overview on the phonology of geminates. Section 2 presents the Tashlhiyt Berber language. Data and methods are described in section 3. In section 4, we present the preliminary results of a longitudinal two-case study, focusing on syllable types and geminates. The phonological representation of geminates will be discussed therein. We will be testing the “ambisyllabicity” hypothesis in relation with the assumption that any segment in Tashlhiyt Berber may function as a syllable nucleus. Section 5 presents an alternative to standard syllabic analysis, using a strict CV approach. Section 6 concludes the paper.
2. The phonology of geminates, an overview

Geminates have been the source of much debate in phonological theory, with regard to their representation and their behaviour in phonological processes (cf. Hayes 1986, 1989; Schein & Steriade 1986; Selkirk 1990; Kenstowicz 1994; Davis 1994; 1999, 2003, 2011). Cross-linguistic studies state interesting implicational relationships, whereby a language having word-initial and word-final geminates also has word-medial ones (cf. Taylor 1985; Thurgood 1993; Kraehenmann 2011). The general finding is that word-medial, specifically intervocalic, is the preferred position for this kind of segments. Among the languages with phonemic geminates, very few contrast them with singleton consonants in word-initial position. Tashlhiyt Berber is one such language in which lexical geminates occur not only in medial position (e.g. agːu “smoke”, urːi “turn back”) but also in initial and final positions (e.g. dːu “go”, fːi “pour”, al: “raise”, uf: “inflate”). In addition to lexical geminates, Tashlhiyt has morphological (e.g. lkm “arrive” / lkːm “arrive-imperfective”), concatenated (e.g. /t-lkm-m/ > [tlkm:] “you (2.MS.PL) arrived”) and assimilated geminates (e.g. /rad tftu/ > [ratːftu] “she will go”).

Taking advantage of autosegmental representations, most phonologists (cf. Guerssel 1977; Dell & Elmedlaoui 1997, 2002, 2011; Lahrouchi 2001; Ridouane 2010) analyse geminates in Berber as a single melodic unit associated to two adjacent skeletal slots, as opposed to singleton consonants which attach to only one slot. This is illustrated below in (2).

(2) a. b.
    x x x
    | \\ /
    C C

This kind of representation inevitably raises the issue of syllabification, leading to cases where the two skeletal positions in (2b) are dominated either by one or by two syllabic constituents. The later case is found particularly in the intervocalic position, where geminates are ambisyllabic, attached to coda and onset positions. Dell & Elmedlaoui’s (1985, 2002, 2011) hypothesis that any segment in Tashlhiyt Berber can be syllabic allows this configuration (see (3a)) as well as the one where geminates branch into nucleus + onset positions (see (3b)). However, it prohibits any branching into onset + nucleus positions, relying on Hayes (1989, 258) idea that onset segments are weightless and that branching
into onset and nucleus positions entails a “flopped structure”.

(3) O, R, N and C respectively stand for Onset, Rime, Nucleus and Coda.

\[
\begin{array}{ll}
\text{a. } \text{idːa “he went”} & \text{b. } \text{islːa “he heard”} \\
\sigma & \sigma \\
R & R & R & R \\
O & O & O & O \\
N & N & N & N \\
x & x & x & x & x & x & x & x \\
i & d & a & i & s & l & a
\end{array}
\]

We will return to this issue in section 4. We will present the frequency of geminates in children productions. Then we discuss some of the problems they raise with regard to syllable structure and so-called “ambisyllabicity”. Meanwhile, we provide an overview of the phonemic system of Tashlhiyt Berber and its syllable structure. We also present the participants in our study as well as the methods.

3. Tashlhiyt Berber: phonemes and syllables

3.1. Phonemic system

Berber belongs to the Afroasiatic family. It is spoken in large parts of North Africa, mainly in Morocco and Algeria, and to a lesser extent in Mali, Niger, Libya, Egypt, Tunisia and Burkina Faso. Berber communities also live in Diasporas, mainly in France, Spain, Netherlands, and Belgium. Tashlhiyt, whose native speakers are estimated at four millions based on the data of the official census of 2004, is spoken in South-west Morocco.

Regarding the phonemic system, Tashlhiyt Berber has only three vowels: /i/, /a/ and /u/. The so-called “transitional vocoids”, which appear in certain consonant clusters (e.g. nkar “wake up, stand up”, lkam “arrive”), have no syllabic status according to Dell & Elmedlaoui (2002), while Coleman (1996, 2001) argues that they are epenthetic, filling syllabic nuclei that would otherwise remain empty. As to consonants,
Tashlhiyt has 33 consonants and two glides /j, w/, which often occur in complementary distribution with the corresponding high vowels /i/ and /u/ (cf. Guerssel 1986; Lahrouchi 2013). The consonantal inventory is given in table (1).

Table 1: The consonantal inventory of Tashlhiyt Berber.

<table>
<thead>
<tr>
<th></th>
<th>Labial</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Uvular</th>
<th>Pharyngeal</th>
<th>Laryngeal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td>d’</td>
<td>k</td>
<td>g</td>
<td>q</td>
<td>q”</td>
</tr>
<tr>
<td>Fricative</td>
<td>f</td>
<td>s</td>
<td>z</td>
<td>l</td>
<td>3̆</td>
<td>3’̆</td>
<td>3”̆</td>
<td>h</td>
</tr>
<tr>
<td>Trill</td>
<td>r</td>
<td>r’</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td>m</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td></td>
<td>n</td>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>w</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each of the above consonants has a geminate counterpart. Table 1 also shows pharyngealized coronals and labialized dorsals, as opposed to their plain counterparts. All of these phonological properties prove interesting in the study of language acquisition. In the remainder of this paper, we will be focusing on geminates in relation with syllable structure.

### 3.2. Syllable structure

One of the most interesting features of Tashlhiyt Berber is that it allows complex consonant clusters, which may result in utterances without any vocalic segment (e.g. *ts:lkmt:nt* “you made them (fem.) arrive”). Based on this characteristic, Dell & Elmedlaoui (1985, 1988, 2002) argued that in Tashlhiyt any segment, even a voiceless obstruent, can be syllabic if it is the most sonorous segment in the string (see also Boukous 1987). In order to regulate the competition of segments for the status of syllable nucleus, the authors (2002, 76) use the following sonority scale where segments appear in order of decreasing sonority: *a > high vocoids > liquids > nasals > fricatives > stops*. For instance, in *krf* “tie” *r* is the
syllable nucleus because it is more sonorous than \( k \) and \( f \). In addition to sonority requirements, the authors use a set of constraints which determine the organization of segments within the syllable. Among these, the prohibition against onset-less syllables (except domain-initially) and the avoidance of complex onsets and codas (unless they contain a geminate). In (4), we list the constraints proposed by Dell & Elmedlaoui (2002), some of which have been slightly modified for the purpose of the present study.

(4)
- Complex onsets are prohibited (Dell & Elmedlaoui 2002, 114). This constraint has been slightly modified in this paper, suggesting that only domain-initial onsets can be complex.
- Any rime contains at most three consonants. If so, the last two consonantal slots host a geminate (Dell & Elmedlaoui 2002, 98).
- The coda position cannot be more sonorous than the nucleus (Dell & Elmedlaoui 2002, 102).
- Every syllable has an onset, except domain-initially where the onset may be empty (Dell & Elmedlaoui 2002, 92).
- Any sequence with a sonority peak must contain a nucleus (Dell & Elmedlaoui 2002, 100).
- A geminate cannot branch into an onset followed by a nucleus (Dell & Elmedlaoui 2002, 102). However, in cases where none of the above constraints leads to the appropriate syllabification such as in words made of one geminate (e.g. \( f \): “eat”) as well as in domain-initial position (e.g. \( k:int \) “cross them (fem.)”), the geminate seems to constitute an onset-nucleus sequence, inevitably running counter Dell & Elmedlaoui’s proposal.

4. Data and methods

Two children acquiring Tashliyi Berber participated in our study: a girl named Imane and a boy named Reda. They live in a village named Azrarag, located twenty kilometres northeast Agadir. At the time we collected the data, they were exposed only to one language; their parents speak Tashliyi at home. Moroccan Arabic is generally acquired later, at school and in urban areas. Both kids were recorded every two weeks in their home from 7 to 24 months (the whole data is downloadable from CHILDES\textsuperscript{2}). Parents were told to follow their normal types of activities with their child. No extra material was used in the recording sessions, so
that samples reflected the children’s typical vocalizations in familiar surroundings. We ended up with 31 sessions of one hour for Reda and 26 for Imane. The data were broadly transcribed using IPA. They were entered into the PHON software designed for describing phonetic patterns (Rose et al. 2006). Then, they were divided in two periods: period 1 from 7 to 12 months and period 2 from 13 to 24 months. Only babbling utterances were considered in the first period and word utterances in the second period. Frequency of segments and syllable types were calculated for both periods.

For the purpose of our study, we distinguished open syllables from closed ones. The consonant clusters, which appear later in both kids productions were syllabified using Dell & Elmedlaoui’s model of syllabification (2002), as presented in the previous section.

5. Results

This section presents the results of our study on syllable types and geminates in Imane and Reda’s productions. The reader is referred to Lahrouchi & Kern (2015) for further results and analysis.

5.1. Ratio of consonants to vowels

Table 2 shows the frequency of segments and utterances.

Table 2: Frequency of occurrence of segments and utterances.

<table>
<thead>
<tr>
<th></th>
<th>Utterances</th>
<th>Consonants</th>
<th>Vowels</th>
<th>C/V ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imane</td>
<td>3339</td>
<td>8924</td>
<td>7261</td>
<td>1,23</td>
</tr>
<tr>
<td>Reda</td>
<td>2986</td>
<td>5537</td>
<td>5823</td>
<td>0,95</td>
</tr>
<tr>
<td>Total</td>
<td>6325</td>
<td>14461</td>
<td>13084</td>
<td>1,09</td>
</tr>
</tbody>
</table>

As illustrated by the C/V ratio, Imane produced more consonants than vowels while Reda produced the reverse pattern. This is relatively different from patterns observed in other languages such as French, Romanian, Turkish, Dutch and Tunisian Arabic. According to Kern & Davis (2009), the number of vowels exceeded consonants in these languages: the C/V of the participants was 0.78 in French, 0.80 in Romanian, 0.83 in Turkish, 0.85 in Dutch, and 0.88 in Tunisian Arabic.
The ratio obtained in Table 2 is probably due to the ambient language effect, namely the fact that Tashlhiyt Berber has common complex consonant clusters.

5.2. Syllable types

Open light syllables have been generally reported as the most predominant types in children productions, often reduplicated or resulting from consonant cluster reduction and coda deletion, especially in first words (Fikkert 1994; Levelt & Van de Vijver 1998; de Boysson-Bardies 1999; Levelt, Schiller & Levelt 2000). Closed syllables are very limited in children babbling and first words, regardless of the input language. However, in certain languages, children tend to produce a relatively high number of closed syllables. According to Fee et al. (1982) and Stoel-Gammon & Cooper (1984), some English-learning children had dominant closed syllables. In Dutch, one child followed by Elbers & Ton (1985) is reported to have no preference for CV syllables in his first words. Another interesting case can be found in languages with phonemic geminates. In intervocalic position, these geminates are generally analysed as ambisyllabic, branching into coda and onset positions. This automatically increases the number of closed syllables as soon as the geminates emerge in children productions. Tashlhiyt Berber is one such language where geminates are produced early in the babbling and first words periods.

Our results show that open syllables and more precisely CV ones were predominant in both kids productions during the whole period of study. Out of 3000 syllables produced by Imane and Reda in the babbling period, 85% were open. In the first words period, Imane produced 2370 syllables, 73% of which were open, whereas Reda had 82% of the syllables open (2082 syllables in total). The results are given in tables 3 and 4 (underlined consonants are syllabic).

Table 3: Syllable types in Reda’s productions

<table>
<thead>
<tr>
<th></th>
<th>CV</th>
<th>CC</th>
<th>V</th>
<th>C</th>
<th>CCV</th>
<th>CCC</th>
<th>VC</th>
<th>CC</th>
<th>CVC</th>
<th>CCC</th>
<th>VCC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babbling period</td>
<td>59,78</td>
<td>0,05</td>
<td>21,85</td>
<td>2,62</td>
<td>0,92</td>
<td>0</td>
<td>4,20</td>
<td>0,81</td>
<td>8,85</td>
<td>0,32</td>
<td>0,54</td>
</tr>
<tr>
<td>Words period</td>
<td>57,10</td>
<td>0,28</td>
<td>22,04</td>
<td>2,01</td>
<td>1,15</td>
<td>0,04</td>
<td>5,42</td>
<td>0,09</td>
<td>11,43</td>
<td>0</td>
<td>0,38</td>
</tr>
</tbody>
</table>
Table 4: Syllable types in Imane’s productions

<table>
<thead>
<tr>
<th></th>
<th>Open</th>
<th></th>
<th></th>
<th></th>
<th>Closed</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CV</td>
<td>CC</td>
<td>V</td>
<td>C</td>
<td>CCV</td>
<td>CCC</td>
<td>VC</td>
<td>CC</td>
<td>CVC</td>
<td>CCC</td>
<td>VCC</td>
</tr>
<tr>
<td>Babbling period</td>
<td>75,77</td>
<td>0,11</td>
<td>8,27</td>
<td>0,23</td>
<td>0,11</td>
<td>0</td>
<td>1,55</td>
<td>0,23</td>
<td>13,54</td>
<td>0,11</td>
<td>0</td>
</tr>
<tr>
<td>Words period</td>
<td>62,78</td>
<td>1,09</td>
<td>8,52</td>
<td>0,12</td>
<td>0,88</td>
<td>0</td>
<td>4,30</td>
<td>3,37</td>
<td>17,80</td>
<td>0,80</td>
<td>0,29</td>
</tr>
</tbody>
</table>

Among open syllables, CV ones are much more frequent in both children productions. CC syllables are far fewer though phonologically speaking they are equivalent to CV syllables in that they contain an onset plus a nucleus, both of which are simplex. The relatively high percentage of closed syllables (in particular CVC) in both children productions is due to the frequency of geminates, which are syllabified as amabisyllabic, typically branching into coda and onset positions. The following section is dedicated to this type of segments.

5.3. Geminates

Both children started producing geminates early in their babbling. Reda produced 1725 geminates during the whole period of study, 53% of which involve coronal consonants, followed by labials (26%), dorsals (13%) and gutturals (8%). Imane produced almost the same amount of geminates (1677), with the same preferences: 54% of the geminates are coronals, followed by labials (29%), dorsals (11%) and gutturals (6%). Figures 5 and 6 present the results for both kids.
Both children preferred geminates in the initial and medial positions. Very few occurred in the final position. Examples in first words are given in (5) for both kids.
Cross-linguistic studies (cf. Taylor 1985; Thurgood 1993; Kraehenmann 2011) in adult language state that geminates most commonly occur in word-medial position, particularly in the intervocalic position. Word-final geminates are also often allowed, while word-initial ones are relatively rare. The low frequency of word-final geminates in children productions can be explained by the late acquisition of consonants in the coda position, as opposed to the onset (cf. Ingram 1978; Fikkert 1994; Demuth 1995; Kirk & Demuth 2006; Gnanadesikan 2004, among others) and the preference for unmarked structures (cf. Jakobson 1968), namely open syllables.

As mentioned earlier in section 2.2, Dell & Elemdlaoui’s (2002) approach to syllable structure in Tashlhiyt Berber allows:

- complex coda but no complex onset
- geminates in the coda position but not in the onset position (except domain-initially). Any geminate that can potentially occur in the onset position is split into an ambisyllabic structure coda + onset or nucleus + onset
- geminates as a nucleus + onset sequence but not the mirror-image onset + nucleus.

In this study, we assume that complex onsets may occur domain-initially for it is difficult to show that forms like kti “remember” and bdu “start” involve an initial syllabic consonant. This is even more difficult in
forms containing sonority-equal segments such as bdg “be wet” and fsx “invalidate”, since sonority hierarchy cannot determine which consonant is syllabic. In the same way, we assume that domain-initial geminates belong to the same syllabic constituent, namely the onset: forms in (5) like tːɛ, bːi and mːæ are analysed as monosyllabic. However, in post-vocalic position geminates are syllabified either as tautosyllabic, associated to the coda position, or as ambisyllabic when followed by a syllabic segment. The latter case is well attested in both child and adult productions. It is also stated in typological studies as the preferred context for ambisyllabic consonants (see section 1 for relevant studies). As to ambisyllabic geminates attached to nucleus and onset positions, they are found only in Tashlhiyt Berber. As far as this assumption holds, the geminates in forms like ħlli and ʕmːi in (5) are syllabified exactly as in (3b). For the sake of convenience, both forms are represented in (6).

(6)  a. ħlli “milk”      b. ʕmːi “my uncle”

\[
\begin{align*}
\sigma & \quad \sigma \\
O & \quad R \quad O \quad R \\
N & \quad N \quad N \quad N \\
x & \quad x \quad x \quad x \quad x \\
\text{ħ} & \quad l & \quad i & \quad \text{ʕ} & \quad m & \quad i
\end{align*}
\]

In the competition for the status of syllabic peak, the liquid /l/ and the nasal /m/ have priority over the less sonorous fricatives /ħ/ and /ʕ/, resulting in bisyllabic forms where the ambisyllabic character of the geminate is reflected in its association to nucleus and onset positions. In case the first consonant is more sonorous than the following geminate, the resulting structures involve more common ambisyllabic geminates, interposed between two nuclei. The forms xtːa and ftːæ illustrate the situation in (7).
Ambisyllabicity therefore holds insofar as it involves a coda + onset or nucleus + onset sequence. One can still ask how relevant is the assumption that a language may have geminates attaching to a nucleus + onset sequence but not their mirror-image, onset + nucleus. The ban of the latter structure in syllable theory generally relies on the argument that onsets do not contribute to weight, and more particularly the avoidance of the so-called “flopped structure” (cf. Hayes 1989, 258, and Topinzi 2008 for an alternative view). However, as far as Tashlhiyt Berber is concerned, there is no phonological evidence for syllable weight, apart from metrics and verse structure which has been argued to rely on fixed alternations of light and heavy syllables (cf. Dell & Elmedlaoui 2002; Dell 2011; the reader is referred to Hammane 2010 for an alternative analysis without any distinction between light and heavy syllables). Furthermore, we are left with forms such as ʃː “eat” (see 5) and kː “cross” which, according to Dell & Elmedloui’s approach, should be analysed as closed syllables only because they eschew prohibited branching into an onset + nucleus structure. This is all the more questionable in child phonology that closed syllables are generally favoured over open ones. Tashlhiyt Berber is no exception to this trend: During the whole period of study, Reda produced 5537 consonants, of which only 909 occurred in the coda position. Likewise, Imane produced 8924, 22.24% of which are coda consonants. The relatively important percentage of coda consonants is due the high frequency of ambisyllabic consonants in their productions. Of the geminates produced by Imane, 90% occurred in the onset position. Only 8% occurred in the coda and 2% in the nucleus. Reda exhibited relatively different patterns: 25% of the geminates he produced occurred in the coda position, as opposed to 73% in the onset. Less than 2% occurred in the nucleus position.

The issues that geminates raise with regard to syllabification in Tashlhiyt may be viewed as by-product of analysis. Dell & Elmedlaoui’s
syllable model, designed to account for complex consonant clusters, allows geminates to display various syllable configurations, of which ambisyllabicity is the most intricate. Based on standard syllable structure, the model allows geminates to branch not only into coda+onset, but also in nucleus+onset. Furthermore, they can constitute a branching onset (domain-initially) as well as a complex coda. These multiples syllables types, some of which are hardly attested cross-linguistically, can be accounted for if we assume that surface complexity results from strict alternation of light syllables in the underlying representation.

The next section presents an alternative analysis within the strict CV approach to syllable structure, also known as CVCV model.

6. An alternative approach

6.1. CVCV model

The CVCV approach (Lowenstamm 1996), which falls within the framework of Government Phonology (Kaye et al. 1990), holds that syllable constituency universally reduces to strict alternations of non-branching onsets and non-branching nuclei, i.e. C positions and V positions, which interact laterally to various syllable types in the surface form. Within this model, a geminate uses the same skeletal material as a long vowel, a closed syllable and branching onset, namely two CV units.

(8)

<table>
<thead>
<tr>
<th>a. geminate</th>
<th>b. branching nucleus</th>
<th>c. branching onset</th>
<th>d. closed syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td>bat</td>
<td>[bra]</td>
<td>[bra]</td>
<td>[bra]</td>
</tr>
</tbody>
</table>

The way segments are associated to these two CV units leads to different surface forms. That is, the geminate in (8a) and the branching onset in (8c) contain an empty V position, while the long vowel in (8b) has an empty C position. The consonant in the coda position in (8d) appears in the onset position of the second syllable whose nucleus is empty.

Skeletal positions which have no phonetic realization are said to be licensed to remain empty by virtue of the government relations they share.
with the neighbouring segments. Proper Government (PG) is one such relation which allows a vocalic position to remain empty if it is followed by a vowel. For instance, the V position between /b/ and /r/ in (8c) remains empty because it is governed by the following vowel. In addition, it is assumed following Kaye (1990, 314) that domain-final V positions are parametrically licensed to remain empty despite being ungoverned. Proper Government proves particularly interesting in accounting for the distribution of epenthetic vowel and the well-known V/Ø alternation found in many languages, including French, Somali, Maghrebi Arabic and Berber varieties other than Tashlhiyt. The examples represented in (9) illustrate the situation in Magherbi Arabic and Kabyle Berber.

(9) Maghrebi Arabic  Kabyle Berber

(a) *ktāb ‘he wrote’

(b) *kētbu ‘they wrote’

(c) *xādm ‘work—imperative 2 sg’

(d) *xādmat ‘work—imperative 2 pl.’

In (9a), the V position between /k/ and /t/, properly governed by the schwa that appears between /t/ and /b/, remains empty. The same position, non-governed, non-governed in (9b) surfaces as schwa. The same situation obtains in Kabyle: the vocalic position between /x/ and /ð/ remains empty in (9c) but not in (9d). Interestingly, this position is properly governed by the following schwa in the first form but not in the second one (see Bendjaballah 2001 : 188). In Tashlhiyt Berber where no epenthetic schwas are found, syllabic consonants arise. Any vocalic position lacking proper government systematically hosts one syllabic consonant (see also Hammane 2010; and Blaho 2004 and Scheer 2008 on Slavic languages). To illustrate this, let us consider the verbs *xādīm and *xādmat just discussed in Kabyle (9c,d).

In Tashlhiyt, no epenthetic vowel appears in the surface form of these verbs. Accordingly, the consonant /m/ fills the V position to its left and then governs the empty V between /x/ and /d/ (10a), whereas in *xdmat it is the consonant /d/ that branches into the preceding ungoverned V position
(10b).

(10) *Tashlhiyt Berber*

a. *xdm* “work – imperative 2 sg”

```
PG
x  d  m
|   |   |
C  V1  C  V2  C  V3
```

b. *xdmat* “work – imperative 2 pl”

```
PG
x  d  m  a  t
|   |   |   |
C  V1  C  V2  C  V3  C  V4
```

The distribution of epenthetic vowels in Kabyle supports the left-branching representation of syllabic consonants in Tashlhiyt. Indeed, we notice that the syllabic consonants in (10a,b) are exactly the ones preceded by an epenthetic schwa in (9c,d). Further evidence is found in genetically unrelated languages such as German where many instances of `aC` are in complementary distribution with syllabic consonants: e.g. *haːbən / haːbm* “to have”, *daŋkən / daŋk̥* “to thank” (see Clark & Yallop 1995, 68).

In addition to phenomena just discussed, the CVCV model allows unifying the representation of geminates and accounts for the frequency of surface syllable types in children productions. This is shown in the following section.

### 6.2. Syllables and geminates

Within CVCV approach to syllable structure, there is no need for ambisyllabicity. The configurations Dell & Elmedlaoui (2002, 2011) assign to geminates are reanalysed as associated to two CV units. The forms in (7), which according to the authors model contain coda+onset and nucleus+onset geminates, are reinterpreted as in (11).
Proper Government regulates the distribution of empty V positions in these forms. In (11a,b), the vowel /a/ properly governs the empty V position interposed between the members of the geminate. The preceding V position, ungoverned, hosts syllabic /x/ in (11a) and /f/ in (11b). The right-branching of syllabic consonants is explained by the impossibility for the geminate to spread into the neighbouring V position. Such an option would indeed result into prohibited ternary branching (the reader is referred to Hammane (2010) for a similar proposal on the representation of syllabic consonants).

The CVCV approach not only allows unifying the representation of geminates, but also accounts for the syllabic patterns observed in children productions, including those presented in this study. As mentioned previously in section 4.1, core syllables, i.e. CV, were predominant in both kids productions, as opposed to closed syllables. This has been generally explained by the late acquisition of the coda position. To give just one example, Smolensky (1996) explains within the Optimality Theoretic framework (Prince & Smolensky 1993) that a child producing [kæ] for cat uses in his or her grammar the markedness constraint NOCODA which outranks the faithfulness constraint MAX(seg). Within the CVCV model, consonants in coda position are analysed as onsets followed by an empty V position (see (8d)). Licensing, a lateral relation which holds between C and V positions (see Scheer 2004, 138; Charette 1990, 1991), allows explaining the asymmetry between consonants in the onset and those in the coda. That is, licensing enhances the segmental content of its target while Proper Government inhibits the content of vocalic positions. The consonants in the onset are licensed by the following vowel whereas those in the coda remain unlicensed for they are followed by an empty V position. For the sake of convenience, we repeat in (12) the form given in (8d).
As can be seen from this representation, /b/ is licensed by the vowel /a/, whereas /t/ remains unlicensed since it is followed by an empty V position. In early children productions, particularly in babbling, such a form is more likely to be reduced to a light syllable, by deleting the second consonant. In other words, children prefer start producing licensed consonants before unlicensed ones. Further examples are found in Imane and Reda’s productions in (5): For instance, at the age of 18 months, Imane produced [hl:i] for /lhlɪb/ “milk”, [b:i] for /b:i/t/ “cut it”, and [iʃaː] for [iʃaːt] “he ate it”. Likewise, Reda pronounced /ij:iːh/ “yes” as [jːi] and /kːiss/ “remove” as [kːikːi], at the age of 23 months. In all these examples, the consonant in the coda is deleted, but the one in the onset.

The asymmetry between onsets and codas, and the late acquisition of the latter, is diagrammed as in (13). The arrows represent the path children tend to follow when acquiring syllables: onsets are generally acquired before codas.

<table>
<thead>
<tr>
<th>Standard syllable structure</th>
<th>Onset</th>
<th>Coda</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVCV approach</td>
<td>Licensed</td>
<td>Unlicensed</td>
</tr>
</tbody>
</table>

Within CVCV model, where “coda” is no longer a valid constituent, the asymmetry is captured by means of licensing: children start producing consonants in the onset position because they are licensed by the following vowel. Those that are followed by an empty vocalic position remain unlicensed.

7. Conclusion

In this paper, we presented preliminary results of a longitudinal study in Tashlhiyt Berber, which focused on the acquisition of geminates in relation with syllable structure. Children who participated in our study started producing geminates early in the babbling period. These geminates
occurred preferably in word-medial position. Word-initial geminates were favoured over word-final ones, departing from the patterns observed in adult language. This was explained by the late acquisition of coda consonants, as opposed to onset consonants. We also examined geminates in relation to syllable structure. We discussed two types of ambisyllabic geminates, using Dell & Elmedlaoui’s syllabic model: Those attaching to coda + onset and those associated to nucleus + onset positions. An alternative to ambisyllabicity was proposed within CVCV model. This model allows unifying the representation of geminates as a sequence of two CV units. It also accounts for the asymmetry between onsets and codas, using licensing relation between nuclei and onsets. The late acquisition of consonants in the coda position is explained by the fact that they are unlicensed, since they are always followed by an empty vocalic position.

Notes

1 Chaker (1992) refers to three million natives at the beginning of the 90s. According to Boukous (2011, 28), 28% of the population speaks Berber, of which 52% are native Tashlhiyt.
2 http://childes.psy.cmu.edu/browser/index.php?url=PhonBank-Phon/

References


Kahn, Daniel (1976), *Syllable-based generalizations in English phonology*. Doctoral dissertation, MIT.


Lahrouri, Mohamed (2001), *Aspects morpho-phonologiques de la dérivation verbale en berbère (parler chleuh d’Agadir)*. Doctoral
Mohamed Lahrouchi

A propos de I et de U en berbère : de la phonologie, de la morphologie et des phases,” in: Ali Tifrit (Ed.), 


Research on the second language acquisition of definite articles by L1 speakers of articleless languages dates back at least four decades (see e.g. Hakuta 1976). Early studies (Huebner 1983; Tarone & Parrish 1988, Thomas 1989) used the typology of definite/indefinite contexts proposed by Bickerton (1981) to analyze the over/underproduction of articles by L2 learners. This typology was based on two binary features, viz. ‘speaker reference’ [+/-SR] and “hearer knowledge’ [+/-HK]. Results were mixed (see Thomas 1989 for a review).

In the early years of this century an experimental paradigm came up that singled out one specific subtype of [+SR; -HK] contexts as particularly problematic for L2 learners, viz. those in which the speaker intends to refer to a specific noteworthy individual. (1) and (2) show how this type of specificity or lack thereof is operationalized:

(1) I’m very sorry, but my sister doesn’t have time to talk right now. She is meeting with a very important client from Seattle. He is quite rich, and she really wants to get his business for our company.

(2) Professor Peterson is meeting with a student, but I don’t know who it is.

(short version of items in Ko et al. (2010) and Ionin et al. (2004))

The operationalization of specificity in (1) lies in the modifiers of *client* as well as in the follow-up sentence suggesting the speaker has a specific noteworthy individual in mind. The operationalization of non-specificity in (2) lies in the addition of “I don’t know who it is’, making sure that no intention of specific noteworthy reference is available. Ionin et al. (2004) show how Korean and Russian L2 learners of English are more likely to overproduce definites in (1) than in (2). The paradigm has generated consistent results in a number of replication studies with
Russian (Ionin et al. 2009), Korean (Ko et al. 2010), Japanese (Hawkins et al. 2006) and Mandarin (Trenkic 2008) L2 learners of English and has become a standard in the literature (cf. Slabakova 2016, though see Trenkic 2008 for early criticism).

The focus on Bickerton’s typology and the role of specificity during the past decades has considerably alienated the L2 literature from the semantic literature on definites. Questions about the role of maximality and familiarity – central concepts in the theoretical literature – have hardly been investigated. This paper briefly introduces the major ingredients of semantic analyses of definites and then continues with a focus on maximality. We summarize the contributions in the L2 literature that focus on it (section 3), introduce new data from a felicity judgement task (section 4) and propose an interpretation (section 5).

2. Definiteness in the semantic literature

Definiteness has attracted attention in the theoretical semantic literature for over a century. Two intuitions stand out. The first is that definiteness is related to uniqueness. The use of the definite in (3) would e.g. entail that there is a single king of France (Russell 1905):

(3) **The king of France** is bald.

Strawson (1950) adds a presuppositional dimension to definiteness by relegating existence and uniqueness to the domain of presuppositions. Hawkins (1978) extends uniqueness to plurals:

(4) **The children of our neighbor** are always very noisy in the evening.

*The children of our neighbor* refers to the unique maximal sum of the children of the speaker’s neighbor. The (singular) concept of uniqueness can thus effectively be extended to plurals. To avoid a singularity bias, the term *uniqueness* is sometimes replaced by more general terms like *maximality* and *exhaustivity*.

The second theoretical intuition about definites is that they mark familiarity. An early precursor of this type of account is Christophersen (1939). It gained popularity in the 1980s in dynamic frameworks like Discourse Representation Theory (Kamp 1981). The gist is that definites
refer back to something given in the discourse context. A standard example would be the one in (5) where the man anaphorically refers back to the previously introduced man.

(5) A man and a woman are walking in the park. The man is whistling.

The exact definition of familiarity varies considerably between authors. E.g., Christophersen (1939) claims that familiarity entails that the hearer is able to “call up the image of the exact individual the speaker is thinking of” whereas Kamp & Reyle (1993) mainly refer to anaphoricity as in (5).

One way of bringing together the uniqueness/maximality and the familiarity perspective is presented in Hawkins (1978). Hawkins’ analysis assumes definites are referential and guide the hearer in identifying their referent within one of the following situations:

(6) Immediate situation:
Pass me the bucket, please.

(7) Discourse situation (anaphoric use):
Fred brought me a bucket, but the bucket had a hole in it.

(8) Larger situation:
[uttered in the UK] The Prime Minister has just resigned.

The way Hawkins (1978) integrates familiarity is by identifying discourse as one of the possible situations that can be used to locate the referent. Uniqueness/maximality comes in as a prerequisite of the act of reference: in order for reference to be successful there should be no more than one individual (or unique maximal sum of individuals) satisfying the descriptive content of the noun within the relevant situation. Presuppositionality enters through the same route: reference can only be successful if the intended referent is part of the speaker’s and hearer’s common ground.

3. Uniqueness/Maximality in the L2 literature

As we indicated in the introduction, there are only a few L2 studies that have zoomed in on theoretical key concepts like
uniqueness/maximality and familiarity. We focus in this paper on uniqueness/maximality. We use this section to briefly summarize three relevant studies: Trenkic et al. (2014), Yang & Ionin (2009) and Ko et al. (2010).

*Trenkic et al. (2014)*

Trenkic et al. (2014) argue on the basis of eye-tracking data using a visual world paradigm that Mandarin L2 learners of English associate the definite followed by a singular noun with uniqueness. In particular they find that L2 learners with an articleless L1 behave like natives when they are asked to *Pick up the cube and put it inside the can*: they are quicker at fixating on a can in a context in which there are two cans but only one that is big enough to fit the cube than in a context in which there are two cans that are both big enough. Put differently: with a definite they expect there to be only one contextually relevant can and they start doubting if there are two. This strongly suggests that they use the uniqueness semantics of the definite article in processing.

*Yang & Ionin (2009) and Ko et al. (2010)*

Yang & Ionin (2009) argue on the basis of an Acceptability Judgement Task with Mandarin L2 learners of English that L2 learners accept definites even when the intended referent is not unique. This happens in particular when the referent has been introduced before as part of a bigger set, either explicitly as in (9) or implicitly as in (10):

(9) Dennis has many interesting books. His cousin borrowed the book from him yesterday.

(10) I went to a wedding yesterday. The guest gave a speech.

The first sentence of (9) implies that Dennis has more than one book. The second sentence consequently refers to one of Dennis’ books rather than to his only one. The same holds for (10) as soon as we make the plausible assumption that weddings involve more than one guest.

Ko et al. (2010) replicate the findings of Yang & Ionin (2009) for Korean L2 learners of English in a forced-choice production task. In particular, they find that L2 learners overuse definites in contexts like (11) as compared to contexts like (12):

(11) Elissa: How is your nephew Aaron doing? He is such a nice little boy!
Robert: He has some good news—his parents finally allowed him to get a pet—just one! So last week, he went to our local pet shop. *This pet shop had five puppies and seven kittens*, and Aaron loved all of them. But he could get only one!

Elissa: Oh, so what did he do?

Robert: Well, it was difficult for him to make up his mind. *But finally, he got (a, the, —) puppy*. Aaron went home really happy!

(12) Elissa: How is your nephew Joey doing? He is such a nice boy!

Robert: Well, he was a bit depressed the last few days. So, his parents decided to get him a pet. So last week, he went to our local pet shop.

Elissa: Oh, so did he buy some animal there?

Robert: No, he did not like the puppies in the pet shop, in fact. But then he was walking home, and *he found (a, the, —) kitten in the street!* So now he has a new pet after all!

The crucial difference between (11) and (12) is that the puppy in (11) has been introduced before (as part of the puppies in the pet shop) whereas the kitten in (12) has not.

Based on the three studies introduced above there seems to be no real consensus in the literature on whether definites for L2 learners with an articleless L1 are unique/maximal or not. Trenkic et al. (2014) find that L2 learners are sensitive to uniqueness/maximality but Yang & Ionin (2009) and Ko et al. (2010) find that this sensitivity does not prevent L2 learners from interpreting/producing definites that are explicitly non-unique/non-maximal. The study we present in section 4 is an offline felicity judgement task that probes the availability of non-maximal readings for those definites whose referents do not have any antecedents in the discourse. In this sense, we complement Yang & Ionin’s (2009) and Ko et al.’s (2010) offline paradigms.

4. The Study

4.1. Materials

The goal of this study was to check whether L2 learners allow for non-maximal readings of definites. In order to do so we designed a felicity judgment task with 25 items in which we opposed definites to indefinites
and to *all*. The basic task of the subjects was to decide whether the bold parts of examples like (13) to (15) were compatible or incompatible with the parts preceding them:

(13) Indefinite condition
Harry smoked 5 cigarettes, **so now he only has 3 left**.

(14) Definite condition
I stole the 2 pens of Mary’s, **so now she only has one left**.

(15) *All* condition
I stole all of Jacky’s books, **so now she only has one left**.

(13) states that Harry smoked 5 cigarettes but makes no claim about whether or not these 5 cigarettes were the only ones he had before he started smoking. The second part of the sentence is thus expected to be compatible with the first. (15) states that the speaker stole all of Jacky’s books and it is consequently expected that Jacky has no books left, making the second part of (15) incompatible with the first. If the 2 pens of Mary’s in (14) receives a non-maximal reading, we expect (14) to pattern with (13). If, however, the 2 pens of Mary’s does not allow for a non-maximal reading, we expect it to pattern with (15). On a standard analysis of definites, (15) is considered an accommodated definite (Roberts 2003) and is expected to be interpreted as referring to all of Mary’s pens, a larger situation use on Hawkins’ analysis.

Next to the experimental conditions, each represented by 3 items, we included a number of filler items. We strived for a balance between “compatible” and “incompatible” responses. The items were pseudo-randomized and are available in the Appendix.

### 4.2. Participants

The L2 participants in this study were 20 L1 Mandarin/L2 English speakers. All were undergraduate students of English at the Beijing International Studies University, aged between 17 years 6 months and 20 years 9 months (average: 19 years 1 month). All of them had followed regular English training from primary school onwards but had not spent any time abroad. They were recruited by a student assistant and offered a small financial compensation for their participation.

This L2 population suits our needs in the sense that they are L1 speakers of a language without articles. Problems with the production and
interpretation of definites have been argued to be relevant for beginner, intermediate and advanced learners (Ionin et al. 2004). To guarantee the level of English that was required for the task we however focused on students studying English at University, which implies they had passed the NMET (Cheng & Qi 2006) and were preparing for TEM-4 (Jin & Fan 2011). As such, they do not qualify as beginners but rather as intermediate learners. This level is comparable to the one of the learners in Trenkic et al. (2014), Ko et al. (2010) and Yang & Ionin (2009). Compared to the participants in Yang & Ionin (2009), students who did not necessarily major in English, the English level of our participants might be a bit higher. Compared to the participants in Ko et al. (2010), mainly international students/foreign workers who had spent between half a year and eight years in the US, their level might be a bit lower. All in all, we expect our participant pool to be comparable to the ones used in the literature up till now.

Next to our L2 participants we recruited 29 adult L1 participants on the online crowdsourcing platform Crowdflower for a reduced version of the experiment. Crowdflower divides its workers into three categories, depending on the reliability they have demonstrated in earlier tasks. We recruited those in the highest category and furthermore limited our participant pool to workers in the UK or in the US. We removed eight participants because they underperformed on the filler items.

4.3. Procedure

For our L2 learners, the test was administered by the student assistant in a quiet environment at the university or sent through email so that the participants could complete the test on their own, on their home or school computers. The instructions as well as the 25 test items were presented in a Word file. The answer options were presented as check boxes. The Word file furthermore contained a small language biography questionnaire. The participants were given no time limit but most of them completed the experiment in under ten minutes.

For our L1 participants, the test was administered through the Crowdflower platform.
4.4. Results

A summary of the results is presented in Table 1. We go through them in detail in the remainder of this section.

The indefinite and the all condition

Our L2 participants behaved as expected on the indefinite condition and the all condition. For the indefinite condition the percentage of “compatible” responses was 100%, for the all condition the percentage of “compatible” responses was only 4%. Our L1 participants generated similar responses: 93.5% of “compatible” responses in the indefinite condition and only 3% for the all condition.

Table 1: Compatible responses per participant group and condition

<table>
<thead>
<tr>
<th></th>
<th>indefinite</th>
<th>Definite</th>
<th>all</th>
</tr>
</thead>
<tbody>
<tr>
<td>L2 learners</td>
<td>100%</td>
<td>81.5%</td>
<td>4%</td>
</tr>
<tr>
<td>L1 participants</td>
<td>93.5%</td>
<td>50%</td>
<td>3%</td>
</tr>
</tbody>
</table>

The definite condition

Our L2 participants chose a “compatible” response in 81.5% of the cases. To determine the significance of this result in comparison to the results on the indefinite and the all condition we ran a mixed effects model with item and participant as random factors and a sequential Bonferroni correction. The analysis revealed a significant effect of condition (p<.001) and pairwise comparisons of the model showed that the definite condition was significantly different from both the indefinite (p<.001) and the all condition (p=.029).

Our L1 participants also had a high degree of “compatible” responses for the definite condition, even though it was lower than that of our L2 participants, viz. 50%. To determine the significance of this result in comparison to the results on the indefinite and the all condition we also ran a mixed effects model. The analysis revealed a significant effect of condition (p<.001) and pairwise comparisons showed that the definite condition was significantly different from both the indefinite (p<.001) and the all condition (p<.001). To determine the significance of the difference between the L2 and L1 participants on the definite condition, we ran an independent samples t-test on the number of “compatible” responses per participant. The difference between the two was significant (p=.036).
5. Discussion

5.1. L1 participants

The responses to the indefinite and the all condition were expected and confirm that participants were sensitive to the experimental manipulation. We had not anticipated the high degree of “compatible” responses for the definite, though. In retrospect, we relate these to the intuition proposed by Brisson (1998) according to which definites are different from all in that they allow for “sloppy” readings. What we mean by this is that we assume the definite is as maximal as all but that native speakers do not always enforce the maximality of the definite to the same extent on their interlocutors’ speech. A closer look at the results of the L1 participants shows that the majority either allowed or rejected non-maximal readings of definites. Each strategy is represented by eight participants. This means participants’ responses were either “strict” or “sloppy”.

5.2. L2 participants

For our L2 participants, the results on the indefinite and all items also show that they were sensitive to the experimental manipulation. As for the definite items, the question that imposes itself is whether there is a difference between the “sloppy” maximal readings of our L1 participants and the “compatible” responses of our L2 participants. We would like to suggest that there is: if participants could be either “strict” or “sloppy”, we would expect a more or less even distribution between “compatible” and “incompatible” answers. This is however not what we get and the independent samples t-test picks up on this. What we get is an overwhelming majority of “compatible” answers that cannot easily be explained away. Our results thus seem to be in line with those of Yang & Ionin (2009) and Ko et al. (2010) in that L2 definites can be non-maximal.

5.3. Are L2 definites really non-maximal?

The final question we turn to is whether or not there is a way to interpret the results of Yang & Ionin (2009), Ko et al. (2010) and the present study in such a way that they are compatible with the results of Trenkic et al. (2014). We would like to suggest that there is by exploiting
the theoretical intuition proposed by Le Bruyn & Dong (2016) according to which L2 definites can be non-presuppositional. We explain how this intuition can account for the different results (5.3.1.) and briefly defend the plausibility of this intuition (5.3.2.).

5.3.1. Non-presuppositionality and (apparent) non-maximality

Schwarz (2009) picks up on the analysis of Hawkins (1978) that we presented in section 2. Next to the immediate, discourse and larger situation he however adds the topic situation as a possible situation for interpretation (see also, from a different perspective, Löbner 1985). Glossing over the semantic details, the topic situation contains those and only those individuals that are introduced in the corresponding proposition. For a proposition like (16), this means that the topic situation contains four and only four individuals: the speaker, the two bird cages and the fence post:

(16) I hung up two bird cages on a fence post.

Now note that the bird cages and the fence post are by definition maximal/unique in the topic situation (they are the only ones that are mentioned) and that the maximality/uniqueness requirements the definite version of (16) comes with are automatically met:

(17) I hung up the two bird cages on the fence post.

If the definites in (17) can be interpreted with respect to the topic situation, we consequently predict the maximality/uniqueness requirements the definites come with to be neutralized and for potential non-maximal readings to pop up. For examples like (18), this would mean that the two pens of Mary’s does not necessarily refer to all pens of Mary’s but could be interpreted as referring to the two pens of Mary’s occurring in the topic situation, i.e. to the two pens of Mary’s the speaker stole.

(18) I stole the two pens of Mary’s.

The crucial twist now is that native speakers cannot use definites in (17) or (18) and evaluate them with respect to the topic situation unless it has been established in the common ground of speaker and hearer that the
topic situation contains two bird cages and one fence post. This is a requirement imposed by the presuppositionality of definites. As Schwarz (2009) notes, this requirement is unlikely to be met given that the topic situation is in general new to the hearer. For native speakers, (17) and (18) are consequently expected to be infelicitous when they try to interpret them with respect to the topic situation. The same does not hold for L2 learners with an articleless L1 if we assume with Le Bruyn & Dong (2016) that definites for these learners are non-presuppositional. For these learners we expect the topic situation to be a felicitous situation for interpretation for (17) and (18) and consequently for the follow-up of (18) in (19) to be felicitous:

(19) So now she only has one left.

If we assume the definite in (18) can be interpreted as referring to the two pens of Mary’s that the speaker stole, there is no reason to expect (19) to be incompatible with (18).

If the analysis we have worked out is on the right track, the high degree of “compatible” responses for our L2 learners does not indicate that they are insensitive to maximality/uniqueness but rather that they can evaluate maximality/uniqueness with respect to more restricted situations than native speakers. The analysis can straightforwardly be extended to the findings of Yang & Ionin (2009) and Ko et al. (2010). The findings of Trenkic et al. (2014) also follow if we assume that visual world paradigms are powerful enough to enforce the same situation of interpretation for both natives and L2 learners with an articleless L1. Given that both populations are sensitive to maximality/uniqueness for definites, aligning the situations of interpretation suffices for them to perform in the same way.

5.3.2. Non-presuppositional readings are not exotic

The classical debate about presuppositionality of definites was an all-or-nothing debate between Russell (1905) and Strawson (1950). Whereas Russell assumed a sentence like (20) states that there is a single king of France and that he is bald, Strawson assumed that it presupposes that there is a single king of France and that it states that this king is bald.

(20) The king of France is bald.
As a consequence, (20) would – in the current state of affairs – be false on Russell’s account and undefined on Strawson’s.

For most of the second half of the 20th century the literature has followed Strawson. Partee (1987) however suggests that we might want to keep both the Russell and Strawson variants and this is reinforced by recent work by Coppock & Beaver (2015) and Le Bruyn, de Swart & Zwarts (2016) who show that certain uses of the do not show the exact presuppositional behavior Strawson would want them to have. This is illustrated in (21):\footnote{5}

\begin{enumerate}
\item[(21)] John doesn’t have the only sweet brother.
\end{enumerate}

If the only sweet brother had its usual presuppositional reading, we would expect (21) to be about the only existing sweet brother and to say about him that he is not John’s brother. The most obvious reading however is one in which the existence of a unique sweet brother is denied: (21) states that John is not the only one with a sweet brother. Given that the existence of a unique sweet brother can be denied, the only sweet brother in (21) cannot be presuppositional.

Examples like (21) show that non-presuppositional readings of definites are real and Le Bruyn & Dong (2016) propose a fine-grained account of why we expect these readings to be more frequent for L2 learners with an articleless L1 than for native speakers. The crucial point for the current paper however is that non-presuppositional readings of definites are not as exotic as one might expect them to be.

\section*{6. Conclusion}

In this paper we have studied the availability of non-maximal readings of L2 definites through a felicity judgement task with intermediate L2 learners of English with Mandarin as their L1. In line with earlier findings by Yang & Ionin (2009) and Ko et al. (2010), we have found that our L2 learners differ from native speakers in allowing for non-maximal readings (section 4). We have however argued that these non-maximal readings might only be apparent and indicative of another difference between native and L2 definites concerning the non-presuppositionality of the latter (section 5). This would allow the findings of the present study and those
from Yang & Ionin (2009) and Ko et al. (2010) to be made compatible with those of Trenkic et al. (2014).

Notes

1 We thank the semanticists in Utrecht and Leiden, Jeannette Schaeffer, our research assistants Xinyuan Wang, Wenzhu Xuan, Zimo Lian as well as all our participants. For useful feedback we also thank audiences at the Amsterdam L2 Acquisition Seminar (2015, 2016), our Utrecht Elitu talk, GALA 2015, LCR 2015, the Amsterdam Colloquium 2015 and our CReLLU talk in York. Special thanks to Yunhua Hu and his colleagues for early support and our family and friends in China for their hospitality. The first author furthermore gratefully acknowledges the support of NWO, grant 275-80-006.

2 We had three possible answers: “compatible”, “potentially compatible” and “incompatible”. For statistical purposes we grouped “compatible” and “potentially compatible”.

3 We originally recruited 30 participants but one indicated that he/she was not a native speaker of English.

4 The items included in this reduced version are indicated with an asterisk in the Appendix.

5 See also Champollion & Sauerland (2010) for another type of non-presuppositional use of definites and Le Bruyn & Dong (2016) for an overview of apparent non-presuppositional uses.

References

Christophersen, Paul (1939), The articles: A study of their theory and use in English. Copenhague: Munksgaard.


**Appendix**

Test items
(the items with an asterisk were part of the reduced Crowdflower version of the experiment)

Experimental items [condition, expected response]
*1. I stole 3 pencils of Harry’s, so now he only has two left. [indef., comp.]
*2. I broke 2 pots of Mary’s, so now she only has 1 left. [indef., comp.]
*3. I killed 3 pigs of Peter’s, so now he only has one left. [indef., comp.]
*4. I stole the 2 pens of Mary’s, so now she only has one left. [def., incomp.]
*5. I broke the 2 vases of Anna’s, so now she only has 2 left. [def., incomp.]
*6. I killed the 5 chickens of John’s, so now he only has 2 left. [def., incomp.]
*7. I stole all of Jacky’s books, so now she only has one left. [all, incomp.]
*8. I killed all of Mike’s pigeons, so now he only has 5 left. [all, incomp.]
*9. I broke all of Harry’s cups, so no he only has 5 left. [all, incomp.]

Fillers [compatible]
*Mary ate 2 biscuits and 3 cupcakes, so she ate 5 pastries.
*I had 5 dogs, I sold 3, so now I have 2 left.

Harry smoked 5 cigarettes, so now he only has 3 left.
Mary gave me 3 eggs and John gave me 4, so now I have 7 eggs.
Mary ate 13 cupcakes, so now she only has 2 left.

I had 5 chickens, I killed 2, so now I have 3 left.

John spent 123 yuan, so now he only has 5 yuan left.

I’ve just bought 15 CDs, so now I have 50.

John killed 2 cows, 5 chickens and 20 flies, so he killed 27 animals.
*Mary had 36 eggs, she lost a third, so now she has 24 left.

Fillers [incompatible]
*Max took his phone, a pen and five coins, so he took 8 objects in total.

*1 bus, 3 bikes and 2 lorries passed, so in total 5 vehicles passed.
I stole 2 pens, 5 pencils and 3 laptops, so I stole 9 items.
Bill broke 1 vase, 2 plates and 35 pots, so he broke 37 items.
Minnie talked to her mother, her father, her brother and her sister, so she talked to 5 people in total.
*Anna first read 6 pages, then 15 and then 16, so she read 38 pages in total.
FACTORS AFFECTING PERFORMANCE IN CHILD HERITAGE PORTUGUESE IN GERMANY

TATJANA LEIN, MONIKA ROTHWEILER AND CORNELIA HAMANN

1. Introduction

The aim of the present study is to evaluate the effects of different input factors on the performance in different language domains of Portuguese as heritage language (HL). The major factor to be studied is age of onset (AoO) of exposure to the majority language (ML) German. Therefore, we investigate the performance in the HL in two groups of heritage speakers with different first exposure ages to the ML. Heritage speakers are defined as bilinguals whose first language (L1) is not the ML of the society in that they grow up (Rothman 2009). In the present study, we compare the language performance of simultaneous bilingual (2L1) children who acquire the HL Portuguese and the ML German from birth to the language performance of sequential bilingual children who acquire the HL Portuguese from birth, but only started to learn the ML German at the age of three or later as a child second language (cL2) in kindergarten (Meisel 2011). We will refer to this second group as cL2 learners, even though the focus of the present paper is the L1 Portuguese of these children.

longitudinal studies it is very difficult “to distinguish properly between attrition (what was acquired and then lost), incomplete acquisition (what was available and simply not acquired) and issues related to input type (what could not be acquired given its absence in the input).” Cabo and Rothman (2012) suggest that the term incomplete acquisition may not be adequate for the group of heritage speakers as a whole because many of them are exposed to an input different from that of L1 learners due to cross-generational attrition or language contact changes. Therefore the competence of heritage speakers might be different from monolingual ML learners, but complete, not deficient. This clarification is important, and the term incomplete acquisition should not be used without a reference to how it is meant in a given context, i.e. as incomplete acquisition compared to a given and perhaps incomplete input or as incomplete acquisition compared to the target language. So, besides the need for longitudinal studies also the input in the HL as well as in the ML has to be taken into account.

In the present paper we want to investigate and compare the role of several input factors for the retention/loss of the HL in 2L1 and in cL2 children. One factor is the AoO of the ML. With kindergarten or school entry, a reduction of input, in extreme cases even an interruption of input of the HL takes place so that the ML becomes the dominant input. Due to this change in input, usually the language dominance changes towards the ML which is then called primary or dominant, while the HL becomes the weaker language. Importantly, 2L1 children are exposed to a reduced input already from birth and the input is divided into two languages from early on. The sequential bilinguals, however, have a full input from birth in the HL up to the point when the second language comes into play. This will be subsumed under the factor early input quantity, a factor that is influenced by the language use of the persons in the environment of the child. While in sequential learners the early input of the HL is expected to be up to 100% until intensive exposure to ML with kindergarten entry starts, the amount of HL input varies in the group of 2L1 learners depending of the language use of the persons in the environment of the child. A related measure is length of exposure to the ML (LoE) which is also different in the two groups and might have an impact on the development of the HL as well. A further important factor might be the input quality in the HL as pointed out by Cabo and Rothman (2012). The parents might be second generation immigrants and already experience some language difficulties or attrition in their HL in contrast to parents that are first generation immigrants who arrived in Germany as adults and have a more robust competence. Finally, the current input quantity might
also play a role. Some children might stop speaking the HL when the ML becomes dominant in the input. Other children, receiving support through schooling in the HL, will still have frequent contact to the HL community and therefore a considerable amount of HL input. Given this background, the central question of our study is whether these differences in HL input lead to differences in terms of language competence in the HL. Montrul's (2008) hypothesis 2 is highly relevant for our study and emphasizes the importance of AoO of a ML.

Hypothesis 2 (Montrul 2008, 60):
If language attrition occurs within early (pre-puberty) bilingualism, it will be more severe in simultaneous bilinguals (exposed to the two languages very early) than in sequential bilinguals (when the L1 was acquired before the L2).

This hypothesis may be extended to the prediction that all factors reducing the HL input of 2L1 speakers may have an impact on language attrition or incomplete acquisition of the HL. Thus, the crucial questions in our study are: Are different language domains affected differently by different input factors (AoO, LoE, language use at home, amount of input in early childhood and at the time of testing)? Which of the different factors is the most important? Which language domain is most affected?

In the present paper we will try to contribute to the discussion of these questions. In the following section we will summarize previous research on the language development in 2L1 and cL2 heritage children. In section 3 we will present our research questions. Section 4 describes the tasks that we used as well as the subjects. Section 5 presents our results on the performance in the different language domains and an analysis of input factors. Discussion and conclusion follow in section 6.

2. Previous Research

In this section we concentrate on results from previous research on child heritage speakers targeting the HL Portuguese or another Romance language and/or the ML German. We divide the studies into three groups depending on the type of bilingualism they deal with i) only cL2, ii) only 2L1 and iii) 2L1 vs. cL2.

Studies on cL2 children concentrating on the period in which the children were already exposed to the ML as well, found a reduced lexicon
in the HL compared to monolinguals and a delayed morpho-syntactic development. For example, Klassert, Gagarina and Kauschke (2014) investigated the lexical abilities in terms of noun and verb naming in 60 Russian-German bilinguals (age 4;0-6;11) with Russian as heritage language and German as cL2 (AoO 30 months). They found that the bilinguals performed worse at naming words in both languages than the respective monolinguals. The performance in German was better than in Russian, which underlines the claim that after German had come into play the dominance shifted to German. Barbosa & Flores (2011) did not explicitly differentiate between 2L1 and cL2 heritage speakers when focusing on the HL Portuguese in Portuguese-German bilinguals (age 7-15). All participants spoke predominantly Portuguese at home and all parents were first generation immigrants. Therefore we assume that the children were cL2 learners. The authors found that in the HL group object expression, e.g. position of object clitics, was deviant from that of age matched Portuguese monolinguals. However, the younger children (age <7) performed significantly lower than the older children (age 7-15) and the error pattern resembled that observed in monolingual development. Therefore the authors conclude that the acquisition may only be delayed.

Most studies on 2L1 heritage children do not focus the HL, but the early development in both languages (age 1;0 - 4;0). These studies generally showed that 2L1 learners develop qualitatively similar to monolinguals in both languages and that they separate their languages from early on. This holds for different linguistic areas (phonetics/phonology, lexicon/semantics and morpho-syntax). Meisel (2001, 40) focused on “the grammatical competence” of 2L1 children claiming that it “does not differ in quality from that of the respective monolinguals”. This view is supported by several studies on 2L1 children that found no qualitative differences in early grammatical core grammar phenomena like verb placement (cf. Tsimpili 2014). Meisel (2011, 243f) pointed out that although some studies found cross-linguistic influence, either as transfer, as acceleration or as delay, these interdependencies are only temporary (see also Hulk & Müller 2000, Müller 1998). Such differences are likely to occur in all linguistic domains. Studies on 2L1 children acquiring phonetics/phonology show early separation of the phonetic systems and acquisition patterns that are qualitatively similar to monolinguals despite subtle differences (Kehoe, Lleó & Rakow 2004, Maneva & Genesee 2002, Polka, Rvachew & Mattock 2007, Vihmann 2002). For example, Kehoe et al. (2004) investigated VOT in four Spanish-German 2L1 children in Germany and found a delayed acquisition for two children (in both languages) and an interaction of the
two phonetic systems for one child producing Spanish stops with German VOT, thus showing a transfer from the ML to the HL. For the lexical domain, for example, Pearson, Fernández and Oller (1993) showed that in 25 Spanish-English 2L1 children (age 0;8-2;6) the same developmental patterns arise at the same ages as in monolinguals, although bilinguals have acquired fewer lexical entries per language than monolinguals. The vocabulary size in both languages combined corresponded to that of monolinguals. Pearson, Fernández and Oller (1995) found for 27 Spanish-English 2L1 children (age 0;8-2;6) that they developed two separate lexicons evidenced by the use of 30% doublets from early on.

For the lexical domain, for example, Pearson, Fernández and Oller (1993) showed that in 25 Spanish-English 2L1 children (age 0;8-2;6) the same developmental patterns arise at the same ages as in monolinguals, although bilinguals have acquired fewer lexical entries per language than monolinguals. The vocabulary size in both languages combined corresponded to that of monolinguals. Pearson, Fernández and Oller (1995) found for 27 Spanish-English 2L1 children (age 0;8-2;6) that they developed two separate lexicons evidenced by the use of 30% doublets from early on. Studies with a focus on older 2L1 children (> age 4;0) found signs of attrition in morpho-syntax, but not in phonology. For instance, Watson (1990) found that 15 French-English 2L1 children (ages 6, 8 and 10) were not different in VOT development from the respective monolinguals and native listeners did not detect any differences in pronunciation from that of monolingual children. Silva-Corvalán (2014a, b) studied two Spanish-English 2L1 siblings (with a three-year age difference) in California longitudinally from age 1 to 6. The early HL Spanish input given by child-directed speech of the father was about 30% for the older sibling and about 27% for the younger brother. From about age 4;0 the Spanish input to both children was then reduced by 25% for the older and 22% for the younger sibling and more English was used instead. This reduction of Spanish and the increase of English input influenced the morpho-syntax development of Spanish in both children so that it differed in several phenomena from the Spanish target grammar. Nevertheless, the Spanish morpho-syntax was more stable in the older child. While gender marking and agreement did not cause any problems, both children had difficulties in choosing the correct copula (ser or estar). Additionally, the siblings had problems with subject use (null or overt) in the appropriate contexts, which, again, was more pronounced in the younger child. Around the age of six, especially the younger child showed additional difficulties with complex verbal morphology as e.g. subjunctive forms or conditional. An explanation for the differences between the two children might be found in the quantitative differences in input.

To summarize, 2L1 speakers might be “fully capable of acquiring two languages” (Montrul 2008: 97), but several studies found differences to monolingual acquisition, especially in the HL. Full linguistic competence is not guaranteed as shown by studies on 2L1 heritage children older than 4;0: The language competence of these children may differ from that of monolingual peers.
As to comparisons of 2L1 with cL2 children, in three studies that compared Spanish-English 2L1 children to bilinguals with English as cL2 an advantage in the HL for cL2 children was found (Anderson 1999, 2001; Silva-Corvalán 2003). Anderson (1999) investigated gender morphology while Anderson (2001) studied verb use and verb morphology in the spontaneous speech of two Puerto Rican Spanish-English bilingual children (age 4;7 and 6;7). The older one of the siblings came to the United states at the age of 3;6 and was considered a cL2 learner of English and the other sibling arrived at an age of 1;6, went to an English kindergarten at the age of 2;6 and was considered a simultaneous bilingual. In gender agreement, which is acquired early (until the age of 3) by Spanish monolinguals, the 2L1 child produced an increasing number of errors over time. At the same time the number of nouns she produced decreased and she started to overuse the masculine form as default. In contrast, her sister with English as early L2 produced almost no errors (Anderson 1999). A similar pattern was observed in the study on verb morphology (Anderson 2001). Although both children produced few verb errors, the error rate was higher in the 2L1 child and increased until the last recording. In contrast, the amount of errors remained stable for the older cL2 child. Most errors were observed in agreement morphology while aspect or tense were less problematic. For errors in person/number distinctions Anderson (2001) supposes that these were a result of language loss, because they had been acquired early. In contrast, subjunctive mood is acquired only at the age of 7 in monolingual children and therefore the errors cannot be associated with language loss, but result from incomplete acquisition. Anderson (2001) attributes the different performances of the sisters to differences in factors such as AoO and use of L1 with peers or at home. Finally, Silva-Corvalán (2003) studied 7 Spanish-English 2L1 and cL2 bilingual children (ages 5;1-5;11) and found that the 2L1 children had more problems with complex tenses in their HL Spanish, for example with the aspectual contrast preterit-imperfect, than the cL2 children who spoke more Spanish at home.

Summarizing the studies on child heritage language, in general heritage children clearly differ from monolinguals with respect to vocabulary size but only slightly in phonology. CL2 children develop qualitatively similar morpho-syntactic abilities in their HL as monolinguals, although quantitative differences may occur. 2L1 children in contrast, develop qualitatively similar to monolinguals with some quantitative differences in morpho-syntax only up to the age of 4;0. After that age signs of attrition or incomplete acquisition become visible depending on the quantity of input. The cL2 children always outperformed
the 2L1 children, leading us to the assumption that AoO and amount of early input are the crucial factors for the acquisition of a HL.

3. Research questions and predictions

Following the second Hypothesis of Montrul (2008), the central aim of the present study is to investigate whether there is a difference in the HL Portuguese of 2L1 bilinguals and cL2 bilinguals and if so, in which domains. The input quantity should be different in the two groups because in the 2L1 group the HL is the only input until about age 3. The difference in AoO of and LoE to the second language in the two groups might be related to the migration status of the parents as first or second generation HL speakers, or to both parents being native speakers of the HL or not. Given this different background for the two groups, more input factors may have - eventually different - impacts on the acquisition of the HL. In the present paper we try to answer the following four questions:

1) Does the HL of 2L1 Portuguese-German children differ from that of cL2 Portuguese-German children (AoO of ML German age 3)? To answer this question we will look at group comparisons of 2L1 vs. cL2 heritage speakers in their performance on a Portuguese standardized test. We expect the cL2 children to outperform the 2L1 children.

2) Which language domains are most affected in Portuguese heritage children in Germany? In an attempt to answer this question we run a quantitative analysis of five different subtasks of a standardized Portuguese test (normed with monolinguals). We expect all areas to be affected but the signs of attrition or incomplete acquisition to be most visible in lexicon and morpho-syntax.

3) How do the children of the two groups differ with respect to input factors? The groups were compared with respect to the factors AoO, LoE, input quality, early input quantity, current input quantity and dominance based on information gained by a parental questionnaire. The group assignment should reflect differences between 2L1 and cL2 children with respect to AoO and LoE. We expect group differences for early input quantity as well as input quality. The current input quantity and dominance might be similar in both groups due to a dominance shift that happened after the children started to attend kindergarten, often resulting in less use of the HL even at home.
4) What is the impact of different input factors on the performance in the HL in different domains? This will be analyzed in terms of a correlation analysis with the aforementioned input factors and the test results in the five subtasks. We expect all factors to have an effect on the lexical tasks and factors related to early input to be most relevant for morpho-syntax. Since the morpho-syntactic tasks are measured as overall scores, no predictions with respect to individual grammatical domains are possible.

4. Method

Data were collected as part of the French-German BiLaD project on bilingual language development and specific language impairment in children with French or German as L2 and Portuguese, Arabic or Turkish as L1. In the project we administered a number of different language tasks including standardized tasks in the L1 and in the L2. The present paper will exclusively focus on Portuguese as heritage language in unimpaired children living in Germany. The analyses will refer to the children's performance in the standardized Portuguese task Palpa-P (Castro, Caló & Gomes 2007) and to results of the parental questionnaire PaBiQ (Tuller 2015). For the statistical analyses we used non-parametric tests (Mann Whitney U) and Spearman correlations for the analysis concerning the influence of background factors.

4.1. Palpa-P

Palpa-P (Provas de Avaliação da Linguagem e da Afasia em Português; Castro et al. 2007) is an adaptation of Psycholinguistic Assessments of Language Processing in Aphasia (Kay, Lesser & Coltheart 1992) and contains 60 different subtasks. It is used as a language assessment tool for adults and children in Portugal. We only used five out of these 60 subtasks for the purpose of the BiLaD-project: 1) subtask number 54 on lexical production 2) subtask number 47 on lexical reception 3) subtask number 12 on morpho-syntactic production 4) subtask number 55 on morpho-syntactic reception 5) and subtask number 8 on phonology. Children of all age groups performed exactly the same tasks. Monolingual norms are available for children from age 5 to 9 (with
some exceptions, see 5.1.) which enabled us to compare the results of the participating children to those of monolinguals in Portugal. It took approximately 50 min to administer these five subtasks. The children were recorded with a digital voice recorder OLYMPUS DM-650. All tasks were administered and transcribed by a linguist and a native speaker of Portuguese.

4.1.1. Lexical Production (LexP)

The subtask on lexical production contains 40 items that are shown two at a time on one page. All items depict Portuguese nouns. Half of the words have high and half of the words have low frequency. The subtask uses the method of picture naming. The experimenter points to a picture and asks in Portuguese: “what is this?” The child responses were recorded and transcribed (e.g. *pente* “comb”, *dedo* “finger”). The scoring is 1 correct, 0 not correct.

4.1.2. Lexical Reception (LexR)

The subtask on lexical reception contains 40 items that include only nouns (as in the expressive subtask). Here, the experimenter reads the Portuguese target word and the child needs to select one out of five pictures that are presented in a book (including several distractors: near semantic, distant semantic, visual, not related, see table 1). The scoring is: 1 correct, 0 not correct.

Table 1: Example of the subtask LexR (test item *cenoura* “carrot” with distractors)

<table>
<thead>
<tr>
<th>Test item</th>
<th>Near semantic distractor</th>
<th>Distant semantic distractor</th>
<th>Visual distractor</th>
<th>Not related distractor</th>
</tr>
</thead>
<tbody>
<tr>
<td>cenoura</td>
<td>couve</td>
<td>limão</td>
<td>serrote</td>
<td>machado</td>
</tr>
<tr>
<td>(“carrot”)</td>
<td>(“green cabbage”)</td>
<td>(“lemon”)</td>
<td>(“saw”)</td>
<td>(“axe”)</td>
</tr>
</tbody>
</table>


4.1.3. Morpho-Syntactic Production (MorSynP)

The task on morpho-syntactic production is a sentence repetition task. The child has to repeat 36 sentences that are read by a native speaker. The items can be divided into three sentence types: 12 sentences in active voice (progressive form), 12 sentences with comparative construction and 12 sentences in passive voice. All sentences are in present tense and contain verbs in third person singular contexts only. Examples for each sentence type are shown in (1-3).

The active voice sentences (1) are all in gerund form which means in Portuguese a construction with the auxiliary estar, the preposition a and the verb as infinitive. The comparative construction (2) contains the copula verb estar + the numeral mais + Adj + a conjunction do que. The sentences in passive voice (3) are formed with an inflected form of the verb ser + past participle + preposition. Note, that static passive is formed with the verb estar + past participle, and the dynamic passive with the verb ser. In these Portuguese sentences we have gender and number inflections on the article, on the adjective and on the past participle. For this task the exact repetition is scored, disregarding phonetic or phonological errors. The scoring is 1 correct, 0 not correct (one or more errors per sentence).

1. sentence in active voice (present progressive):
   O homem está a seguir o cão.
   The man is following the dog.

2. comparative construction
   O cão está mais pequeno do que a mulher.
   The dog is smaller than the woman.

3. sentence in passive voice (verbal passive)
   O cavalo é caçado pela (= por a) mulher.
   The horse is hunted by the woman.
4.1.4. Morpho-syntax Reception (MorSynR)

In the morpho-syntactic reception task, the child has to select one out of three pictures matching a sentence read by the experimenter. The sentence types are the same as in the morpho-syntactic production task. In addition, the task contains two different infinitive constructions. It has 60 items that are divided in 20 active (progressive), 12 comparative, 12 passive and 16 infinite constructions (Adj. + Infinitive and Verb + Infinitive). The scoring is: 1 correct, 0 not correct.

4.1.5. Phonology (Phon)

The phonology task is a pseudo-word repetition task with 30 one- to three-syllable words. These items are similar to Portuguese real words and are divided into 10 monosyllabic words (e.g. neite), 10 disyllabic words (e.g. ardol) and 10 trisyllabic words (e.g. ijuda). The productions of the children were analyzed acoustically by a native speaker, including correct repetition regarding the number of syllables, native Portuguese pronunciation of vowels, consonants (including voicing differences) and word stress. In the evaluation every dialectal possible variation was considered correct. The scoring is: 1 correct, 0 not correct.

4.1.6. Parental questionnaire (PaBiQ)

All background information on the circumstances of language acquisition and use, as AoO, LoE, amount and quality of input, was inferred from the parental questionnaire PaBiQ (COST Action IS0804 2011, see also Tuller 2015) that was conducted as an interview with the mother. The PaBiQ asks for information about many factors that might influence the language development of the bilingual child. For the present study we concentrated on six factors: age of onset to the ML (AoO), lengths of exposure to the ML (LoE), input quality, early input quantity, current input quantity, and language dominance. The first two of these factors can be taken directly from the PaBiQ. The other four factors are composite values that we calculated on the basis of ratings and answers to various questions. This procedure and the weighting for each of these factors are explained in 5.2. The language dominance factor is a composite factor using a calculation over different PaBiQ information.
items which were developed in the BiLaD-project by the research team in Tours (see footnote ii, and De Almeida et al. 2015). For the calculation of the dominance factor, for both languages the scores for AoO and LoE, of early exposure, early estimated language use, current language use at home and language use in other contexts are added up. Then the sum of the HL is subtracted from that of the ML resulting in a so-called dominance index (see 5.2 and table 7 for further details).

4.2. Subjects

14 Portuguese-German children aged between 5;9 and 8;7 participated in the study. These 14 children can be divided into two groups (see table 2). Seven children are simultaneous bilinguals (2L1) and all acquired both languages from birth. The other seven children are sequential bilingual children and all started to learn German around the age of three. Consequently, the length of exposure to the ML German is longer for the 2L1 children than for the cL2 children (see 5.2). Even though the groups have the same age ranges, the mean age of the 2L1 children is higher than that of the cL2 children. However, there is no statistical difference in age between the two groups (Mann-Whitney-U, \( p=.318 \)). The 2L1 children were 5 boys and 2 girls, the cL2 children were 4 girls and 3 boys.

Table 2: Subject information (AoO = Age of Onset, LoE= Length of exposure).

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Age in months</th>
<th>Mean age in months (SD)</th>
<th>AoO German in months</th>
<th>LoE German in months (SD)</th>
<th>L1</th>
</tr>
</thead>
<tbody>
<tr>
<td>2L1</td>
<td>7</td>
<td>70-101</td>
<td>86 (13)</td>
<td>0</td>
<td>86 (13)</td>
<td>Portuguese &amp; German</td>
</tr>
<tr>
<td>cL2</td>
<td>7</td>
<td>69-103</td>
<td>77 (11)</td>
<td>35,6</td>
<td>41 (11)</td>
<td>Portuguese</td>
</tr>
</tbody>
</table>
5. Results

In this section we first present the results in the five different subtasks of the Palpa-P comparing the groups (5.1) and in 5.2. we present an analysis of the input factors and the possible interactions with the test results to see how input factors contribute to the performance in the heritage language of the children.

5.1. Results by different domains and groups

Figure 1 shows the results of the simultaneous and sequential bilinguals in the five subtasks of the Portuguese standardized task Palpa-P. They are presented in z-scores which enable us to see how much the children differ from the monolingual group norms. Using these standardized scores, the effects of age differences between the children should disappear. For monolinguals, a z-score of -1.25 SD would be the cut-off for language impairment used by clinicians (Tomblin et al. 1996; Thordardottir 2012, 2015). We suppose that the children’s results are deviant from the monolingual norm if they have z-scores below -1.25 SD. As we compare bilinguals to a monolingual norm, we cannot draw any conclusions on language impairment. Another reason to be careful with the results is that there were some age gaps in the Palpa-P norms in some of the subtasks. In each of LexR, Phon and MorSynP norms are missing for two year-groups. In these cases we decided to apply the age norm of the age group that was closest to the age of the child, or, in case of equal intervals we applied the norms of the next younger age group. A highly problematic case was the subtask MorSynR which is normed on children age 3-5 and on adults, thus, lacking norms for the age range from 6 to 9. In this case we decided to use the norm of age 3-5 for all children. Another important point is that the relevant subtasks were normed on few children (10-20 per task) so far except for the MorSynR which is normed on 134 children in the younger group (age 3-5). Consequently, the z-scores appear to be extreme in some cases. Despite these problems, which will disappear once the authors have fully developed the test, we choose to use this task because it is linguistically well constructed and is used by clinicians.

As a first result, we see that in two subtasks both groups performed in the range of unimpaired monolingual children, i.e. in morpho-syntactic reception and in phonology. This suggests that these child heritage speakers have acquired the phonology of Portuguese to a competence
similar to monolingual children of the same age and comprehend morpho-
syntactic structures as well as monolingual children aged 3-5. As far as
these subtasks are able to capture deviance from monolingual norms, no
signs of attrition or incomplete acquisition were visible. There is no
difference between the 2L1 children and the cL2 children in these tasks
(for MorSynR $p=.456$ and for Phonology $p=.902$).

A different picture emerges from the three remaining subtasks, which
target the lexicon and morpho-syntactic production. In the LexR task only
the successive bilinguals performed in the range of the monolingual norm
whereas the 2L1 children have a mean z-score of -4.05. In contrast, the
tasks MorSynP and LexP were difficult for both groups (both groups with
z-scores lower than -1.25 SD). However, this difficulty was more
pronounced in the 2L1 group. In both lexical subtasks the cL2 children
performed significantly better than the 2L1 children (for LexP, $p=.038,$
for LexR $p=.007$). In MorSynP the difference was only marginally
significant ($p=.053$).

Summarizing the group results in the Palpa-P in terms of z-scores, we
identify two unaffected grammatical domains in both groups, i.e. morpho-
syntactic reception and phonological production. In contrast, two domains,
namely production in morpho-syntax and production in the lexicon, seem
to be vulnerable for all heritage speakers. Lexical reception was
problematic only for 2L1 children, not for cL2 children. However lexical
reception will be considered as third vulnerable domain. In these three
vulnerable domains the cL2 children outperformed the 2L1 children.

![Figure 1. Group results in z-scores.](image-url)
5.2. Input factors

In the present paper we use six factors identified on the basis of the mother's answers and ratings in the PaBiQ: age of onset to the ML (AoO), length of exposure to the ML (LoE), early input quality, early input quantity, current input quantity, and language dominance.

Information on the first two factors AoO and LoE is presented in Table 3. As the AoO of the ML German was the factor separating the groups, it is no surprise that the groups are significantly different which is also true for LoE to ML (for both factors, Mann-Whitney-U, p=.001). As mentioned before, the 2L1 children all started to acquire German from birth and the cL2 children, in contrast, around the age of three.

Table 3: Age, AoO=Age of Onset and LoE= Length of Exposure to ML.

<table>
<thead>
<tr>
<th>Group</th>
<th>Code</th>
<th>Age</th>
<th>Age in months</th>
<th>AoO of the ML</th>
<th>LoE to the ML</th>
</tr>
</thead>
<tbody>
<tr>
<td>2L1</td>
<td>TD1BP</td>
<td>5;10</td>
<td>70</td>
<td>0</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>TD5BP</td>
<td>7;4</td>
<td>88</td>
<td>0</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>TD6BP</td>
<td>6;5</td>
<td>77</td>
<td>0</td>
<td>77</td>
</tr>
<tr>
<td></td>
<td>TD7BP</td>
<td>8;3</td>
<td>99</td>
<td>0</td>
<td>99</td>
</tr>
<tr>
<td></td>
<td>TD9BP</td>
<td>8;5</td>
<td>101</td>
<td>0</td>
<td>101</td>
</tr>
<tr>
<td></td>
<td>TD11BP</td>
<td>8;1</td>
<td>97</td>
<td>0</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>TD12BP</td>
<td>5;10</td>
<td>70</td>
<td>0</td>
<td>70</td>
</tr>
<tr>
<td></td>
<td>TD4BP</td>
<td>5;10</td>
<td>70</td>
<td>36</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>TD13BP</td>
<td>6;2</td>
<td>74</td>
<td>36</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>TD14BP</td>
<td>6;2</td>
<td>74</td>
<td>36</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>TD16BP</td>
<td>6;5</td>
<td>77</td>
<td>36</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td>TD18BP</td>
<td>6;0</td>
<td>72</td>
<td>36</td>
<td>36</td>
</tr>
<tr>
<td></td>
<td>TD19BP</td>
<td>8;7</td>
<td>103</td>
<td>36</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>TD20BP</td>
<td>5;9</td>
<td>69</td>
<td>33</td>
<td>36</td>
</tr>
</tbody>
</table>

The third factor early input quality (see Table 4) includes two types of information. On the one hand, it includes the Portuguese level of the mother / father which is based on self-estimated competence in Portuguese. The interviewed parent (here, the mother) had to rank the competence of both parents (mother and father) according to five possible scores from 0-4 (0= only a few words, 1= gets along, but with difficulty, 2= basic abilities (gets along), 3=well, 4=very well). In the cL2 group the
mothers of all children reported that both parents had very good abilities in Portuguese. In contrast, in the 2L1 group this was only true for two children. For the other five children the mothers reported that only one parent spoke Portuguese very well, whereas the Portuguese level of the other parent ranged from 0-3. On the other hand, in the input quality value we included a migration score ranging from 1 to 3. Here, 1 means German monolingual (L2 speaker of Portuguese), 2 means second generation migrant (the parent was born in a Portuguese family in Germany) and 3 means first generation migrant (born and raised in Portugal, migration to Germany after age of 12). The 2L1 group included more mixed couples with one of the parents being raised monolingually in German with only basic abilities in Portuguese. In the cL2 group only the twin siblings (TD13BP & TD14BP) had a German native father. Nevertheless, the mother reported that he spoke Portuguese as L2 very well and the family had lived in Portugal for a long time until the children were 5 years old. The 4 scores were added, resulting in a composite value input quality with a maximum of 14 which we expect to be the best input quality. All cL2 children reach a score close to the maximum (≥12). In contrast, more than half of the 2L1 children have an input quality score ≤10 and the group is more heterogeneous than the cL2 children. However, the groups do not differ significantly with regard to early input quality (Mann-Whitney-U, p=.097).
The fourth factor is the *early input quantity* (table 5). This composite score consists of two types of information. The first sub-score is based on the *mother’s estimation on amount of Portuguese input* to the child before the age of 4:0 (0= never, 1= seldom, 2= half of the time, 3 = often, 4 = always). For all but one child in the cL2 group the mother reported that the children were exposed to Portuguese often. In contrast, in the 2L1 group this is only true for one child. In all other cases, the mothers reported an exposure to Portuguese half of the time or seldom. Note, that for none of

<table>
<thead>
<tr>
<th>Group</th>
<th>Code</th>
<th>Português level mother (max. 4)</th>
<th>Português level father (max. 4)</th>
<th>Migrat. mother (max. 3)</th>
<th>Migrat. father (max. 3)</th>
<th>Composite value input quality (max. 14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2L1</td>
<td>TD1BP</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>TD5BP</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>TD6BP</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>TD7BP</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>TD9BP</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>TD11BP</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>TD12BP</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>cL2</td>
<td>TD4BP</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>TD13BP</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>TD14BP</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>TD16BP</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>TD18BP</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>TD19BP</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>TD20BP</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>12</td>
</tr>
</tbody>
</table>
the families the child had been always (4 points) exposed to Portuguese < age 4;0 (this is not surprising, because in the cL2 group all of the children went to a German kindergarten from the age of 3).

Table 5: Early input quantity Portuguese.

<table>
<thead>
<tr>
<th>Group</th>
<th>Code</th>
<th>Early estimated Port. use (max. 4)</th>
<th>Early contexts Port. use (max. 3)</th>
<th>Composite Port. input quantity (max. 7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2L1</td>
<td>TD1BP</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>TD5BP</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>TD6BP</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>TD7BP</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>TD9BP</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>TD11BP</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>TD12BP</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>cL2</td>
<td>TD4BP</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>TD13BP</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>TD14BP</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>TD16BP</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>TD18BP</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>TD19BP</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>TD20BP</td>
<td>2</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

The second sub-score of early input quantity is the *early context score for Portuguese*. Out of 7 possible contexts (mother, father, grandparents, babysitter, other adult, siblings, day care and kindergarten) the mother indicated which of these contexts was present before the age 4;0 and which language was spoken in this context to the child. As not all the contexts applied for each child (e.g. some children had no siblings) we calculated the relative use of Portuguese adjusted to the individual number of applicable contexts. Note that the percentage refers to the number of contexts where Portuguese was used and gives no information on the amount of Portuguese in each context. This percentage was converted into a context score for Portuguese from 1-3 ($1 \leq 33.3; 2 = 34-66.7; 3 \geq 67$). The cL2 children used Portuguese in almost all contexts (80-100%). In contrast, the 2L1 children used Portuguese only in 29-67% of the contexts.
The composite score early input quantity sums up the language use and the context score with a maximum of 7 points. A group comparison shows that this early input quantity score is significantly lower in the 2L1 than in the cL2 group (Mann-Whitney-U, p=.001).

The fifth factor, current input quantity (table 6), is divided into Portuguese use at home, Portuguese use with friends, Portuguese media use and schooling in Portuguese. Portuguese use at home includes 4 possible contexts (with mother, father, other adult and siblings) where a score from 0 to 4 is possible (0=never, 4=always), resulting in a maximum score of 16 points. As not all contexts were applicable for each child, we calculated the percentage and then assigned a value from 1 to 4 (1≤25, 2≤50, 3≤75, 4≤100). In the cL2 group Portuguese was used at home in more than 50% of the contexts. Only for two children the current Portuguese use at home was less than 20%. In contrast, most of the 2L1 children used Portuguese in less than 51% except for two children. Portuguese use with friends includes two contexts (with peers and family friends) with a score from 0-4 (0=never, 4=always). To get an equal weighting for input at home and input from friends the sum was then divided by two, resulting in a maximum of 4 points. On average, Portuguese was only seldom used with friends, in some cases half of the time, however. The two bilingual groups look very similar in this respect. Portuguese media use includes language activities like reading, watching films, storytelling, resulting in a maximum of 6 points which were divided by two as well so that it had not more weight than the other scores. Here, only three children got a score higher than 2. Schooling in Portuguese has three categories: 0= no schooling, 1= heritage school once a week (several hours) or 2= bilingual Portuguese-German school (several subjects are taught in Portuguese). Interestingly, the only three children who went to a Portuguese-German bilingual school since preschool were in the 2L1 group. Half of the children in each group had no schooling in Portuguese at all. The rest of the children went to a Portuguese heritage school once a week. The three values were combined to the composite value current input quantity with a maximum of 13 points that might be reached by a monolingual child. Both groups show low values with only 3 children getting more than 7 points which mirrors the status of Portuguese as minority language in Germany. The groups show similar current input quantity scores (Mann-Whitney-U, P=.902).
The final factor number six *language dominance* is a broader composite score based on *PaBiQ* information as well (see De Almeida et al. 2015). It comprises several of the aforementioned factors, except for *input quality* and information on schooling. It includes information on *age of onset, length of exposure, early input quantity* (estimated amount of input before age 4 and number of contexts for early exposure before age 4) and *current input quantity* (language use at home and language use in other contexts) for both languages, combined to so called *exposure indices* for the ML (German) and for the HL (Portuguese). The exposure index of Portuguese is then subtracted from that of German, resulting in a *Language Dominance Index* (>5 = German dominant, < -5 = L1 dominant, ≤5 and ≥ -5 = balanced). In the 2L1 group most children were German dominant or balanced with only one exception. In contrast, in the cL2 group most children were L1 (Portuguese) dominant (see table 7). This difference does not reach significance (Mann-Whitney-U, p=.073).

Table 6: Current input quantity Portuguese.

<table>
<thead>
<tr>
<th>Group</th>
<th>Code</th>
<th>Port. use at home (max. 4)</th>
<th>Port. use with friends (max. 4)</th>
<th>Port. media use (max. 3)</th>
<th>Schooling in HL (max. 2)</th>
<th>Composite value current input quantity (max. 13)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2L1</td>
<td>TD1BP</td>
<td>3</td>
<td>0,5</td>
<td>1,5</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>TD5BP</td>
<td>2</td>
<td>0,5</td>
<td>1</td>
<td>1</td>
<td>4,5</td>
</tr>
<tr>
<td></td>
<td>TD6BP</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>TD7BP</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>TD9BP</td>
<td>3</td>
<td>2,5</td>
<td>2,5</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>TD11BP</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>TD12BP</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>cL2</td>
<td>TD4BP</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>TD13BP</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>TD14BP</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>TD16BP</td>
<td>3</td>
<td>2,5</td>
<td>3</td>
<td>0</td>
<td>8,5</td>
</tr>
<tr>
<td></td>
<td>TD18BP</td>
<td>3</td>
<td>2,5</td>
<td>1</td>
<td>0</td>
<td>6,5</td>
</tr>
<tr>
<td></td>
<td>TD19BP</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>TD20BP</td>
<td>1</td>
<td>0,5</td>
<td>0,5</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>
Table 7: Language dominance.

<table>
<thead>
<tr>
<th>Group</th>
<th>Code</th>
<th>Language dominance index</th>
<th>Dominance</th>
</tr>
</thead>
<tbody>
<tr>
<td>2L1</td>
<td>TD1BP</td>
<td>0</td>
<td>Balanced</td>
</tr>
<tr>
<td></td>
<td>TD5BP</td>
<td>13</td>
<td>German</td>
</tr>
<tr>
<td></td>
<td>TD6BP</td>
<td>24</td>
<td>German</td>
</tr>
<tr>
<td></td>
<td>TD7BP</td>
<td>18</td>
<td>German</td>
</tr>
<tr>
<td></td>
<td>TD99BP</td>
<td>-8</td>
<td>L1</td>
</tr>
<tr>
<td></td>
<td>TD11BP</td>
<td>2</td>
<td>Balanced</td>
</tr>
<tr>
<td></td>
<td>TD12BP</td>
<td>5</td>
<td>Balanced</td>
</tr>
<tr>
<td>cL2</td>
<td>TD4BP</td>
<td>-9</td>
<td>L1</td>
</tr>
<tr>
<td></td>
<td>TD13BP</td>
<td>-12</td>
<td>L1</td>
</tr>
<tr>
<td></td>
<td>TD14BP</td>
<td>-12</td>
<td>L1</td>
</tr>
<tr>
<td></td>
<td>TD16BP</td>
<td>-17</td>
<td>L1</td>
</tr>
<tr>
<td></td>
<td>TD18BP</td>
<td>-11</td>
<td>L1</td>
</tr>
<tr>
<td></td>
<td>TD19BP</td>
<td>16</td>
<td>German</td>
</tr>
<tr>
<td></td>
<td>TD20BP</td>
<td>18</td>
<td>German</td>
</tr>
</tbody>
</table>

To see an influence of the aforementioned input factors (including the group defining factors AoO and LoE) we ran correlation analyses with the results in z-scores. This enables us to compare the children despite age differences. This correlation was run on all children without separating the groups (see table 8).

Table 8: Spearman correlations of input factors and results in z-scores (AoO= Age of Onset, LoE= Length of Exposure).

<table>
<thead>
<tr>
<th></th>
<th>LexP</th>
<th>LexR</th>
<th>MorSynP</th>
<th>MorSynR</th>
<th>PhonP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AoO German</td>
<td>r = .601, p = .023*</td>
<td>r = .747, p = .002**</td>
<td>r = .581, p = .029*</td>
<td>r = .317, p = .269</td>
<td>r = .115, p = .694</td>
</tr>
<tr>
<td>LoE German</td>
<td>r = -.702, p = .005**</td>
<td>r = -.681, p = .007**</td>
<td>r = -.696, p = .006**</td>
<td>r = .060, p = .840</td>
<td>r = .143, p = .625</td>
</tr>
</tbody>
</table>
We find a positive correlation of age of onset and a negative correlation of length of exposure with LexP, LexR and MorSynP. The input quality is positively correlated with LexP and MorSynP. The early input quantity is highly positively correlated with both lexical tasks and moderately with both morpho-syntactic tasks. The current input quantity has almost no effect but is moderately correlated with PhonP (in this task the groups do not differ and perform within the monolingual range, therefore this result can be disregarded). Finally, we have a high negative correlation of dominance with LexP, LexR and MorsynP. The results of this correlation analysis maybe summarized as follows:

i) the later the contact to the L2 German, the better the performance in the Portuguese lexical production and reception and in morpho-syntactic production (conform with Montrul 2008)

ii) the longer the exposure to the L2 German, the worse the performance in lexical production and reception and in morpho-syntactic production in Portuguese

iii) the better the early input quality, the better the performance in lexical production and morpho-syntactic production

iv) the higher the amount of input in Portuguese before age 4, the better the results in LexP, LexR, MorSynP

v) the higher the language dominance index (dominance in direction of the ML German), the worse the performance in LexP, LexR and MorSynP

<table>
<thead>
<tr>
<th></th>
<th>( r = .749, )</th>
<th>( r = .395, )</th>
<th>( r = .562, )</th>
<th>( r = .157, )</th>
<th>( r = -.196, )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( p = .002** )</td>
<td>( p = .162 )</td>
<td>( p = .037* )</td>
<td>( p = .592 )</td>
<td>( p = .502 )</td>
</tr>
<tr>
<td>Early Input Quantity</td>
<td>( r = .707, )</td>
<td>( r = .741, )</td>
<td>( r = .539, )</td>
<td>( r = .537, )</td>
<td>( r = .090, )</td>
</tr>
<tr>
<td></td>
<td>( p = .005** )</td>
<td>( p = .002** )</td>
<td>( p = .047* )</td>
<td>( p = .048* )</td>
<td>( p = .760 )</td>
</tr>
<tr>
<td>Current Input Quantity</td>
<td>( r = .185, )</td>
<td>( r = .145, )</td>
<td>( r = .160, )</td>
<td>( r = .525, )</td>
<td>( r = .628, )</td>
</tr>
<tr>
<td></td>
<td>( p = .527 )</td>
<td>( p = .621 )</td>
<td>( p = .585 )</td>
<td>( p = .054 )</td>
<td>( p = .016* )</td>
</tr>
<tr>
<td>Dominance</td>
<td>( r = -.754, )</td>
<td>( r = -.647, )</td>
<td>( r = -.723, )</td>
<td>( r = -.416, )</td>
<td>( r = -.370, )</td>
</tr>
<tr>
<td></td>
<td>( p = .002** )</td>
<td>( p = .012* )</td>
<td>( p = .003** )</td>
<td>( p = .139 )</td>
<td>( p = .193 )</td>
</tr>
</tbody>
</table>

We find a positive correlation of age of onset and a negative correlation of length of exposure with LexP, LexR and MorSynP. The input quality is positively correlated with LexP and MorSynP. The early input quantity is highly positively correlated with both lexical tasks and moderately with both morpho-syntactic tasks. The current input quantity has almost no effect but is moderately correlated with PhonP (in this task the groups do not differ and perform within the monolingual range, therefore this result can be disregarded). Finally, we have a high negative correlation of dominance with LexP, LexR and MorsynP. The results of this correlation analysis maybe summarized as follows:

i) the later the contact to the L2 German, the better the performance in the Portuguese lexical production and reception and in morpho-syntactic production (conform with Montrul 2008)

ii) the longer the exposure to the L2 German, the worse the performance in lexical production and reception and in morpho-syntactic production in Portuguese

iii) the better the early input quality, the better the performance in lexical production and morpho-syntactic production

iv) the higher the amount of input in Portuguese before age 4, the better the results in LexP, LexR, MorSynP

v) the higher the language dominance index (dominance in direction of the ML German), the worse the performance in LexP, LexR and MorSynP
We can conclude that several input factors affect the performance in the heritage Portuguese of the bilingual children, except for the current input quantity which seems to be irrelevant for the domains tested here.

A correlation analysis with the factor age (in months) and absolute values reached in the tests, results in only one significant correlation, i.e. age with PhonP (see table 9). Thus, in the given data set, the test results did not improve with age, except for the performance in the pseudo-word repetition task, in which the children performed within the monolingual range (so it can be disregarded).

Table 9: Spearman correlations, age and results in absolute numbers.

<table>
<thead>
<tr>
<th></th>
<th>LexP</th>
<th>LexR</th>
<th>MorSynP</th>
<th>MorSynR</th>
<th>PhonP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>$r=.493$, $p=.073$</td>
<td>$r=.169$, $p=.564$</td>
<td>$r=.333$, $p=.245$</td>
<td>$r=.495$, $p=.072$</td>
<td>$r=.619$, $p=.018^*$</td>
</tr>
</tbody>
</table>

As dominance seems to be an important factor that includes several relevant sub-factors, we regroup the children according to dominance resulting in a group of 5 German dominant, a group of 6 Portuguese dominant and a group of 3 balanced children. We see that the German dominant group has the lowest z-scores, the Portuguese dominant group has the highest z-scores and the balanced children are in the middle for all subtasks. We find no statistic difference for any of the tasks between German dominant and balanced children and between balanced and Portuguese dominant children ($p>.05$). However, if we compare German dominant to Portuguese dominant children we have a significant difference in LexP and MorSynP (for both $p=.017$) and a marginal difference in LexR ($p=.052$).
Summary & Conclusion

Summarizing the results, we review them for each of the research questions separately:

1) Does the HL of 2L1 Portuguese-German children differ from that of cL2 Portuguese-German children (AoO of ML German age 3;0)?

To answer this question we compared the performance on the Palpa-P, a Portuguese standardized test, between the two groups of heritage speakers, i.e. simultaneous and sequential bilinguals. We expected the cL2 children to outperform the 2L1 children. The results of the Palpa-P demonstrate a clear difference between the two groups. With respect to morpho-syntactic production as well as lexical production and reception the cL2 children outperformed the 2L1 speakers. We can confirm the Hypothesis 2 of Montrul (2008) that the HL is more affected in simultaneous than in sequential bilinguals.

2) Which language domains are most affected in Portuguese heritage children in Germany?

To answer this question we ran a quantitative analysis of five different subtasks of the Palpa-P. We expected lexicon and morpho-syntact to be more affected than phonology. Actually, we found that production in
morpho-syntax and the lexicon are most affected in both groups, while lexical reception seems to be problematic only for 2L1 children. This confirms other studies which found problems in the areas of the lexicon and morpho-syntax (Silva-Corvalán 2014 a, b; Anderson 1999, 2001, Klassert et al. 2014). In contrast, the area of phonology is not affected in either group. This corroborates the claim of Montrul (2010) that phonetics/phonology might be less affected. It does not confirm the studies who found a phonetic influence during childhood. This might be due to different methods, however.

3) How do the children of the two groups differ with respect to input factors?

The groups were compared with respect to the six factors AoO, LoE, input quality, early input quantity, current input quantity and language dominance based on information from the parental questionnaire PabiQ. AoO of German was the criterion defining the two groups which also results in a crucial difference in LoE to German. Based on this difference between the groups we also expected a relevant difference in early input quantity and quality. However, the current input quantity and dominance might be similar in both groups due to a shift to more German input after the children have started to attend kindergarten. As expected, the two groups differ in the early input quantity. The input quality differed in that the children of the 2L1 group reached scores from 8 to 14, and the cL2 group reached 12 to 14 (out of 14). We find a qualitative (although not significant) difference indicating that the Portuguese input of at least some of the 2L1 children was qualitatively poorer than that of all of the cL2 children. As expected, the groups did not differ in current input quantity and current language dominance. However, in the cL2 children there were more Portuguese dominant children than in the 2L1 group, where most children were German dominant or balanced. This might indicate that the dominance shift towards the ML did not happen yet.

4) What is the impact of different input factors on the performance in the HL in different domains?

We ran correlation analyses between the input factors and the test results in the five subtasks. We expected all input factors to have an effect on the lexical tasks in Portuguese. As expected, we found a highly positive correlation between age of onset to the ML and the performance in lexical perception and production which means that, given a later AoO of
German, the performance in these Portuguese tasks on lexicon was better. A longer period of monolingual acquisition thus allowed for the elaboration of the lexicon in the HL. The same is valid for the results in morpho-syntactic production. Again, a longer monolingual period supported the elaboration of productive syntactic competence in the HL. Additionally, the early input quantity affected all tasks except for phonology. The cL2 children were exposed more to Portuguese and used Portuguese in more contexts (early input quantity) than the simultaneous bilingual children and showed higher scores. The quality of the input (parents were first or second generation migrants with differences in Portuguese competence) did only affect expressive lexical and morpho-syntactic skills. In other words, all factors reducing the early input had an impact on at least the expressive skills in lexicon and morpho-syntax. However, the comparison of groups by dominance showed that the Portuguese dominant children outperformed the German dominant children in the three aforementioned vulnerable areas. This composite factor included the current input quantity as well.

We investigated bilingual acquisition in a setting where the HL and the ML differ with respect to AoO. To conclude, the results support the claim by Montrul (2008, 60) that attrition will be more severe in simultaneous bilinguals compared to sequential bilinguals. This holds especially for the lexical tasks. These were clearly more problematic for the 2L1 than for the cL2 children. This result may be attributed to the input differences induced by the divergent AoO. Whether the results can be interpreted as signs of attrition or as incomplete acquisition cannot be decided on the basis of our data. Our study has shown that there are crucial differences in the input situation between the two types of bilingual child learners. The group defining differences (in terms of AoO and LoE to the ML German) first result in clear quantititative differences of input. Furthermore, the groups differ in homogeneity with respect to the input quality. The children in the cL2 group who were raised monolingually for the first years received input of high quality, the parents themselves being competent speakers of Portuguese and born in Portugal (except for one father of twins who was an advanced L2 learner of Portuguese and spoke Portuguese to the children). This was not the case for all families of the simultaneous bilinguals. In this group, we find more mixed marriages, i.e. one parent was German, or the parents were both born in Germany and were now German dominant. In contrast, this result is only preliminary because of the small groups of children, and the difference between the two groups with respect to input quality might be an artifact of the composition of the small groups.
In contrast to the results of previous 2L1 studies, our results indicate some disadvantages of the 2L1 children in comparison to monolinguals and cL2 speakers. Does this speak against the “earlier is better”-view (e.g. Abrahamsson & Hyltenstam 2009) because cL2 speakers have advantages in the development of the HL? It has to be noted that early acquired phenomena such as phonology were unproblematic in the bilinguals investigated in the present paper. In contrast, the morpho-syntactic task, including late acquired structures such as passives, was especially difficult for some of the children. Vieira (2011) showed that Portuguese monolinguals at age 5;6-6;0 performed still below 10% accuracy on verbal passives. In a future study a detailed error analysis in all subtasks has to be done to search for eventually existing differences between early core grammar and late acquired other grammatical phenomena (cf. Tsimili 2014).

To establish whether AoO or LoE to the ML is more relevant for the maintenance of the HL, future studies should investigate 2L1 and L2 groups with the same LoE to the ML. It could be the case that sequential learners show signs of attrition at a later point in development whereas 2L1 children show a more severe incomplete acquisition at a younger age. So we would find more cases of attrition in cL2 children and more cases of incomplete acquisition in 2L1 children. We can presume that L2 children and 2L1 children acquire the L1 successfully until age 3-4 (in consent with Silva-Corvalán 2014 a, b), with the entry to kindergarten marking a point of dominance shift. However, as there might be a delay in 2L1 acquisition, already at an earlier age there might be quantitative differences between cL2 and 2L1 children because of differences in input quality and quantity. As mentioned above, the data presented in our study are not suited to answer questions dealing with the difference between attrition and incomplete acquisition.

Do the present findings have an impact on language policy and recommendations for parents? Should parents better raise their children monolingually until the age of 3;0 to avoid language loss in the heritage language? Or would this then be counterproductive for the acquisition of the second language? To find answers, in future research we will compare the results in the Portuguese tasks to the results in tasks in the ML German.
Notes

1 Many studies on adult 2L1 speakers do not differentiate between 2L1 or cL2 learners. Here we did not include studies on late L2 learners, L2 loss or returnees.
2 BiLaD: “Bilingual Language Development” is a joint project with C. Hamann, M. Rothweiler and S. Chilla as leading members in Germany and L. Tuller, P. Prévost and colleagues at the University of Tours in France. The project is supported by DFG (Germany) and ANR (France).
3 These words are categorized as monosyllabic by the authors, even though the final Schwa might be produced.
4 Z-scores represent the standard deviation from the mean of the corresponding monolingual age group.
6 Missing age groups in the Palpa-P norms: age 6 and 7 for LexR, and age 6 and 8 in MorSynP and Phon.
7 In Portugal the twins went to a German kindergarten between age 3;0 and 4;0.

References


ACQUIRING THE DENOTATION OF OBJECT DENOTING NOUNS

SUZI LIMA, PEGGY LI AND JESSE SNEDEKER

1. Acquisition of countability

The debate about whether the conceptual content of a noun determines what will count as an atom in grammar is a topic of interest for both formal semantics and developmental psychology. A series of studies suggests that natural atomicity is neither required nor necessary for grammatical counting. Acquisition studies suggest that until 7 years of age children count parts of individuals of a certain kind (e.g. pieces of forks) as if they were themselves individuals of that kind (e.g. individual forks; cf. Shipley & Shepperson 1990). The literature reports that this pattern persists in other tasks that do not necessarily involve counting, such as interpreting quantifiers (“both”, “every”, “more”, etc.). For example, children take a fork broken into three pieces to be “more forks” than two whole forks, and choose to touch each half of a broken fork when asked to “Touch every fork” (Brooks, Pogue & Barner 2011; Srinivasan et al. 2013). It is also reported in the literature that functionality affects children’s responses. For example, Giralt and Bloom (2000) have shown that 4-year-olds tend to not count parts as wholes if those parts have unique functions (wheels of a bicycle, ears of rabbits, wings for butterflies, etc). Srinivasan et al. 2013 also observed this effect in a task that involved novel words: they observed that the proportion of children counting parts as whole objects is much higher when they are not introduced with functions for the parts of the novel objects.

Srinivasan et al. (2013) argue that nouns do not fully specify units for quantification. Instead, under their analysis nouns provide only “partial conditions that are enriched pragmatically.” (Srinivasan & Barner 2016). According to the authors, children stop treating parts as wholes grammatically once they manage to access alternative descriptions (pseudopartitives and measure words: piece, part, half, slice, bit, portion, among many others). This proposal is partially motivated by the
Acquiring the Denotation of Object Denoting Nouns

pragmatics literature on scalar implicatures. Scalar implicatures studies have shown that children are more likely than adults to accept a description such as “some of the horses jumped over the fence” in truth value judgment tasks even when all horses jumped over the fence. However, when the same children are presented with a more precise alternative description – All of the horses jumped over the fence – in a forced choice task, the same children prefer the more precise description (Chierchia, Crain, Gualmini & Meroni 2001; Foppolo, Guasti & Chierchia 2012). These results have suggested that children might have difficulty with some forms of pragmatic inference (they do not activate spontaneously a contrast between a less and a more informative description). Srinivasan and colleagues propose a parallel between accessing pragmatic inferences for scalar implicatures and for interpreting notional count nouns such as chair and banana. Consistent with their proposal, they first showed that children begin to understand pseudopartitives at 3 and 4 years old. Next, they showed that priming 4- and 5-year-olds, slightly older children who have firmer grasp of the language, with descriptions contrasting parts and wholes (“There is a fork and a piece of a fork”) decreases children’s likelihood of counting parts when asked “How many forks?” Thus, similarly to scalar implicatures for constructions with quantifiers, being able to access some pragmatic inferences for notional count nouns possibly emerges late in acquisition – when children can spontaneously access alternative descriptions.

Based on the results of studies with children when interpreting object-denoting nouns, two hypotheses will be explored in this paper. First (H1), we argue that a proper semantic analysis of aforementioned acquisition facts requires the adoption of a theory of countability in which not only natural atoms but also their material parts belong to the extension of count nouns. To illustrate, both a whole banana and a part of a banana belong to the extension of the noun “banana”. Secondly (H2), with Srinivasan et al. (2013), we argue in favor of a blocking mechanism that prevents speakers to refer to parts of individuals using an unmodified count noun when pseudopartitive constructions or measure phrases are available to refer to these parts. Crucially, this mechanism only applies to count nouns that have natural atoms in their extensions, i.e. nouns for which atoms are stable across contexts. For these nouns, the adoption of blocking mechanism makes utterances with count nouns more informative, since it allows speakers to distinguish reference to natural atoms from reference to their parts. For nouns that have no natural atoms in their extension (e.g. fence) blocking does not increase informativity and therefore does not apply. One prediction of Srinivasan et al. (2013) is that in a language that
does not have productive partitives speakers will continue using bare nouns to describe pieces (and even groups) into adulthood. In this paper we explore evidence for this mechanism based on experimental studies with speakers of Yudja, a Tupi language spoken in Brazil that seemingly has no (pseudo)partitive constructions and measure phrases.

2. Yudja Community and Linguistic Background

2.1. The Yudja language

Yudja is a Tupi language spoken in the Xingu Indigenous Territory by approximately 300 people in Mato Grosso, Brazil. Yudja speakers only speak the Yudja language in the community (with few exceptions of families in an interracial marriage) and most adults are bilinguals or multilinguals. In the Yudja communities, the use of Brazilian Portuguese is restricted to interactions with people that are not part of the community or that do not speak Yudja. Yudja children are raised as monolinguals and start to learn Brazilian Portuguese around 6 years of age.

Yudja is a bare noun language, i.e., nouns are unspecified for number (singular, plural) and unspecified for definiteness (definite or indefinite). Therefore, nouns can be interpreted as singular or plural, definite or indefinite depending on the context:

(1) Ali ba’î i xu
    child paca eat
    ‘The child(ren) eat(s)/ate the/a/some paca(s)’
    Lit.: an undefined number of children eat(s)/ate an undefined number of pacas.

Second, unlike other languages such as English where determiners are sensitive to the count/mass distinction, in Yudja, to the best of our knowledge, there are no quantifiers that are restricted to either count or mass nouns. Take for example the quantifier itxibî ‘many’ that can be combined with notional count (2a) and notional mass (2b) nouns and is always interpreted as referring to a number of individuals:

(2) a. Itxibî pîza dju wî
    many  canoe  bring
    ‘(Someone) brought many canoes’
b. Itxibī asa dju wī
much flour bring

‘(Someone) brought many (portions of) flour’

#‘Someone brought a big portion of flour’

Third, only a small class of nouns ([+ human] nouns) can be optionally pluralized in the language:

(3) a. senahī ‘man/men’
    b. ijd’ai ‘women’
    c. kota ‘snake/snakes’
    d. pīkahai ‘chair’
    e. yukidī ‘salt’
    f. pakua ‘banana’

Fourth, as in Inuttut (Gillon 2012), Nez Perce (Deal in press) and Ojibwe (Mathieu 2012) all nouns can be directly combined with numerals (unlike most languages described in the literature where a container/measure phrase or classifier is required in order to combine a numeral with a mass noun):

4. Txabīu ba’ī wāna
   three pacain ran
   ‘Three pacas ran’

5. Txabīu ali wāna
   three child ran
   ‘Three children ran’

6. Txabīu Maria pīkahai īwa
   three Maria chair buy
   ‘Maria bought three chairs’

7. Maria txabīu yukidī apa
   Maria three salt drop/fall
   ‘Maria dropped three (portions of) salt’
Apeta ‘blood’ (substance; bodily fluid)
(8) Txabũ uda apeta wî
three someone blood bring
Context: ‘Someone brought three (portions of) blood’

Y’a ‘water’ (substance; liquid)
(9) Maria yauda y’a dju wî
Maria two water bring
Context: ‘Maria brought two (portions of) water’

It is important to note that Yudja has neither classifiers nor measure words (kilo, liter). Measure words, when used, are borrowed from Brazilian Portuguese. As for container phrases (e.g. two bottles of water), they are characterized by a container noun followed by a postposition he ‘in’, just as in locative phrases in Yudja:

Yukïdî ‘salt’ (substance; granulated)
(10) Maria txabũ yukïdî xãã he apa
Maria three salt bowl in drop/fall
‘Maria dropped three bowls of salt’
Lit.: Maria dropped three (portions of) salt in bowls

Apeta ‘blood’ (substance; body fluid)
(11) Txabũ uda apeta xãã he wî
three someone blood bowl in bring
‘Someone brought three bowls of blood’
Lit.: Someone brought three (portions of) blood in bowls

Y’a ‘water’ (substance; liquid)
(12) Maria yauda y’a karaha he dju wî
Maria two water bottle in bring
‘Maria brought two bottles of water’
Lit.: Maria brought two (portions of water) in bottles

Comprehension studies with Yudja children and adults have shown that these container constructions may be interpreted as locatives (cf. Lima 2012, Lima 2014).

The fifth and last property of the Yudja nominal system we will explore in this paper are the constructions with pseudopartitives. In Yudja there are two morphemes that describe a part/whole relation: atxa ‘round piece’ and akuata ‘long piece’. Based on recordings and elicitation tasks,
it was observed that only two nouns productively require these morphemes in order to describe a particular part/whole relation:

\[
\text{(13) a. Kania 'animal' Kania atxa ' (piece of) meat' }
\]

\[
\text{b. Epa 'tree' Epa atxa/akuata/itxukĩ ' (piece of) wood'.}
\]

In an elicitation questionnaire, Lima (2012, 2014) tested whether these morphemes could be combined with nouns beyond the nouns listed in (Table 1). Two participants were presented with 22 nouns along with the morphemes in study (atxa, akuata) as well as two other morphemes that do not indicate parts, but the state of an object/substance (itxa ‘liquid’, itxukĩ ‘granulated’). Then, they had to evaluate whether those combinations were acceptable or not:

Table 1: Classifier questionnaire: nouns tested

<table>
<thead>
<tr>
<th>Type of the noun</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Humans</td>
<td>ali ‘child’; senahĩ ‘man’</td>
</tr>
<tr>
<td>Animals</td>
<td>ba’ī ‘paca’; amia ‘monkey’; api: dog</td>
</tr>
<tr>
<td>Objects</td>
<td>pikaha ‘chair’; abeata ‘clothes’; wâ’ê ‘pan/ceramics’</td>
</tr>
<tr>
<td>Liquid substances</td>
<td>iya ‘water’; apeta ‘blood’; uã ‘oil’; awïla ‘honey’; yakuha ‘porridge (Yudja’s traditional drink); ikaha ‘fat’.</td>
</tr>
<tr>
<td>Nature substances</td>
<td>amana ‘rain’; eta ‘sand, beach’</td>
</tr>
<tr>
<td>Grains</td>
<td>puju ‘beans’; awatxi ’i ‘rice’</td>
</tr>
<tr>
<td>Granulated substances</td>
<td>asa ‘flour’; yukidĩ ‘salt’; asuka ‘sugar’</td>
</tr>
<tr>
<td>Massy substances</td>
<td>makua ‘cotton’</td>
</tr>
</tbody>
</table>

This elicitation questionnaire has shown that most combinations with atxa ‘round piece’ and akuata ‘long piece’ were rejected by both speakers. As such, these results suggest that these morphemes are lexicalized and productively used only with a few roots, but that overall are rarely used with most nouns (cf. Lima 2014, chapter 2 for details).

Thus, the Yudja pseudopartitive morphemes differ in key ways from ordinary classifiers as they are not obligatory in constructions with
numerals and they do not occur with most nouns in the language. This could explain their low frequency in texts, narratives and dialogues.

2.2. Summary

In this section we have presented a few properties of the Yudja nominal phrases. In sum, we saw that Yudja is a generalized bare argument language where only [+human] nouns can be optionally pluralized. We also saw that all nouns can be directly combined with numerals and count quantifiers (itxībī ‘many’), including notional mass nouns. Yudja is characterized as a non-classifier language where container phrases are optional in constructions with numerals. When they do occur, they do not determine the individuation units for counting in constructions with mass nouns. Instead they are optional in these constructions and may be interpreted as locatives (indicating where a partition of a kind is located). Likewise, the pseudopartitive-candidate words (atxa and akuata) seem to be lexicalized and are only productive with a few nouns.

The properties of Yudja listed above make this language an ideal candidate to test both Rothstein (2010)’s hypothesis on the context-dependency when defining atoms and the predictions of Srinivasan et al. (2013) proposal. First, if mass nouns can interact with the counting system, then contextually determined portions of kinds can be grammatically counted in Yudja. Considering that there is no grammatical distinction between count and mass nouns in Yudja, we could hypothesize that in the same way that parts of a mass-kind can be considered as atoms, parts of count-kinds would also be available for counting. Second, Yudja does not present a productive, grammaticalized system of blocking the use of bare nouns in order to refer to portions or parts: this language has no classifiers, container phrases are interpreted as locatives and pseudopartitives are highly unproductive. Given these facts, in the next section we test Hypothesis 1 (H1) and Hypothesis 2 (H2), repeated below:

H1: natural atoms and their material parts belong to the extension of count nouns. Thus, atoms are contextually determined (Rothstein 2010, Lima 2014).

H2: a blocking mechanism prevents speakers to refer to parts of individuals using an unmodified count noun when pseudopartitive constructions or measure phrases are available to refer to these parts
Srinivasan et al. 2013). If pseudopartitives or measure phrases are unproductive in a language, this blocking mechanism will not apply.

3. Studies

3.1. Methods

3.1.1. Counting task (Study 1): participants and protocol

The study was advertised in the village amongst adults and children. We administered the task to whoever wanted to participate. There was no particular motivation for selection by age as we expect children’s behavior to count parts to persist into adulthood. A total of 25 children (nine 5-to-6-year-olds, six 7-to-8-year-olds, six 9-year-olds, and four 10-year-olds) and 19 adults participated in this study. Participants were tested individually in a room in the Yudja’s local central school in the Tuba Tuba village. A local teacher known by the children was present to facilitate the testing when children were involved. Consent was obtained from all participants. In the case of children, besides children volunteering themselves, their parents had to provide formal consent. Parents were welcome to stay in the room along with the researcher and research assistant (Chadaha Juruna or Tawaiku Juruna).

3.1.2. Felicity judgment task (Study 2) and forced choice task (Study 3): participants and protocol

A new group of participants was recruited and tested individually in the same manner as Study 1. A total of 46 Yudja speakers participated in this study. Nineteen children were under 10 years old (five 3-to-5-year-olds, five 6-to-7-year-olds, and nine 8-to-9-year-olds; M=6.9 years old, stdev=1.9), nineteen were 10 to 16 years old (nine 10-to-11-year-olds, six 12-to-13-year-olds, four 13-to-16-year-olds; M=12 years old, stdev=1.9), and eight were 18 years old or older. Each participant was tested in two tasks and in the same order: felicity judgment task and forced choice task.
3.1.3. Counting task: methods

Study 1 tested whether Yudja children and adults differentiate broken vs. whole objects. Each participant answered 2 questions (one that included a question about an aggregate (shoes, clothes) and another that included a question about an object, fruit or animal (fish, banana, spoon). Children and adults had to count the total number of objects that was presented in a picture. In the example below (Figure 1) the participants were invited to say how many bananas the two men had.

Figure 1: Example of stimuli for counting task.

The goal of the task was to verify whether Yudja speakers would count parts as wholes. For example, in a context where the two men had one banana each (a total of two bananas in the context), one would say four bananas if a bare noun could be used to refer to parts. If not, then they would say two. We would expect that they would only count parts as wholes productively if parts can be contextually defined as atoms.

3.1.4. Felicity judgment task: methods

Participants were asked to act as “language teachers” by judging whether a non-native speaker appropriately described pictures with
sentences of Yudja. The participants were always encouraged to provide a better alternative description, if he/she judged there was a better description.

On each trial, participants were always asked the question “are there two x?” where x could be one of four nouns: paca (animal), canoe (artifact), shirt (artifact), and banana (fruit). With four nouns, there were thus four different “are there two x” questions. Each question was paired with four types of pictures (see Figure 2): control ‘no’, whole, group, and piece. The control no pictures depicted a number mismatch to the number mentioned in the question (e.g., one canoe). Participants were expected to reject these pictures. The whole pictures depicted whole objects that matched in number (e.g., two canoes) and served as control ‘yes’ pictures. The group pictures depicted individuals in groups, where the number of groups matched the number in the question (e.g., two groups of three canoes). Finally, the piece pictures depicted two pieces of one object (e.g., two halves of a canoe). If nouns provided individuation criteria that would restrict reference to natural atoms, speakers should only accept these descriptions for the whole pictures and reject the group and piece pictures.

Are there two canoes?

Control No | Whole Objects | Groups | Pieces
---|---|---|---

Figure 2: Example of critical stimuli for the Felicity Judgment Task: control no (condition 1), whole objects (condition 2), groups (condition 3), and cut pieces (condition 4).

### 3.1.5. Forced choice task: methods

Study 3 tested Yudja’s interpretation of possible pseudopartitive words using a picture-choice task. Each participant answered 16 questions, 2 questions per condition and 2 control questions. Participants heard either bare nouns (condition 1), or a noun followed by one of three
candidate-partitives (N atxa ‘round’ (condition 2); N akuata ‘long’ (condition 3); N itxukî ‘granulated’ (control)). As presented in Section 2, the morphemes atxa and akuata appear in a handful of lexicalized nouns and are the best candidates for (pseudo)partitives in the language. They were combined with four nouns (e.g., paca, potato, fish, tree). After the participants heard the nouns, they were asked to identify its referents from among 4 photos of an object in different states: whole; round piece; long piece; granulated.

(a) A’i de ba’î where paca ‘Where is the paca?’ (answer: top left picture)
(b) A’i de ba’î atxa where paca round piece ‘Where is the round piece of paca?’ (answer: top right picture)
(c) A’i de ba’î akuata where paca long piece ‘Where is the long piece of paca?’ (answer: bottom left picture)
(d) A’i de ba’î itxukî where paca granulated ‘Where is the granulated paca?’ (answer: bottom right picture)
Acquiring the Denotation of Object Denoting Nouns

(a) A’i de pitxa
where fish
‘Where is the fish?’
(answer: top right picture)

(b) A’i de pitxa atxa
where fish round piece
‘Where is the round piece of fish?’
(answer: bottom right picture)

(c) A’i de pitxa akuata
where fish long piece
‘Where is the long piece of fish?’
(answer: bottom left picture)

(d) A’i de pitxa itxukï
where fish granulated
‘Where is the granulated fish?’
(answer: top right picture)

Figure 3: Example of stimuli for the Forced Choice Task

In this task we tested whether the apparent unproductive use of the pseudopartitive words in Yudja (atxa and akuata) would be confirmed in a task where those morphemes were combined with nouns that they are not usually combined with such as atâu ‘potato’ and pitxa ‘fish’. If participants do not correctly associate the meaning of these morphemes to the expected picture, we would confirm that these morphemes are not productive in the language, which could just be an effect of the fact that bare nouns can denote parts.
3.2. Results

3.2.1. Counting task (Study 1): results

Most children counted parts as wholes (77% for aggregates, 73% for objects). A large number of adults counted pieces as wholes, without differentiating the whole object (e.g., the whole banana) from the pieces (the single banana cut in three pieces). One hypothesis to explain this pattern comes from the overall observation that partitive words are optional and bare nouns can be used to refer to both wholes and parts.

In the adults’ group, most participants counted objects as wholes (21% for count and 16% for aggregates). That is, in a photo where two men had two bananas, they said *yauda* ‘two’. Other adults (21% for count and 21% for aggregates) described the same scenario as someone having a whole object (which was the case for all pictures) and someone having a broken object. Crucially, only one participant out of this group used a ‘partitive’ to describe a shirt that was cut in three pieces:

(14) Akuata dju txabïu
    long piece with three
    ‘Three long pieces’

Most of the participants that counted the pieces and recognized them as such emphasized that the object was broken:

(15) Txabïu lakïrikïrï yãhã
    three tear NOM
    ‘Three torn [pieces]’
3.2.2. Felicity judgment task (Study 2)

Figure 4: Felicity Judgment Results. Adult and children’s acceptance of the numeral noun phrase for a given picture. Pictures were spatially temporally discrete individuals or groupings that matched in number and kind of object mentioned in the queried phrase (e.g., *yauda pîza*; “two canoe”). The “individual” pictures depicted individual whole objects (two canoes), the “piece” pictures depicted an individual cut into pieces (two pieces of a canoe), and the “group” pictures depicted groups of whole objects (two groups of three canoes).

How often participants accepted the numeral in combination with the noun for the four types of pictures is depicted in Figure 4. For simplicity, participants were divided into three groups by age. However, we treated age as a continuous variable in our analyses. The percentage of trials in which participants accepted the picture as matching the number+N question was entered into a generalized linear model with condition (control ‘no’, wholes (control ‘yes’), pieces, and groups) as a within-subjects variable and age as a covariate. There was a significant effect of Condition (*Wald χ²(3) = 34.67, p < .001*), indicating that acceptance varied across the four types of picture. All pairwise Wilcoxon signed rank comparisons indicate that with the exception of Group being almost significantly different than Control No (*p = .06* with Bonferroni corrections), all other pairwise comparisons were statistically significant.
In other words, whereas Control No and Group were low in acceptance, they contrasted with Whole and Piece. Whole and Piece in turn differed from each other. That is participants had a greater likelihood of accepting Whole pictures than Piece pictures. The generalized linear model also revealed that Age (Wald $\chi^2(1) = 8.81$, $p = .003$) and Condition by Age interaction (Wald $\chi^2(2) = 18.84$, $p < .001$) were both significant. The effects indicate that acceptance changed with age and the changed varied by condition. Specifically, with age, participants were more likely to decrease in their acceptance of the Control No, Group, and Piece pictures, but not of the Whole pictures.

The fact that children and even some adults accepted pieces as valid units is consistent with the claim that nouns themselves do not necessarily provide criteria of individuation and that context might affect the interpretation of nouns. Furthermore, without productive partitives to differentiate or indicate units of quantification, speakers are more flexible and willing to accept pieces as valid units.

### 3.2.3. Forced choice task (Study 3): results

When presented with options to choose from, adults and children preferred to associate the whole object with bare nouns. That is, they did not choose to match the bare argument with the photos were the object was broken or granulated. Crucially, when the participants heard a noun with any of the two pseudopartitive-candidate words, they frequently did not match them to the expected picture (Figure 5), possibly because these forms are neither systematically used nor obligatory. That is, the participants matched these morphemes with the objects in a transformed stage (in pieces or granulated), not in its unmodified version (whole), but did not systematically associated ‘atxa’ with the round piece of x (as expected) and did not systematically associated ‘akuata’ with the long piece of x (as expected). These results are compatible with the hypothesis that these morphemes are not productive alternatives for describing parts as their interpretation varied within and across nouns. Consistent with Study 2, this study is compatible with the hypothesis that in languages that lack productive alternatives for non-whole units, speakers could count parts using bare nouns, unlike speakers whose language does.
Figure 5: Felicity Judgment Results. Participants did not systematically favor the expected picture, suggesting that these pseudopartitives do not have meanings that can be productively combined.

3. General discussion

3.1. Semantic hypothesis on nouns’ denotations

Semantic theories diverge on the characterization of the denotation of count nouns. Some theories do not distinguish natural atomicity from semantic atomicity while others claim that, in grammar, atoms can be contextually specified. Proposals that claim that semantic atoms do not necessarily coincide with natural atoms (such as Rothstein 2010 and Lima 2014) are compatible with proposals that argue in favor of a pragmatics account for the acquisition and interpretation of nouns (cf. Srinivasan et al. 2013). In Srinivasan et al. (2013) studies, count nouns could be interpreted as referring to parts until children learn that there was a better, more precise description. In this paper we explored one prediction of Srinivasan et al. proposal’s: in the absence of productive pseudopartitives, nouns can be used to refer to parts into adulthood. We tried to correlate this proposal with the formal semantics literature debate on the basic
denotation of nouns. The two goals of this paper can be summarized in terms of two hypotheses that were tested in three studies with Yudja speakers:

**H1**: natural atoms and their material parts belong to the extension of count nouns. Thus, atoms are contextually determined (Rothstein 2010, Lima 2014).

**H2**: a blocking mechanism prevents speakers to refer to parts of individuals using an unmodified count noun when pseudopartitive constructions or measure phrases are available to refer to these parts (Srinivasan et al. 2013). If pseudopartitives or measure phrases are unproductive in a language, this blocking mechanism will not apply.

The first two studies showed that Yudja children and adults can use bare nouns in order to refer to parts and wholes but children were more likely to use bare nouns to refer to parts than adults. Alternative descriptions were provided by some participants but they did not necessarily include pseudopartitive-candidate morphemes. The non-inclusion of pseudopartitive words in the description of parts is directly correlated with the fact that pseudopartitive morphemes are evaluated by Yudja speakers as incompatible with most nouns. This was also attested in elicitation tasks (cf. Lima 2014). In Study 3 we tested whether children and adults correctly interpreted these morphemes (associating them with pictures in a forced choice task) and it was observed that participants did not systematically favor the expected picture, suggesting that these pseudopartitives do not have meanings that can be productively combined.

These results support theoretical semantic analyses such as Rothstein (2010) and Lima (2014) that claim that atoms can be contextually defined. More precisely, the authors argue that concrete portions of any kind can be counted as atoms in the grammar, as long as they don’t overlap physically.

The analysis that predicts that concrete portions can be counted as semantic atoms in Yudja was introduced in Lima (2014) in order to account for the fact that in this language notional mass nouns can be combined with numerals (as presented in the introduction). Lima (2014) has shown that we cannot explain those facts as coercion cases, since counting with notional mass nouns like *apeta* (‘blood’) is possible even when the counting unit is not conventional, and even when the atoms that are being counted differ in shape and size.
In Lima’s analysis, the basic denotation of nouns in Yudja was hypothesized to be kinds (16). Under this analysis, the property of being an atomic part of a kind depends on an operation (AT*) that maps an individual x, a world w and a kind k to the truth value 1 if and only if x is an atomic part of k(w) or is the sum of atomic parts of k(w):

(16) \[[\text{apǐ}]\] = \(\lambda w. \text{DOG}(w)\)
\[[\text{apǐ}]\] = DOG

KO (Kinds to Objects)
(a) \(\text{KO} = \lambda k: k \in K. \lambda x. \lambda w. \text{AT}^*(w)(x)(k)\)
(b) \(\text{KO}([\text{apǐ}]) = \lambda x. \lambda w. \text{AT}^*(w)(x)(\text{DOG})\)

Lima argues that the atomic members of a kind are defined as maximal self-connected\(^1\) portions (aka ‘concrete portions’) of the kind described by the root in the world of evaluation. Saying that an entity is a maximal self-connected portion of a kind k in a world w means that this entity is a self-connected portion of k in w that is not a proper part of any self-connected portion of k in w. Thus, the condition of atomicity under this view can be established below:

(17) Condition on atomicity:
An entity x is an atomic portion of a kind k in a world w only if x is a maximal self-connected part of k(w).

The definition in (17) states that being a maximal self-connected part of a kind in a world of evaluation is a necessary condition of being an atomic portion of that kind in that world. This condition has two important consequences. First of all, for any kind k and world w, the mereological fusion of two disconnected parts of k(w) can never be treated as an atom of k(w).

In order to exemplify these notions, consider the following example. When a Yudja speaker is presented with the following picture, one possible description for it is going to be the following:
Each puddle in this scenario is taken as a maximal self-connected portion of water cannot be counted as an atom. That is, a part of these parts of the kind of water will never count as an atom in this scenario. This won’t be described as “five (portions) of water”, for example. This analysis predicts that parts of a kind could be considered as an atom as long as these parts do not overlap with other parts. Data from Yudja (Lima 2014) confirms this prediction for notional mass nouns as we have seen that mass nouns in Yudja can be directly combined with numerals and given that portions of substance-denoting kinds that do not overlap can be counted as individuals.

This analysis predicted that parts of any kind could be considered as atoms as long as these parts do not overlap with other parts. The three studies presented in this paper supported this proposal as non-overlapping parts of objects would be available for counting as an individual for children and part of adults.

### 3.2. Typological considerations

Taking into consideration the results for English-speaking children and Yudja in similar tasks, we could hypothesize that languages like English and Yudja are not so different: we could argue that in all languages context-dependency is a relevant feature in order to determine atoms in the grammar. A similar study in Brazilian Portuguese (Lima in progress) shows that Brazilian Portuguese adults accept the use bare count nouns (64% of the trials) to describe parts. However, as in other studies (Srinivasan et al. 2013), if participants have a pseudopartive description option available, they will prefer that description to the bare noun option.
Under this perspective, the difference between English and Yudja adults when it comes to the grammar of countability is not a difference in the denotation of nouns or their syntactic make-up, but rather a difference in the availability of a blocking mechanism, which in turns depends on the availability of pseudopartitive constructions and measure phrases in these languages. To conclude, with a proposal that argues in favor of an analysis where atoms are contextually determined, we do not want to ignore the role of functionality and a cognitive bias for whole objects. We predict that if we ask a Yudja speaker to draw 3 canoes, it is most likely that they will draw three whole canoes for two reasons: 1) a cognitive bias that favors ‘natural atoms’, i.e., wholes/prototypical objects. By hypothesis, individuals that are more likely to appear in different forms will be more likely to overrule cognitive bias. That is the case for food items (as previously suggested by Frisson and Frazier 2005); 2) functionality: parts of individuals are less functional than wholes. Functionality has been shown to be a relevant feature in determining how to interpret object mass nouns (Grimm & Levin 2011) and it will be explored in future studies.

4. Final Remarks

The results of the studies in Yudja combined with the results of the studies with English speaking children support the hypothesis that counting may rely on atoms that are contextually determined (with Rothstein 2010 for English, Lima 2014 for Yudja). These results are also compatible with theories such as Srinivasan et al. (2013) according to which nouns’ denotations provide only partial conditions in order to define individuals and that pragmatic inferences are required in order to access a more informative description for parts. Natural atomicity can bias what counts as an atom in the grammar (we are more likely to consider a whole chair as ‘chair’ in contrast with a piece of the chair), but a natural atom is not necessarily a semantic atom (cf. object mass nouns in English that are grammatically mass nouns) and non-natural atoms can be considered as semantic atoms (cf. substance-denoting nouns in several languages that can interact directly with the counting system and parts of individuals).

We claimed that in languages that lack productive alternatives for non-whole units, speakers never cease to count parts or spatio-temporal groups for bare nouns. Thus, the difference between languages like English and Yudja when it comes to the grammar of countability is not a
difference in the denotation of nouns but rather a difference in the availability of a blocking mechanism (pseudopartitives and measure/container phrases).

Notes

1 Varzi (2007) defines self-connectedness as the following:
   
   (i) Self-connectedness:
   
   $$\text{SC}(x) =_{\text{def}} \forall y \forall z [ \forall v [ O(v)(x) \leftrightarrow (O(v)(y) \lor O(v)(z))] \rightarrow C(y)(z) ]$$
   
   According to definition, saying that an entity is self-connected means that whenever we partition this entity in two parts, these two parts are connected to each other.

   Also according with Varzi (2007), Finally, a maximal self-connected portion of a kind \( k \) in a world \( w \) can be defined as follows:
   
   (ii) Maximal self-connected portion of a kind in a world of evaluation:
   
   $$\text{MSC}(x)(k)(w) =_{\text{def}} \text{SC}(x) \land x \leq k(w) \land \exists y [ x < y \land \text{SC}(y) \land y \leq k(w) ]$$
   
   According to the definition, saying that an entity is a maximal self-connected portion of a kind \( k \) in a world \( w \) means that this entity is a self-connected portion of \( k \) in \( w \) that is not a proper part of any self-connected portion of \( k \) in \( w \).

Acknowledgement

We thank the Yudja communities, Chadaha Juruna and Tawaiku Juruna (research assistants in the field) and the families of the children in the Tuba Tuba village (Xingu Indigenous Territory, Brazil). This work was supported by DRCLAS ECD Faculty Grant/ Fundação Universitária José Bonifácio (UFRJ). All usual disclaimers apply.

References


deal, Amy Rose (To appear), “Countability distinctions and semantic Variation,” *Natural Language Semantics*.


lima, Suzi. Ms. *Bare nouns and partitives in Brazilian Portuguese*.


RESOLVING PRONOUN AMBIGUITY IN EUROPEAN PORTUGUESE: CHILDREN VS. ADULTS

MARIA LOBO AND CAROLINA SILVA

1. Introduction

Previous studies show that the interpretation of clitic forms (reflexive and non-reflexive) is acquired early (Cristóvão 2006; Silva 2014, 2015a). The reflexive clitic forms are not subject to ambiguous interpretations (cf. (1a)) and the contexts in which the interpretation of non-reflexive clitic forms was tested are usually contexts in which the interpretation of the clitic is not ambiguous (cf. (1b)). There have been less studies on contexts in which a pronoun can recover more than one antecedent, although there are clear preferences in the adult grammar for a specific reading. Silva (2015a) is one of the studies which investigates the development of the interpretation of strong pronouns in non-categorical contexts (cf. (2)) and of null and overt pronouns in complement clauses with the indicative and the subjunctive moods (cf. (3)), showing that these contexts are acquired later.

(1) a. O pai do menino lavou-se.
   the father of-the boy washed CL-REFL-3SG
   *The boy’s father washed himself.

   b. O pai do menino lavou-o.
   the father of-the boy washed CL-him
   *The boy’s father washed him.

(2) A fada está a apontar para ela.
   the fairy is to point at her
   *The fairy is pointing at her.

(3) a. O bombeiro disse ao avô que pro/ele emagreceu. (Indicative)
   the fireman said to-the grandpa that pro/ele lost-weight
   *The fireman told grandpa that he lost weight.
b. O bombeiro pediu ao avô que pro/ele saltasse. (Subjunctive)
the fireman requested to-the grandpa that pro/ele jumped

*The fireman asked grandpa to jump.*

In the current research, we seek to contribute to a better understanding of the development of the interpretation attributed to pronominal forms, considering another context in which the interpretation of pronominal forms is not categorical, but in which there are clear preferences in the adult grammar – null and overt subjects in adverbial clauses. Here we present the results of a new experimental study that aims to investigate the preferential interpretation given by preschool and school age children to null and overt embedded pronominal subjects in adverbial clauses in ambiguous contexts, compared to a group of adults.

In section 2, we refer to previous studies that contemplate the interpretation of null and overt subjects in different populations; in section 3, we present the study we conducted with monolingual children who are native speakers of European Portuguese and with a control group of adults, including the methodology (3.1.) and results (3.2.); in section 4, we discuss the results and present the main conclusions.

## 2. Background

The interpretation of null and overt pronouns has been studied for several languages, considering processing, first and second language acquisition. The interpretation of these forms is an interesting topic to investigate the syntax-discourse interface, since it is constrained both by syntactic and discourse factors. In a non-null subject language, an overt subject pronoun may recover both a subject and an object antecedent (4); in contrast, in a null subject language null subjects and overt pronominal subjects usually have different interpretations. A null subject prefers to recover a subject antecedent, which usually is also a topic, whereas an overt subject pronoun prefers to recover an antecedent that is not the subject, and signal a topic shift (5).

(4) The boy/greeted the grandfather, when he/ came home.

(5) a. O rapaz/cumprimentou o avô/ quando pro/ chegou a casa.
the boy greeted the grandfather when pro arrived at home
b. O rapaz cumprimentou o avô quando ele chegou a casa.
the boy greeted the grandfather when he arrived at home

In the Government and Binding framework the preference for null subjects in some syntactic contexts was formulated as an economy principle, which was not very well understood: the Avoid Pronoun Principle (Chomsky 1981; Montalbetti 1984). Choosing a null form was considered more economical than choosing an overt form. An overt pronoun would only be used when necessary. Cardinaletti & Starke (1999) propose another economy-based principle called Minimize Structure (or Economy of Representations) that favors, whenever possible, the choice of the most deficient pronominal form according to their typology of pronouns (clitic, weak and strong pronouns). The choice of a null embedded subject instead of an overt pronoun follows from their different properties: a null subject is a weak pronoun whereas an overt pronoun is a strong pronoun. The overt pronoun is thus considered a marked form that will only be used when other forms are not appropriate. Since by default an embedded subject will recover the most prominent antecedent, and typically the subject is more prominent than the object, the overt pronoun will only be used to signal that the antecedent is different from the subject.

Although the division of labor between null and overt pronominal subjects is found in all null subject languages, there is some crosslinguistic variation concerning mostly the interpretation of overt pronouns. These are more sensitive than null pronouns to semantic and discourse factors. Even though the reason for these differences is not very clear, one factor that may play a role is the grammatical status of the pronoun, which is subject to crosslinguistic variation: in some languages it is a personal pronoun, as in Spanish (Filiaci, Sorace & Carreiras 2014), in other languages, as in Basque, it is a demonstrative (Iraola, Santesteban & Ezeizabarrena 2014).

Contrary to Montalbetti (1984)’s predictions, the differences between null and overt subject pronouns in null subject languages are not as clear as we would expect. In contexts where the subject pronoun has a bound variable reading, speakers sometimes accept that an overt pronominal subject be bound (Alonso-Ovalle, Férnandez-Solera, Frazier & Clifton 2002 for Spanish). Also Filiaci, Sorace & Carreiras (2013) show that although Spanish and Italian behave alike in what concerns the interpretation of null subjects, they differ in the processing of overt pronouns, visible also in different reading times for each language. In
Spanish overt pronouns recover a subject antecedent more easily than in Italian.

The interpretation of null and overt subjects is more complex than we might think, since different factors play a role (Costa, Faria & Matos 1998) and different grammar components may be involved. Among other factors, the following have been shown to condition the interpretation of the pronominal subject:

i) Syntax. The structural position of the antecedents conditions the interpretation of the pronoun: a null subject preferentially recovers an antecedent that is a structural subject and an overt pronominal subject recovers preferentially a non-subject antecedent. In Carminati’s (2002) study on Italian, the differences in interpretation between null and overt pronouns are analyzed as processing constraints. According to her Position of Antecedent Hypothesis, null subjects preferentially recover a more prominent antecedent, typically the syntactic subject in Spec, TP, while overt pronominal subjects prefer to recover a less prominent antecedent, usually a constituent that is not the syntactic subject. Luegi (2012) also shows that different factors compete in the interpretation of pronominal subjects. These include the syntactic function (subject vs. oblique) and structural position of the antecedent (preverbal vs. postverbal subject; postverbal oblique complement vs. topicalized oblique complement). The author confirms differences between null and overt pronominal subjects in the selection of the antecedent in European Portuguese: in off-line tasks, a null subject prefers to recover the constituent that is the syntactic subject and an overt subject pronoun tends to recover a non-subject constituent; this effect is not as clear in on-line tasks.

ii) Prosody. Prosodic emphasis on the overt pronoun may favor a subject antecedent, as observed by Papadopoulou et al. (2015).³

iii) Morpho-syntax. The morpho-syntactic status of the pronoun plays a role: null subjects, which are analyzed as deficient pronouns, differ in interpretation from overt subject pronouns, which are analyzed as strong pronouns.

iv) Semantics. Different types of pronouns display different restrictions in animacy features: an overt pronoun usually recovers animate entities, whereas a null pronoun is not sensitive to the animacy features of the antecedent – it can recover both animate and inanimate antecedents. Morgado (2014) shows, for example, in a pilot study with European Portuguese adult speakers, that, although an overt pronoun prefers to recover a non-subject antecedent, this effect disappears when the subject antecedent is the only candidate with [+ animate] features.
Therefore, in a sentence like (6), most participants choose the subject antecedent, although the subject pronoun is an overt form:

(6) Enquanto o pintor segurava o escadote, ele partiu a janela.
(preference for ele = o pintor)

while the painter held the ladder, he broke the window
*While the painter was holding the ladder, he broke the window.*

The importance of animacy features in the distribution of null and overt pronouns has also been found in a study based on a corpus comparing Brazilian Portuguese and European Portuguese: Barbosa, Duarte & Kato (2005) show that null subjects are overwhelmingly preferred when the referent is not animate. Semantic differences between null and overt (strong) pronouns had already been signaled by Cardinaletti & Starke (1999). The authors state that null pronouns (as well as clitics) are not restricted to human entities, whereas strong pronouns usually refer to human entities. Although these effects are not as clear as the authors defend (see for instance Zribi-Hertz 2000), there are interpretative differences between null and strong pronouns to which adult speakers are sensitive.

v) Discourse. Null and overt pronouns differ in their ability to refer to a previous topic, to signal topic-shift or to be focused: a null subject recovers a topic more easily than an overt pronoun does, and, as mentioned above, a focused overt pronoun tends to accept a subject antecedent more easily.

The interpretation of subject pronouns is thus an interface area, that involves at least syntactic factors (structural position of the antecedents), morphosyntactic factors (morphosyntactic status of the pronoun), semantic factors (features associated with different types of pronouns), and discourse factors (the ability to signal topic maintenance, topic shift or focus) and we expect to find crosslinguistic differences depending on the status of the overt pronoun in each language (as a weak pronoun, a strong pronoun or a demonstrative).

The idea that interface areas are acquired later and are more vulnerable in second language acquisition has been defended by several authors (Sorace 2000; among others). Therefore, we expect there to be a slower development in the interpretation of pronouns in first language acquisition. In fact, data from reading comprehension tasks point to specific difficulties in the interpretation of referential dependencies involving different types of pronouns (Batalha 2014).
It has been shown that syntactic properties of null subject languages are acquired early, such as the ability to license a null subject both in matrix and subordinate clauses, and the ability to invert subjects (e.g. Hyams 2012; and, for Portuguese, Gonçalves 2004). In contrast, there seems to be a late development of interpretative differences between null and overt pronouns, which may follow from processing factors, according to some authors (Carminati 2002; Alonso-Ovalle et al. 2002; Papadopoulou et al. 2015).

Although there have been several studies on the use and interpretation of null and overt pronominal subjects in the fields of second language acquisition, bilingual acquisition and language attrition (Pérez-Leroux & Glass 1999; Sorace & Filiaci 2006; Margaza & Bel 2006; Serratrice, Sorace & Paoli 2004; Sorace, Serratrice, Filiaci & Baldo 2009; Madeira, Xavier & Crispim 2010, 2012; Katsa, Tsimpli & Rothman 2015; Pirkmayr 2015; among others), the studies on L1 monolingual acquisition of subject pronouns are less numerous, and the results provide conflicting evidence. In Croatian children seem to behave adult-like (Kras & Stipéc 2013). For other languages, children tend to behave as adults in the interpretation of null subjects, but overaccept subject antecedents for overt pronouns (Serratrice 2007 for Italian; Iraola, Santesteban & Ezeizabarrena 2014 for Basque). In a study on the processing of pronouns in ambiguous contexts with children and adult Greek speakers, Papadopoulou et al. (2015) show that both groups are sensitive to differences between null and overt subject pronouns, but the mechanisms of pronoun resolution are not completely developed at age 10-11. For European Portuguese, Silva (2012, 2015a, 2015b) studied the interpretation of null and overt subject pronouns in different types of complement clauses (with the indicative and with the subjunctive) by children aged 3 to 6 years-old, using a truth-value judgement task. The results show that in indicative complement clauses with two antecedents, children’s interpretation of embedded subject pronouns differs from the adult one: children overaccept object antecedents for null subjects and subject antecedents for overt pronouns.

There have not been previous studies in Portuguese L1 acquisition on the interpretation of embedded subject pronouns in adverbial clauses, such as the ones included in the abovementioned studies on Basque, Spanish, Italian and Greek. Our study considers precisely those contexts. In adverbial clauses, unlike in complement clauses, the interpretation of the subject does not depend on the main clause predicate. Thus, we do not expect differences in subject interpretation according to lexical knowledge of the main verb (e.g. dizer ‘say’, perguntar ‘ask’, querer
‘want’ or pedir ‘ask’ select complement clauses with different properties). In the present study, we wanted to investigate whether, in adverbial clauses, which are not dependent on selectional properties of the main predicate, children’s interpretation was convergent with the adults’ one.

3. Study

Taking into account the mentioned studies, we sought to determine how children interpret null and overt pronominal subjects of adverbial clauses in ambiguous contexts, that is, with two possible antecedents in the main clause. We wanted to verify, on the one hand, if children distinguish null subjects from overt pronominal subjects, preferring a subject antecedent for a null subject and a non-subject antecedent for an overt subject, and verify, on the other hand, if the linear position of the antecedent was relevant.

Accordingly, we considered the following questions:
  i) Do children distinguish null subjects from overt pronominal subjects, preferring subject antecedents for null pronouns and object antecedents for overt pronouns?
  ii) Does the position of the anaphoric expression (null or overt pronominal subject) with regard to potential antecedents condition their interpretation?
  iii) Is there development in the interpretation of null and overt subjects? Do children’s performances differ from adult ones?

Considering the studies carried out for other languages, we expect that the interpretation of these forms is acquired late. We further expect that the interpretation of null subjects is acquired earlier than that of overt pronouns, and that the anaphoric contexts are easier than the cataphoric contexts.

3.1. Methodology

The study included three groups of preschool and school age children and a group of adults. The information on the participants is shown in Table 1:
Table 1: Participants

<table>
<thead>
<tr>
<th>Age range</th>
<th>Mean age</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 year-olds</td>
<td>5;6</td>
</tr>
<tr>
<td>6-7 year-olds</td>
<td>7;1</td>
</tr>
<tr>
<td>8-9 year-olds</td>
<td>8;8</td>
</tr>
<tr>
<td>Adults</td>
<td>31</td>
</tr>
</tbody>
</table>

We applied to each participant an interpretation test with selection of images. The participant heard a sentence and saw two images. He was, then, asked to choose the image he considered most adequate as an illustration of the test sentence. Each pair of images corresponded to a possible interpretation of the subject pronoun (null or overt) of a temporal adverbial clause: the (null or overt) pronominal subject could be interpreted as coreferential with the subject or the object of the main clause. Therefore, the dependent variable was the selection rate of the image corresponding to the choice of a subject antecedent or an object antecedent. All test sentences had a temporal adverbial clause initiated by *quando* ‘when’ with a situation involving a single human argument and a main clause with a transitive verb and two human arguments, a subject and a direct object. The gender of the two arguments was always the same, so that there was ambiguity in the interpretation of the pronoun, which could recover the subject or the object of the main clause.

The test manipulated the variable type of pronoun (null or overt) and the variable position of the adverbial clause (initial or final), with four conditions, each one with 6 items, in a total of 24 items:

i) Final adverbial clause and null subject in the adverbial clause (anaphoric null subject) – 6 items;

ii) Final adverbial clause and overt pronominal subject in the adverbial clause (anaphoric overt subject) – 6 items;

iii) Initial adverbial clause and null subject in the adverbial clause (cataphoric null subject) – 6 items;

iv) Initial adverbial clause and overt pronominal subject in the adverbial clause (cataphoric overt subject) – 6 items.

Here are examples for each condition, following the order above:

(7) A mãe cumprimentou a avó *quando pro* entrou na cozinha. (anaphoric null subject)  
the mother greeted the grandma when *pro* entered in-the kitchen
The mother greeted grandma when she entered the kitchen.

(8) O avô fotografou o menino quando ele saiu da garagem, the grandpa photographed the boy when he came-out of-the garage

Grandpa photographed the boy when he came out of the garage.

(9) Quando pro chegou a casa, o avô cumprimentou o menino, when pro arrived at home, the grandpa greeted the boy

When he came home, grandpa greeted the boy.

(10) Quando ela saiu da garagem, a bruxa molhou a princesa, when she came-out of-the garage, the witch wet the princess

When she came out of the garage, the witch wet the princess.

The order of presentation of the test items and the position of the images in each pair were randomized (see Appendix with all the stimuli according to their presentation order) and the pronouns were produced with a neutral intonation.

Below we give the example of the pair of images for sentence (7):

Image 1: Example of images used in the test
As we can see, for the sentence *A mãe cumprimentou a avó quando entrou na cozinha* (‘The mother greeted grandma when she entered the kitchen’), if the participant chooses the first image, he will be preferring the interpretation in which the null subject recovers the subject of the main clause (the mother). If he chooses the second image, he will be choosing the interpretation in which the null subject recovers the object (the grandma). Each participant was asked to choose one and only one image. For this reason, we did not consider the hypothesis of a participant simultaneously selecting the two images presented on screen. Differently from other studies, in our experiment we did not consider a third referent, not mentioned in the clause, but present in the context, which may be a possible interpretation, particularly in cataphoric contexts.

### 3.2. Results

The results were coded by checking if the participant chose the subject antecedent or the object antecedent. In Table 2, we present the option rates for the subject antecedent and the object antecedent in each of the conditions for each group of participants:

Table 2: Results for the selection of subject and object antecedents by condition and by group

<table>
<thead>
<tr>
<th></th>
<th>Anaphoric null subject</th>
<th>Anaphoric overt subject</th>
<th>Cataphoric null subject</th>
<th>Cataphoric overt subject</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>5 year-olds</strong></td>
<td>109/180</td>
<td>61%</td>
<td>71/180</td>
<td>39%</td>
</tr>
<tr>
<td><strong>6-7 year-olds</strong></td>
<td>100/150</td>
<td>67%</td>
<td>50/150</td>
<td>33%</td>
</tr>
<tr>
<td><strong>8-9 year-olds</strong></td>
<td>150/210</td>
<td>71%</td>
<td>66/210</td>
<td>29%</td>
</tr>
<tr>
<td><strong>Adults</strong></td>
<td>188/210</td>
<td>90%</td>
<td>22/210</td>
<td>10%</td>
</tr>
</tbody>
</table>

As we can observe, adults clearly prefer that a null subject recovers a subject antecedent, independently of this antecedent occurring before or after the pronoun (90% of option for the subject antecedent in the anaphoric and cataphoric null subject conditions). For the overt pronominal subject, adults prefer a non-subject antecedent when it is anaphoric (85%), but there is a lot of hesitation when the pronoun is
Maria Lobo and Carolina Silva

In the children’s groups, we found that the 5 year-olds have results close to chance level with overt subjects (in both anaphoric and cataphoric conditions) and with cataphoric null subjects. However, this group is already sensitive to the distinction between null subjects and overt subjects in the anaphoric context ($\chi^2 = 6.43, p = 0.01$). In the cataphoric context, this difference is not significant for the 5 year-old children ($\chi^2 = 0.91, p = 0.34$). The 6-9 year-old children distinguish more clearly null subjects from overt pronominal subjects, but are still distant from the control group. In the anaphoric condition, the comparison between null subjects and overt subjects is always significant for the 6-7 year-old group ($\chi^2 = 21.41, p < 0.01$), for the 8-9 year-olds ($\chi^2 = 41.99, p < 0.01$) and for adults ($\chi^2 = 229.33, p < 0.01$). Regarding the cataphoric condition, this comparison also showed to be significant for these groups of participants (for the 6-7 year-olds $\chi^2 = 9.36, p < 0.01$; for the 8-9 year-olds $\chi^2 = 22.29, p < 0.01$; for adults $\chi^2 = 119.45, p < 0.01$). Comparing the selection of the subject antecedent observed in null subjects with that observed in overt subjects, all children’s groups have results that are still far from the adults’ behavior and even the oldest group (8-9 year-olds) remains significantly distant from the control group (8-9 year-olds vs. adults: for the anaphoric null subject $\chi^2 = 20.75, p < 0.01$; for the cataphoric null subject $\chi^2 = 13.60, p < 0.01$; for the anaphoric overt subject $\chi^2 = 29.94, p < 0.01$; for the cataphoric overt subject $\chi^2 = 9.19, p < 0.01$).

In all groups of participants, there seems to be a difference between anaphoric and cataphoric contexts for the overt pronouns: in anaphoric contexts, the preference for disjoint reference (recovery of a non-subject antecedent) is evident; in cataphoric contexts, children’s results are in general at chance level and even the adults’ group has variable results. Making comparisons between the choice of the subject antecedent with anaphoric overt pronouns and with cataphoric overt pronouns, the 5 year-old group displayed results at the edge of a significant difference ($\chi^2 = 4.02, p = 0.045$), the 6-7 year-olds did not show a significant difference ($\chi^2 = 3.45, p = 0.06$) and the 8-9 year-olds achieved a very significant difference ($\chi^2 = 8.60, p < 0.01$); adults had the highest significant difference ($\chi^2 = 28.91, p < 0.01$). In the comparisons between the option for the subject antecedent with anaphoric null pronouns and with cataphoric null pronouns, there were no significant differences in any age group ($p > 0.05$ in all of them).

If we look at the individual results, we can see which percentage of participants, in each group, has results above chance in the choice of the
subject antecedent for null subjects and of the object antecedent for overt subjects. We considered 5 responses in 6 as above chance.

Table 3: Percentage of participants with results superior to chance level for the preferential interpretation (null subject = subject antecedent; overt subject = object antecedent) by group and by condition

<table>
<thead>
<tr>
<th></th>
<th>Anaphoric null subject</th>
<th>Anaphoric overt subject</th>
<th>Cataphoric null subject</th>
<th>Cataphoric overt subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 year-olds (n=30)</td>
<td>27%</td>
<td>17%</td>
<td>13%</td>
<td>3%</td>
</tr>
<tr>
<td>6-7 year-olds (n=25)</td>
<td>40%</td>
<td>20%</td>
<td>36%</td>
<td>4%</td>
</tr>
<tr>
<td>8-9 year-olds (n=35)</td>
<td>46%</td>
<td>20%</td>
<td>57%</td>
<td>0%</td>
</tr>
<tr>
<td>Adults (n=35)</td>
<td>89%</td>
<td>71%</td>
<td>86%</td>
<td>29%</td>
</tr>
</tbody>
</table>

We can observe that the identified tendencies are common to most adults and are more robust for null subjects than for overt subjects. The adult group showed, however, a high rate of chance performance in the overt subject cataphoric condition (60%), which may be attributed to the fact that the task did not include a third referent, possibly the preferred choice for adult participants. It was only in this condition (cataphoric overt subject) that we found adults (4/35, that is, 11%) with a response pattern contrary to what was expected. In the children’s groups, the interpretation of null subjects seems to stabilize earlier than that of overt pronominal subjects and the anaphoric contexts tend to be acquired earlier than the cataphoric contexts, even though the 8-9 year-old group is still distant from the adults’ performance. Even in this group, there is only a slight majority of participants (57%) with similar results to most adults (86%) in the cataphoric null subject condition. This means that most children have results at chance level (and, in some cases, interpretations that are contrary to those of adults) still at 9 years old. We can, consequently, identify the following developmental scale towards the target grammar:

(11) anaphoric null subjects >> cataphoric null subjects >> anaphoric overt subjects >> cataphoric overt subjects

4. Discussion and conclusions

As we anticipated, the results showed that, in European Portuguese, anaphora resolution is an aspect of late development compared to other
dependencies such as binding or wh-dependencies. It is not strange that this phenomenon develops late, because it involves non-categorical contexts, which present variation also in the adult grammar and require integration of syntactic, semantic and pragmatic information.

In line with what was found for other languages, we found that the interpretation of null pronouns is acquired earlier and is more stable than the interpretation of overt pronouns. The preferential association of overt pronouns to non-subject antecedents is more difficult to acquire. This developmental pattern is expected since, on the one hand, the interpretation of overt pronouns is less categorical even in the adult grammar and, on the other hand, it is constrained by a higher number of factors. Overt pronouns are more sensitive than null pronouns to semantic and discourse factors and are more marked forms.

Hence, it appears that in null subject languages, overt pronouns are sensitive to the animacy of the antecedent, preferentially recovering animate antecedents. This does not happen with null pronouns, which recover both animate and non-animate antecedents (Costa, Faria & Matos 1998; Cardinaletti & Starke 1999; Barbosa, Duarte & Kato 2005). Furthermore, overt pronouns are sensitive to prosodic factors: their interpretation varies depending on whether they receive an emphatic stress/prosodic focus or not. A stressed overt pronoun recovers a subject more easily than an unstressed pronoun. A null pronoun, on the contrary, due to the lack of phonetic content, cannot be stressed. Finally, the interpretation of overt pronouns is dependent on discourse factors, in particular on topic maintenance or topic shift: we will have an overt pronoun in topic shift contexts more easily.

These factors put a burden in the processing of overt pronouns, since not only the position of the antecedent has to be taken into account, but also its feature content, as well as the discourse status/prosodic prominence of the pronoun. Consequently, extra-syntactic factors have to be considered.

In contrast to overt pronouns, null pronouns, in accordance with the typology of Cardinaletti & Starke (1999), are defective forms, syntactically more economical, which do not require the integration of so many factors for their interpretation. We note that the interpretation of deficient forms – in this case, null pronouns (weak forms) – is acquired earlier than the interpretation of strong pronouns. The morphosyntactic status of the pronoun thus conditions to a certain extent the rhythm of development of its interpretation.

As for the position of the pronoun in relation to the expression which sets its reference, we found that the occurrence of the antecedent after the
pronoun makes it less accessible: children tend to resolve the ambiguity as soon as possible by choosing the first linear antecedent (the subject). In this case, processing limitations appear to be involved: when the pronoun occurs in an anaphoric context, the two possible antecedents are already available and the participants only need to evaluate which one of them (subject or object) is the best candidate; when the pronoun occurs in a cataphoric context, its interpretation is suspended until an antecedent is available, without necessarily anticipating the presence of a second antecedent.

Studies on processing show that a similar effect exists in adult grammar (Fedele & Kaiser 2014; among others). Sorace & Filiaci (2006) and Serratrice (2007) also show that the preference for a subject antecedent is stronger in cataphoric contexts, independently of the morphosyntactic status of the pronoun. It seems that in those contexts choosing a subject antecedent is an option with less processing costs than choosing a non-subject antecedent. This factor competes with the division of labor between null and overt subject pronouns and this has as a consequence that the effects are less clear in cataphoric contexts. However, since our experiment did not consider a third possible reading, that in which the pronoun refers to an entity not mentioned in the clause, we do not know what would have been the participants’ choice in that case.

Notes

1 This research was funded by Fundação para a Ciência e a Tecnologia (FCT), through the Strategic Project Pest-OE/LIN/UI3213/2014.
2 There is an additional possibility: the subject pronoun can refer to a third referent that is not mentioned in the clause. We have not considered this possibility in our experimental study, which may be seen as a possible shortcoming.
3 Several studies have shown that there are other contexts where emphasis on a pronoun may trigger different interpretations. In English, emphasis on a complement pronoun may allow anaphoric readings (e.g. Alladin sees HIM = himself) (see Wexler & Chien 1985; Grodzinsky & Reinhart 1993, for instance).
4 Our experimental design did not include a third referent as a possible antecedent for the subject pronoun. This interpretation seems to be available mostly in cataphoric contexts. Previous studies that consider a third referent (e.g. Sorace & Filiaci 2006: 354; Serratrice 2007) show that, in cataphoric contexts, the choice of a referent that is not mentioned in the clause but that is available in the context is always minoritary compared to the choice of a subject or object antecedent, When
the subject is an overt pronoun, adults prefer this third referent, but children do not. We do not know what the behavior of Portuguese adults and children would have been if we had included a third referent. In the studies of Sorace & Filiaci (2006) and Serratrice (2007), although there was a third referent available, the choice of the subject antecedent in cataphoric contexts was always higher than in the anaphoric contexts, We leave the effect of the presence of a third referent for future work.

5 The statistical comparisons were made through Pearson’s chi-square tests with Yates’ continuity correction. In this study, a difference is considered significant when $p < 0.05$.

6 Comparison between the context of anaphoric null subjects and the context of cataphoric null subjects: for the 5 year-olds $\chi^2 = 2.21$, $p = 0.14$; for the 6-7 year-olds $\chi^2 = 0.06$, $p = 0.81$; for the 8-9 year-olds $\chi^2 = 1.24$, $p = 0.27$; for the adults $\chi^2 = 0.03$, $p = 0.87$.

References


Madeira, Ana, Maria Francisca Xavier, and Maria de Lourdes Crispim (2010), “Interpretação semântica e/ou pragmático-discursiva de sujeitos na aquisição de português L2,” *Textos Seleccionados, XXV*
Encontro Nacional da Associação Portuguesa de Linguística, 513-529, Porto, APL.


and objects in English–Italian bilingual and monolingual acquisition,” *Bilingualism: Language and Cognition* 7(3), 183–205.


**Appendix: Experimental sentences used in the test**

1. O pai fotografou o menino quando se sentou.
   the father photographed the boy when CL-REFL-3SG sat
   *The father photographed the boy when he sat down.*
2. A avó cumprimentou a menina quando ela chegou a casa.
the grandma greeted the girl when she arrived at home
Grandma greeted the girl when she came home.

3. Quando subiu à árvore, a menina chamou a mãe.
when climbed to-the tree, the girl called the mother
When she climbed the tree, the girl called her mother.

4. Quando ela saiu da garagem, a bruxa molhou a princesa.
when she came-out of the garage, the witch wet the princess
When she came out of the garage, the witch wet the princess.

5. A princesa viu a bruxa quando subiu à árvore.
the princess saw the witch when climbed to-the tree
The princess saw the witch when she climbed the tree.

6. Quando caiu na rua, o menino chamou o polícia.
when fell on-the street, the boy called the policeman
When he fell on the street, the boy called the policeman.

7. Quando ele subiu à árvore, o polícia viu o ladrão.
when he climbed the tree, the policeman saw the thief
When he climbed the tree, the policeman saw the thief.

8. Quando acordou, a avó viu a menina.
when woke-up, the grandma saw the girl
When she woke-up, grandma saw the girl.

9. A menina viu a bruxa quando ela começou a correr.
the girl saw the witch when she started to run.
The girl saw the witch when she started running.

10. Quando ele tropeçou, o avô chamou o médico.
when he tripped, the grandpa called the doctor
When he tripped, grandpa called the doctor.

11. A mãe cumprimentou a avó quando entrou na cozinha.
the mother greeted the grandma when entered in-the kitchen
The mother greeted grandma when she entered the kitchen.

12. Quando chegou a casa, o avô cumprimentou o menino.
when arrived at home, the grandpa greeted the boy
When he came home, grandpa greeted the boy.

13. O menino viu o pai quando ele acordou.
the boy saw the father when he woke-up.
The boy saw his father when he woke up.

14. Quando ela se sentou, a mãe fotografou a menina.
when she CL-REFL-3SG sat, the mother photographed the girl
When she sat down, the mother photographed the girl.

15. Quando começou a correr, o menino viu o ladrão.
when started to run, the boy saw the thief
When he started running, the boy saw the thief.

16. A menina chamou a professora quando tropeçou.
    the girl called the teacher when tripped.
    The girl called the teacher when she tripped.

17. Quando ele entrou no escritório, o pai cumprimentou o avô.
    when he entered in-the office, the father greeted the grandpa
    When he entered the office, the father greeted grandpa.

18. O avô fotografou o menino quando ele saiu da garagem.
    the grandpa photographed the boy when he came-out of-the garage
    Grandpa photographed the boy when he came out of the garage.

    the fireman wet the boy when came-out of-the garage
    The fireman wet the boy when he came out of the garage.

20. O menino chamou o avô quando ele subiu à árvore.
    the boy called the grandpa when he climbed to-the tree
    The boy called his grandpa when he climbed the tree.

21. Quando saiu da garagem, a mãe fotografou a menina.
    when came-out of-the garage, the mother photographed the girl
    When she came out of the garage, the mother photographed the girl.

22. Quando ela se levantou, a menina chamou a professora.
    when she CL-REFL-3SG got-up, the girl called the teacher
    When she got up, the girl called the teacher.

23. A menina chamou a mãe quando ela caiu na rua.
    the girl called the mother when she fell on the street.
    The girl called her mother when she fell on the street.

24. O policia chamou o bombeiro quando se levantou.
    the policeman called the fireman when CL-REFL-3SG got-up
    The policeman called the fireman when he got up.
1. Introduction

This paper contributes to the body of literature on second language (L2) phonology by discussing the results of a study into the acquisition and production of English voiceless stops by Japanese learners of English.¹

It is well known that second language acquisition works quite differently from first language acquisition. This is evident from the fact that L2 learners—and late L2 learners in particular—find it extremely difficult to develop a native-like L2 pronunciation. Almost without exception, the speech of learners displays some amount of interference from their first language, giving the impression of a non-native or foreign accent. But why is L2 pronunciation always acquired imperfectly? Some scholars see it as a perceptual problem, while others argue that it has more to do with speech production. Whichever view is taken, it seems that our knowledge of L2 acquisition is not yet able to explain why learners cannot avoid acquiring a non-native pronunciation. To help enrich our understanding of second language learning, the present study examines the impact of phonetics training on L2 pronunciation.

For native speakers of Japanese, the pronunciation of English presents a number of difficulties. The most challenging aspect of English pronunciation is undoubtedly the accurate production of vowels—something not unexpected, given the obvious differences in the size and shape of the vowel systems of Japanese (5 vowels) and English (20 vowels). But there are also issues with consonant production. Well-known problems include the difficulty in producing distinctions such as those between [l] and [r] (e.g. lock vs rock), [s] and [ʃ] (e.g. seat vs sheet), and [h] and [f] (e.g. hold vs fold), where the failure to distinguish between these consonants can lead to miscommunication because each pair
involves a sound difference that is lexically contrastive in English but not in Japanese. For example, Japanese phonology allows the CV sequence [ʃi] but not [si], so any loanword containing [si] is automatically perceived and reproduced as [ʃi] by most native Japanese speakers. As a result, pairs of words which are lexically distinct in English can become homophones when they are borrowed into Japanese, e.g. *seat* and *sheet* are both pronounced [ʃiːt] by Japanese speakers, so listeners must rely on context to disambiguate them.

Because sound distinctions such as [s] versus [ʃ] are linguistically significant (i.e. contrastive) in English, pronunciation textbooks make a point of drawing learners’ attention to them. By contrast, little attention has been paid to Japanese speakers’ production of other English consonants such as the voiceless stops *p* (e.g. *pen*), *t* (e.g. *attack*) and *k* (e.g. *cool*). One reason for this is that Japanese speakers’ pronunciation of English *p, t, k* is generally assumed to be uninteresting from a phonological point of view, since communication is not usually compromised when English *p, t, k* are produced inaccurately, i.e. using a typical ‘Japanese-English’ pronunciation. The stop series *p, t, k* is present in both English and Japanese, the only difference between them being one of phonetic realisation: in the onset of a stressed syllable native English speakers produce *p, t, k* with aspiration, i.e. *pen* [pʰɛn], *attack* [əˈtʰæk], *cool* [kʰuːl], whereas the same consonants are unaspirated in Japanese. Unlike the distinction between consonants such as [s] and [ʃ], however, this difference between aspirated and unaspirated stop release presents no problems for communication, because aspiration does not, by itself, serve as a cue to lexical contrasts either in Japanese or in English. So although Japanese speakers tend to transfer their native pronunciation rules onto English words, resulting in the inaccurate production of unaspirated *p, t, k*, this rarely causes problems for listeners.

And yet, while the failure to produce aspirated forms of *p, t, k* in stressed positions may not affect communication, it does result in an unnatural L2 pronunciation. In fact, this tendency towards non-native pronunciation is just one characteristic of the so-called ‘Japanese-English’ variety of English mentioned above, which is typically used by native Japanese speakers who have learnt English largely by studying the written language, i.e. via the English curriculum taught at junior-high and high schools in Japan. Among Japanese people it is widely believed that Japanese-English pronunciation gives listeners a negative impression of a speaker’s overall English ability. After all, no matter how competent an L2 speaker’s use of grammar and vocabulary may be, listeners find it difficult to ignore poor pronunciation and they cannot help making hasty
judgments about speakers on the basis of pronunciation alone. Not surprisingly, therefore, many Japanese speakers are keen to eliminate Japanese-influenced characteristics from their spoken English, including errors such as the failure to aspirate voiceless stops. In this paper we take up this point by investigating aspects of the production of English voiceless stops by Japanese L2 learners of English. In sections 3 and 4 we present data to demonstrate the level of accuracy with which Japanese university students produce aspirated $p, t, k$ in English words. We also test whether the same speakers can improve their accuracy as a result of receiving pronunciation training. Then in section 5 we discuss the significance of our findings for theories of L2 phonological acquisition and for the teaching of L2 pronunciation, focusing on the use of compromise realisations as a means of maintaining (i) a phonetic difference between L1 and L2 pronunciations and (ii) the relevant phonological oppositions in the L2. This is preceded in section 2 by an overview of the phonetic characteristics of VOT and its role in creating language typology distinctions.

2. Voice onset time and language typology

Many languages have a two-way laryngeal contrast in obstruents, where $b$ is distinct from $p$, $v$ from $f$, and so on. And traditionally, such contrasts have been described as voiced-voiceless pairs. Recently, however, phonologists have come to realize that not all voiced-voiceless pairs actually involve a voicing contrast, at least not in the phonetic sense. In fact, the terms ‘voiced’ and ‘voiceless’ can be misleading, because languages differ as to whether voicing is really the distinguishing property or not. An alternative to this traditional view of laryngeal contrasts has developed into a line of research known as Laryngeal Realism (Honeybone 2005; Iverson & Salmons 1995), which makes a typological split between two groups of languages. In one group, the group of so-called ‘voicing’ languages, there really is a contrast between voiceless and fully voiced obstruents, in which case it makes sense to use the distinctive feature [voice]—or alternatively, expressed in Element Theory terms (Nasukawa 2005; Backley 2011), the laryngeal element $|L|$—to represent the contrastive property: fully voiced obstruents contain [voice]/$|L|$ in their structure while their voiceless equivalents lack this feature altogether. The voicing group includes languages such as Japanese, Russian and French. Then in the other group, the group of so-called ‘aspiration’ languages, the contrast between sounds such as $b$ and $p$
does not, strictly speaking, refer to the presence or absence of voicing. Instead, the distinguishing property is aspiration, where \( p \) can be aspirated (in strong syllables) whereas \( b \) cannot, and importantly, both are phonetically voiceless. So in these languages the feature [voice] or the element \( \text{|L|} \) is not appropriate. To capture the relevant distinction accurately, advocates of Laryngeal Realism prefer to use the feature [spread glottis] or the element \( \text{|H|} \), since these refer to the presence of contrastive aspiration or voicelessness. The aspiration group includes the languages English, Swedish and German.

In the study to be described below, we tested whether it is possible for native Japanese learners of English to improve their L2 English pronunciation through regular phonetics training. Specifically, the training focused on one particular aspect of English pronunciation, the use of aspiration. As was noted above, aspiration functions as a contrastive property in English, an aspiration language, but not in Japanese, a member of the group of voicing languages. As a phonetic property, aspiration can be measured in terms of voice onset time (VOT)—an aspect of spoken language which speakers manipulate in order to express the phonological distinction between different laryngeal categories among stops. VOT is a physical measurement which refers to the period of time from the release of a stop closure until the onset of vocal fold vibration in a following sonorant, usually a vowel (Lisker & Abramson 1964). It is perceived as a short period of voicelessness, which English listeners interpret as stop aspiration. To reiterate, English (but not Japanese) employs this voicelessness or aspiration for contrastive purposes. Figure 1 shows the three values for VOT that languages regularly use to define phonological (laryngeal) categories.\(^1\)
**VOT value**

- a. positive VOT (voicing lag)  
  representation: [spread glottis] or |H|
- b. zero VOT (negligible voicing lag/lead)  
  representation: unspecified
- c. negative VOT (voicing lead)  
  representation: [voice] or |L|

Figure 1: VOT values as cues to phonological categories

Zero VOT in Figure 1(b) is cross-linguistically the default setting for VOT; all languages have a ‘neutral’ series of stops in which VOT is close to zero. Zero VOT is achieved when a speaker releases a stop closure and at the same time initiates voicing by allowing the vocal folds to begin vibrating. It produces a neutral (i.e. unaspirated, voiceless) stop, which may function in one of two ways in consonant systems. In aspiration languages such as English and Swedish, neutral stops constitute the lenis series and are informally (and inaccurately) labelled ‘voiced’, e.g. [b d g] in English. By contrast, in voicing languages such as French and Japanese, the same neutral stops form the fortis series and are labelled ‘voiceless’, e.g. [p t k] in Japanese. Note that the terms ‘voiced’ and ‘voiceless’ are employed here as labels for phonological categories, not as phonetic descriptions (where ‘category’ is defined by its role in the phonology of a specific language and not necessarily by the phonetic characteristics of the sounds in question). In purely phonetic (i.e. VOT) terms, the English ‘voiced’ stops b, d, g are very similar to the Japanese ‘voiceless’ stops p, t, k (Iverson 2007). It is unfortunate—and potentially confusing—that there is an established convention of using the symbols b, d, g for the neutral stops in an aspirated language like English, even though they are phonetically similar to the neutral stops in a voicing
language like Japanese, which are written \( p, t, k \). Clearly this state of affairs is not ideal. Regrettably, however, conventional practice appears to be winning out, and for now at least we are burdened with the traditional but inaccurate labels ‘voiced’ and ‘voiceless’. Note that this is in spite of the fact that the Laryngeal Realism approach, which contradicts traditional phonological descriptions by highlighting the need for a typological split between aspiration and voicing languages, is able to eliminate altogether this confusion over terminology.

Positive VOT in Figure 1(a) is relevant to aspiration languages such as English, in which there is a contrast between neutral obstruents (i.e. so-called ‘voiced’, with zero VOT) and aspirated obstruents (i.e. so-called ‘voiceless’, with positive VOT). In (a) the label ‘h’ marks a period of voicelessness which begins with stop release and ends with the onset of vocal fold vibration (indicated by the wavy line). By contrast, negative VOT shown in (c) is relevant to voicing languages such as Japanese, in which the obstruent system shows a contrast between neutral consonants (i.e. so-called ‘voiceless’, with zero VOT) and fully voiced or pre-voiced consonants (i.e. so-called ‘voiced’, with negative VOT). In a fully voiced stop, the vocal fold vibration associated with voicing begins at or before the release of the stop.

So although English and Japanese both have a two-way laryngeal contrast in obstruents, and in both languages this contrast is informally expressed using identical symbols such as \( b, p, d, t \) and so on, the relevant categories are not the same: English distinguishes between neutral and aspirated stops whereas Japanese distinguishes between neutral and voiced stops. And as figure 1 shows, the difference between the two systems involves different values for VOT. Indeed, because the length of VOT in stops is noticeably different in English and Japanese, it is a useful way of distinguishing native from non-native pronunciation in the two languages. And for Japanese learners of English who aim to improve the accuracy of their pronunciation of L2 stops, it seems that the focus should be on learning how to shift VOT from their native Japanese values to the values that characterise spoken English. Below we describe how phonetics training can help learners to improve their level of accuracy, although our findings suggest that this is effective only up to a certain point. Before discussing the details, however, we provide some background on the study itself.
3. Procedure

To measure the degree of accuracy with which Japanese speakers produce L2 English stops, we collected spoken language data from 11 native speakers. All were undergraduate students majoring in English at a university in Tohoku, northern Japan. They performed two rounds of speaking tasks involving reading aloud in Japanese (L1) and English (L2). The two rounds of tasks were separated by an interval of five months, during which time the participants received phonetics training, one session every three weeks. During the training sessions they completed self-study tasks using audio-visual tools: online tutorials\textsuperscript{2} to study the articulation of voicing and stop articulation, and speech analysis software to analyse the phonetic properties of vowels after English aspirates. In addition, they received instruction on the acoustic differences between English and Japanese VOT, the physiology of VOT production, and vowel quality after VOT.

After completing the phonetics training, the participants performed the second round of speaking tasks using the same test material that was used in the first round tasks. To elicit enough pronunciations of the target forms (i.e. pre-vocalic stops), the test material made use of familiar personal names in carrier sentences, as shown in (1). Each participant was asked to pronounce all the sentences twice, so the test generated a total of 132 tokens of the target stops (6 words, 2 repetitions, 11 subjects).

(1)  

<table>
<thead>
<tr>
<th>Target: English VOT</th>
<th></th>
<th>Target: Japanese VOT</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[p]</td>
<td>He said, Pete is a puppy.</td>
<td>[p]</td>
<td>kare ga yuu ni wa,</td>
</tr>
<tr>
<td></td>
<td>[t] He said, Ted is a teacher.</td>
<td>[t] kare ga yuu ni wa,</td>
<td>he subj. <em>say</em> nominaliser topic</td>
</tr>
<tr>
<td></td>
<td>[k] He said, Kate is a cook.</td>
<td>[k] tetsuya wa tomato ga suki desu.</td>
<td><em>Peter</em> topic <em>potato</em> subj. <em>like</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>He said, <em>Peter</em> likes <em>potatoes</em>.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>He said, <em>Tetsuya</em> likes <em>tomatoes</em>.</td>
</tr>
</tbody>
</table>
VOT values were collected using software from the Speech Filing System (SFS) suite of speech analysis tools created by faculty at University College, University of London. VOT was measured from the start of the release burst to the onset of voicing (i.e. the first positive peak of F2 in the periodic portion of waveform).

4. Results

As discussed in section 2, the laryngeal (voiced vs voiceless) contrast in Japanese may be compared directly with that in English by referring to the phonetic property of VOT. The average VOT values in milliseconds for native English and native Japanese speakers are shown in Figure 2 (data from Harada 2007). Zero VOT (dotted line) marks the point at which stop release and the onset of vocal fold vibration occur simultaneously.

![Figure 2: VOT values for English and Japanese stops](image_url)

Figure 2 shows that VOT varies according to place of articulation: labials have the shortest VOT, velars the longest, with coronals in between. However, this pattern of place variation, though interesting in its own right, will be ignored here since we are mainly concerned with typological differences in VOT between English and Japanese. In the remainder of this paper, the two languages are compared using labial stops to illustrate the relevant differences. Still referring to the VOT
values in figure 2, it can be seen that native English speakers produce short voicing lag (on average, 7ms) in neutral \( b \) and long voicing lag (68ms) in aspirated \( p \). Both sounds are produced with voicing lag, so strictly speaking, both are voiceless. In the case of native Japanese speakers, they also produce \( p \) with a relatively short lag (24ms) whereas \( b \) is truly voiced (minus 27ms of voicing lead). Based on this data from Harada (2007), therefore, we are justified in claiming that there is a clear phonetic difference between the two systems. And it is unsurprising to find that this difference leads to inaccuracies in L2 pronunciation. To measure these inaccuracies, we analysed our own data to reveal patterns in the production of VOT by Japanese university students speaking native Japanese and L2 English.

Table 1 shows the VOT duration in milliseconds for \( p, t, k \) produced by the 11 native Japanese participants. The results are shown in three blocks, corresponding to three different testing conditions: Japanese sentences (left), English sentences recorded prior to receiving phonetics training (centre), and English sentences after training (right).

Table 1: VOT values for \( p, t, k \) in Japanese and English

<table>
<thead>
<tr>
<th>participants</th>
<th>Japanese</th>
<th>English (pre-training)</th>
<th>English (post-training)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( p )</td>
<td>( t )</td>
<td>( k )</td>
</tr>
<tr>
<td>1</td>
<td>54</td>
<td>43</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>37</td>
<td>30</td>
<td>57</td>
</tr>
<tr>
<td>3</td>
<td>35</td>
<td>31</td>
<td>47</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>29</td>
<td>49</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>50</td>
<td>65</td>
</tr>
<tr>
<td>6</td>
<td>31</td>
<td>54</td>
<td>72</td>
</tr>
<tr>
<td>7</td>
<td>57</td>
<td>37</td>
<td>66</td>
</tr>
<tr>
<td>8</td>
<td>65</td>
<td>48</td>
<td>44</td>
</tr>
<tr>
<td>9</td>
<td>27</td>
<td>44</td>
<td>58</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>49</td>
</tr>
<tr>
<td>11</td>
<td>38</td>
<td>40</td>
<td>54</td>
</tr>
<tr>
<td>average</td>
<td>34.9</td>
<td>36.9</td>
<td>56.4</td>
</tr>
</tbody>
</table>

As expected, when pronouncing the Japanese sentences, all the participants made a true voicing distinction in stops (i.e. fully voiced \( b, d, g \) versus voiceless \( p, t, k \)) although their average VOT for \( p \) (35ms) was slightly higher than the value reported by Harada (24ms) and shown in figure 2. Turning to their stop production when pronouncing the English sentences, even before receiving phonetics training the participants did show a marginal distinction between Japanese \( p \) (35ms) and English \( p \)
Phonological Categories of VOT and Phonetic Compromise in L2 Acquisition

(40ms), which indicates that they already had an awareness of the phonetic difference between the realisation of p, t, k in the two languages. Admittedly, their average score of 40ms of voicing lag for English p is not particularly close to the VOT produced by native English speakers, who average 68ms. Nevertheless, it is longer than the VOT they produce for native Japanese p. The question we address here, however, is whether phonetics training had any impact on the Japanese speakers’ production of English VOT. As table 1 shows, the participants did make improvements to the accuracy of their L2 pronunciation, since their average voicing lag for English p increased from 40ms before training to 51ms after training. Once again, the VOT values produced by our participants do not reach the levels expected of native English speakers pronouncing p. However, after running a paired difference test (Wilcoxon signed-rank test) it was found that the improvement in VOT accuracy was significant, at least in the case of the labial stop p. This is summarised in (2).

(2) Comparative VOT values for L2 English (tasks 1 and 2)
   a. \( \text{<p> } \hat{Z} = -1.87, p = .06, r = .40 \) significant improvement
   b. \( \text{<t> } \hat{Z} = -1.33, p = .18, r = .28 \) no significant improvement
   c. \( \text{<k> } \hat{Z} = -0.61, p = .54, \text{ns, } r = .13 \) no significant improvement

On the basis of the raw figures in table 1, we make the tentative claim that phonetics training has indeed caused the production of VOT in p by L2 learners of English to move in the right direction. This is illustrated in figure 3, which also shows how a more modest improvement was also recorded for t and k.

[Figure 3: Comparative mean VOT values for English stops by English L2 learners]
Our findings do raise a number of questions, however, such as the following. First, if Japanese native speakers are aware that there is a VOT difference between their native stops and L2 English stops, why do they not reproduce this difference more accurately? Our data show that Japanese are capable of producing different degrees of VOT in their L1 and L2, but the fact remains that VOT production in their L2 English never comes close to that of native English speakers, i.e. the degree of aspiration they produce is not strong enough to replicate native English pronunciation. This is true even after they have received phonetics training. We are forced to conclude, therefore, that training can help English learners to improve the phonetic accuracy of their L2 stops, but evidently there is a limit to the degree of improvement that can be achieved.

This finding leads us to a second question, which asks what factors could be responsible for this upper limit on the degree of improvement in VOT accuracy. We have shown that native Japanese speakers can produce L2 voicing lag that is longer than the voicing lag in their L1, yet it appears that something prevents them from lengthening VOT beyond a certain point. This is one of the reasons why they never manage to develop a native-like pronunciation. So how can this upper limit on VOT accuracy be accounted for? The simplest explanation is one which makes appeal to L1 interference. But if this is a case of interference from the native language, then further questions arise. For example, does this affect the phonetic level or the phonological level? And if it turns out to be a phonological issue, is the age of the L2 learner important? This latter question becomes relevant in view of the suggestion that the Critical Age Hypothesis may also apply to L2 acquisition (Oyama 1976). On the other hand, if the issue is a purely phonetic one, then should the focus of investigation be on L2 production or on L2 perception? It is clear that we must continue to seek a better understanding of the nature of the problem, before we are in a position to develop a more effective approach to phonetics/pronunciation training. Once this happens, however, the goal will be to achieve even greater improvements in the accuracy of L2 pronunciation.

5. Compromise phones and imperfect acquisition

Besides appealing to L1 interference, we claim that the notion of compromise phones (Williams 1980) can also help to make sense of our findings. It appears that the pattern of English VOT production by
Japanese learners is nothing unusual, since it follows a similar pattern that has been reported for native speakers of other voicing languages learning L2 English. For example, (3) summarises the results obtained by Caramazza et al. (1973), who investigated VOT production in native French Canadians speaking L2 English.

(3) $p$ voicing lag

<table>
<thead>
<tr>
<th></th>
<th>English (L2)</th>
<th>39ms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>French (L1)</td>
<td>20ms</td>
</tr>
</tbody>
</table>

On average, their speakers produced English $p$ with 39ms of voicing lag, compared with 20ms for native French $p$. So again, we observe a lengthening of VOT in L2 English beyond the VOT values that characterise the L1. But as noted above, these speakers fail to produce an English-like aspirated stop (which Caramazza et al. claim has an average of 61ms of voicing lag). These findings, together with our own results, show that L2 learners typically settle on a so-called compromise phone—a stop with a VOT value lying somewhere in between those of the L1 and L2 targets.

To native speakers of the L2 this compromise pronunciation sounds foreign, because it is not identical to the native L2 sound. But presumably, the L2 speakers in question consider it to be close enough to the L2 target sound because it is sufficiently distant from their own L1 realisation. Flege (1990) also reports on a compromise VOT produced in L2 English by native Spanish speakers—Spanish being, like French, another voicing language. The tentative conclusion given in section 4 may now be recast in a way which incorporates this notion of phonetic compromise. We claim that learners can improve the accuracy of a compromise phone following L2 phonetics training, although the improved pronunciation is still a compromise—it rarely becomes a native-like realisation.

In the L2 acquisition literature it has been pointed out that not all L2 phones are produced inaccurately—rather, some of them sound more native-like than others. The Speech Learning Model developed by Jim Flege and colleagues describes how L2 phones with no close counterpart phone in the speaker’s native language are acquired more accurately than L2 phones that are phonetically similar to L1 sounds. Flege & Hillenbrand (1984) use the term ‘interlingual identification’ to describe this kind of phonetic proximity. For example, if the sound system of the L2 contains a series of clicks which are not present in the L1, these will probably be acquired accurately for the simple reason that they are quite
different from anything in the L1 sound system. On the other hand, if the L2 has a vowel length contrast which does not exist in the L1, vowel length may not be acquired with the same degree of accuracy because short vowels and their long vowel counterparts are phonetically fairly similar, even confusable. And we suggest that this latter point is relevant to the present study, since our own data indicate that the same effect is taking place in the case of stops with differing VOT values: if the L2 has stops with a long VOT while L1 has stops with a shorter VOT, then the likelihood is that the L2 sounds will not be acquired perfectly.

The reasoning behind such a claim runs as follows. If the L2 phone is similar to a segment that already exists in the learner’s L1 phonology, then the learner is likely to treat this L2 phone as just another phonetic realisation of his or her existing L1 phonological category. (Here, ‘category’ refers to a contrastive class of sounds, as identified by a particular distinctive feature (e.g. [voice], [spread glottis] or element (e.g. [L], [H]).) In other words, when L2 sounds are acquired they get filtered through the L1 phonological system. For example, when a native Japanese speaker learns the English word *pint* [pʰaint], the aspirated *p* cannot be stored in the lexicon as a true aspirated stop because in the L1 (Japanese) there is no category of stops containing [H] in their structure, and consequently, no stops characterised by long voicing lag (i.e. aspirates). Instead, it will be assigned to the closest L1 phonological category—that is, the category of plain stops, which are also voiceless and which also have some degree of voicing lag (see Japanese *p* in figure 2)—but importantly, which have no [H] in their structure. We have already established that the native Japanese participants in our study are sensitive to a phonetic difference between Japanese *p* and English *p*, so for those speakers English *p* amounts to nothing more than an atypical or exaggerated realisation of their own native *p*. In this case VOT is the property which is exaggerated, in order to create a phonetic distance between L2 *p* and native *p* and thus keep the two sounds distinct.

However, because this L2 phone belongs to a Japanese phonological category which is characterised by short voicing lag and represented without the presence of a laryngeal element [L] or [H], there seems to be a limit to how much VOT can be extended. Specifically, a fully aspirated stop with [H] in its representation is apparently beyond the phonetic scope of this L1 category of short lag stops which lack [H]. Consequently, the resulting realisation ends up as a compromise phone: it displays properties of the L2 sound but also certain (VOT) properties of sounds that are typical of the L1 category. The result is, in effect, a phonetic merger, as described in Flege (1981). And because VOT is a measurable,
gradient property, we can express the outcome of this merger in fairly precise terms: Japanese speakers produce an English aspirated *p* with a VOT value which lies at around the mid-point between the VOT values of native English *p* and native Japanese *p*.

6. Conclusion

What our data have revealed is that (i) native Japanese speakers can be trained to lengthen VOT beyond the range of values employed in their L1, but (ii) there remains an upper limit to the extent of this lengthening because their L2 speech production is constrained by the phonetic boundaries imposed by their existing L1 phonological categories. By interpreting our data in this way, we arrive at the conclusion that non-native pronunciation is the result of interference from prior language experience, since L2 sounds are processed with reference to the phonological categories established for the L1. As a consequence, if learners persist in identifying L2 phones as tokens (or perhaps, as variants) of L1 sounds, then they are unlikely to ever reach the stage of forming separate L2 sound categories in their stored mental representations of L2 words.

One question to consider is whether this interference affects the perception or the production of L2 sounds. Our instinct is to see it as a perceptual problem, assuming that the accuracy of the L2 sounds we produce is directly related to the accuracy with which we first perceived them. In other words, imperfect perception leads to imperfect production. On this basis it is doubtful that Japanese learners of English can ever produce a native-like degree of VOT in their L2 pronunciation unless they succeed in developing two different perceptual targets, one for Japanese stops and another for English stops. In principle, this could happen if the learner in question were to start learning L2 English at a young age—say, before the age of five, when prior language experience is limited and L1 interference is not so intrusive. In reality, however, most Japanese speakers begin learning English as adolescents, by which time their L1 perceptual targets have become firmly established in the mental grammar. As is predicted by the Native Language Magnet model developed by Kuhl (2000), learners have a ‘neural commitment’ to L1 category mappings which influences the processing – and at a later stage, the production – of L2 speech sounds.

If it is the case that non-native pronunciation boils down to the inaccurate perception of L2 sounds, then this indicates that L2 teachers
should place a greater emphasis on phonetics training than they currently do. The issue of how L2 learners can be trained to effectively develop perceptual targets that are distinct from their existing L1 targets will be the focus of the next stage in our project.

Notes

1 Image taken from [http://image.healthhaven.com](http://image.healthhaven.com).
2 The online tutorial material used in this study was designed by phonetics teaching staff at London University (UCL) and is available at [http://www.phon.ucl.ac.uk/resource/tutorials.html#phon](http://www.phon.ucl.ac.uk/resource/tutorials.html#phon).

Acknowledgement

An earlier version of this paper was presented at the 12th Generative Approaches to Language Acquisition (GALA 12) conference hosted by University of Nantes, France, on 10-12 September 2015. We are grateful to a GALA reviewer and to members of the audience for their insightful comments. This research was partially funded by grants (no. 26284067 and no. 15K02611) from the Japanese government (Grants-in-Aid for Scientific Research (B) and (C), Ministry of Education, Culture, Sports, Science and Technology (MEXT)).

References


ACTIONALITY SPEAKS LOUDER THAN FELICITY: ENGLISH-SPEAKING CHILDREN’S COMPREHENSION OF PASSIVES

EMMA NGUYEN, DIANE LILLO-MARTIN AND WILLIAM SNYDER

1. Introduction

In many studies, English-speaking children do not seem to understand actional passives with a ‘by’-phrase (henceforth, long passives) such as (1) until about age 4 or 5 (Bever 1970; Horgan 1978; de Villiers & de Villiers 1978). These studies use verbs with ‘reversible’ interpretation, meaning that the arguments used are equally plausible as subject or object. Furthermore, it has been claimed that children do not begin to understand non-actional long passives such as (2) until as late as age 7 (e.g., Maratsos et al. 1998). However, although many studies have found children to have delayed acquisition, others have found adult-like performance at an early age. It is this discrepancy in the literature regarding children’s acquisition of the English reversible verbal passive that we investigate in the present research.

(1) The dog was hugged by Ernie.
(2) The dog was liked by Ernie.

The idea that children are delayed in their understanding of long passives is long-standing, leading some researchers to propose that young children lack an adult-like grammar for the verbal passive. For example, Borer & Wexler (1987, 1992) give an account whereby the ability to form A-chains, a necessary component of the verbal passive, is biologically inaccessible to young children.

Contrary to the studies finding delays, other studies have found early understanding of verbal passives. Fox & Grodzinsky (1998), for example, found that many of the children in their study (10 out of their 13 subjects, age 3;06-5;05 years) performed perfectly on short non-actional passives as well as on long actional be- and get-passives. Many of these same
children (8 of the 10) had difficulty on non-actional passives if and only if the passive included a by-phrase. Furthermore, as will be discussed below, O’Brien, Grolla, & Lillo-Martin (2006) found adult-like performance on both long actional and long non-actional passives, under certain testing conditions, in both 3- and 4-year-old children. The question leading to our study was why children showed adult-like performance on long passives in some studies, but not in others. In particular, we wanted to test the generalizability of O’Brien et al.’s (2006) results to a new, larger sample of children.

The paper is structured as follows: Section 2 will describe O’Brien et al.’s (2006) experiments and their results. Section 3 will describe the results from our first experiment testing the generalizability of O’Brien et al.’s results. Section 4 will describe a follow-up study. Section 5 contains a general discussion of our findings, including future directions.


O’Brien, Grolla, & Lillo-Martin (2006) started from the observation by Crain & Fodor (1993) that long passives have a low frequency of occurrence in the speech of adults as well as (older) children. Crain & Fodor suggested that this low frequency may be due to long passives being "marked" forms that are appropriate only in certain discourse situations. O’Brien et al. hypothesized that children generally have difficulty with long actional and long non-actional passives precisely because the contexts in which they are presented in experiments are pragmatically inappropriate for the use of a by-phrase.

3- and 4-year-old children were tested to see whether their performance on long passive sentences would improve if the sentences were presented in contexts that made the by-phrase more pragmatically appropriate. For the authors, a story was deemed to be pragmatically appropriate for the by-phrase if, in addition to the character corresponding to the actual agent or experiencer, there was another character who could have been the agent or experiencer. The contrast between the potential and the actual agent/experiencer motivated the use of a by-phrase, because otherwise this information was unavailable to the listener. Note that this manipulation motivated the inclusion of a by-phrase, but may or may not have been sufficient to motivate the use of the passive voice in the first place.

O’Brien et al. used a Truth-Value Judgment Task (Crain & McKee 1985) to see whether children would accept passive sentences as
descriptions of stories acted out using toy props. So while one experimenter told a story to child and acted it out with toys, another experimenter delivered the test sentence while manipulating a puppet, “Gobu”. The child was asked to indicate whether Gobu’s utterance was an appropriate description of the story by either rewarding or correcting the puppet.

Each child saw stories that contained either an actional verb (hug, chase) or a non-actional verb (see, like). O’Brien et al. tested the 3- and 4-year-olds separately. 4-year-old children (N=11) only received test stories that contained the additional character, whereas 3-year-olds (N=7) received some stories that did, and some stories that did not, have the extra character. Crucially, the presentation of the stories was blocked so that the three-year-olds always saw the stories without an extra character before the stories that included an extra character.

Taken from O’Brien et al. (2006), a sample story containing two potential agents/experiencers is shown in (3).

(3) Long Actional Passive with 2 Potential Agents
(Matched Sample Story)

Exp1: Bart, the gorilla, and the cheetah were relaxing in the jungle one day, when Bart found a bunch of bananas.
Bart: Hey, cool! Look what I found!
Gorilla: Would you mind sharing some of those with me?
Bart: No way, dude, these are mine, all mine! Hee, hee. If you want some, you’re gonna have to chase me.
Cheetah: I could chase him, but I’m not all that fond of bananas.
Gorilla: Well bananas are my favorite, so watch out Bart, here I come!!!! (Gorilla chases Bart)
Exp1: Gobu, can you tell me something about the story?
Gobu: Well, let’s see. In that story, Bart was chased by the gorilla.

O’Brien et al. found that when 3- and 4-year-old children were presented with stories that contained two potential agents/experiencers, they were significantly above chance for both types of long passives, even though long non-actional passives had been reported to be difficult for children that young (Maratsos et al. 1998; Fox et al. 1995). Additionally, when the 3-year-olds were shown stories with only one potential agent/experiencer, they performed at chance on long actional and long non-actional passives.

The authors interpreted their results as evidence that young children
actually have adult-like comprehension of English long passives, as long as the sentences are presented in contexts that properly motivate the use of a *by*-phrase.

In Experiment 1 we tested the generalizability of O’Brien et al.’s (2006) findings to a new, larger sample of children. Specifically, we checked whether performance improved when the children were presented with stories that featured two potential agents/experiencers. In contrast to the earlier study, however, Experiment 1 intermixed test items with long and short passives.

### 3. Experiment 1

In this study, a different mode of presentation was used, as described below; otherwise, the study shared with O’Brien et al. (2006) the general approach of contrasting stories with or without an extra character as a potential agent/experiencer. 3- and 4-year-olds were tested to see whether the presence of the additional character improved their performance on long passives. Children were also presented with short passives in order to check whether the new methodology would replicate previous findings that children often perform better on passives without a *by*-phrase (Horgan 1978; Fox et al. 1995).

#### 3.1. Subjects & Procedure

25 monolingual English-acquiring children were tested (3;00-5;03, mean age=4;03) using a modified version of the Truth Value Judgment Task (henceforth laptop-TVJT) (Crain and McKee 1985). Instead of having two experimenters act out the stories with toys and manipulating a puppet, the materials were presented through a laptop. A child would watch previously recorded videos of the stories. After each story, a puppet, named “Oscar”, would come on the screen as part of the video, but portrayed to the child as via “webcam”, and deliver the test sentence. After each test sentence, children were asked to indicate whether the puppet was “right” or “silly” by stamping a report card.

In order to be included in the data analysis, the child had to have answered correctly to at least three out of the four control items (i.e., scoring at least 75% correct), and not exhibit a bias towards either answer
(i.e., a child who gave the same answer to 90% or more of the test items was removed from further analysis).

3.2. Materials

The stories that were created were generally similar in plot but varied in the number of characters: either there was only a single potential agent/experiencer, in addition to the patient/theme (henceforth a “2-Character Story”), or there were two (henceforth a “3-Character Story”). Similar to O’Brien et al. (2006), two actional and two non-actional verbs were tested: *hug, chase, see*, and *like* respectively. For each verb, three stories were created: a Long Passive in a 3-Character Story, a Long Passive in a 2-Character Story, and a Short Passive in a 2-Character Story.

Whether the test sentence matched the story was counterbalanced. After two initial active controls, the stories were presented in an intermixed and pseudo-randomized order so as to discourage children from developing a bias towards any particular pattern of answers.

In total, each child watched 16 stories on the laptop: four active controls and 12 test items. Four sample stories are given in (4-7).

(4) Long Non-actional Passive in a 2-Character Story (Mismatched Sample)

**Narrator:** This is a story about Santa and a lion. The lion is mean and grumpy. He doesn’t like anyone. But I wonder if there’s anyone that likes the lion. Here’s Santa. I wonder if Santa likes the lion.

**Santa:** I know that the lion doesn’t like me but he has such nice hair and a nice long tail. So yes, I like the lion very much!

**Experimenter:** Oscar, can you tell me something about the story?

**Oscar:** Santa was liked by the lion.

(5) Long Non-actional Passive in a 3-Character Story (Matched Sample)

**Narrator:** This is a story about Snow White, a dwarf, and a cow. The cow doesn’t like a lot of people. But I wonder if anyone likes the cow? Here’s Snow White. I wonder if Snow White likes the cow?
Snow White: No way! I don’t like the cow because he has such silly spots!
Narrator: Okay, well here’s the dwarf. I wonder if the dwarf likes the cow?
Dwarf: I know that the cow doesn’t like me but I like the cow. I like his spots and his horns very much. I like you, cow!
Experimenter: Oscar, can you tell me something about the story?
Oscar: The cow was liked by the dwarf.  

(6) Short Actional Passive in a 2-Character Story
(Mismatched Sample)
Narrator: This is a story about Lisa and Fancy Lady. Lisa and Fancy Lady are playing one day.
Lisa: Fancy Lady, let’s play a chase game. I dare you to chase me around that trashcan.
Fancy Lady: But Lisa, I wanted to play with my box by the house!
Lisa: Come on, Fancy Lady. You can play with your box after you chase me.
Fancy Lady: Okay, I’m going to chase you now, Lisa!
Experimenter: Oscar, can you tell me something about the story?
Oscar: Fancy Lady was chased.

(7) Short Actional Passive in a 2-Character Story
(Matched Sample)
Narrator: This is a story about a giraffe and Daisy Duck. The giraffe is a lonely creature.
Giraffe: [Sigh] I am so lonely!
Narrator: Oh, here’s Daisy Duck. The giraffe is always nice to Daisy Duck.
Daisy Duck: Giraffe, why are you alone? Maybe I can give you a hug. Yes, I’m going to give you a hug because you are always so nice to me!
Experimenter: Oscar, can you tell me something about the story?
Oscar: The giraffe was hugged.
3.3. Results

The percentage of accurate responses across conditions with both the long and short passives is presented in Figure 1. Children’s performance on short passives was significantly above chance (by one-sample $t$-test, two-tailed $p < .05$). Children were not significantly better than chance on long non-actional passives, regardless of story type. Children were significantly above chance on long actional passives (by one-sample $t$-test, two-tailed $p < .05$) when presented with stories that only had one potential agent/experiencer, but their performance was only marginally better than chance on these passives when presented with stories that had two potential agents/experiencers (by one-sample $t$-test, two-tailed $p < .05$). A summary of children’s accuracy as compared to chance is reported in Table 1.

Actionality had a marginal effect on children’s performance on long passives when 2- and 3-character stories were collapsed ($W = 71$, two-tailed $p = .0688$). The pragmatic appropriateness of the by-phrase, however, had no effect on their performance (Wilcoxon Signed-Rank Test, $W = -37$, two-tailed $p = .298$).

![Figure 1: Comprehension of passives by 3- and 4-year-olds](image-url)

(Error bars indicate standard error)
Table 1: Comparisons to Chance, for Experiment 1

<table>
<thead>
<tr>
<th>TYPE</th>
<th>Percent Correct</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actional, long, 2-Char</td>
<td>70%</td>
<td>3.10</td>
<td>24</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Actional, long, 3-Char</td>
<td>64%</td>
<td>1.90</td>
<td>24</td>
<td>.0696</td>
</tr>
<tr>
<td>Actional, short</td>
<td>76%</td>
<td>5.10</td>
<td>24</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Non-actional, long, 2-Char</td>
<td>58%</td>
<td>1.07</td>
<td>24</td>
<td>.294</td>
</tr>
<tr>
<td>Non-actional, long, 3-Char</td>
<td>50%</td>
<td>0</td>
<td>24</td>
<td>1</td>
</tr>
<tr>
<td>Non-actional, short</td>
<td>66%</td>
<td>2.32</td>
<td>24</td>
<td>.0293</td>
</tr>
</tbody>
</table>

It is important to note that O’Brien et al. (2006) had a different inclusion criterion from the one we applied: O’Brien and colleagues only included the participants who passed the active controls and answered more than 50% of the test items correctly. It is possible that applying the same inclusion criterion to our sample would render our results more similar to theirs. When we did so it led to the exclusion of three additional participants, and this changed the results slightly. Children performed significantly above chance on all actional passives, regardless of length, and on the short non-actional passives. But as in the original sample, children performed at chance on the long non-actional passives, regardless of story type. Table 2 shows this new group’s accuracy, across conditions, as compared to chance by one-sample t-test.

Interestingly, there was no longer a main effect of actionality ($W = 39$, two-tailed $p = .177$). Furthermore, there was still no effect of story type ($W = -16$, two-tailed $p = .542$). Figure 2 shows the percentages of accurate responses from this new sample.

Table 2: Results of Experiment 1 (Modified Sample)

<table>
<thead>
<tr>
<th>TYPE</th>
<th>Percent Correct</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actional, long, 2-Char</td>
<td>72%</td>
<td>3.17</td>
<td>21</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Actional, long, 3-Char</td>
<td>70%</td>
<td>2.881</td>
<td>21</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Actional, short</td>
<td>75%</td>
<td>4.583</td>
<td>21</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Non-actional, long, 2-Char</td>
<td>64%</td>
<td>1.82</td>
<td>21</td>
<td>.083</td>
</tr>
<tr>
<td>Non-actional, long, 3-Char</td>
<td>57%</td>
<td>.720</td>
<td>21</td>
<td>.480</td>
</tr>
<tr>
<td>Non-actional, short</td>
<td>70%</td>
<td>2.881</td>
<td>21</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>
3.4. Discussion

While children generally had better performance on short passives than on long passives, regardless of verb type, as expected from Horgan (1978) and Fox et al. (1995), the results for long actional and long non-actional passives were more unexpected. The results of Experiment 1 were different from those of O’Brien et al. (2006): children performed similarly on 3-character stories and 2-character stories.

It is important to note that the materials and design of this experiment were not identical to those of O’Brien et al. (2006). Experiment 1 presented young children with short as well as long passives, and intermixed the test items instead of blocking them. It is conceivable that these differences made the task harder for children and thus led us to obtain different results. We therefore conducted a follow-up study in which the task was modified so as to resemble more closely that of O’Brien et al. (2006).

4. Experiment 2

One notable difference between our Experiment 1 and the studies conducted by O’Brien et al. (2006) was that we intermixed the test items
in a pseudo-random order. Focusing on children’s performance on long passives, Experiment 2 aimed to test whether blocking the test items of Experiment 1 would yield different results.

4.1. Methods

19 children were tested (3;05-5;02, mean age=4;03) using the laptop-TVJT methodology described in Section 3.1. The materials were slightly modified in this version to correspond more closely to the methods used by O’Brien et al. (2006). Instead of intermixing short and long passives, children were only presented with long actional and non-actional passives. Crucially, these long passives were presented in a blocked order: all the stories with only a single potential agent/experiencer (i.e. “2-Character Story”) were presented prior to the stories with two potential agents/experiencers (i.e. “3-Character Story”).

For each of the four verbs (see, like, chase, and hug), two passive sentences were presented (one long passive in a 2-Character Story and one long passive in a 3-Character Story). Due to the new focus on long passives, the short passives used in Experiment 1 were eliminated, leading to an overall reduction in total test items. In addition, two training items in the active voice were added at the beginning of the experiment in order to help children get better acquainted with the methodology. Children were presented with 14 stories in all (2 training, 4 active controls, 8 test items).

Children were included and analyzed based on the same criteria as outlined in Section 3.1.

4.2. Results

Figure 3 shows the results of a single-sample t-test for each of the experimental conditions. Children's performance was not significantly better than chance in any of these conditions. There was no effect of story type (W = -6, two-tailed p = .873) by the Wilcoxon Signed-Rank Test. Furthermore, the marginal effect of actionality found in Experiment 1 was absent (W = 33, two-tailed p = .308). Figure 3 shows children’s accuracy on long passives, across conditions.
Table 3: Results of Experiment 2

<table>
<thead>
<tr>
<th>TYPE</th>
<th>Percent Correct</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actional, long, 2-Char</td>
<td>61%</td>
<td>1.29</td>
<td>18</td>
<td>.215</td>
</tr>
<tr>
<td>Actional, long, 3-Char</td>
<td>55%</td>
<td>.698</td>
<td>18</td>
<td>.494</td>
</tr>
<tr>
<td>Non-actional, long, 2-Char</td>
<td>50%</td>
<td>0</td>
<td>18</td>
<td>.50</td>
</tr>
<tr>
<td>Non-actional, long, 3-Char</td>
<td>53%</td>
<td>.294</td>
<td>18</td>
<td>.772</td>
</tr>
</tbody>
</table>

Figure 3: Comprehension of long passives by 3- and 4-year-olds
(Error bars indicate standard error)

4.3. Discussion

Blocking of the test items did not lead to improved performance. There was no effect of story type, and there was no longer the marginal effect of actionality seen in Experiment 1. It is not clear why this is the case but it is possible that the laptop-TVJT methodology was somehow unsuitable for this study. It would be interesting to test within subjects to see whether children perform differently on a laptop-TVJT and a traditional TVJT when it comes to understanding passive constructions. It is unclear at the moment whether the laptop-TVJT utilized in Experiment 1 and Experiment 2 had a negative impact on children’s performance. So while we did not find a similar pattern to O’Brien et al.’s results in Experiment 1 or 2, it is still possible that their results would replicate if the experimental method were entirely identical to theirs.
5. Conclusions

We conducted two experiments that tested the generalizability of O’Brien et al.’s (2006) findings. But while O’Brien et al. found adult-like performance in young children on long actional and non-actional passives, our results were different. In Experiment 1 we found no improvement in children’s performance when a second candidate for the agent / experiencer role was introduced. We did find a marginal effect of verb type, however, in which children performed better on the passives of actional verbs than non-actional verbs.

In Experiment 2, we sought to reduce potential difficulties in Experiment 1 by eliminating the short passives, and by blocking the test items. These modifications did not improve children’s performance. In fact, it seems that children performed worse than they had when the items were intermixed. It is unclear why children performed so poorly on this task and why their performance was not significantly better than chance on the actional passives, as it had been in Experiment 1. For this reason, we will focus on the results from Experiment 1 for the remainder of the discussion.

Our results were substantially different from those of O'Brien et al. One possible explanation is that a laptop-TVJT is not appropriate for testing children's knowledge of passives. O’Brien et al. (2006) conducted a traditional Truth-Value Judgment Task, in which children were presented with stories acted out in front of them using toy props. It is possible that this difference yielded different results. We are now conducting a study to test whether a traditional version of the Truth-Value Judgment Task yields results more like those of O'Brien et al.

Another possible explanation is that O’Brien et al.’s (2006) findings are not generalizable, perhaps because their study happened to include high-performing children not representative of the larger population. In this case, it is possible that children are not in general influenced by the presence of another potential agent/experiencer. Or, it could be that other factors override this in most studies, such as the discourse motivation for the use of passive versus active sentences in the first place.

Aside from the question of generalizing the results of O’Brien et al., the results from this study show, along with previous research, that children do seem to be influenced by the presence of a by-phrase. This is evidenced by children’s significantly better-than-chance performance on short non-actional passives, and their worse performance on long non-actional passives, in Experiment 1. Children also seem to be influenced by the type of verb that is used in the passive. We found a marginal effect
of verb type in Experiment 1, where children were better at actional passives than non-actional passives. Differences in performance on actional vs. non-actional passives have also been found in other studies (Hirsh & Wexler, 2006; Maratsos et al., 1987). This suggests that any viable account of the acquisition of the English verbal passive must account for this asymmetry.

Our study adds to the large body of research that has been conducted on English-speaking children’s knowledge of the passive. We hope to continue this line of research and gain a better understanding of the differences between our own results and those of O’Brien et al. (2006).

Notes

1 In the O’Brien et al. (2006) study, the 4-year-olds were also tested with the verb hear but this verb was found to be more difficult for children than any of the other verbs. Because there have been other reports in the literature of children’s general difficulty with this verb, the verb hear was taken out from subsequent analysis (Maratsos et al. 1985; Fox et al. 1995). When 3-year-olds were tested, the verb like was used for the second non-actional verb.

2 This modification was not expected to affect performance given that a within-subjects experiment conducted by Conroy (2008) (not on passives) found no difference between live and laptop presentation of a TVJT.

3 The mismatch response would be “The dwarf was liked by the cow.”

4 Several people, including an anonymous reviewer, have reported that they do not think the test sentences for the Short Passive condition sound felicitous. This is because there is no manipulation serving to motivate the use of passive rather than active voice (following O’Brien et al.). However, these materials have been tested on undergraduate students at the University of Connecticut, and the majority gave the expected answers.

5 We found good performance on short non-actional passives as compared to long non-actional passives. It is important to note that there have been a few studies that also found a difference between short and long non-actional passives, such as O’Brien et al. (2006) and Fox et al. (1995). Other studies have found poor performance on non-actional passives, regardless of length (Maratsos et al. 1987; Gordon & Chafetz 1990; Hirsh & Wexler 2004).

6 Specifically, we are modeling our materials on Experiment 2 of O’Brien et al. (2006). This is because their Experiment 1 tested both short and long passives, and only used 3-Character Stories. It is Experiment 2 where children were presented with blocked test items.
Acknowledgements

We'd like to thank Susi Wurmbrand, Lyn Tieu, and an anonymous reviewer for extensive comments and discussion on the paper, Garrett Smith for assistance with stimuli creation, and the audiences at AcquiLab and the UMass-UConn-Smith Language Acquisition Workshop. We wanted to further extend appreciation towards UConn K.I.D.S and all of the daycares, parents, and children for participating in this project. And lastly, this project was funded by a National Science Foundation IGERT (Integrative Graduate Education and Research Training) grant at the University of Connecticut and approved by the UConn Institutional Review Board, protocol #H14-212.

References


1. Background: Heritage Speakers

Heritage languages (HLs) are minority languages in a specific socio-cultural context in which the home language differs from the dominant language of the greater community. Benmamoun et al. (2013: 132) define the term "heritage speaker" as typically referring to children of immigrants who live in bilingual contexts. Such children are generally exposed first to their family’s home language (their heritage language) and have strong cultural connections with the language (Fishman 1964) that often persist into adulthood. Then in early childhood, these children begin to receive increased input in the majority language of the greater community. Under these circumstances they may experience a switch in their language dominance, especially once they enter formal education, which is typically delivered in the majority language. Due to the language input patterns just described, heritage speakers acquire their heritage language naturalistically and look similar to monolingual native-speaker comparison peers in some aspects of their language development, but divergent in others (Rothman 2009: 156). In some domains, the HL may not be completely acquired because of the dominance shift (Polinsky & Kagan 2007: 369). In still other cases, it is possible that some feature of the HL had once been acquired, but due to insufficient use and/or the influence of the more dominant language, there is attrition or degradation in the HL (Polinsky 2011). As is typical in most cases of bilingual acquisition, heritage speakers display wide variation in proficiency due to complex interactions of input factors such as parents’ language background, contexts of use, attitude towards the heritage language in the local community, etc. Relatively restricted input, compared to that
received by non-heritage speakers, may have significant effects on the acquisition and maintenance of heritage languages (see Benmamoun et al. 2013; Scontras et al. 2015).

Heritage language effects can be noted in different linguistic domains. In this paper, we summarize evidence for HL characteristics in bimodal bilingual children in the domains of phonology, syntax, and discourse. For evidence of HL effects in the lexicon and morphology of bimodal bilingual adults from the US and Brazil, see Quadros et al. (2016, in preparation). First, we set the stage by summarizing research on HL effects in these domains for spoken language bilinguals.

Heritage speakers may seem most comparable to monolingual native speaker peers in the domain of phonology. In both production and perception, they significantly outperform second language learners (e.g., Lukyanchenko & Gor 2011), for whom phonological abilities pose a notable difficulty. Heritage speakers may also maintain native-like phoneme discrimination (Oh et al. 2003). Nevertheless, in phonological production heritage speakers may differ from monolingual peers (Au et al. 2002; Godson 2004), for example by producing vowels more similar in quality to those in their second language.

There have been numerous studies of HL effects in syntax, finding a degree of proficiency but some overall simplification patterns such as a reliance on basic word order (Scontras et al. 2015), and some difficulties with more complex structures such as relative clauses (Polinsky 2011) and long-distance dependencies in binding (Kim et al. 2009). Other syntactic effects are related to the strong tendency for morphology to be affected, particularly in the areas of case marking (Montrul et al. 2012; Polinsky 1997) and gender (Polinsky 2008).

At the discourse level, many issues are observed for heritage speakers, and some characterize the syntax-discourse interface as one of the most vulnerable domains (along with morphology) in heritage language grammars (e.g., Benmamoun et al. 2013). This vulnerability is apparent in the expression of referents across discourse, including the use of overt versus null pronominals (Sorace 2004, 2011). Researchers have observed that speakers use fewer null subjects in their HL, across a wide range of languages, possibly reflecting processing demands of using and interpreting null pronouns (Sorace & Serratrice 2009). Discourse-pragmatic features are also commonly affected in HL, including aspects of scope interpretation and implicatures (Benmamoun et al. 2013, Scontras et al. 2015).

This brief overview of HL effects demonstrates that although there is great variation in language development, interaction and use among
heritage speakers, researchers have identified a number of characteristic patterns that apply to bilinguals speaking a wide variety of heritage languages. Presumably, these patterns might also apply in contexts where the heritage language is not a spoken language, but a sign language. With this in mind, we introduce our research on sign+speech bilinguals, also known as bimodal bilinguals.

2. Bimodal Bilinguals: Codas and DDCI

Broadly speaking, the term bimodal bilingual applies to any individual who uses both a signed language and a spoken language. Most commonly, this term includes hearing individuals who are raised in signing Deaf families, known as Codas (for Child of Deaf Adults; or sometimes Kodas, for children and teenagers, although in this paper we will use Codas to designate either adults or children), Deaf signers who access a spoken language through a visual form, hearing L2 signers, children (Deaf or hearing) whose hearing parents use a sign language, and Deaf children of Deaf parents who access spoken language through cochlear implants (henceforth abbreviated DDCI, in contrast to DHCI, or cochlear implanted children from hearing families). This chapter will focus only on Codas and DDCIs (see Chen Pichler and Koulidobrova 2015 for an overview of L2 sign language research).

The proportion of signed versus spoken language input that bimodal bilingual children receive varies according to several factors, such as whether one or both parents are Deaf and whether they have regular contact with other sign language users, such as Deaf siblings or extended family members. The same factors also lead to variation in age of exposure to the majority spoken language; children who have a hearing parent and/or hearing older siblings may be exposed to regular spoken language input from birth, while others may have only minimal contact with the spoken language until they enter hearing (pre-)schools. For DDCI children, age of implantation (and activation) as well as unilateral versus bilateral implantation are also important factors affecting children’s language input. Setting aside for the moment variations in language input and corresponding language proficiency, we will assume the following broad definitions of Codas and DDCI:

(1) Codas: Audiologically hearing individuals with at least one Deaf parent and regular exposure to a sign language and a spoken language within the first six years of life
Demographically, the frequencies with which Codas and DDCI occur across Deaf families differ dramatically. Researchers estimate that between 80-90% of the children born to Deaf, signing families are Codas (Mitchell et al. 2006; Singleton & Tittle 2000). In contrast, only an estimated 5% of deaf children born in the US are born into Deaf, signing families (Mitchell & Karchmer 2004), and of those children, only a small (but growing) percentage receive cochlear implants (Paludneviciene et al. 2011; Mitchiner 2015). Codas and DDCI also typically differ in age of exposure to spoken language and explicit training in that language. Whereas Codas are born with normal hearing and are thus able to access spoken language from birth, the U.S. Food and Drug Administration (FDA) has approved cochlear implantation only for children 12 months of age or older, and after the surgery the recipient must wait another two to four weeks for the device to be activated and calibrated before they can begin to access speech. DDCI children also receive intensive and explicit training in speaking and listening as part of the services supporting their cochlear implant, much more than Coda children typically receive (although see Chen Pichler et al. 2014 for a discussion on the frequent misdiagnosis of Codas as language delayed, on the assumption that their Deaf parents do not provide adequate spoken language input for normal development). Due to these differences in early speech input, we consider Codas and DDCI to be potentially distinct subpopulations within the category of bimodal bilinguals. This distinction also reflects the prevalent view in the Deaf community that even DDCI with excellent spoken language development are still fundamentally Deaf, while Codas are hearing.

Although there are important differences between the groups, in terms of their bilingual development, Coda and DDCI children often pattern together. In a series of longitudinal and experimental studies conducted as part of our research project, Development of Bimodal Bilingualism, we have documented numerous similarities in the spoken language and sign language development of Coda and DDCI children in the US and Brazil. Both groups emerge as successful spoken language users, by a variety of measures. Davidson et al. (2014) report similarly high scores for American Codas and DDCI on standardized tests of English expressive
vocabulary, articulation, phonological awareness, productive syntax and general linguistic development. With respect to sign language development there is more variation across children, among both Codas and DDCI, a point to which we will return below. However, in a test of receptive ASL skills (Enns & Herman 2011), most American Codas and DDCI between the ages of 4;0 and 7;0 scored within or above the range reported for native signing Deaf children without cochlear implants (Davidson et al. 2014; Palmer 2015), broadly confirming that like their (unimplanted) Deaf counterparts, both Codas and DDCI are acquiring ASL as an L1. However, detailed analysis of specific aspects of Coda and DDCI sign language development reveals notable divergences from previously reported developmental patterns for native signing Deaf children. We turn to these divergences next, focusing in particular on code-mixing, phonology, word order and referent cohesion. These developmental patterns, along with sociolinguistic and cultural aspects of Coda and DDCI relationships with their native sign language, support the proposal that bimodal bilingual children are heritage signers, comparable in many ways to heritage speakers.

3. Bimodal bilingual children as heritage signers

3.1. Identity, input, and variation

Heritage speakers are characterized by particular linguistic features, as well as unusual patterns of language input, great variation in proficiency in the heritage language, and a strong familial or cultural connection to that language. Compton (2014: 275), arguing that ASL qualifies as a heritage language for many child signers, highlights strong familial ties to ASL as the primary hallmark of heritage signers, whatever their signing competence. Similarly, Pizer (2008) notes that many of the adult Codas whom she interviewed feel strong emotional connections to ASL, even if ASL is not necessarily their dominant language, as illustrated by the following quote from one of Pizer’s interviewees, reprinted in (3).

(3) “...even though, when I speak English I don’t have to think, it’s still something that I have to do. ASL is just something that I am.” (Pizer 2008: 35).
In terms of language input, Coda and DDCI children also resemble heritage speakers in that their heritage language is a minority language used in restricted contexts, mainly centered around their home and Deaf community (e.g. Deaf churches and other social gatherings). Similar to heritage speaker siblings who prefer to speak the dominant language with each other rather than their heritage language, bimodal bilingual siblings tend to interact in spoken language rather than in sign language (Pizer 2008). With the exception of a small segment of Coda and DDCI children living in areas with large Deaf populations who may enroll for part of their educational careers in daycare programs, preschools or schools established for signing children, where sign language is used as a language of instruction, school-aged bimodal bilingual children spend the vast majority of their waking hours in a spoken language environment. This dramatic imbalance in the levels of input they receive in their sign language and their spoken language is a key factor underlying wide variations in signing proficiency across Coda and DDCI signers, another hallmark of heritage speakers (Benmamoun et al. 2013).

In contrast, native (unimplanted) Deaf signers, especially those attending schools for the Deaf or specialized programs for Deaf students in hearing schools, typically receive more sign language input throughout the day than their Coda and DDCI counterparts. They also socialize regularly with other signing Deaf children; socialization with in-group peers sharing a common ethnic language has been identified as a factor correlating with increased heritage language proficiency (Phinney et al. 2001). Indeed, in studies of L1 sign language acquisition, native-signing Deaf children raised in Deaf families and attending Deaf schools are presented as a homogenous group, acquiring their sign language fully and consistently. In the following subsections, we detail several studies comparing Coda and DDCI sign language patterns with those of native Deaf signers. In several of the studies, the native-signing Deaf participants cluster as a group, while the bimodal bilingual participants display more individual variation. In many cases, the sign language production of the bimodal bilingual children diverges from what has been reported for native-signing Deaf children at comparable ages. Taken together, the results from these studies support the view that the term *heritage language* applies equally well to spoken languages and sign languages.
3.2. Code-blending

As bilinguals, Codas and DDCI are prone to code-mixing, a common occurrence among spoken language bilinguals. However, code-mixing among bimodal bilinguals rarely manifests as code-switching, although it is physically possible to switch from signing to speech and back again. Bimodal bilinguals are much more likely to code-blend, simultaneously producing elements of their sign language and spoken language (Emmorey et al. 2008; Lillo-Martin et al. 2016). Frequent and sustained code-blending, often involving whispering rather than full phonation (Petroj et al. 2013), is regarded by many as a hallmark of bimodal bilingual production, constituting a striking and immediately noticeable contrast with signing by Deaf individuals who access the dominant language of the surrounding hearing community through print rather than speech. Although many Deaf signers also code-blend in certain contexts (e.g., Deaf parents at home with their hearing children), many of them choose not to use speech for personal or cultural reasons. Deaf code-blending has thus received far less attention to date in the literature than bimodal bilingual code-blending.

Combining two languages in a way that respects the distinct grammatical and lexical properties of both languages requires fairly sophisticated bilingual knowledge (Poplack 1980; Van den Bogaerde & Baker 2005). Quadros et al. (in preparation) and Quadros (in press) report that balanced adult bimodal bilinguals can use their two languages interchangeably, depending on the context, as the primary or the secondary language. The primary language provides more of the grammatical structures in evidence, although aspects of each language can appear in the same sentence. According to this approach, the derivation of a sentence starts with grammatical features from one or both languages, with lexical and morphological output from either or both simultaneously (Lillo-Martin et al. 2010; Koulibrova 2012; Lillo-Martin et al. 2012; Lillo-Martin et al. 2016). Bimodal bilinguals can also switch from using one language as the primary language to the other, or they may use both languages as primary languages simultaneously, which requires special skill to select grammatical structures from each language that do not violate the grammar of the other.

Preference for code-blending over code-switching is a robust bimodal bilingual pattern that has been repeatedly documented for Codas, from toddlers (Petitto et al. 2001) to adults (Emmorey et al. 2008), and in a variety of sign languages (e.g., Kanto et al. 2016 for Finnish Sign Language; Fung 2011 for Hong Kong Sign Language; Petitto et al. 2001
for Langue des Signes Québécoise). Longitudinal recordings of young bimodal bilinguals collected by our research group in the U.S. and Brazil reveal that the children generally match their language choices to those of their interlocutors, producing more signed utterances when interacting with Deaf interlocutors, and more spoken utterances when interacting with hearing interlocutors (Lillo-Martin et al. 2014; see also Petitto et al. 2001; Kanto 2016). However, they show differential code-blending patterns, producing the most code-blends when interacting with Deaf interlocutors who cannot hear speech. We propose that code-blending in these cases occurs for the benefit of the bimodal bilingual children, perhaps as a mechanism for easing the cognitive costs of suppressing their dominant language (spoken English or Portuguese) while signing (Lillo-Martin et al. 2014; Petroj et al. 2013; Emmorey et al. 2008). Experimental studies show that bimodal bilinguals make faster semantic judgements with code-blends than they do with ASL or English alone, providing additional evidence for the facilitative nature of code-blending for language processing (Emmorey et al. 2012). While language suppression and switch costs are common among speech bilinguals, the ability to mitigate these costs by producing both of one’s languages at the same time is unique to bimodal bilinguals, and underscores the importance of studying this population for a comprehensive understanding of bilingual language development.

3.3. Phonology

Sign language phonological development among Coda and DDCI signers is still a very understudied area. An early study by Kantor (1978) reports that native adult Deaf signers asked to view video samples of anonymous signers and identify them as native Deaf, L2 Deaf, Coda or L2 hearing, had more difficulty identifying Coda signers than native Deaf or hearing L2 signers. Although Kantor’s study does not include any systematic comparison of the phonological features of her Coda or native Deaf signers, the relative difficulty of native Deaf ASL users in identifying Coda signers on the basis of their sign production is in line with general assumptions that heritage language users are relatively successful in their phonological development.

More recently, we have compared phonological accuracy of Coda, DDCI and native-signing Deaf children between 4;0 and 8;0 in their reproduction of invented pseudosigns consistent with phonological constraints in ASL or Libras (depending on the sign language of the
children we test). Children watched short videos of a Deaf adult modeling each pseudosign, then were required to repeat the sign they had just seen. Responses are coded as target-like if they match the target form in handshape, movement, location and palm orientation; errors in one or more of these phonological parameters results in the response being coded as nontarget-like. Kozak (in preparation) reports that Coda and DDCI children performed at slightly lower accuracy than native Deaf controls in pseudosign repetition, although the difference was not significant. Parallel testing in Brazil comparing accuracy of pseudosign reproduction across Brazilian Codas, Deaf cochlear implant users (mostly with hearing parents in Brazil) and native Deaf controls reveals similar patterns, with the control group scoring highest, followed closely behind by Coda signers and Deaf cochlear implanted signers (Cruz et al. 2014). Across all groups, signs with marked features were reproduced with higher error rates than those with less marked features. Signs involving a combination of a marked handshape and a complex movement (path + internal) were universally the hardest for the children to reproduce in target-like fashion. Due to the small percentage difference across all groups, it appears that despite anecdotally-reported differences between Coda and DDCI phonology from that of native-signing Deaf comparisons, all three groups of children in our study performed similarly, consistent with current characterization of heritage speaker phonology.

3.4. Syntax: Canonical and noncanonical word order

Our analyses to date on syntactic development of Coda and DDCI children focus mostly on word order patterns. Palmer (2015) analyzed the early word orders produced by four bimodal bilingual children (two Codas and two DDCI) during longitudinal spontaneous production. The goal of his study was to determine how frequently each child used canonical versus noncanonical word order and to determine when each child satisfied a repeated-use criterion (Stromswold 1996) as a measure of mastery. The results were compared with word order patterns produced by native signing Deaf children, as reported by Chen Picher (2001). ASL is a variable word order language; canonical SVO word order occurs in pragmatically neutral contexts, typically with verbs without complex morphology, but noncanonical orders also frequently occur. Noncanonical VS word order is permissible when the post-verbal subject is a pronoun and most frequently used for emphasis or confirmation (Coulter 1979; Padden 1983). Object-verb (OV) order is permitted with a host of verbs
bearing complex morphology (Liddell 1980; Chen Pichler 2001) encoding spatial, aspect, instrumental or handling information, among others. Finally, certain lexical verbs (e.g. WANT and HAVE) seem to permit post-verbal objects as well (Chen Pichler 2001; Lillo-Martin et al. 2003). Noncanonical orders with preverbal objects and postverbal subjects are quite common in adult signing.

Palmer (2015) analyzed all two-word utterances with a verb (including partially and fully bimodal utterances) and at least one overt argument from 34 hours of video recordings spanning the ages of 20 and 40 months for his four bimodal bilingual participants, following criteria for identifying canonical and grammatically acceptable noncanonical word orders from Chen Pichler (2001). For canonical word orders (SV and VO), the bimodal bilingual children met the repeated-use criterion at 23 months, the same age reported by Chen Pichler (2001) for native signing Deaf children. A mixed effects two-way linear regression reveals that over time there is no difference between the production of canonical SV ($\beta = 0.271$; s.e. = 0.440, $t = 0.616$) and canonical VO: ($\beta = -0.272$; s.e. = 0.313, $t = -0.868$) between the two groups. This suggests that similarly to native signing Deaf children, bimodal bilingual children exhibit control and acquisition of canonical word orders early as expected.

For noncanonical word orders (VS and OV), the native signing Deaf children satisfy the repeated-use criterion at 23 months, confirming Chen Pichler’s reports that these children use both canonical and grammatical noncanonical orders from their earliest multisign combinations. Noncanonical word orders constitute a large portion of the overall utterances produced by the native signing Deaf children: 54% of the utterances containing a verb and an object appear in OV order, and 37% of the utterances containing a subject and verb appear in VS word order. The bimodal bilinguals, however, display a radically different acquisition pattern for noncanonical word orders. For VS word order, two of the participants never reached criterion while the other two participants only did so at 36 months, more than a year after the Deaf controls. As for OV word order, none of the participants met the repeated-use criterion by 40 months of age. In general, noncanonical word orders represent a very small percentage (less than 1%) of the bimodal bilinguals’ total utterances. Linear regression shows that the native signing Deaf children produce significantly more noncanonical OV utterances ($\beta = -6.81$; s.e. = 1.35; $t = 5.03$) and noncanonical VS utterances ($\beta = 5.32$; s.e. = 1.35; $t = 3.93$) than the bimodal bilingual children. This suggests that the bilingual bimodal children are developing quite differently from the Deaf controls with respect to noncanonical word order. Figure 1 illustrates this
difference graphically. The right panel shows that native Deaf children appear to be developing both canonical (SV and VO) and noncanonical (VS and OV) word orders at a steady rate. Although the development of canonical orders (SV and VO) appears typical for the bimodal bilingualism (left panel of Figure 1), the trend lines representing noncanonical VS and OV utterances for the bimodal bilinguals remain flat, indicating there was not much change during the time period study.

![Figure 1: Word Order trend lines for native Deaf (control) and bimodal bilingual (Coda and DDCI) children](image)

Although Palmer (2015) does not explicitly track development of ASL morphology in the signing of his Coda and DDCI participants, he notes a conspicuous absence of the complex morphological forms that license OV word order in ASL and proposes that these bimodal bilinguals diverge in their developmental path from Deaf controls for this aspect of ASL grammar. This proposal is consistent with observations from spoken language research that morphology is a particularly vulnerable feature of heritage grammars and there can be an over-reliance on basic word order.
3.4.1 Syntax: Wh-questions

Word order for Wh-questions in both ASL and Libras is more variable than in English or Brazilian Portuguese, which both place the Wh-element in sentence-initial position. This word order option is grammatical in ASL and Libras, but those languages additionally allow Wh-elements to appear sentence-finally, or both sentence-initially and sentence-finally. This variation is illustrated for ASL in Table 1. In both American and Brazilian contexts, the possible word orders for Wh-questions in the spoken language are a subset of the possible word orders in the sign language.

Table 1: Word orders for Wh-questions in English and ASL

<table>
<thead>
<tr>
<th>English</th>
<th>ASL</th>
</tr>
</thead>
<tbody>
<tr>
<td>What did John buy?</td>
<td>WHAT JOHN BUY?</td>
</tr>
<tr>
<td></td>
<td>JOHN BUY WHAT?</td>
</tr>
<tr>
<td></td>
<td>WHAT JOHN BUY WHAT?</td>
</tr>
</tbody>
</table>

Previous research eliciting Wh-questions from native Deaf signing children (ages 4;0- 6;0) by Lillo-Martin (2000) reported that even the youngest participants produced all three possible ASL word orders, with 5- and 6-year olds producing proportionally more Wh-questions in final and doubled positions than the 4-year olds. In contrast, the same elicitation methodology administered to American Coda and DDCI children reveals a strong preference for sentence-initial Wh-questions, especially for object Wh-questions; Brazilian Coda and cochlear implanted children (including mixed DDCI and DHCI children) behaved similarly to their American counterparts (Lillo-Martin et al. 2012; Quadros et al. 2013), as illustrated in Figure 2 (due to restrictions in space, results for only ASL object Wh-questions are shown here).
Figure 3: Word orders for object Wh-questions elicited from native Deaf controls (top) and Coda and DDCI children (bottom) in the US.

For Wh-questions, as was the case for the word order phenomena investigated by Palmer (2015), the preferred word order for bimodal bilingual children’s spoken language is a subset of the possible word orders for their signed language. In both cases, Coda and DDCI children seem to follow the basic pattern of selecting structures that will be acceptable in both grammars. The strong tendency for bimodal bilingual children to code-blend, mentioned earlier, may lead to choosing matching syntactic structures in the children’s simultaneously activated signed and spoken languages.
3.5. Discourse

Differential developmental patterns at the discourse-pragmatic level have been observed for heritage speakers of null subject languages, as noted earlier. Heritage speakers frequently produce overt forms in contexts where null forms are used by monolinguals (Montrul 2004; Polinsky et al. 2007; Silva-Corvalan 1994), such as in contexts when a previously-mentioned referent is maintained in discourse. This tendency for bilinguals to be “overly overt” surfaces even when both languages allow null subjects, as illustrated by the finding that child Spanish-Italian bilinguals still overproduce overt forms in their non-dominant (presumably heritage) language (Sorace et al. 2009). Some preliminary support for our initial prediction of over-reliance on overt forms by bimodal bilingual children comes from adult L2 learners of sign language, who demonstrate an over-reliance on overt subjects in their signed narratives (Bel et al. 2014 for L2 Catalan Sign Language signers; also, to some extent, Frederiksen et al. 2014 for L2 ASL signers).

Reynolds (2016) elicited ASL narratives at two points in time (Time 1 from 5;2-6;9; Time 2 from 6;8-8;2) for six bimodal bilingual children (three Codas and three DDCI) and compared them to six native-signing Deaf children of roughly similar ages (from 5;5-7;10). ASL allows a variety of null elements, including both pro and Chinese-style null topics (Lillo-Martin 1986). Also, sign language narratives frequently involve depicting verbs, sign-specific forms that show actions or events in a highly iconic fashion (Liddell 2003), often in conjunction with null arguments. All the children produced signed narratives displaying some degree of depiction and null arguments, but the five- to eight-year-old bimodal bilinguals used increasingly fewer null arguments between Time 1 and Time 2 for reference maintenance and reintroduction than their age-matched Deaf controls (Figures 3 and 4). Instead of null forms, the bimodal bilingual children used full noun phrases and a variety of reduced nominal forms including pronouns and fingerspelled entities.
Figure 4: Distribution of narrative referent overt and covert forms in ASL narratives by bimodal bilingual (Bibi) children showing significantly different proportion of overt vs. null forms compared to native-signing Deaf controls (Bibi T1 vs. Deaf, overt vs. null, $\chi^2 8.999$, $p<.003$; Bibi T2 vs. Deaf overt vs. null, $\chi^2 22.650$, $p<.000$)

Figure 5: Distribution of referent forms by function in ASL narratives by bimodal bilingual (Bibi) children at Time 1 (T1) and Time 2 (T1), compared to native-signing Deaf controls
The Coda and DDCI children studied by Reynolds (2016) pattern more like heritage speakers than their native-signing Deaf peers in their use of null arguments for reference cohesion in their signed narratives. At the same time, Coda and DDCI narratives feature innovative and skillful use of depiction and other grammatical devices associated with successful ASL narratives, all of which are reportedly difficult for L2 learners of sign language to master. The spontaneity and ease with which young bimodal bilinguals employ these aspects of ASL grammar in their signed narratives highlights their status as native signers, even if their use of null arguments diverges from expected patterns.

4. Conclusions and implications

The developmental patterns reported for Coda and DDCI signers in the studies discussed above echo general patterns observed for heritage speakers of a variety of languages in minority contexts. Like their native-signing Deaf counterparts, Codas and DDCI receive early exposure to a natural sign language from Deaf parents and develop strong emotional ties to that language. Their receptive skills and phonological production are both quite good, resembling those of native-signing Deaf children. Coda and DDCI also control many other aspects of their sign language grammar with much greater facility than is typically observed for hearing L2 learners. Yet unlike their native-signing Deaf counterparts, Coda and DDCI are typically dominant in their non-sign language and display notable language synthesis effects (transfer) from that language into their sign language. At the syntactic level, we have observed decreased use of word orders that are normally licensed by verbal morphology or discourse factors, both domains that are reported to be difficult for heritage speakers. At the discourse level, Coda and DDCI narratives are characterized by over-reliance on overt forms, mirroring similar patterns in heritage speakers of null argument languages. Taken collectively, the results of the preliminary studies reported here suggest very strongly that the same types of heritage language effects observed for spoken languages exist for sign languages, and that Coda and DDCI children can be viewed as heritage signers.

Compton (2014) points out that given the unusual intergenerational language transmission patterns for sign languages, “the majority of native signers are not deaf but rather hearing,” (276). Accordingly, a comprehensive research program on the grammar and development of sign language must consider data not only from Deaf signers, but from
Codas (and DDCI) as well. Research on heritage signer development is still in its infancy, but already, discussions about the implications and applications of these studies are occurring throughout the Deaf community. For instance, to the extent that the needs and experiences of Codas learning ASL as heritage learners differ from those of typical L2 sign language learners, there is a need for more interpreter training programs tailored to Codas (cf. Williamson 2015, Isakson 2016) such as the COMPASS program (http://www.vrsii.com/compass). Language maintenance is another topic of interest in the Deaf community, as Deaf parents struggle to help their children continue developing their heritage sign language once they become dominant in their spoken language. Researchers point to the importance of studying bimodal bilingual populations to understand how languages in different modalities develop, attrite and interact; some linguistic phenomena, such as code-blending, are unique to bimodal bilingualism and offer insights that would not be available from the study of spoken language bilingualism. Similarly, inclusion of heritage signers in future research will allow a broader understanding of the complex variety of possible heritage language effects.

Notes

1 The extent to which language development/attainment/processing across these populations is parallel remains to be more fully investigated.
2 Following convention, “Deaf” written with a capital D differs from “deaf” in that the former term designates individuals who identify as culturally Deaf and use ASL as their primary mode of communication, while the latter refers to an audiological status. Deaf individuals vary widely in their level of hearing and use of spoken language; what unifies them as a group is their perception of themselves as members of a Deaf community, rejecting a pathological view of deafness associated with disability that needs to be “fixed.”

Acknowledgements

Warm thanks to the bimodal bilingual children and their families who participate in our research and to our research assistants and collaborators. We also are grateful for financial support from: The Gallaudet Research Institute; CNPQ (Brazilian National Council of Technological and Scientific Development) Grant #200031/2009-0 and #470111/2007-0; and
award number R01DC009263 from the National Institutes of Health (National Institute on Deafness and Other Communication Disorders). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDCD or the NIH.

References


Lillo-Martin, Diane, Ronice Müller de Quadros, Helen Koulidobrova, and Deborah Chen Pichler (2010), “Bimodal Bilingual Cross-Language Influence In Unexpected Domains,” in: João Costa, Ana Castro, Maria Lobo, and Fernanda Pratas (Eds.), Language Acquisition and
Heritage Signers: Bimodal Bilingual Children from Deaf Families


Pizer, Ginger Bianca (2008), Sign and Speech in Family Interaction: Language Choices of Deaf Parents and Their Hearing Children. Doctoral dissertation, University of Texas at Austin.


1. Introduction

Since the 80s, following the development of phonological multilinear models, some research has been conducted on the acquisition of different types of consonant groups in several languages of the world, particularly in the context of the observation of how children acquire their mother tongue(s)’ syllabic structure (among others Fikkert 1994; Freitas 1997, 2003; Rose 2000; Freitas & Rodrigues 2003; Goad & Rose 2004; Fikkert & Freitas 2004; Almeida 2011). The identification of the different natures of the various consonant groups studied has been possible thanks to studies that focus on the research of different orders of acquisition and of different repair strategies. Similarly, multiple studies have provided information on the learning of orthography in the early years of schooling (for data on the Portuguese, see Pinto 1997; Alves Martins & Nisa 1998; Miranda 2007; Veloso 2010), usually being the segmental dimension more crafted than the syllabic one. However, possible correlations between phonological development and the consolidation of orthographic knowledge have not been explored. In fact, for European Portuguese, no study has analyzed the relationship between phonological development, focused on implicit knowledge, and learning of writing.

In the context of language acquisition studies, several studies have been published on the acquisition of different syllabic constituents, empirically proving Jakobson’s proposal (1941: 68) that the first structures to be acquired are the non-marked ones, present in all languages of the world and the most frequent, hence the first syllabic structure to be acquired has the CV format (Blevins 1995; Fikkert 1994; Freitas 1997). In European Portuguese, aware of what happens in other languages, branching onsets, which are more complex than the non-branched onsets, are the last syllable structure to stabilize in the child’s phonological system (Fikkert 1994; Freitas 1997; Bernhardt & Stemberger 1998;
Lamprecht et al. 2004). In some children, its productions may not be stabilized yet when going to school (Sim-Sim 1997, 1998; Freitas 1997, 2003; Gonçalves et al. 2011; Mendes et al. 2009, 2013). Additionally, there are some other consonant groups designated in the literature as problematic in what concerns its acquisition, which, being rare in the target system, are also rarely selected by children in an acquisition phase of the European Portuguese (Freitas 1997). In Freitas (1997), in a total of about 18,500 spontaneous productions, only the following lexical targets were registered: *helicóptero* “helicopter” (9 productions, *gnu* “gnu” (1 output), *subtil* “subtle” (2 productions), *pneu* “tyre” (1 output), *Batman* “Batman” (21 productions), *Simpson* “Simpson” (2 outputs). However, these productions were not always produced in accordance with the targets. Its acquisition is thought, therefore, to be even more problematic and delayed than the acquisition of branching onsets, although there are no hitherto experimental data to support this evidence.

Branching onsets are formed by sequences of two consonants. However, due to strong adjacency constraints to the consonants which may occur in this position (The Sonority Sequencing Principle; The Dissimilarity Condition; concerning these principles, see, e.g., Selkirk 1982, 1984; Blevins 1995 for data on Portuguese; see Vigário & Falé 1994; Mateus & d’Andrade 2000; Bisol 2005; Veloso 2006), Portuguese does not accept all consonant clusters as onsets.

The Sonority Sequencing Principle predicts the segment sequences which are permitted in a language taking into account the sonority hierarchy; for the European Portuguese it is set, in Mateus et al. (2003: 1040), as follows: “The sonority of the segments that constitute the syllable increases from the beginning till the nucleus and decreases till the end”1. The hierarchy of the segments is defined by the sonority scale, presented by decreasing sonority: vowels > liquids > nasals > fricatives > plosives (Vigário & Falé 1994). Thus, in onset position, only sequences formed by a stop consonant followed by a lateral /l/ or a vibrant /ɾ/ - ([br]aço “arm”; [pl]uma “plume”) and sequences formed by a fricative consonant followed, likewise, by a lateral or a vibrant - (li[vr]o “book”; [fl]amingo “flamingo”), or obstruent + liquid sequences can occur. These sequences are considered to be tautosyllabic combinations and, as such, they are represented under the domain of the same syllabic node of the onset ([[C1C2]onset[V]nucleus]syllable).

In problematic consonant sequences (*afita* “aphtha”), both consonants are considered to be heterosyllabic as they mandatorily require vocalic epenthesis in Brazilian Portuguese and optional vocalic epenthesis in European Portuguese (*afita* [afita] in Brazilian Portuguese and [afita] in
European Portuguese) as well as they violate phonotactic principles: the Sonority Sequencing Principle and the Dissimilarity Condition. Consequently, Mateus and d’Andrade (2000) suggest that the two consonants are represented under the domain of two non-branched onsets of different and adjacent syllables: the first consonant being the simple onset of a first syllable with an empty Nucleus, this one hosting the epenthetic vowel; the second consonant being the onset of the adjoining syllable, on the right ([C]onset[Ø]nucleus)syllable + [C]onset[V]nucleus)syllable).

(1) Syllabic representation of problematic consonant groups.

The representation above contrasts with the representation adopted for the sequences formed by obstruent+liquid, which do not violate phonotactic principles. They are, as such, analyzed as tautosyllabic and are represented under the domain of the same syllabic node of the onset ([C\textsubscript{1}C\textsubscript{2}]onset[V]nucleus)syllable). In (2), it is given the syllabic representation of branching onsets, in the form of a tree diagram.

(2) Syllabic representation of branching onsets.
The table presented below, adapted from Veloso (2003), sets out briefly the distinguishing characteristics between true branching onsets and problematic consonant groups in European Portuguese.

Table 1: Summary of the distinguishing characteristics of branching onsets and problematic consonant groups (adapted from Veloso 2003: 121)

<table>
<thead>
<tr>
<th></th>
<th>TYPE I Consonant sequences</th>
<th>TYPE II Consonant sequences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Typical segmental combinations</strong></td>
<td>Obstruent+Liquid</td>
<td>Obstruent+Obstruent Obstruent+Nasal</td>
</tr>
</tbody>
</table>
Given the complex nature of both types of consonant clusters, children often resort to the activation of repair strategies, which arise from the phonological and orthographic development stage in which they are. Children’s oral productions are a crucial empirical evidence to understand better the phonological processing (Fikkert 1994). Similarly, children's written productions can provide clues to that same processing as Treiman states (1998: 291): “Children’s spellings also provide an excellent window into their knowledge of phonology and orthography”.

Often used between the two consonants which make up a heterosyllabic consonant cluster, vowel epenthesis is one of the main criteria to distinguish branching onsets (Type I consonant sequences) from heterosyllabic consonant clusters (Type II consonant sequences) in European Portuguese (Mateus & d’Andrade 2000). However, that strategy is rare and seldom attested in other languages of the world (Bernhardt & Stemberger, 1998). Hence, to carry out the research reported in this article, we define the following goal: to identify the level of productivity of epenthetic productions in oral and written outputs of both types of consonant clusters in primary education.
2. Methodology

In this investigation, both spoken and written infantile productions of isolated words containing consonant clusters of both types were observed. 56 children of both sexes participated in this study divided into two experimental groups: (i) 27 1st graders; (ii) 29 4th graders; the couple attending two public primary schools. The distribution of children observed is shown in Table 2:

Table 2: Sample Distribution

<table>
<thead>
<tr>
<th>School</th>
<th>Escola Primária de Turquel</th>
<th>Centro Escolar da Benedita</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1&lt;sup&gt;st&lt;/sup&gt; grade</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1&lt;sup&gt;st&lt;/sup&gt; grade</td>
<td>23</td>
<td>4</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>41,07%</td>
<td>7,14%</td>
<td></td>
</tr>
<tr>
<td>4&lt;sup&gt;th&lt;/sup&gt; grade</td>
<td>23</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>41,07%</td>
<td>10,71%</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>46</td>
<td>10</td>
<td>56</td>
</tr>
<tr>
<td></td>
<td>82,1%</td>
<td>17,9%</td>
<td></td>
</tr>
</tbody>
</table>

Data for this investigation was collected during individual sessions with an instrument specifically created for this study. The tasks applied to the 1<sup>st</sup> graders were composed by 20 pictures that served as incentives for the production of 20 lexical items; in turn, the tasks applied to the 4<sup>th</sup> graders included 26 pictures. Note that we have opted to include fewer lexical stimuli in the tests aimed to the first group of children due to their lower attentional capacities. All pictures were displayed on a computer screen, which was placed in front of both the child and the researcher, through the Windows 2007® PowerPoint<sup>™</sup> program. In both tasks (oral production task and written task), the pictures used for both tasks (oral production and written production) were the same, for each grade level.

For the tests applied to the subjects attending the 1<sup>st</sup> grade of primary education, two-syllable words with branching onsets (globo; “globe”) and also problematic groups consisting of plosive+nasal consonant (pigmeu; “pygmy”) and fricative+plosive (afta; “aphtha”) were used. For the tests
applied to the subjects attending the 4\textsuperscript{th} grade, one used polysyllabic branching onsets \textit{(radiografia; “x-ray”, eg.)} and also problematic groups composed by \textit{plosive+plosive} \textit{(pictograma; “pictogram”)}, \textit{plosive+nasal consonant} \textit{(algoritmo; “algorithm”)}, \textit{plosive+fricative} \textit{(fricção; “friction”)} and \textit{fricative+plosive} \textit{(oftalmologista; “ophthalmologist”). The distribution of the various types of structures tested in the instrument created for the collection of data is shown in Table 3, below. Note that given the nature of children's lexicon, it was not possible to include \textit{stimuli} composed by \textit{nasal consonant+nasal consonant}.

Table 3: Structure of the instrument used for collecting the data.

<table>
<thead>
<tr>
<th></th>
<th>1\textsuperscript{st} grade</th>
<th>4\textsuperscript{th} grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Branching onset</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{plosive+vibrant}</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>\textit{plosive+lateral}</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>\textit{fricative+vibrant}</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>\textit{fricative+lateral}</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Problematic consonant group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>\textit{plosive+plosive}</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>\textit{plosive+nasal}</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>\textit{plosive+fricative}</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>\textit{fricative+plosive}</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>\textit{nasal+nasal}</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The oral production task \textit{(naming task done after pictures)} was followed by the writing task \textit{(written production task)}. In the instruction given for the execution of the \textit{naming task}, children were asked to visualize the image that appeared on screen and to name it. It was mentioned that it was intended that the child only used a word for what the child visualized. In some situations, the researcher had to resort to the use of the definition of the item \textit{(semantic clue)}. This semantic clue was a part of the test application protocol and it was used in order to encourage a
response from the child (see Santos 2013; Appendix 3). As a last resort, when the child could not name a certain item, repetition was used – the researcher verbally produced the stimulus for the child to repeat it. For data analysis purposes, whenever the researcher had to fall back upon “imitation”, the child’s production was, then, encoded as such. The written production task was in turn registered by the child in an A4 white sheet of paper with all the coloured pictures arranged in the order they were displayed on the computer screen throughout the naming task. Note that there was a training phase to ensure that both tasks were understood by the child.

The researcher audio recorded the entire process of the execution of the tasks using a Philips digital recorder LFH0635/0, which was placed near the computer so as not to interfere with the implementation of the tasks. The data were then transferred to a laptop HP Pavilion dv6500 Notebook PC. The children’s verbal productions were later heard, transcribed and analyzed. In the research reported here, all productions where the target structures (either a branching onset or a problematic consonant group) were produced in accordance with the adult form, even though other word structures did not meet the target format, were considered correct productions.

3. Description and discussion of the results

In this study both branching (Type I sequences formed by obstruent+liquid) and non-branching onsets (Type II sequences) prove to be problematic for the children: children do not completely master them. The graphic below (Figure 1) illustrates the success rates for branching onsets and problematic consonant groups as well as the overall success values for each structure studied in this paper.
Analyzing Figure 1, branching onsets arise, then, as less complex structures than problematic consonant groups: branching onsets always record higher success values than the problematic consonant groups. Considering the total sample, branching onsets have an overall average of success of 78%; problematic consonant groups have an average of success of 66%.

Data in Figure 1 also show that the children who are attending the 1st grade reveal a symmetrical behaviour towards the two types of consonant sequences covered in this research: levels of success for both sequences are high in oral productions (above 79%) and low in written productions (below 32%). In both cases, the success in oral productions is not accompanied by success in writing. In the 4th grade, the difference between success rates in oral productions and written productions is lower for branching onsets and higher for problematic consonant groups, registering therefore an asymmetrical behaviour between the two types of consonant sequences.

When dealing with complex structures, there are several resources used by children in order to tailor their productions to target forms, namely, children often resort to repair strategies, which arise from the phonological and orthographic development stage in which they are. In this study, in order to produce either branching onsets or problematic consonant groups, children used the repair strategies displayed in the tables below.
Table 4: Strategies used by 1\textsuperscript{st} graders to produce both consonant clusters

<table>
<thead>
<tr>
<th>Type of production</th>
<th>Used strategies</th>
<th>Frequency per structure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Branching Onset</td>
</tr>
<tr>
<td>Oral production</td>
<td>Vowel epenthesis</td>
<td>29%</td>
</tr>
<tr>
<td></td>
<td>Deletion</td>
<td>59%</td>
</tr>
<tr>
<td></td>
<td>Metathesis</td>
<td>12%</td>
</tr>
<tr>
<td></td>
<td>Segmental substitution</td>
<td>--</td>
</tr>
<tr>
<td>Written production</td>
<td>Vowel epenthesis</td>
<td>62%</td>
</tr>
<tr>
<td></td>
<td>Deletion</td>
<td>32%</td>
</tr>
<tr>
<td></td>
<td>Metathesis</td>
<td>4%</td>
</tr>
<tr>
<td></td>
<td>Segmental substitution</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Other productions</td>
<td>2%</td>
</tr>
</tbody>
</table>

Table 5: Strategies used by 4\textsuperscript{th} graders to produce both consonant clusters

<table>
<thead>
<tr>
<th>Type of production</th>
<th>Used strategies</th>
<th>Frequency per structure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Branching Onset</td>
</tr>
<tr>
<td>Oral production</td>
<td>Deletion</td>
<td>--</td>
</tr>
<tr>
<td>Written production</td>
<td>Vowel epenthesis</td>
<td>60%</td>
</tr>
<tr>
<td></td>
<td>Deletion</td>
<td>12%</td>
</tr>
<tr>
<td></td>
<td>Metathesis</td>
<td>20%</td>
</tr>
<tr>
<td></td>
<td>Segmental substitution</td>
<td>8%</td>
</tr>
<tr>
<td></td>
<td>Homophone spellings</td>
<td>--</td>
</tr>
</tbody>
</table>

One of the main criteria to distinguish branching onsets from heterosyllabic consonant clusters in European Portuguese is the strategy of vowel epenthesis. This strategy is often used between the two consonants which make up a heterosyllabic consonant cluster but it seldom occurs
between members of a tautosyllabic consonant cluster (see Table 1. adapted from Veloso 2003). Common and documented productions of *vowel epenthesis* amidst the problematic consonant groups seem to suggest that the children’s phonological knowledge recognizes the existence of a different phonological nature between those clusters and branching onsets. Nevertheless, apart from the strategy *homophone spellings*, which only occurs in the production of problematic consonant groups, children in this study use the same repair strategies when dealing with the two types of consonant clusters.

The insertion of an epenthetic vowel is found, in this study, not only in the middle of problematic consonant sequences – which is legitimised by the presence of an empty nucleus in the analysis of Mateus & d’Andrade (2000) (see the representation in (1)) – but also in between the two consonants of branching onsets (see Table 4), evidencing thus the complexity of both structures. Figure 2 illustrates the frequency of use of the strategy *vowel epenthesis* in both types of consonant sequences.

![Figure 2: Frequency of the *vowel epenthesis*’ strategy in the two different syllabic structures.](image)

Despite being common to all the productions above, *vowel epenthesis*’ strategy is used in different manners by the children of this study. When it comes to oral productions, children insert no more than the vowel [ɨ] between tautosyllabic clusters (29% of *vowel epenthesis*’s productions, see Graphic 2 supra), hence corroborating the findings of Freitas (2002). This author reports some uses of [ɨ] before the stabilization of branching onsets (*fralda* “diaper” /ʃɾɐlda/ → [ʃɪ'rawdə] Luis 2;6, eg.). Thus, if we assume
that epenthetic vowels are unmarked segments in linguistic systems, the use of [i] (or <e> in written productions) as epenthetic vowel, in the collected corpus, gives no empirical evidence to Mateus and d'Andrade (2000) hypothesis that in European Portuguese the unmarked vowel is [i].

Epenthetic vowel by excellence, [i] is considered a prosodic filler in European Portuguese, operating also as a filler for an empty nucleus in problematic consonant groups (Freitas 2002, 2004; Veloso 2010, 2012). Consequently, it is expected [i] to be the chosen segment by the children to be orally produced between the two consonants that make up the problematic groups. However, when children showed vowel epenthesis with heterosyllabic clusters, present in 46% of oral productions (see Figure 2), in addition to the insertion of the vowel [i] – which represents 36% of the cases of epenthesis (pneu “tyre” → [pi’new]⁵, eg.) – children in this study also inserted the segments [ɐ] and [u] 11% of the times they produced epenthesis (afta ”aphtha” → [’afetə]⁶; pneu “tyre” → [‘punew]⁷ and pigmeu “pygmy” → [pigu’mew]⁸). The heterogeneity and the amount of different epenthetic vowels inserted between the clusters of these problematic groups seem to confirm that there is indeed a distinct phonological nature among the two types of structures studied in this paper. Such procedure leads us to consider the possibility for the children to be following a trend of assimilation of the properties of a vowel or of a surrounding consonant, in these oral productions. Note that the production of [’afetə] can be considered as a result of vowel harmony. In the case of the insertion of [u] ([’punew] and [pigu’mew]), there is the assumption that the [w] could legitimize the harmony, although it never occurs in gnomo “gnome”, which has two labial vowels.

In written productions, the insertions of the epenthetic vowel <e>, which may be interpreted as a record of [i], are predominant: see the values of 61% for branching onsets (69% in the 1st grade and 53% in the 4th grade) and 65% for problematic consonant groups (69% in the 1st grade and 61% in the 4th grade). Such behaviour is expected as it establishes a parallel with the oral data, where [i] is the epenthetic vowel by excellence. Nonetheless, in their written productions, children choose to use other vowels besides the vowel <e> when inserting segments between both types of consonant clusters. In obstruent+liquid sequences, the insertion of vowels <a>, <o> or <u> occurs in 31% of the productions of children attending the 1st grade; 4th graders, on the other hand, insert the vowels <a>, <i>, <o> and <u> in 47% of the cases. In problematic consonant sequences, in turn, 1st graders resort to the use of <a>, <o> or <u> in 31% of the times, and the older children opt to use <a>, <i> or <o> 39% of the times. The following table shows some examples of these productions.
Table 5: Examples of vowel harmony in the written productions of the 1st and 4th graders of this study.

<table>
<thead>
<tr>
<th>Branching onset</th>
<th>Problematic consonant group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EV</strong></td>
<td><strong>Target production</strong></td>
</tr>
<tr>
<td>&lt;e&gt;</td>
<td>planta</td>
</tr>
<tr>
<td>&lt;e&gt;</td>
<td>flores</td>
</tr>
<tr>
<td>&lt;e&gt;</td>
<td>globo</td>
</tr>
<tr>
<td>&lt;e&gt;</td>
<td>atlântico</td>
</tr>
<tr>
<td>&lt;e&gt;</td>
<td>biblioteca</td>
</tr>
<tr>
<td>&lt;e&gt;</td>
<td>confront</td>
</tr>
<tr>
<td>&lt;e&gt;</td>
<td>insufável</td>
</tr>
</tbody>
</table>

*EV: Epenthetic vowel*

Children seem to be assimilating the vocalic features of the vowels from contiguous syllables when they spell epenthetic vowels other than the non-marked (<e> / [i]). In fact, when 1st graders note obstruent+liquid sequences, in 20% of their productions, they resort to the assimilation of the vocalic features of the following vowel. When 1st graders write heterosyllabic clusters, the frequency of vowel harmony increases up to 27%. 4th graders, in turn, seem to resort to this phonological process in 47% of their productions of obstruent+liquid clusters and 25% of their heterosyllabic clusters. Thus, by revealing traces of adjacent syllable vowels’, the insertions of vowels that are different from the epenthetic vowel par excellence, can be interpreted as examples of vowel harmony in writing.

4. Final considerations

In this paper, we described the productivity of vowel insertions between complex consonant clusters in oral and written productions of monolingual children attending the 1st and 4th grades of primary education in two public schools. In European Portuguese, branching onsets are formed by sequences of two consonants (an obstruent as C1 – either a plosive or a fricative – followed by a liquid – which can be either a lateral or a vibrant). These clusters obey phonotactic principles: The Sonority Sequencing Principle and The Minimal Distance Principle and are,
therefore, considered being tautosyllabic and, as such, they are represented under the domain of the same syllabic node of the onset (\([C_1C_2]_\text{onset}[V]_\text{nucleus}[_\text{syllable}])\). In problematic consonant sequences (pneu “tyre”), both consonants are considered to be heterosyllabic as they mandatorily require vocalic epenthesis in Brazilian Portuguese and optional vocalic epenthesis in European Portuguese, as well as they violate phonotactic principles. Given the complex nature of both types of consonant clusters studied in this paper, children often resort to the activation of repair strategies in order to attain the target. Consequently, it was found that the vowel epenthesis strategy occurs in greater numbers: (i) in the 1st graders’ productions, which confirms the expectation that the group of older children show a more advanced phonological development and orthographic knowledge than the group of younger children, since children attending the 4th grade are more advanced in their phonological development path, have a longer school career – which results in a more stable knowledge of spelling rules – and they are also more familiar with the existing lexicon in European Portuguese; (ii) in the productions of problematic consonant clusters, which shows a distinct phonological nature of the two types of structures studied in this paper and which can also be justified by the low frequency of lexical items with this complex structure in European Portuguese (among others, Vigário & Falé 1994; Freitas 1997); (iii) in written productions, which, in our opinion, is related to their quality of apprentices of the spelling rules and it simultaneously corroborates the existence of processing differences regarding the type of record (oral and written record).

Children assessed in this study activate vowel harmony especially when they have to write down the clusters (whether they are tautosyllabic or heterosyllabic); however, during their oral productions, the same children show a preference for the insertion of [i] between both consonant clusters. Aforementioned results contribute to the discussion of the non-marked vowel in European Portuguese. According to Mateus & d’Andrade (2000), [i] is the default vowel in the system; however, the low activation of vowel harmony in the oral collected corpus and the frequent use of [i] as epenthetic vowel gives no empirical evidence to Mateus & d’Andrade (2000) hypothesis that in EP the unmarked vowel is [i].

Notes

1 “A sonoridade dos segmentos que constituem a sílaba aumenta a partir do início até ao núcleo e diminui desde o núcleo até ao fim” (translation by the author).
For a more detailed description of the methodology of this study, as well as a description and discussion of the broader data, see Santos (2013).

Primary school of Turquel.
Primary school of Benedita.
Child 141.
Child 221.
Child 221.
Child 111.

References


Aquisição da linguagem: estudos recentes no Brasil. Porto Alegre: EDIPUCRS.
Santos, Rita Nazaré (2013), Aquisição de grupos consonânticos e seu impacto nos desempenhos escritos no 1.º Ciclo do Ensino Básico. Dissertação de Mestrado apresentada à Faculdade de Letras da Universidade de Lisboa.


THE ACQUISITION OF RECURSIVE LOCATIVE PPS AND RELATIVE CLAUSES IN CHILD ENGLISH

ANCA SEVCENCO, TOM ROEPER AND BARBARA ZURER-Pearson

1. Introduction

Recursion is a core property of grammar, evident in the basic operation of Merge from the two word stage on (Hauser, Chomsky & Fitch 2002). It also arises in classic environments of prepositional phrases (PPs), relative clauses (RCs), possessives, and adjectives, among others, as a form of Indirect self-embedding structure, as we shall clarify. However not every type of recursion is found in every language, so evidence must be present for the child in each language that various specific constructions allow recursion (Snyder & Roeper 2004a). Now several other acquisition questions arise:

(i) Is there a recursion trigger?
(ii) Does one form of recursion trigger another?
(iii) Is there an acquisition path for recursion?

We argue that Direct Recursion is a default form of conjunction:

(1) XP => X (and XP)

which creates a flat structure and is always available to any child.\(^1\)

We can define the classic Indirect Recursion in traditional phrase-structure terms as:

(2) XP => X YP
    YP => Y XP

which at the categorical level becomes:
(3) \[ PP \Rightarrow P \ NP \]
\[ NP \Rightarrow N \ (PP) \]

di Sciullo (2015) proposes that all recursive nodes are marked by a Functional Category (FC), but it may be overt or covert, as we illustrate below.

The level of abstraction in (2) leads to an important empirical consequence. If it is correct, it could serve as a trigger across recursion types. Do individual children learn recursive PPs and RCs together?

It is useful to lay out even more abstract options that future work can address. Chomsky (2013) proposes that there is no word order in abstract rules. If the rules are stated as sets, then this system (Chomsky, pc) is insensitive to left- or right-branching: \[ XP \Rightarrow \{X, YP\} \]. Such a template would allow a right-branching rule of PP-recursion to trigger a left-branching rule, like possessive recursion. It remains to be systematically explored.

How could a child recognize Indirect Recursion? We argue that the mechanism of Minimal Search can be syntactically extended to capture recursive connections across phases. Covert FCs create a challenge for Minimal Search when one has to give an invisible Feature to a recursive Node.

If such an abstract formalism as (2) underlies all forms of recursion, we can now ask if a language will lead to a consistent path for recognition of recursion since not every language has every type. For instance, German has prenominal possessives (Maria’s Haus) but, unlike in English, it is not recursive. What factors could determine acquisition order or cause delay when a structure like the prenominal possessive is recursive?

Does semantic interpretation constrain the recognition of recursion? It could: recursive structures resemble other operator constructions, such as negative polarity items, which have multiple marking of a single negative meaning (I didn’t buy anything for anyone). Likewise, the recursive locative in:

(4) put the ball in the kitchen in the corner in a box

semantically fulfills a single LOCATIVE argument with multiple PPs. In addition, recursive structures also cross phase boundaries and simultaneously allow independent references (each in independently modifies an NP). Pérez-Leroux et al. (2013) has shown how the combination of syntax and semantics, for example restrictiveness for RCs
and PPs, can be a delaying factor. Our approach seeks to isolate the syntactic challenge of recursion.

Finally, does morphology play a role? We will argue that overt recursive morphemes play a key role.

Recent studies on the acquisition of recursive structures by young children (around 3-5 years of age) consistently show these participants have difficulty with producing and comprehending recursive structures that involve various levels of recursive embedding, i.e. from 2-level up to 4-level recursive phrases (Pérez-Leroux et al. 2013; Hollebrandse & Roeper 2014). However, the average age for acquiring recursive interpretations varies across structures and across languages. Studies of the acquisition of PP and RC recursion in Romanian (Sevcenco et al. 2015) and Japanese (Roeper et al. 2012) point to age 5 as a time of transition to comprehension of recursion for these structures. Building upon these results, we propose three hypotheses for English-speaking children.

**Hypothesis I:** We predict that most responses for children younger than 5 will be direct recursion (or conjunction), and most responses for older children and adults will show a preference for Indirect recursion (or “recursive” interpretations). Thus a sentence like:

(5) the dog next to the cat next to the horse [dog cat horse]

will be repeated and acted out by the younger children as:

(6) the dog next to the cat and next to the horse [cat dog horse]

We predict that conjunction will appear first for every structure in young children, and will re-appear when structures are very complex. Determining the age at which English-speaking children make the transition from a preference for direct recursion to Indirect recursion is an empirical question this study is designed to explore.

**Hypothesis II:** Following di Sciullo (2015) that languages have to mark recursion by overtly or covertly realized FCs, we predict that overtly marked recursive configurations (relative clauses + that) will be more transparent than covertly marked ones (modification by means of locative PPs). Therefore, they could emerge at earlier ages. We also predict that children will give more recursive responses overall for RCs than PPs.

**Hypothesis III:** Younger children will find it significantly easier to understand recursive structures that involve fewer levels of embedding.
Although in principle, recursive sequences are of indefinite length, performance factors could make comprehension and production of longer sequences less accurate, especially in early phases.

The paper is organized as follows. Section 2 presents the experimental study, Section 3 summarizes the main results, Section 4 offers a discussion of these results and a proposal about how Minimal Search can be adapted to the acquisition model and Section 5 outlines the conclusions of the study.

2. Method

The present study investigates the acquisition of Indirect recursion with adnominal locative PPs and subject relative clauses in child English. It is designed to explore at what ages (in the 4-9-year range) young monolingual speakers of English prefer to assign a direct recursion syntactic representation (see 1) or an Indirect recursion syntactic representation (see 2) to experimental items that permit indirect recursion.

Thirty-seven monolingual English speaking children in western Massachusetts took part in the study (see Table 1). The children came from middle-class neighborhoods, and were typically developing. Ten adults made up the control group.

Table 1: Participants by age

<table>
<thead>
<tr>
<th>Age group</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 years (n= 7)</td>
<td>4;3 (.232)</td>
</tr>
<tr>
<td>5 years (n= 6)</td>
<td>5;4 (.386)</td>
</tr>
<tr>
<td>6 years (n= 6)</td>
<td>6;1 (.361)</td>
</tr>
<tr>
<td>7 years (n= 7)</td>
<td>7;3 (.302)</td>
</tr>
<tr>
<td>8 years (n= 7)</td>
<td>8;2 (.261)</td>
</tr>
<tr>
<td>9 years (n = 4)</td>
<td>9;3 (.359)</td>
</tr>
<tr>
<td>Adults (n = 10)</td>
<td></td>
</tr>
</tbody>
</table>

The experiment consists of an act out task. Participants were asked to rearrange animals in an array on an iPad in response to a prompt, and then to describe what they had done. General conversation during and after the experiment was transcribed and analyzed. As a warm-up activity, there was a repetition section with five sentences containing relative clause or PP recursion. The testing sessions took place in a quiet room either at
their school or at the Children’s Museum in Holyoke, MA, USA. Prompts are illustrated in (7) to (10):

“Please show me”:

(7) the lion next to the zebra under the crocodile
(8) the gorilla by the lion by the bear by the crocodile
(9) the lion that is next to the bear that is next to the zebra
(10) the lion that is on the zebra that is on the bear that is on the crocodile

The two arrangements shown in Figure 1 represent different structural interpretations coded as recursive (a) and conjunctive (b) for item 7:

(11) a. Recursive: a lion \([next \ to \ [\text{zebra \ [under \ a \ crocodile]\}]]\]
b. Conjunctive: a lion \([\text{[next \ to \ a \ zebra]} \ (\text{and}) \ [\text{under \ a \ crocodile}]]\]

Figure 1:

[Images of lion, zebra, and crocodile]

a. Recursion  
b. Conjunction

In the recursive structure (11a), the second PP under phrase is attached to zebra inside the first next to phrase ([lion next to zebra] +PP) and then next to phrase that includes the under phrase is attached to lion. In the conjunctive structure (11b), however, the two PPs are independently associated with lion as they are linked by and between them (Lion + [next to zebra] + [under the crocodile]). Children showed this overtly in repetitions: repeat “… the lion next to the zebra next the crocodile” came out as “… the lion next to a zebra AND a crocodile”.

The main study included 16 recursive items: 4 test items each with two PPs, with three PPs, with two relative clauses and with three relative
clauses. The prepositions used were: *in front of, next to, by, under,* and *on.* There were also four fillers. In addition, there were two exploratory comprehension items with a slightly different format that were used as control items.

Independent within-subjects variables were level of embedding (2 vs. 3 phrases or clauses) and recursive structure type (locative PP vs. RC). The between-subjects variable was age. The dependent variable was the type of response, with recursive answers coded as 1, conjunctive answers as 0.

3. Results

Results pertaining to Hypothesis I.

There was a significant age effect for the recursive answers in the act-out items (*F*(5,495) = 30.5, *p* < .0001), as shown in Figure 2.

![Figure 2: Recursive vs. Conjunctive answers by age group (%)](image)

Tukey post hoc tests showed the existence of three homogeneous age groups according to the overall act out results: (i) 4 and 5, (ii) 6, 7 and 8 and (iii) 9 year olds. The 9-year-old group performed significantly better than the other two groups and the 5-7 year olds were significantly better than their 4-5-year-old peers. We also identified a significant age effect for the repeat answers (*F*(5, 166) = 8.07, *p* < .0001). In this case, in the Tukey post hoc test, only the 4yr-olds were significantly different from all older groups, who did not differ from each other. A cross-tabulation of
repeat versus act out percentages by child showed 13 children with high recursion scores on both tasks, and 7 with low recursion scores on both tasks. 13 children had repeat scores higher than act out scores. Only two children (a 6-year-old and a 7-year-old) gave fewer recursive answers in the repetition task than in the act out, suggesting that children could repeat accurately before they could do the act out, but if they could not repeat, then they generally could not do the act out either.

Adult percentages of recursive responses were 93% consistently across all conditions. Thus, in these data, the transition away from mostly conjunctive responses appeared to develop around 6yrs of age and was approaching adult levels of recursive responses at age 9.

Results for Hypothesis II.

A comparison of the indirect recursive answers for the test items with prepositions and relative clauses suggests that there was only a marginally significant difference between them ($F (1,469) = 3.58, p = .059$). See Figure 3.

![Figure 3: Recursive vs. Conjunctive answers by structure type and age group (%)](image)

There was no significant interaction with age. Still, there was one age group (the 5-year-olds) that found indirect recursion to be significantly easier with relative clauses than with PPs (that is, PP and RC means were outside the 95% confidence intervals for each other).
Results for Hypothesis III.

Since the difference in percentages of recursive responses between PP and RCs was slight, we report level of embedding results for the two structures combined. Level of embedding significantly influenced the percent of indirect recursive answers at most ages, as seen in Figure 4. Overall, there were 65% recursive answers for the test items with 2 levels of embedding and 55% recursive answers for the structures with 3 levels ($F(1,636) = 17.14, p < .0001$). No interaction with age was found overall, but post hoc tests indicated that items with 2- and 3-level recursion were significantly different for 5-, 6- and 7-year-olds.

Alternative strategies in the acquisition path. The child participants were observed to use a couple of strategies to avoid recursive structures: conjoined sequences (as in Figure 1a and b above and 12), alternation of PP and relative clauses (13), conjunctive iteration of single-level relationships (14), and recursive iteration of single-level relationships (15):

(12) the zebra next to the crocodile and the bear

(13) the lion next to the crocodile that is next to the bear

(14) there is a lion next to the bear, there is a lion next to the gorilla

(15) there is a lion next to the bear; there is a bear next to the gorilla.
Another type of answer that avoided having one PP embedded inside the other PP we called “high attachment.” The child projects another higher DP. For the lion next to the zebra under the crocodile, the first PP next to phrase is attached to lion first and then the second PP is attached to it through the DP. These “avoidance” strategies are important: they indicate that Indirect recursion must be more difficult conceptually. That is, recursion is not just a form of repetition.

In summary, we see that there are steps on an acquisition path that hold across children.
1) The 5-, 6- and 7yr-old children in the study gave more indirect recursion readings for 2-level embeddings than for 3-level embeddings.
2) Children, below 6, were observed to go through a stage when they preferred to assign a conjunctive syntactic representation to an Indirect recursive structure. Older children and adults used Default Conjunction as a way to reduce very complex forms.
3) For production: 5-, 6- and 7yr-olds were better for 2-level embedded recursive structures than for 3-level ones.
4) Consistent with di Sciullo’s argument, 5yr-olds comprehended more recursive relative clauses than recursive PP modifiers.
5) Adult controls, unlike children, gave target answers 92.85% of the time. When they failed to give the correct response, none of them made the errors in (13)-(15), which pertain thus exclusively to children.

In the next section, we propose a theory could account for all of this.

4. Discussion

We propose that the problems children experience with recursion have two sources: (i) the structure in the syntactic derivation of recursive nodes and (ii) which structures in each language have overt or covert FCs for Indirect recursion.

In di Sciullo (2015)’s approach she analyzed pairs like N-N (Hausarbeit (homework)) compounds and N-s-N compounds (Transformation-s-grammatik (transformational grammar)), as the same, but one has an overt marker –s- while the other does not. The overt /s/ carries neither possessive nor plural meaning, but marks recursive compositionality. In Japanese no marks Indirect recursion with possessive structures (Roeper et al. 2012); English marks recursive relative clauses overtly by means of the complementizer that, or covertly in sequences of
recursive PPs. The possessive marker 's is obligatorily overt in English (but not African-American English). For PPs, while English covertly marks indirect recursion (see (16), Romanian has the overt functional preposition de/’of’ to indicate recursion with adnominal locative PPs (Giurgea 2015). For instance, the sequence in (17), where de/’of’ is present, has a recursive interpretation whereas (18), without de, can only be interpreted as a conjunctive structure (as indicated in the translation of the example).

(16) a. put the hat in the bag in the box in the corner
   b. put+LOC in+LOC in+LOC

(17) pune pisica pe zebra de pe gorila de pe porc
      put cat.the on zebra.the de on gorilla.the de on pig
   Put the cat on the zebra on the gorilla on the pig

(18) pune pisica pe zebră pe gorilă pe porc
      put cat.the on zebra on gorilla on pig
   Put the cat on the zebra and on the gorilla and on the pig

When there is overt morphology, recognition is easy. How does the English-speaking child decide that a covert FC is present? They must carry out a semantic/pragmatic/syntactic mapping reconstruction using contextual support. Thus for:

(19) the dog next to the cat next to the horse

the child must see (via context) that the second PP modifies the cat not the dog (which would be the conjunctive option), and therefore attaches the PP below the NP, which itself is inside a PP. At this point, an invisible FC is necessary. That is, the child must allow the first PP [the dog [next to the cat]] to “see” inside the next PP [next to the horse] and decide that the PP modifies the NP inside itself, [cat], and not [dog] which then causes the child to project a recursive FC which can be linked by Minimal Search to the previous PP. If FC is overtly marked with a unique recursive marker, then recognition is facilitated. If not, an invisible head above the PP must be projected, which can be extended to all PPs recursively within other PPs if context shows that the conjunctive reading is wrong. This is a substantial amount of implicit reasoning.

Chomsky (2013) argues that lexical items merge, followed by a Minimal Search for the head of the newly formed syntactic object such that the projection can be labelled (and ultimately interpreted at the
interfaces) at the phase level. For an indirect recursive structure to be interpreted, an Extended Minimal Search proceeds all the way down the syntactic tree so all the relevant identical heads of projections are identified, crossing phase boundaries (20).

(20) Indirect recursion → Find (or contextually project) identical heads across phase transfer sites

In this framework, we suggest that the younger children in our experiment either fail to execute the Extended Minimal Search, or fail to project invisible recursive FCs. Note again that recursive input for classical recursion can be rare and must be independently present in each language since languages vary (Snyder & Roeper 2004b).

When children fail to interpret a structure as recursive, they assign it the conjunctive reading—i.e. the PPs/RCs in our test items are predicated individually of the noun they modify. For example, the conjunctive interpretation for the test item in (21) creates (22), where children merge PPs in a non-hierarchical syntactic structure.

(21) the lion by the zebra by the bear

(22) conjunctive structure

Assigning an indirect recursive interpretation to (21) involves choosing a different computation route from (22): the Ps and DPs merge, the labels for the PP and DP projections are correctly found, the covert marker of
recursion (LOC) is identified and so are all the relevant P heads down the entire syntactic tree, as shown in (23):

(23) indirect recursive structure

The syntactic account for the failure of our younger participants to give recursive responses also explains why our results indicate that there seems to be a small headstart for RC recursion over PP recursion for 5-year-olds in English. We propose that the difference between these structures has to do with the realization of the functional marker of recursion. In relative clauses the marker is overt—the complementizer *that*—whereas in locative PPs, it is covert. The convergence of slight indications in English that RCs and PPs come in together, where only the relative is overtly marked sometimes, points toward support for di Sciullo’s theory that the overt/covert distinction has an impact on acquisition.

This proposal receives stunning and precise support from cross-linguistic data. Sevcenco et al. (2015) shows that in Romanian, a language that overtly marks indirect recursion both in the PP and clausal domain, there is no significant difference in answers with recursive PPs vs. RCs in a task that is an adaptation of our English experiment and in
which only 5-year-old children participated. The straightforward comparison between English and Romanian 5-year-olds leads to the conclusion that the difference in overt vs. covert FCs has a significant impact upon time of acquisition: 5-year-olds in Romanian (overt) gave a higher percentage of recursive answers than 6yr-olds in English (covert). This observation neatly follows up the claim made earlier in Pérez-Leroux et al. (2012) that recursion learning is parameterized because it is subject to language specific labeling and selectional requirements.

The data begin to define an abstract recursion path for two structures that pivots on the overt/covert nature of FCs. Children start with the conjunctive stage at the first stage—3-5-years—and then move to structures with overt recursion, and then to those with covert recursion. Our analysis shows again in acquisition the unique strength of cross-linguistic analysis under the Universal Grammar hypothesis. Needless to say, much more detailed comparative research on the emergence of Indirect recursion is needed to fill out the full picture.

5. Conclusions

This study has explored the acquisition of Indirect recursion in the phrasal (adnominal locative PPs) and clausal domain, focusing on data from English.

The close coupling of the acquisition of PP and RC recursion argues for their similarity at the abstract level. The nature of that connection remains an important topic. Classically it has been suggested that PPs derive from relative clauses. They clearly are similar at the semantic level.

We have found that English-speaking typically developing children go through a stage in which they fail to assign a recursive syntactic representation to the test items they were presented with, preferring instead to give responses that rely on a conjunctive (possibly non-hierarchical) representation. This stage lasts from 4-, 5- years up to 6-years, when a turning point appears. The difficulty younger children go through is also shown by the fact that the 5-7 year olds still did better on the test items with 2-level recursive embedding than 3-level embedding, while older children appear able to carry on indefinitely. Adult-like performance on the interpretation of Indirect recursion became apparent at 9 years old, when both 2- and 3-level recursive embeddings could be successfully acted out.
We have ascribed the difficulty with the interpretation of Indirect recursion to two factors: the syntactic derivation of recursive nodes and language specific variation in the realization of the functional categories associated with Indirect recursion (overt versus covert marking). For the Indirect recursion interpretation to come through, Extended Minimal Search has to apply all the way down the syntactic tree and the overt/covert functional category associated with recursion needs to be identified. Overt marking on the functional category appears to help with establishing the recursive interpretation. This is why the participants in our English experiment performed differently than their peers who speak other languages studied to date.

This result suggests that one cause for the problems with the production and comprehension of recursive structures lies in parametric variation, a point made in Pérez-Leroux et al. (2012). Finally, our data bring evidence for the existence of an acquisition path for recursion, which begins at the conjunctive stage and moves towards a fully recursive stage across the particular variety of recursive structures each language displays.

Notes

1 It is well-known that the grammar of co-ordination is far more complex than such a default articulates. It does not state how sequences of and-structures arise or varieties of interpretation of coordination (Winter 1996), and interaction with case, binding theory, and ellipsis. We proceed with the assumption that a default conjoined meaning is available because children frequently supply it, but greater precision awaits further results.

2 We will not articulate the null hypothesis—which is nonetheless a valuable contrast—that there is no connection among how different forms of Indirect recursion are acquired. That is, each child would learn them in a different order depending upon experience.

References

THE LACK OF AGENT-ORIENTATION FOR
ZIBUN AND THE EPP PARAMETER
IN CHILD JAPANESE

HIROYUKI SHIMADA

1. Introduction

In this paper, I demonstrate that antecedents of the Japanese anaphor zibun in child-directed speech are virtually agents only. Given the use of zibun in child-directed speech, it is quite possible that zibun in child Japanese is agent-oriented despite being subject-oriented in adult Japanese. The current study, however, demonstrates that at approximately age five Japanese children correctly select theme arguments as zibun’s antecedent despite the lack of counterexamples to the agent-orientation of zibun in the input they receive from adults. This observation indicates that the possibility of agent-orientation for zibun is excluded from possible grammars due to some guidance from innateness. The current study also provides new evidence that the EPP parameter is correctly set in child Japanese by at least age 5, despite the lack of relevant cues in the input data from adults (Alexiadou & Anagnostopoulou 1998; Doner 2014).

2. Subject-orientation of zibun and earlier research

According to Wexler and Manzini (1987), anaphors such as reflexive pronouns in natural languages can be classified into two types. In some languages such as Japanese, only subjects can be antecedents of anaphors (Kuroda 1965); in other words, such anaphors are subject-oriented. The sentence given in (1) demonstrates the subject-orientation property of zibun.

(1) Taro-\text{ga} \ Hanako-\text{ni} \ zibun-\text{no} \ koto-o \ hanasi-ta.
Taro-NOM Hanako-DAT zibun-GEN thing-ACC tell-PAST
Taro told Hanako (things) about zibun.
As shown in (1), the subject *Taro* can be *zibun*'s antecedent, but the indirect object *Hanako* cannot.

In contrast to Japanese anaphors, antecedents of English anaphors are not restricted to subjects. Let us observe the following example.

(2) Ken\textsubscript{i} told John\textsubscript{j} some gossip about himself\textsubscript{i/j}.

As shown in (2), the English anaphor *himself* can take not only the subject *Ken* but also the indirect object *John* as its antecedent. Given this cross-linguistic variation, an interesting question arises: How do Japanese children acquire the adult-like knowledge of *zibun*? Is it possible for children to learn this knowledge from input data?

In the previous literature, Otsu (1997) has demonstrated that at approximately age 3-5, Japanese children demonstrate adult-like performance in selecting *zibun*'s antecedent in sentences such as (3).

(3) Taro\textsubscript{-ga} Hanako\textsubscript{j}-ni zibun\textsubscript{i/*j}-no e-o mise-ta.
Taro-NOM Hanako-DAT zibun-GEN picture-ACC show-PAST
Taro showed Hanako a picture of zibun.

As mentioned above, *zibun* is subject-oriented, and thus, the subject *Taro* can be *zibun*'s antecedent but the indirect object *Hanako* cannot. According to Otsu (1997), even 3-year-old children seem to know that *zibun* has the subject-orientation property.

One might think that *zibun* is not subject-oriented but rather nominative-oriented since Japanese is a nominative-accusative language; thus, only nominative-marked DPs can be *zibun*'s antecedent. However, *zibun* in Japanese is not nominative-oriented (Shibatani 1977; Takezawa 1987). Let us examine the example of transitive adjectives given in (4), which is a slightly modified version of the example given in Shibatani (1977).\textsuperscript{1}

(4) Taro\textsubscript{-ni-wa} Hanako\textsubscript{j}-ga zibun\textsubscript{i/*j}-no kumi-no-naka-de
Taro-DAT-TOP Hanako-NOM zibun-GEN class-GEN-in
itiban omosiro-i.
the most interested-PRES
*Hanako is interested the most in Taro in zibun’s class.*

In Japanese, when a predicate is stative, subjects can be marked with the dative Case-marker *-ni* instead of the nominative Case-marker *-ga*, and objects can be marked with the nominative Case-marker *-ga* instead of the
accusative Case-maker -o. In (4), the predicate is a stative predicate *omensiroi*, and the object is marked with the nominative Case marker -ga. The nominative-marked object *Hanako* cannot be *zibun*’s antecedent. Thus, this example demonstrates that *zibun* is not nominative-oriented.

Although Otsu’s finding suggests that Japanese children at approximately age 3-5 have already acquired the subject-orientation property of *zibun*, as pointed out by Sano, Shimada & Fujiwara (2014), it is possible that in Otsu’s experiment children selected the nominative-marked DP as *zibun*’s antecedent. In other words, children might have selected *Taro* in (3) not because it is the subject, but because it is marked with the nominative Case-maker -ga. They called this “nominative-orientation” for *zibun*. To address this issue, Sano, Shimada & Fujiwara (2014) investigated children’s interpretation of *zibun* in (5).

(5) *Inu*-ni-wa *buta*-ga *zibun*-no kumi-no-naka-de dog-DAT-TOP pig-NOM *zibun*-GEN class-GEN-in itiban omosiro-i. the most interested-PRES

The dog is interested the most in the pig in *zibun*’s class.

In (5), which also provides an example of transitive adjectives, the dative-marked DP *Inu* can be *zibun*’s antecedent but the nominative-marked DP *buta* cannot in adult Japanese. If *zibun* in child Japanese is not subject-oriented but nominative-oriented, it is predicted that they incorrectly allowed the nominative-marked DP to be *zibun*’s antecedent in (5). According to Sano, Shimada & Fujiwara (2014), however, Japanese children at approximately age 5 do not select the nominative-marked DP as *zibun*’s antecedent. Thus, their finding suggests that *zibun* in child Japanese is not nominative-oriented.

3. Agent-orientation for *zibun* and child-directed speech

According to Otsu (1997) and Sano, Shimada & Fujiwara (2014), Japanese children seem to have acquired the subject-orientation of *zibun*. However, there is a possibility that *zibun* in child Japanese is neither subject-oriented nor nominative-oriented; children may depend on the thematic roles of DPs in selecting *zibun*’s antecedent. The subject of test sentences such as (3) in Otsu (1997) is an agent. In addition, the subject of test sentences such as (5) in Sano, Shimada & Fujiwara (2014) is an experiencer. These thematic roles are ones that are assigned to external
arguments; therefore, there is a possibility that children select zibun’s antecedent depending on whether the DP has a thematic role that is assigned to an external argument. In the current study, I call this “agent-orientation” for zibun. One might then query to what extent it is likely that zibun in child Japanese is agent-oriented. To answer this question, I examined the corpora of four Japanese-speaking children (Akira M, Asato, Nanami, and Tomito) aged approximately 3 to 5 in the CHILDES database (MacWhinney 2000). For the corpora of the four children, I examined antecedents of zibun in child-directed speech. In total, I found 149 mothers’ utterances with zibun out of 98,823 utterances. However, out of the 149 mothers’ utterances with zibun, there was no instance of zibun with a non-agentive antecedent.

This investigation reveals that the antecedents of the Japanese anaphor zibun used in child-directed speech are virtually agents only. In other words, there is no counterexample to the agent-orientation for zibun. Therefore, if the knowledge regarding zibun is acquired by experience-based learning, it is likely that zibun in child Japanese is not subject-oriented but agent-oriented.

If zibun in child Japanese is not subject-oriented but agent-oriented, then it is predicted that children do not select an internal argument as zibun’s antecedent even if it is a subject. In adult Japanese, the internal argument behaves as a subject in a passive construction such as (6).

(6) Dorobou-ga keisatsukan-ni zibun-ni no ie-de tsukamae-rare-ta. house-in catch-PASS-PAST

The thief was caught by the policeman in zibun’s house.

In (6), the internal argument dorobou can be zibun’s antecedent, but the ni-phrase (by-phrase in English) cannot. This example demonstrates that the internal argument in passives behaves as a subject, and thus, it can be zibun’s antecedent. This example also indicates that zibun in adult Japanese is not agent-oriented since the agentive phrase keisatsukan cannot be zibun’s antecedent. If zibun in child Japanese is not subject-oriented but agent-oriented, then children incorrectly disallow the internal argument from being zibun’s antecedent in (6).

Hence, at first glance, a passive construction such as (6) seems to be an apt test case for examining this prediction. However, it has been reported that the acquisition of Japanese passive constructions such as (6) is delayed (Sugisaki 1999; Minai 2000; Sano, Endo & Yamakoshi 2001).
In order to avoid this problem, I used one of the unaccusative constructions.

(7) Dorobou\textsubscript{i} - ga keisatsukan\textsubscript{j} - ni zibun\textsubscript{i\textsuperscript{e} j\textsuperscript{-no th}ief-NOM policeman-by zibun-GEN ie-de tsukamat-ta. house-in catch-UNACC-PAST

*The thief was caught by the policeman in zibun’s house.*

In (7), the predicate is an unaccusative verb *tsukamaru* (be-caught), and it also has the *ni*-phrase (agentive phrase) like full passives such as (6). Sano, Endo & Yamakoshi (2001) call sentences such as (7) “full unaccusative” constructions. In this paper, I use the same terminology. The internal argument *dorobou* in (7) also behaves as a subject, and thus, it can be *zibun*’s antecedent. According to Sano, Endo & Yamakoshi (2001), Japanese children at approximately age 5 demonstrated good performance in comprehending full unaccusatives although they had difficulty comprehending full passives. Therefore, using full unaccusatives such as (7) is more suitable for testing whether or not *zibun* in child Japanese is agent-oriented. The prediction is as follows: If *zibun* in child Japanese is not subject-oriented but agent-oriented, then it is predicted that children wrongly disallow the internal argument to be *zibun*’s antecedent in full unaccusatives such as (7). ³

### 4. Experiment

In order to test this prediction, I conducted an experiment using the Truth Value Judgement Task (Crain & McKee 1985; Crain & Thornton 1998). I examined 20 Japanese children aged from 5;3 to 5;11 (mean age 5;7). I tested two types of target item and one type of control item. First, let us examine the target items.

(8) Zousan\textsubscript{i} - ga butasan\textsubscript{j} - ni zibun\textsubscript{i\textsuperscript{e} j\textsuperscript{-no elephant-NOM pig-by zibun-GEN niwa-de tsukamat-ta. garden-in catch-UNACC-PAST

*The elephant was caught by the pig in zibun’s garden.*

(9) Butasan\textsubscript{j} - ni zusan\textsubscript{i} - ga zibun\textsubscript{i\textsuperscript{e} j\textsuperscript{-no pig-by elephant-NOM zibun-GEN
Example (8) shows full unaccusatives containing *zibun* with canonical word order. The internal argument *zousan* behaves as a subject, and can be *zibun*’s antecedent, but the *ni*-phrase (butasan-*) cannot. Example (9) also shows full unaccusatives containing *zibun*, but the word order is scrambled. Thus, the second DP, which is the subject, can be *zibun*’s antecedent, but the first DP, which is not the subject, cannot. Target sentences such as (9) were tested in order to investigate whether or not the participants would select *zibun*’s antecedent depending on word order. If the participants selected the subject as *zibun*’s antecedent in (8) on the grounds that it is the first DP, then it was predicted that they would incorrectly select the *ni*-phrase as *zibun*’s antecedent in (9). Regarding target sentences such as (8) and (9), two trials were conducted for each child.

I also tested control items such as (10), which is an example of a full unaccusative containing an R-expression instead of *zibun* in the phrase “X’s garden” with canonical word order.

(10) Zousan-ga butasan-ni zousan-no
     elephant-NOM pig-by elephant-GEN
     niwa-de tsukamat-ta.
     garden-in catch-UNACC-PAST

*The elephant was caught by the pig in the elephant’s garden.*

The control items were tested to confirm that children did not ignore “X’s garden” in these stimulus sentences. Regarding the control items, two trials were also conducted for each child.

The following is a sample story for (8) and (9). A pig and an elephant are playing with a net. First, the pig tries to catch the elephant in the pig’s garden, but fails. After that, the pig finally manages to catch the elephant in the elephant’s garden.
(11) The final scene of the story for (8) and (9)

As mentioned above, the internal argument zusan can be zibun’s antecedent in (8) and (9). Therefore, in the situation given in (11), the target sentences (8) and (9) are true.

Next, let us examine a sample story for (10). A pig and an elephant are playing with a net. First, the pig tries to catch the elephant in the elephant’s garden, but fails. After that, the pig finally manages to catch the elephant in the pig’s garden.

(12) The final scene of the story for (10)

In the situation given in (12), the control item (10) is false. Let us examine the results shown in the following table.
Table 1: Results of the experiment

<table>
<thead>
<tr>
<th>Correct Answer</th>
<th>Target (8) Canonical word order with zibun</th>
<th>Target (9) Scrambled word order with zibun</th>
<th>Control (10) Canonical word order with an R-expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct Response Rate</td>
<td>Accept</td>
<td>Accept</td>
<td>Reject</td>
</tr>
<tr>
<td>32/40 (= 80.0%)</td>
<td>37/40 (= 92.5%)</td>
<td>38/40 (= 95.0%)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>69/80 (= 86.2%)</td>
<td>38/40 (= 95.0%)</td>
<td></td>
</tr>
</tbody>
</table>

As mentioned above, target sentences such as (8) and (9) were given for matching situations and control items such as (10) were given for mismatching situations. Therefore, the correct answer for target items such as (8) and (9) was acceptance, and the correct answer for control items such as (10) was rejection. For target sentences such as (8), out of the 40 trials in total, 32 responses were correct acceptances. Thus, the correct acceptance rate of (8) was 80.0%. In addition, regarding test sentences such as (9), 37 responses were correct acceptances. The high acceptance rate of (9) indicates that children did not select zibun’s antecedent based on word order. In other words, it was not the case that participants selected the theme argument as zibun’s antecedent because it was the first DP. In total, the acceptance rate was 86.2%. Furthermore, regarding control items such as (10), 38 responses were correct rejections. Hence, the correct rejection rate was 95.0%. The high correct rejection rate of (10) indicates that the acceptance rates of the target sentences were not due to yes-bias; it was not the case that children ignored the phrase in “X’s garden” in the test sentences. To summarize, Japanese children at approximately age 5 correctly selected internal arguments of unaccusatives as zibun’s antecedent.

5. Discussion

As shown in the previous section, the current study has revealed that Japanese children at approximately age 5 correctly select internal
arguments of unaccusatives as *zibun*’s antecedent. Do Japanese children learn the adult-like knowledge from input data? As discussed in section 3, the answer is “no.” This examination of the four children’s corpora has revealed that the antecedents of *zibun* in child-directed speech are virtually agents only. Thus, if grammatical knowledge regarding *zibun* is acquired from experience-based learning, it is quite possible that *zibun* in child Japanese is not subject-oriented but agent-oriented. As discussed above, if *zibun* in child Japanese is agent-oriented, then it is predicted that children do not select internal arguments of unaccusatives as its antecedent since the thematic roles of internal arguments of unaccusatives are generally theme or patient. In contrast to this prediction, children correctly selected the theme argument as *zibun*’s antecedent in the experiment of the current study. This result suggests that the agent-orientation for *zibun* is excluded from possible grammars due to some guidance from innateness.5

The current study has also demonstrated that the internal argument of full unaccusatives is a subject not only in adult Japanese but also in child Japanese. Assuming that subjecthood is assigned at the specifier of TP (Ura 2011), this result indicates that internal arguments of unaccusative verbs move from the complement of V to the specifier of TP. In the Probe-Goal Theory framework (Chomsky 2000, 2001), this movement occurs due to the EPP feature of T. In both child and adult Japanese, I assume the structure given in (13) for unaccusatives, including full unaccusatives.6

\[(13)\]

\[
\text{TP} \quad \text{DP}_1 \quad \text{T'} \quad T \\
\text{vP} \quad \text{v} \\
\text{t}_i \quad \text{VP} \quad \text{V}
\]

The internal argument of unaccusatives is base-generated in the complement of V. The Nominative Case of the internal argument is valued by T via Agree, and moves to the specifier of TP in order to satisfy the EPP feature of T.

However, it has been claimed that there are cross-linguistic variations regarding how to satisfy the EPP feature (Alexiadou & Anagnostopoulou 1998; Doner 2014).7 This prompts an interesting question here: How do Japanese children correctly set the value of this parameter? In other words, how do Japanese children know that internal arguments of unaccusatives must move to the specifier of TP? As discussed in Ura
Hiroyuki Shimada

(2011), unlike other languages such as English, Japanese seems to lack obvious evidence for the EPP feature of T. First, Japanese lacks expletives such as English *it* or *there*. In addition, the word order of unaccusatives cannot act as decisive input for setting the EPP Parameter in Japanese (a V-final language), unlike in SVO languages, since the word order does not change whether or not unaccusatives involve DP-movement.

Let us compare the word order of SVO languages such as English with that of SOV languages such as Japanese.

(14) a. Unaccusatives in English
\[ [TP \, DP_i \, [vP \, [VP \, V \, t_1 ]]] \]

b. Unaccusatives in Japanese (with DP-movement)
\[ [TP \, DP_i \, [vP \, [VP \, t_i \, V ]]] \]

c. Unaccusatives in Japanese (without DP-movement)
\[ [TP \, [vP \, [VP \, DP \, V ]]] \]

In SVO languages, an internal argument always appears to the right of the verb. However, as shown in (14a), when DP-movement occurs, the internal argument of unaccusatives appears to the left of the verb. Thus, the DP-movement changes the position of internal arguments of unaccusatives, and the evidence of this movement seems to be quite obvious for children acquiring an SVO language such as English. In contrast, in Japanese, which is an SOV language, an internal argument always appears to the left of the verb. Thus, as shown in (14b) and (14c), the word order of unaccusatives in Japanese does not change regardless of the existence of DP-movement. Considering these facts, it seems that Japanese lacks obvious cues for DP-movement.

Thus, Japanese seems to lack obvious evidence for the EPP feature of T and for DP-movement in unaccusatives. For these reasons, it may be expected that the EPP parameter is set very late in Japanese; in other words, it may be difficult for Japanese-speaking children to observe the DP-movement in unaccusatives. In fact, according to Miyamoto et al. (1999), Watanabe (2008) and Murasugi & Watanabe (2009), the internal arguments of unaccusatives initially remain in-situ.\(^8\) However, if the internal arguments of unaccusatives in child Japanese initially remain in the object position, and children do not receive decisive evidence for DP-movement until at least age 5, it would be expected that the internal arguments of unaccusatives in child Japanese continue to remain in-situ.
until age 5. However, despite the lack of relevant cues, this study provides new evidence that the internal arguments of unaccusatives do not remain in-situ at approximately age 5; rather, the internal arguments of full unaccusatives in child Japanese behave as subjects. Given that Japanese children acquire adult-like knowledge regarding the DP-movement (i.e., the EPP parameter) without decisive evidence, there may be a possibility that unaccusatives in child Japanese involve DP-movement from the beginning, contrary to the claim of previous studies (Miyamoto et al. 1999; Watanabe 2008; Murasugi & Watanabe 2009). I would like to leave this issue for future research.

6. Conclusion

If knowledge regarding zibun is acquired through experience-based learning, it is likely that zibun is not subject-oriented but agent-oriented in child Japanese, since there is no counterexample to agent-orientation for zibun. However, Japanese children at approximately age 5 correctly select the internal arguments of unaccusatives as zibun’s antecedents. This result indicates that the possibility of agent-orientation for zibun is excluded from possible grammars due to some guidance from innateness.

Assuming that subjecthood is assigned at the specifier of TP, internal arguments of (full) unaccusatives move from the complement of V to the specifier of TP in order to satisfy the EPP feature of T in Japanese. This study provides new evidence that the EPP parameter is correctly set at approximately age 5 in Japanese although decisive cues for the EPP feature seem to be extremely rare in the input data.

Notes

1 In (4), the dative-marked DP can be the antecedent of zibun. This is not because the dative-marked DP is marked with the topic marker -wa.
(i) Taro-ni Jiro-ga zibun-gen-no bou-de tatak-e-ta koto.
    Taro-DAT Jiro-NOM zibun-Gen stick-with hit-CAN-PAST thing

The fact that Taro could hit Jiro with zibun’s stick.
As shown in (i), even the dative-marked DP without the topic marker can be the antecedent of zibun. Therefore, topichood is not related to subjecthood. Thus, the fact that the dative-marked DP in (4) can be zibun’s antecedent indicates that it is the subject.
Although I use the term “agent-orientation,” it does not mean that only agentive phrases can be zibun’s antecedents. Other thematic roles that are assigned to external arguments (e.g., experiencer) are involved, but I use “agent-orientation” for simplicity.

It is also predicted that Japanese children incorrectly allow the ni-phrase to be zibun’s antecedent. I would like to leave this issue for future research.

The difference between the acceptance rates of (8) and (9) is not statistically significant at $p \leq .01$ by Mann-Whitney U test.

One might think that children can easily observe the subject-orientation property of zibun on the basis of the nominative Case-maker –ga, since Japanese is a nominative-accusative language; subjects are generally marked with the nominative Case-maker –ga. However, the existence of the nominative Case-marker would not be helpful for the generalization of subjecthood. First, as discussed in section 2, nominative marked DPs are not always subjects, and such DPs cannot be zibun’s antecedents. Moreover, Japanese is a pro-drop language. Thus, subjects are often omitted in colloquial speech. In addition, the nominative Case-marker -ga and topic marker -wa can also be omitted in colloquial speech even when subjects appear overtly. Therefore, it would not be easy for children to generalize subjecthood on the basis of morphological markers.

The position of the ni-phrase in full unaccusatives is beyond the scope of the current study.

According to Doner (2014), cross-linguistic variations concerning how to satisfy the EPP feature of T are classified into four types. It is satisfied by a DP, $D^0$, vP, or $V^0$. I assume that the EPP feature of T in Japanese is satisfied by a DP (Ura 2011). See Doner (2014) for detailed discussion.

See these references for detailed discussion about unaccusatives in child Japanese.

Acknowledgement

I would like to thank the audience at the 12th Generative Approach to Language Acquisition and members of Tokyo Psycholinguistic Laboratory for their helpful comments and discussion. I would also like to thank an anonymous reviewer for helpful comments. All remaining errors are my own.

References


THE ROLE OF EXECUTIVE FUNCTIONS IN THE ACQUISITION OF REFERENCE: THE PRODUCTION OF DEMONSTRATIVE PRONOUNS BY GERMAN MONOLINGUAL CHILDREN

JACOPO TORREGROSSA

1. Introduction

This paper deals with the acquisition of reference by German monolingual children between the age of 8 and 10. In particular, we will focus on the acquisition of the use of demonstrative pronouns (der, die, das – d-pronouns, henceforth) as compared to the use of personal pronouns (er, sie, es). We analyze the production of these referential expressions (REs, henceforth) in the context of story-telling.

When telling a story, speakers have to keep track of different referents, introducing them, maintaining reference to them across two (or more) adjacent discourse units, or reintroducing them after a hiatus. The degree of the referent’s activation (alias accessibility, salience – cf. Arnold 2010) varies throughout the story based on these discourse functions (introduction, maintenance and reintroduction). More specifically, various factors contribute to determining the referent’s activation. For example, recently mentioned referents are usually associated with a relatively high activation (cf. Ariel 1990). The grammatical role and the syntactic position of the referent’s previous mention (antecedent) play a significant role, too. The referent’s activation is enhanced if the antecedent is a subject (vs. object) and occurs in a main clause (vs. subordinate clause). The referent’s activation is also affected by the occurrence of competitor referents in discourse (cf. Arnold 2010 for a discussion of these factors and Kibrik 2011 and Torregrossa & Bongartz, to appear for a multi-factorial approach). Other factors that have an impact on activation include the perceptual availability of the referent in the discourse context and the extent to which both interlocutors are able to attend to the referent in question (Allen et al. 2008). In our study, we control for these additional factors, since the production of REs is elicited by means of a picture-based
story-telling task, in which only the child has visual access to the pictures (Section 3). In this context, the assessment of the referent’s activation in the discourse model – shared between the child and the investigator – should be insensitive to extra-linguistic sources (such as occurrence in the perceptual environment) and only based on discourse factors.

From the cognitive point of view, the ability to keep track of story referents requires attentional resources and executive functions (EFs, henceforth). The referent’s previous mention has to be stored and maintained in memory along with its linguistic features (e.g., argument role, syntactic position and distance) and retrieved at the point of the story in which the referent is mentioned again. Thus, the referent’s activation varies as a function of decay and retrieval history (Lewis et al. 2006). While working memory (WM, henceforth) is involved in the retention of information, the updating of this information in the unfolding discourse requires EFs (see, e.g., de Cat 2015).

The production of REs depends on the activation of the corresponding referent (Arnold 2010; Kibrik 2011). The more active a referent, the less explicit (or less informative) the corresponding RE. For example, personal pronouns in German (e.g., *er*, ‘he’) tend to pick up referents that are highly active, while full noun phrases (e.g., *der Hase* ‘the rabbit’) usually encode a low degree of a referent’s activation. However, the mapping between the referent’s activation and the use of a certain RE is not categorical, but rather subject to individual variation. For example, some speakers may rely on ‘overprotective’ strategies, using a full noun phrase in association with a high degree of the referent’s activation, to avoid mistakes leading to ambiguities (Kibrik 2011). Cognitive constraints contribute to variation of reference production, too. For example, Rosa & Arnold (2011) argue that under cognitive load, speakers tend to produce REs that encode a lower degree of the referent’s activation (i.e., full nouns vs. pronouns). Likewise, Torregrossa & Bongartz (to appear) show that a reduced processing speed correlates with the production of ambiguous forms (cf. also Hendriks 2016). In this paper, we investigate to what extent differences in EFs contribute to variation in reference production, given the role of WM in particular and EFs in general in maintaining and updating information held in mind.

From this short overview, it can be concluded that the acquisition of reference is a complex task, since it involves the integration of linguistic competence and cognitive abilities. Children have to acquire the repertoire of REs available in their language and the syntactic constraints regulating their distribution (cf. Torregrossa et al. 2015). Furthermore, they have to keep track of the varying activation of referents in the discourse model and
learn how to adequately map the referent’s activation into the use of a certain RE. All these learning processes are supported by cognitive systems which are not fully developed. For example, at the age at stake in this paper (i.e., from 8 to 10 years), EFs have still not reached the adult level (Gathercole et al. 2004). Some studies on the acquisition of reference have shown that in early spontaneous productions, children are sensitive to the distinction between given and new information from early on (i.e., from 3 y.o), which is reflected in the production of less explicit REs for given referents and more explicit REs for new ones (Skarabela et al. 2013; Hickman et al. 2015 for discussion). However, when taking into account more complex tasks, such as picture-based narrative production, adequate form-function mappings do not seem to be well mastered until 7-10 years. These tasks require the integration of visual and verbal information into a coherent discourse and the dynamic updating of the discourse model (de Cat 2015), and therefore impose a cognitive load. The literature indicates different patterns of acquisition of reference, depending on the discourse function that the child intends to express. While reference maintenance is mastered early on (from 4-5 years, Hickmann & Hendriks 1999), the acquisition of reference introduction and reintroduction emerges later (Koster et al. 2011; Hendriks et al. 2015). The short narrative in (1) – taken from Hendriks (2016: 2) – was told by a 6 y.o. Dutch child (Hendricks reports the English translation).

(1) A pirate\textsubscript{1} with the football. Then he\textsubscript{1} kicks it. Then it is in the water. Then the knight\textsubscript{2} goes to catch it. And he\textsubscript{2} has caught the ball in a net. Now he\textsubscript{1} has his ball back again.

To reintroduce the pirate in the last sentence, the child produces a pronoun. The use of a pronoun results in ambiguity for the interlocutor, since he can refer to both the pirate and the knight. Van Rij et al. (2011) and Hendriks (2016) argue that the production of underspecified, ambiguous REs is motivated by insufficient WM-capacity. Children have difficulty in managing (i.e., maintaining and updating) discourse information (distance, intervening referents, etc.) to determine the discourse topic, to which pronouns usually refer.

In this paper, we investigate how the acquisition of reference depends on the development of WM and EFs, by analyzing the production of d-pronouns by German monolingual children. In Section 2, we will show that the felicitous use of d-pronouns involves maintaining grammatical information encoded in previous discourse and updating the discourse model following a topic shift.
2. The phenomenon: The use of d-pronouns in German

German has demonstrative pronouns that are inflected for number, gender and case, and can refer to persons, as is the case of personal pronouns (Bosch et al. 2003). Some studies have shown that personal pronouns and d-pronouns exhibit a complementary pattern of anaphoric reference resolution: while personal pronouns tend to be resolved to antecedents that are topical, d-pronouns refer to non-topical entities (Bosch et al. 2003). This claim is supported by the corpus analysis carried out by Bosch et al. (2003). The authors show that more than 75% of d-pronouns refer to non-nominative antecedents, while this is the case of only 13% of personal pronouns. However, more recent studies have reconsidered the complementary hypothesis formulated by Bosch et al. (2003) and argued that while d-pronouns usually refer to non-topical entities, personal pronouns can be resolved to antecedents that are either topical or non-topical (Bosch & Umbach 2007 and Hinterwimmer 2015). The different distribution of d-pronouns and personal pronouns is illustrated by the sentences (2b)-(2d) below – taken from Patel-Grosz & Grosz, to appear. (2c) and (2d) show that d-pronouns can only refer to the object (i.e., non-topical) antecedent Paul. On the contrary, er (he) and ihn (him) in (2a) can pick up both Hans and Paul.

(2) a. Hans_1 wollte_3SING.PAST Paul_2 besuchen. 
   Hans_1 want.3SING.PAST Paul visit.INF
b. aber dann hat_3SING.PRES. er_1/2 ihn_2/1 angerufen. 
   but then AUX.3SING.PRES. he_1 him call.PPT
c. aber dann hat_3SING.PRES. der_2 ihn_1 angerufen. 
   but then AUX.3SING.PRES. D-PRON him call.PPT
d. aber dann hat_3SING.PRES. er_1 den_2 angerufen. 
   but then AUX.3SING.PRES. he_1 D-PRON call.PPT

The pattern shown in (2) is confirmed by the corpus analysis presented in Torregrossa (submitted). Based on narrative production data elicited from adult speakers, the author analyzes the distribution of personal pronouns and d-pronouns in four different contexts – which we will refer to again in Figure 1 under Section 4: i) the referential expression and the antecedent are in subject position (SUBJ-SUBJ); ii) the referential expression is in subject position and the antecedent is non-subject (NONSUBJ-SUBJ); iii) the referential expression is in non-subject position and the antecedent is subject (SUBJ-NONSUBJ); iv) both the referential expression and the antecedent are non-subjects (NONSUBJ-
The role of executive functions in the acquisition of reference

NONSUBJ). The corpus analysis shows that the great majority of d-pronouns (almost 70%) appear in contexts where the antecedent is a non-subject (i.e., NONSUBJ-SUBJ and NONSUBJ-NONSUBJ), in compliance with (2c) and (2d). Crucially, the study in Torregrossa (submitted) relies on the same methodology for data elicitation that is used in this paper (i.e., ENNI stories, cf. Section 3.2). Therefore, the results of the two papers are readily comparable with each other.

The example in (2c) also shows that d-pronouns may be associated with a topic-shift function: Hans is the topic in (2a), while Peter is the topic of (2c). Topic-shift has the effect of updating the activation of the referent and, more in general, the discourse model.

Finally, it should be noticed that the distribution of personal and d-pronouns complies with the abovementioned principle, according to which the more complex (i.e., the more explicit) the DP, the less activated is the corresponding referent in discourse. (3a) and (3b) illustrate the structure of personal pronouns and d-pronouns, respectively. While personal pronouns consist only of a functional projection, d-pronouns have a null NP and a DP-shell (cf. Wiltschko 1998 and Patel-Grosz & Grosz, to appear, from which the two structures are taken). The more complex structure of d-pronouns reflects their preference to pick up less activated antecedents (i.e., in object position)³.

(3a) Personal pronouns

\[
\begin{array}{c}
\text{FP} \\
\downarrow \\
F^\circ \\
\downarrow \\
er
\end{array}
\]

(3b) D-pronouns

\[
\begin{array}{c}
\text{DP} \\
\downarrow \\
D^\circ \\
\downarrow \\
d- \\
\downarrow \\
F^\circ \\
\downarrow \\
er \\
\downarrow \\
NP \\
\downarrow \\
\emptyset
\end{array}
\]

Coming back to the acquisition issue, our analysis aims to investigate whether children’s production of d-pronouns complies with the adult pattern described in this section. In particular, since d-pronouns tend to be used to reintroduce referents in discourse, we will verify the claim that the acquisition of reference in contexts of reintroduction is vulnerable (Section 1). Furthermore, we will verify to what extent WM and EFs contribute to the acquisition of reference. In particular, the children’s
adequate use of d-pronouns should be predicted by their performance in EF task.

3. Methods

3.1 Participants

The study is based on a sample of 21 German monolingual children ranging in age from 8.10 to 10.6 (mean age: 9.4; SD = .72). At the time of testing, the children attended the third or the fourth grade of a public primary school in the North Rhine Westphalia in Germany. The teachers reported that none of these children had a history of language delay or disorder, or socio-emotional problems.

3.2 Materials and procedure

The production of REs (personal pronouns and d-pronouns, in particular) was observed in the context of a story-retelling task. Retellings were elicited by using the Edmonton Narrative Instrument (ENNI) designed by Schneider et al. (2005). ENNI includes six stories, divided into three groups of increasing complexity. For our task, we used the stories of the greatest complexity (A3 and B3). Each of them consists of 13 pictures representing a series of events involving two major characters (an elephant girl and a giraffe boy in A3 and a dog girl and a rabbit boy in B3) and two minor ones (of different gender, too). The two stories have been designed to be structurally equivalent. The retelling task is used to facilitate the decoding of the pictures and the comprehension of the story (see Gagarina 2016 for discussion). Furthermore, it allows to establish the gender of the characters from the beginning of the experiment and avoid confounding effects caused by the fact that the elephant and the dog (which are masculine in German, i.e., *der Elefant* and *der Hund*) have visual appearance consistent with female stereotypes (e.g., a long dress, a skirt), while the giraffe and the rabbit (which are feminine in German, i.e., *die Giraffe* and *die Hase*) have visual appearance consistent with male stereotypes (e.g., short trousers and a hat). In the story prompt, the dog and the elephant are referred to by means of the feminine gender, i.e., *Hundina* and *Elefantina* respectively, while the giraffe and the rabbit by means of the masculine gender, i.e., *Giraffo* and *Haso*. Accordingly, the children were able to associate consistently throughout the story the
masculine gender to one character and the feminine gender to the other character.

The task was administered as a sequence of Power Point slides on a computer screen. The children had to choose one of three envelopes. Although all the envelopes contained only one of the two target stories (i.e., either A3 or B3), the children were told that the envelopes contained different stories (Serratrice 2007). Then, the participants looked at the story pictures two by two, while listening to the model story on the headphones. Finally, once the thirteen picture synopsis had appeared on the screen, they were asked to tell the story to the investigator, who feigned ignorance of the plot.

Each child was asked to tell one narrative and hence, the materials of the study consist of 21 narratives. The stories were audio-recorded and then transcribed into CHAT format (MacWhinney 2000) by a German native speaker. We refer to Andreou et al. (2015) for further details concerning the procedure of the experiment.

At the end of the experiment, we tested each child for WM and EFs, by using the backward digit recall test, which consists in listening to and recalling sequences of digits in reverse order. This test is classified as a complex memory span test, since it involves both storage (digit recalling) and processing (inverting order), and hence taps the phonological loop and the central executive, respectively (Baddeley & Logie 1999; Gathercole et al. 2004).

3.3 Analysis

We divided each narrative into clausal units, defined by the occurrence of a verbal form. To control for animacy effects, we analyzed only referential expressions denoting animate characters. We labelled each pronoun and d-pronoun for its grammatical role (distinguishing between subjects and non-subjects, see the third column in Table 1) and for the grammatical role of its antecedent (see fourth column in Table 1). For instance, the antecedent of the d-pronoun der in subject position in (2) is the definite DP der Hund (the dog) in subject position in the previous clause (see Table 1). Then, we counted the instances of personal and d-pronouns referring to subject or non-subject antecedents.
Table 10: Coding of referential expressions for: i) type (column 1); ii) grammatical role (column 2); and iii) grammatical role of the antecedents (column 3).

<table>
<thead>
<tr>
<th>UNITS</th>
<th>TYPE</th>
<th>GRAMMATICAL ROLE (RE)</th>
<th>GRAMMATICAL ROLE (A)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[... und dann kam der <em>Hund</em> [and then came the dog]]</td>
<td>DEFINITE</td>
<td>SUBJECT</td>
<td>...</td>
</tr>
<tr>
<td>...</td>
<td>DP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) und <em>der</em> wollte ja auch einen neuen Luftballon [and he also wanted a new balloon]</td>
<td>D-PRONOUN</td>
<td>SUBJECT</td>
<td>SUBJECT</td>
</tr>
<tr>
<td>(3) und <em>der</em> hatte auch kein Geld dabei. [and he also had no money with him]</td>
<td>D-PRONOUN</td>
<td>SUBJECT</td>
<td>SUBJECT</td>
</tr>
<tr>
<td>(4) und dann kam <em>die Mutter</em> von dem Hasen [and then came the mother of the rabbit]</td>
<td>DEFINITE</td>
<td>SUBJECT</td>
<td>INTRO</td>
</tr>
<tr>
<td>[and then came the mother of the rabbit]</td>
<td>DP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4') und dann kam <em>die Mutter</em> von dem Hasen [and then came the mother of the rabbit]</td>
<td>DEFINITE</td>
<td>NON-SUBJECT</td>
<td>SUBJECT</td>
</tr>
<tr>
<td>[and then came the mother of the rabbit]</td>
<td>DP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) und <em>der</em> hat die eingeholt. [and he reached her]</td>
<td>D-PRONOUN</td>
<td>SUBJECT</td>
<td>NON-SUBJECT</td>
</tr>
<tr>
<td>[and he reached her]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5') und <em>der</em> hat <em>die</em> eingeholt. [and he reached her]</td>
<td>D-PRONOUN</td>
<td>NON-SUBJECT</td>
<td>SUBJECT</td>
</tr>
<tr>
<td>[and he reached her]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) und <em>er</em> fragte (...) [and he asked (...)][and he asked (...)]]</td>
<td>PRONOUN</td>
<td>SUBJECT</td>
<td>SUBJECT</td>
</tr>
</tbody>
</table>

To account for individual variation in the (felicitous vs. infelicitous) use of d-pronouns, we took into account the instances of d-pronouns referring to a subject antecedent (which corresponds to the non-adult-like pattern, see Section 2) produced by each child, and normalized them for the square root of the total number of instances of d-pronouns. Finally, we ran a correlation between the results of this normalization and the scores obtained in the backward digit recall test (which amount to the number of correct trials) on the one hand, and with the participants’ age on the other. In this way, we could investigate whether the non-adult like production of d-pronouns was an effect of age or limited EF abilities (or both).
4. Results

The children produced 584 units in total. The narratives ranged in length from 20 to 38 units (M: 29.79; SD = 4.65). The analysis is based on 114 pronouns and 67 d-prouns.

Figure 1 reports the raw numbers concerning the occurrence of pronouns and d-prouns in each condition introduced in Section 2 (following the examples in (2)), i.e., SUBJ-SUBJ, NONSUBJ-SUBJ, SUBJ-NONSUBJ and NONSUBJ-NONSUBJ.

Figure 3: Distribution of pronouns and d-prouns across the four conditions SUBJ-SUBJ, NONSUBJ-SUBJ, SUBJ-NONSUBJ, NONSUBJ-NONSUBJ.

The two referential forms do not differ from each other in their distribution ($\chi^2(2) = 6.9$, $p > .05$). The great majority of d-prouns (either in subject or in non-subject position) is used to refer to a subject in previous discourse (55 instances, amounting to 82% of the total number of d-prouns). For example, in (4) the d-proun $\text{die}$ in U2 picks up the referent $\text{Elefantina}$ in subject position in U1. (5) shows a case in which the dative d-proun $\text{dem}$ refers to a subject antecedent, i.e., the young rabbit $\text{Haso}$ (written in parentheses because the referential expression $\text{Haso}$ appears in the first conjunct).

(4) U1: $\text{Elefantina hat ihn angeschaut}$

$\text{Elefantina AUX-3SG.PRES PRON-3SG.MASC.SING see-PPT.}$

$\text{Elefantina has looked at him}$
U2: weil die kein Spielzeug hatte.  
because D-PRON,FEM,SING. no toy have,3SG,PAST  
because she had no toy  

(5) U1: und (Haso) fragte nach dem  
and Haso ask,3SING,PAST for ART,DAT,MASC,SING.  
schönsten Ballon  
beautiful,ADJ,SUP,DAT,MASC,SING. balloon  
and (Haso) asked for the most beautiful balloon  

U2: aber der Hase sagt dem  
but the rabbit say,3SING,PRES. D-PRON,DAT,MASC,SING.  
but the rabbit told him  

(4) and (5) are instances of non-adult like use of d-pronouns. The children produce only few d-pronouns (either in subject or in non-subject position) referring to a non-subject antecedent (12 instances, i.e., 18%) in accordance with the adult pattern, as exemplified in (6), where the d-pronoun die (in U3) picks up the dative argument der Mutter (the mother) occurring in U1.  

(6) U1&2: und der Mutter erzählt was passiert  
and the mother,DAT,FEM,SING. tell,PPT. what happen,PPT.  
ist.  
AUX,3SING,PRES.  
and he told the mother what happened  

U3: und dann hat die uhm [//]  
and then AUX,3SING,PRES. D-PRON,NOM,FEM,SING.  
war die einverstanden.  
AUX,3SING,PAST D-PRON agree,PPT.  

(CH1; 10.5)  

The correlational analysis reveals that the (infelicitous) use of d-pronouns referring back to subject antecedents correlates negatively with the scores obtained in the digit backwards task ($r = -52.3$, $p < .05$) – see Figure 2: the lower the score in the EF task, the greater the frequency of d-pronouns referring to subject antecedents. Furthermore, to control for age effects, we correlated the use of d-pronouns with age and, as expected, we found no significant effect ($r = -28.4$, $p > .05$).
5. Discussion and conclusive remarks

Reference production and comprehension requires EF resources, which are necessary for calculating the referent’s activation and updating it in the unfolding discourse. German d-pronouns provide a relevant empirical domain to assess the role played by EFs in reference production. The adequate use of d-pronouns involves “backwards-processing” of the features associated with the referent in previous discourse (the grammatical role of the antecedent, in particular) and “forwards-processing” related to discourse updating due to topic shift (Schumacher et al. 2015).

With respect to the acquisition of reference, the aim of this study was to investigate if children use d-pronouns in an adult way (as described in the studies reviewed in Section 2) and to determine if possible differences are motivated by limited EF abilities. In particular, we formulated the hypothesis that better executive functions should be reflected in a more adult-like production of d-pronouns.

The results of our study suggest that children tend to use d-pronouns to refer back to subject antecedents and do not distinguish personal and d-pronouns in their conditions of use (Figure 1 in Section 4), contrary to what has been found for adults (cf. Section 2 and Torregrossa, submitted). Furthermore, we found that the children’s performance in EFs is a good
predictor of accuracy in the production of d-pronouns: the better the performance in EF tasks, the more felicitous the use of d-pronouns. Age correlates neither with the scores in the EF task nor with the (adequate) use of d-pronouns. This result confirms the appropriateness of our choice of the age group: in the age span at issue, there seems to be no significant development progression in EFs, which is reflected in the absence of a developmental trend in the felicitous use of d-pronouns.

Our study argues in favor of the hypothesis that the development of EFs plays a crucial role in the acquisition of reference. More in general, it supports the idea that referential resolution processes rely on attentional and EF resources.

Notes

1 German has several types of demonstrative pronouns (in addition to the demonstrative pronoun der), e.g., dieser (this), jener (that), derjenige (the one who), derselbe (the same), etc. (Bosch et al. 2003). In this paper, we take into account only the acquisition of der, given its pattern of complementary distribution with respect to the personal pronoun er (Section 2).

2 Bosch et al. (2003) identify the notion of topic with the notion of subject (i.e., a constituent that is nominative marked). However, other types of constituents can be topics, e.g., left-dislocated accusative-marked constituents. In this paper, we follow the analysis of Bosch et al. (2003), since the narratives of our corpus mainly comprise SVO sentences, in which the informational category of topic aligns with the grammatical function of subject.

3 (3a) and (3b) can easily account for the structure of singular masculine and neuter personal pronouns and d-pronouns in the nominative, dative and accusative case, abstracting away from some spelling differences (Wiltschko 1998: 149). The other forms in the paradigm exhibit some idiosyncrasies, which Wiltschko discusses in depth (ibid.:150-156). For example, the morpheme s- occurring in the singular feminine personal pronoun sie in the nominative and accusative case is analyzed as a support morpheme (and crucially not as D°, as is the case of the corresponding d-pronoun die). Also, the apparent irregularities of the genitive paradigm (personal pronouns: seiner/ihrer/??seiner; d-pronouns: dessen/deren/dessen) can be attributed to the ‘defective’ nature of the genitive agreement ending. In the case of d-pronouns, the genitive agreement marker –es in F° cannot license the empty NP in its complement, and the agreement ending –en is added to save the structure. With personal pronouns, the expected genitive form es/er/es (which cannot be spelled out by itself) attaches to the possessive determiner, resulting in a full DP. Finally, Wiltschko analyzes the –en suffix which appears in the dative plural forms of personal pronouns (ihnen) and d-pronouns (denen) respectively as a cliticization on F°. In all cases, the analysis of Wiltschko
is also supported by diachronic data, with special reference to the transition from Middle High German to New High German.

Many studies have shown that the use of referential expressions is sensitive to the animacy of the corresponding referent (a.o., Fukumura & van Gompel 2010). Therefore, children may tend to associate one type of referential expression (e.g., pronouns) to animates and another type (e.g., d-pronouns) to inanimates. By considering only animates, our analysis eliminates animacy as an interfering factor.

The correlational analyses concern only 17 children among the 21 included in the study, since 4 children did not produce d-pronouns.

References


Referring Expressions: Bridging the gap between computational, empirical and theoretical approaches.


LIST OF CONTRIBUTORS

Philip BACKLEY (Tohoku Gakuin University)

Anamaria BENTEÁ (University of Geneva)

Solveig CHILLA (University of Education Heidelberg)

Michael CLAUSS (UMass Amherst)

Hamida DEMIRDACHE (Université de Nantes)

Ronice Müller DE QUADROS (Universidade de Santa Catarina)

Xiaoli DONG (UiL-OTS)

Yoshiki FUJIWARA (Meiji Gakuin University)

Maria Teresa GUASTI (University of Milano-Bicocca)

Cornelia HAMANN (University of Oldenburg)

Shenai HU (University of Verona)

Koji KAWAHARA (Nagoya University of Foreign Studies)

Natasa KNEZEVIC (Université de Nantes)

Viola Laura KOZAK (Gallaudet University)

Mohamed LAHROUCHI (Université Paris 8 –CNRS)

Bert LE BRUYN (UiL-OTS)

Tatjana LEIN (University of Bremen)

Peggy LI (Harvard University)
Diane LILLO-MARTIN (University of Connecticut)
Suzi LIMA (Universidade Federal do Rio de Janeiro)
Maria LOBO (FCSH/Nova University)
Kuniya NASUKAWA (Tohoku Gakuin University)
Hitomi ONUMA (Tohoku Gakuin University)
Emma NGUYEN (University of Connecticut)
Jeffrey Levi PALMER (Gallaudet University)
Deborah Chen PICHLER (Gallaudet University)
Barbara PEARSON (UMass Amherst)
Wanette REYNOLDS (Gallaudet University)
Tom ROEPER (UMass Amherst)
Monika ROTHWEILER (University of Bremen)
Rita SANTOS (FLUL, University of Lisbon)
Anca SEVCENCO (University of Bucharest)
Hiroyuki SHIMADA (Meiji Gakuin University)
Carolina SILVA (FCSH/Nova University)
Jesse SNEDEKER (Harvard University)
William SNYDER (University of Connecticut)
Jacopo TORREGROSSA (University of Cologne)