Considering the whole paradigm: Preschoolers' comprehension of agreement is not uniformly late

Hannah Forsythe

Language Sciences, University of California, Irvine

2211 Social Science Plaza B, UCI, Irvine, CA 92612, hforsyth@uci.edu

Cristina Schmitt

Linguistics and Languages, Michigan State University

Considering the whole paradigm: Preschoolers' comprehension of agreement is not uniformly late

Picture-selection studies testing the comprehension of subject-verb agreement across a wide variety of languages have found that children fail to use these markers to infer the cardinality of a null or masked subject, even long after spontaneous production of agreement becomes adult-like (Johnson et al. 2005, Pérez-Leroux 2005, Gxilishe et al. 2009, Rastegar et al., 2010). This production/comprehension asymmetry has since been called into question by researchers using more sensitive comprehension measures (Brandt-Kobele and Höhle 2010, Legendre et al. 2014, Verhagen & Blom 2014, González-Gómez et al. 2017). Here, we continue to use a picture selection task but expand our focus to include 1st and 2nd person markers, which are more phonologically salient and less complex to interpret than 3rd person markers. We find that preschoolers comprehend Spanish 1st and 2nd person agreement markers early (2;3-2;11) and just as accurately as clitic pronouns, showing that comprehension of agreement is not uniformly difficult. In the 3rd person, accuracy was low not only for agreement but also for clitics, and both adults and children allowed these expressions to refer to a prominent referent from the immediately preceding discourse—even if that referent was the speaker or hearer. We argue that the production/comprehension asymmetry originally observed for 3rd person agreement was driven not just by task demands but also by children's sensitivity to the discourse dependence of 3rd person referring expressions. Future experimental investigations of agreement markers should take their discourse properties into account.

Keywords: verbal morphology, clitics, discourse prominence, Spanish, phi-features, person features, number features

1. Introduction

Abstract agreement relations are part of the core syntax of natural languages. Although the presence of overt agreement morphology is not universal in human language, overt marking of the subject-verb agreement relation is widespread (see Corbett 2006, among others), and children acquiring languages with rich agreement morphology begin producing it especially early on in life (Phillips 1995, and references therein). While it may be argued that the existence of agreement relations is provided by UG, their morphological realization varies widely across languages and must therefore be learned. This paper takes up the question of how children acquire such morphological realizations.

Acquiring agreement morphology requires the child to establish a mapping between a morphophonological form, the formal syntactic features realized by that form, and the semantic interpretation of these features, which in the case of verbal agreement typically denote the subject's person, number, and possibly also gender or class. Complete acquisition can thus be reasonably defined as the ability to reliably traverse this mapping in both directions—from phi-features to morphology (production) and from morphology back to phi-features (comprehension). Children's early target production of agreement markers would seem to suggest that this mapping is established early, yet a substantial number of studies have found that children fail to use agreement markers in comprehension tasks until as late as 5 or 6 years old (English: Johnson et al. 2005, Farsi: Rastegar et al., 2012, Spanish: Pérez-Leroux 2005, Xhosa: Gxilishe et al. 2009). Why would this be? If children produce agreement, they must be able to successfully link the underlying semantic features

of agreement to their surface forms. What would prevent them from making the same link in the other direction?

Researchers have responded to this puzzle in different ways: one impulse has been to link children's failure to use verbal agreement to the fact that the verb itself is not the source of person, number, and gender features, and thus there is no direct link between verbal morphology and the locus of interpretation of these features. In minimalist terms, verbal agreement features receive an interpretation in the phonological component but not in the semantics (Chomsky 1995 and much subsequent work). This means that associating a semantic interpretation to verbal agreement is an inferential process; the presence of a feature on the verb should be interpreted as indicating that there is an interpretable feature on the subject, which could potentially explain why children fail to assign agreement markers the proper interpretation, despite producing them correctly (Johnson et al. 2005). In other words, children should interpret phi-features on verbal forms less readily than phi-features on nominal and pronominal forms.

There are at least two problems with this approach. First of all, the boundary between verbal agreement, which is in principle uninterpretable, and pronominal agreement, which is in principle interpretable, is not always very clear. The status of verbal morphology varies across languages, depending on its morpho-phonological properties and one's theoretical commitments. For example, in canonical null subject languages like Spanish and Greek, some analyses propose that verbal agreement is actually the instantiation of the pronominal subject (ex. Alexiadou & Anagnostopoulou 1998 and much subsequent

literature; see Barbosa 2009 and Roberts 2010 for review). So-called agreement markers are actually enclitic pronouns attached to the verb and are therefore the source of interpretable phi-features. Accepting this widely adopted proposal means losing the explanation for children's difficulty with agreement. The second problem with this approach is empirical. More recent work finds that comprehension of so-called verbal agreement forms is not equally delayed across languages. For instance, Legendre et al. (2010a) argues that French pre-verbal agreement markers[1] are comprehended by children as young as 2;6.

A more recent response to the puzzle of children's difficulty with agreement, at least in some languages, has been to argue that task demands play a large role in performance. The studies that originally reported the comprehension delay all employed a picture-selection paradigm requiring the child to point to one of two pictures. For example, Johnson et. al. (2005) presented a singular picture (ex. one duck swimming) and a plural picture (ex. a pair of ducks swimming) and asked the child to "Show me the picture where…the duck swim**s** in the water" or "…the ducks swim_ in the water." (The plural /s/ on the subject noun was always masked by a following /s/-initial verb.) One problem with this paradigm is that pointing arguably places higher demands on executive control than spontaneous production. Another problem is that singular phrases like "the duck swims" are technically compatible with a portion of the plural picture, since each individual duck independently

---

[1] These authors assume, based on Legendre et al. (2010b), that pre-verbal subject clitics in child-directed French have the status of agreement markers, though not all authors agree (ex. deCat 2002). This disagreement only serves to underscore the lack of consensus about where to draw the line between pronominal (interpretable) and verbal (uninterpretable) agreement.

performs the target action. Subsequent studies evened the playing field between comprehension and production tasks by using preferential looking instead of pointing (Brandt-Kobele and Höhle 2010, Legendre et al. 2010a), by comparing picture selection to elicited rather than spontaneous production (Verhagen & Blom 2014), and by substituting static pictures with videos (Legendre et al. 2014, González-Gómez et al. 2017). Other studies have discouraged the distributive interpretation of plural pictures by including semantically general objects (*agarran el objeto* '(they) grab-PL the object') as well as nonce objects (*agarran el micho* '(they) grab-PL the micho') within the carrier phrase (Legendre et al. 2014, González-Gómez et al. 2017).

These modifications generally reveal earlier comprehension than was reported in the original picture-selection studies. However, it is still not clear exactly how much of the originally reported comprehension delay is directly attributable to task effects, since each of these modifications was made in a different language. This is important because cross-linguistic work reveals that children's performance also varies by language (Legendre et al. 2014), leading some researchers to suggest that phonological and/or morphological properties of the agreement markers that children are exposed to play an important role in their comprehension of those markers (Pérez-Leroux 2005, Legendre et al. 2014).

While both of these approaches have contributed to a better understanding of children's acquisition of verbal agreement, it is important to note that they are based on only a subset of the agreement paradigms of these languages. Both the original picture-selection studies (Johnson et al. 2005, Pérez-Leroux 2005, Gxilishe et al. 2009, Rastegar et al., 2010) and the

subsequent task modifications (Brandt-Kobele and Höhle 2010, Legendre et al. 2014, Verhagen & Blom 2014, González-Gómez et al. 2017) focused exclusively on the number contrast between 3rd person singular and 3rd person plural agreement. This was done to maintain consistency with results from languages like English that lack 1st and 2nd person agreement markers, but it has meant that our understanding of the acquisition of agreement is shaped by only a portion of the paradigm—a portion that may present particular challenges to the child listener. Cross-linguistically, 3rd person morphology tends to be more phonologically reduced than 1st and 2nd person morphology (Harley & Ritter 2002), potentially making 3rd person agreement markers more difficult to perceive or identify during an experimental task. Moreover, 3rd person pronouns can be complicated to interpret because of the way that they depend on the surrounding discourse. While 1st and 2nd person pronouns pick out the speaker or the addressee (or a group of speakers or addressees, in the case of plurals), 3rd person pronouns pick out a referent that not only matches their phi-features but that is also sufficiently prominent in the discourse for the listener to recover, either from the physical context or the preceding linguistic context (Ariel 1988, Gundel, Hedberg & Zacharski 1993). Consequently, listeners must consider not only the features of competing potential referents, but also their prominence. In sum, in the case of 3rd person agreement markers, children may find comprehension more challenging than production because of their reduced phonological content and greater dependence on the surrounding discourse. Studying the comprehension of 1st and 2nd person agreement markers would

allow us to observe children's developing comprehension of agreement markers without either of these complications.

   In this paper, we therefore examine how preschoolers acquiring Spanish in Mexico City interpret not only 3rd person singular and plural agreement markers, but also 2nd person singular[2] and 1st person singular and plural markers, using a picture-selection task. In addition, we compare children's comprehension of agreement to their comprehension of accusative clitics, which encode the same person and number features (in addition to gender, see Table 1). This will allow us to probe the extent to which phonological and semantic differences may have contributed to children's reported difficulty with agreement in the original picture-selection tasks. If low phonological salience interferes with children's ability to perceive agreement markers, then comprehension of not only 3rd person but also 1st and 2nd person singular agreement markers should show a delay vis-à-vis clitics, which are more phonologically robust in that they are comprised of a full syllable. If on the other hand, the problem is due to the inherent complexity of interpreting semantically 3rd person expressions, then comprehension of not only 3rd person agreement but also 3rd person clitics should show a delay vis-à-vis 1st and 2nd person agreement and clitics.

---

[2] We set aside 2nd person plural pronouns and their accompanying agreement because these forms are subject to a rather complex pattern of regional variation, with some dialects differentiating between formal and informal register and others not (see Lipski 1994 for details). Mainstream Mexican Spanish, the variety studied here, makes no formality distinction in the 2nd person plural. The pronominal subject *ustedes* triggers 3rd person plural agreement and the accusative clitic is syncretic with 3rd person plural. To avoid the ambiguity caused by this syncretism, we restrict out attention to 2nd person singular informal forms only.

Table 1. Spanish regular present-tense verbal agreement and accusative object clitics

	singular	plural	singular	plural
1st person	-o	-mos	me	nos
2nd person	-s	--	te	--
3rd person	-Ø	-n	lo *(masc.)*	los *(masc.)*
			la *(fem.)*	las *(fem.)*

Our results reveal early comprehension of 1st and 2nd person agreement markers, with similar accuracy scores across agreement and clitic conditions, showing that even in a picture-selection task, children do not necessarily struggle to comprehend agreement markers *per se*. Within the 3rd person, we replicate the low accuracy scores reported in other picture-selection tasks testing 3rd person agreement, but accuracy is low for clitics as well, suggesting that children's comprehension problems may have more to do with the semantics of 3rd person expressions than with the properties of agreement. More importantly, however, adults also performed significantly worse in 3rd person conditions relative to 1st and 2nd person conditions. Further analysis of adults' non-target answers suggests that they allow 3rd person expressions to refer to a recently mentioned referent from the preceding linguistic discourse—even if that referent is the speaker or hearer. Children show the same behavior, indicating that they are aware of the discourse dependency of the 3rd person.

2. Research questions and background

Before asking whether the interpretation of verbal person and number agreement is difficult for semantic reasons or for morpho-phonological reasons, the most basic research

question we address is simply whether children experience any difficulty comprehending 1st and 2nd person subject agreement markers in picture-selection tasks, since this has yet to be tested. The second question we address is whether children's difficulty comprehending 3rd person subject agreement in this type of task is due to their low phonological salience relative to clitics, or to the greater complexity of interpreting 3rd person expressions, relative to 1st and 2nd person expressions.

The literature provides compelling evidence for the role of phonological salience in determining the time course of children's acquisition of inflectional morphemes. Still other researchers attribute differences in the rate of acquisition of these morphemes to their semantic properties. We review this literature briefly before spelling out the predictions for children's comprehension of Spanish agreement and clitics.

In his pioneering work on the acquisition of morphology, Dan Slobin was one of the first to suggest that the acquisition of inflectional morphemes is conditioned by their phonological salience. This type of explanation is supported by a recent study by Davies, Rattanasone & Demuth (2016) on the acquisition of English plural nominal markers showing that English-acquiring children have earlier comprehension of the plural allomorph [s], relative to the more frequent but less acoustically salient [z] allomorph, which has a much shorter average duration. Aside from duration, what exactly makes a morpheme "salient"? Polišenska (2010) offers a general operationalization of phonological salience, defined as "how easy it is to detect the root and the affix in a given word form" (pg. 23). She proposes that phonological salience is determined by both its content and its

environment: inflectional morphemes with more phonological material, that are word-initial (or at least, not word-internal), that are stressed, and that are not subject to reductive processes like shortening, are predicted to be acquired earliest.

This type of explanation has been argued to account for asymmetries in the acquisition of verbal inflection, specifically, both within and across languages. In a video selection task, Legendre et al. (2014) found evidence of very early comprehension of French singular and plural pre-verbal markers /il/ and /ilz/ at around 30 months and somewhat later comprehension of Spanish plural agreement marker /-n/, at 30-47 months, while in a (supposedly less taxing) preferential looking task they found no evidence of any comprehension of the English singular agreement marker /-s/ among any of the children in their sample, aged 28-46 months. They propose that these differences—especially the very early comprehension found in French—can be attributed to cross-linguistic differences in the perceptual salience of these agreement markers, complemented by what they refer to as cue reliability[3]. In French, the segment [z] which distinguishes plural *ils* from singular *il*, is a strident and hence more acoustically salient than the segment [n] that distinguishes the Spanish plural (ex. *agarran* '(they) grab') from the singular (ex. *agarra* '(s/he) grabs'). And unlike the English strident allomorphs [s] and [z], which distinguish singular (ex.

---

[3] Cue reliability refers to the strength of the association between a marker and its meaning. In French, the segment [z] which distinguishes plural *ils* from singular *il* also distinguishes singular from plural nouns and is thus a reliable cue to plurality across verbal and nominal domains. In Spanish, word-final [n] is associated to plurality when it appears on verbs, but on nouns it has no meaning, making it a somewhat weaker cue to plurality. And in English, word-final /s/, realized as [s] or [z] or [əz], is associated to singularity when appearing on a verb but to the exact opposite meaning when found on nouns, making this segment a downright unreliable cue to singularity. This paper will focus on testing the effects of perceptual salience.

grabs) from plural (ex. *grab*), the French [z] appears in an environment that enhances the perceptibility of consonantal features: due to a process known as liaison, this segment only appears when the following word is a vowel (with which it resyllabifies), a position that has been argued to enhance consonant perceptibility (Redford & Diehl 1999, Benki 2003). These characteristics—coupled with the fact that [z] is associated with plurality not only in the verbal domain but also the nominal domain—make the French marker a very reliable cue to plurality.

Phonological salience is not the only dimension along which agreement markers differ; they also encode different phi-features whose mapping to the semantics is not always straightforward. How much do these semantic differences affect acquisition? Literature on the semantics and acquisition of person suggests that morphemes instantiating 1st and 2nd person features should be simpler to comprehend than their 3rd person counterparts for both formal and pragmatic reasons.

Since at least the work of Benveniste (ex. Benveniste 1971), 3rd person is considered to be formally less marked than 1st and 2nd persons. In Distributed Morphology accounts, 3rd person exponents are unmarked in the sense that they either lack a participant feature or just carry fewer features than other persons, thus capturing the fact that 3rd person tends to pattern differently from 1st and 2nd person cross-linguistically (see Harley & Ritter 2002 for an implementation in which 3rd person lacks all participant features, but see Nevins 2007 for arguments that 3rd person exponents cannot be completely unmarked). On a semantic level, 3rd person is unmarked in the sense of being underspecified: whereas 1st and 2nd

person denote the speaker and addressee, respectively[4], 3rd person imposes no restrictions

on the potential referent. As long as other features like person and gender are satisfied, 3rd

person expressions may refer to anyone, even the speaker or addressee as in (1) below.

(1)   Every one of us loves **her** mother. That's why we are here at this mother-daughter brunch

The developmental implications of this formal difference are most clearly spelled out

by Legendre & Smolensky (2012), who show that even in French, where children

comprehend pre-verbal agreement markers very early, children still comprehend the 3rd

person subject marker *elle* ('she') later than they produce it and, crucially, later than they

comprehend 1st and 2nd person subject markers *je* and *te*. Adopting an OT framework, they

show how this production/comprehension asymmetry, which only affects the 3rd person, is

a direct result of its formal underspecification: in a nutshell, it causes ambiguity for the

listener but not for the speaker. In production, children must decide which surface form best

expresses reference to a non-speaker, non-addressee referent. Since *je* and *te* both carry

features that clash with this interpretation, the only grammatical choice is a 3rd person form.

In comprehension, however, children must decide which interpretation is best suited to a

surface form like *elle*, which carries no person features whatsoever and is therefore a

grammatically licit choice for *any* feminine referent, including the speaker and addressee.

Thus, the interpretation of *elle* and other 3rd person expressions requires children to choose

---

[4] Alternatively, 1st and 2nd person features introduce presuppositions that restrict the value of the pronoun (ex. Sauerland 2008, Charnavel 2017, and many others).

between multiple options, whereas production is not ambiguous in this way. Adults

approach this problem by assuming that their interlocutor follows the convention of using

maximally specified forms like *je* and *te* wherever possible and reserving underspecified

forms like *elle* for only those contexts where all other forms would be ungrammatical (a

maxim first proposed by Heim 1991 and dubbed 'Maximize Presupposition'). Thanks to

this assumption, adults typically infer that the use of 3rd person expressions implies

reference to *non*-speaker, *non*-addressee referents, an implication referred to as an

'Implicated Presupposition' (see Sauerland 2003). Children, on the other hand, do not

initially assume 'Maximize Presupposition.' Therefore they initially fail to calculate

Implicated Presuppositions, leading to non-adult-like interpretations of 3rd person

expressions like *elle*.

There is evidence from other languages as well that 3rd person is more difficult for

children to comprehend relative to 1st and 2nd person. In English, Brener (1983) found that

preschoolers comprehend 3rd person pronouns in overheard speech later than 1st and 2nd

person pronouns in the same context. An even more interesting and relevant case comes

from a study on the comprehension of agreement in Spanish (Miller & Schmitt 2014, Expt.

3). In Chilean Spanish, the 2nd person agreement marker /-s/ is variably aspirated ([h]) or

deleted (ø) thanks to a process of variable lenition that targets syllable-final /s/ segments.

Miller & Schmitt used an act-out task to test the comprehension of the 2nd person singular /-

s/, the 3rd singular /ø/, and the 3rd plural /-n/ by Chilean 3-to-5-year-olds and by Mexican

children of the same age, who are not exposed to /s/ lenition. The Chilean children

performed significantly worse in the 2nd singular /-s/ condition compared to the Mexican children, as expected. Interestingly however, both groups of children still achieved greater-than-chance accuracy in this condition (80% and 88% accuracy, respectively), compared to much lower performance in both the 3rd singular condition (10% and 42% accuracy) and 3rd plural condition (63% and 73% accuracy). This suggests that *even in a language with /s/ reduction*, children comprehend the 2nd singular /-s/ agreement marker earlier than both singular and plural 3rd person agreement markers.

Even when listeners do infer that a 3rd person expression must refer to a non-speaker, non-addressee referent (i.e., they successfully calculate the Implicated Presupposition associated to 3rd person), the intended referent may still not be immediately obvious, since there can be *a priori* any number of available non-speaker, non-addressee referents. The listener must rely on additional information to infer which of these potential referents is the one intended by the speaker. If the pronoun carries number and gender features, this can narrow down the set of potential referents, but there is no guarantee that this will unambiguously identify a single referent. Often, listeners must rely on contextual information, such as the relative prominence of referents in the surrounding discourse, to infer which is the one intended by the speaker. For example, in example (2) from Arnold et al. (2000), both referents are grammatically and pragmatically compatible, but *Donald* is more prominent by virtue of being the first-mentioned antecedent, which is why this interpretation tends to be processed more readily by adults.

(2)   Donald is bringing some mail to Mickey while a violent storm is beginning. **He**'s carrying an umbrella, and it looks like they're both going to need it.

In contrast, 1st and 2nd person pronouns are much easier to disambiguate, since in most situations there is only one speaker and one addressee, and both referents are highly salient.

3. Hypotheses and predictions

The first hypothesis we test is that phonological salience conditions children's comprehension of agreement in Mexican Spanish. Generally speaking, for any two morphemes with the same features, the one with more phonological material, more acoustically salient material, and a more perceptually enhancing phonological environment will be comprehended earlier. Since Spanish accusative clitics instantiate the same person and number features as agreement, and since both agreement markers and accusative clitics can appear in a prosodically dependent, post-verbal position as in (3), we can test this hypothesis by comparing children's comprehension across the two paradigms.

(3)   a. *Muéstrame la foto en donde baila**mos/o/s/n/ø**.*
            Show-me the photo in where dance-1P/1S/2S/3P/3S
         "Show me the photo where we/I/you/they/(s)he dance(s)."

        b. *Muéstrame la foto en donde Nemo está tapándo**nos/me/te/las/la**.*
            Show-me the photo in where Nemo is covering-1P/1S/2S/3P/3S
         "Show me the photo where Nemo is covering us/me/you/them/her."

Every agreement marker except the 1st person plural has less phonological material than the corresponding accusative clitic (which does not combine with the verbal stem directly).

We therefore predict significantly poorer and/or later comprehension of each agreement marker relative to its clitic counterpart (except for 1st person plural).

The second hypothesis we test is that the formal and pragmatic differences between 3rd person, on the one hand, and 1st and 2nd person, on the other, condition children's comprehension of agreement. This hypothesis predicts that, for any two morphemes with the same number features (and gender features, if applicable), the one that encodes 1st or 2nd person will be easier to comprehend than the one that encodes 3rd person. We test this hypothesis by comparing 3rd person to 1st and 2nd person forms with the same number features, within each paradigm. We predict poorer and/or later comprehension of each 3rd person clitic or agreement marker relative to its 1st person and 2nd person counterparts. These predictions are summarized in Table 2.

Table 2. Comprehension asymmetries predicted by the phonological hypothesis and the semantic hypothesis

	Agreement	Clitics	Semantic hypothesis
	1st plural *–mos*	1st plural *nos*	
	1st singular *–o*	1st singular *me*	earlier
	2nd singular *–s*	2nd singular *te*	
	3rd plural *–n*	3rd plural *los, las*	-------
	3rd singular *–ø*	3rd singular *lo, la*	later
Phonological hypothesis	later <<	earlier	

4. Methods
4.1. Subjects

We tested 46 native Spanish-speaking children ages 2;3-6;7 (mean 4;2) from a daycare

in Mexico City, Mexico; 4 were excluded from the final analysis due to poor performance

in filler trials (see below). Adult subjects included 25 adults (12 women) recruited from

among the teachers and administrators at the daycare, or from the Latinx community of

Michigan State University. All adults were born and raised in Mexico, with Spanish as their

first language and were younger than 40 years old, to ensure the felicitous use of the

informal 2nd person pronoun *tú* during the task. 2 adults were excluded because they used

the formal 2nd person pronoun *usted* when addressing the experimenter. The primary

experimenter in Mexico was a female teacher from the school, and the primary

experimenter in the U.S. was a female native Spanish-speaking undergraduate at Michigan

State University. Adult subjects in Mexico were not compensated, while those in the U.S.

received $15 for their participation.

4.2. Experimental stimuli and fillers

Subjects' interpretation of present-tense agreement and object clitics was tested using a

picture-selection task consisting of 30 test items (15 agreement, 15 clitics), as in (4)-(5).

Subjects were presented with an array of 5 photos depicting (i) the subject, (ii) the

experimenter, (iii) the subject and experimenter together, (iv) an unrelated teacher, and (v)

two unrelated teachers together. Everyone in the photos was depicted performing the same

action so that the target photo could only be identified by interpreting the agreement or

clitic morpheme used in the prompt. Only feminine clitics were tested, and all

experimenters were female.

    (4)   *Muéstrame la foto en donde salta**mos/o/s/n/Ø***
           show-me the photo in which jump-**1P/1S/2S/3P/3**
           "Show me the photo where we/I/you/they/she is/are jumping."

    (5)   *Muéstrame la foto en donde Nemo está besándo-**nos/me/te/las/la***
           show-me the photo in which Nemo is kissing-CL.**1P/1S/2S/3P.fem/3S.fem**
           "Show me the photo where where Nemo is kissing us/me/you/them/her."

14 out of 30 experimental trials were preceded by fillers like those in (6)-(7). Subjects

were presented with an array in which each of the same 5 person(s) performed a different

action, so that the target photo could be identified by interpreting the lexical content of the

verb. 14 out of 30 experimental trials were preceded by a completely unrelated distractor

asking the subject which of two cartoon characters had more of various objects and

substances. 2 out of 30 experimental trials appeared at the beginning of the block and were

not preceded by anything.

    (6)   *Muéstrame la foto en donde hay alguien saltando/bailando/…*
           show-me the photo in which there-is someone jumping/dancing/…
           "Show me the photo where someone is jumping/dancing/etc."

    (7)   *Muéstrame la foto en donde Nemo está besando/peinando/… a alguien.*
           show-me the photo in which Nemo is kissing/combing…A someone.
           "Show me the photo where Nemos is kissing/combing/etc. someone."

For the agreement condition, the following de-transitivized verbs were used: *saltar* ('jump'), *aplaudir* ('clap'), *dormir* ('sleep'), *dibujar* ('draw') and *bailar* ('dance'). In the clitic condition, photos showed a puppet named Nemo performing the following actions on each of the 5 person(s): *besar* ('kiss'), *peinar* ('comb'), *lavar* ('wash'), *tapar* ('cover') and *tocar* ('touch'). Fillers used the same ten actions, plus an additional four items in which the target action was sitting (*sentada*) and lying down (*acostada*).

4.3. Design and procedure

Agreement and clitics were presented in separate blocks, with the agreement block first. Subjects were randomly assigned to one of two different versions of the task, each with a different random ordering of test items. Fillers and distractors followed the experimental items in alternating order. Fillers were ordered so that each filler would depict a different action from the immediately preceding or following test item, and this ordering was re-shuffled after every other subject, in order to mitigate the possibility that any particular order would bias the interpretation of the test items.

Photos not including the subject were taken beforehand and pre-inserted into the arrays in random order. Photos including the subject were taken no more than one week before testing and then inserted into the arrays in random order.

Testing was preceded by a familiarization phase in which the subject was asked to identify each of the actors by name (him/herself, the primary experimenter, and the other

two adult women) and any errors were corrected. Next, the primary experimenter

introduced the task and obtained consent through the following:

(8)   *Vamos a ver algunas fotos de personas hacienda varias cosas y tú me vas a
      señalar la foto que yo te diga, ¿te parece? Pero sólo me vas a señalar una foto
      nada más, ¿bien?*
      We're going to see some photos of people doing different things and you're going
      to point out the one I tell you, sound good? But you can only pick one photo,
      okay?

Halfway through each block, there was a short break in which children were given a

sticker. After the task was complete, child subjects received a piece of candy and adult

subjects were debriefed and/or given compensation. Any child who refused or repeatedly

displayed unwillingness to participate in any part of the test was excused. The entire

procedure, including the photo shoot and testing, lasted approximately 30 minutes.

4.4. Coding and exclusions

Responses were recorded on a sheet of paper by the secondary experimenter and then

transferred to a spreadsheet for coding and analysis in R (R Core Team 2013). Any photo

containing the target referent was considered a target response, regardless of whether it also

included another referent as well. This means that in singular conditions, plural answers

containing the target referent were marked as correct.

Three children were excused before completing the task and one child was excluded

due to an extremely low score on filler questions (50% correct of all intelligible responses),

leaving a total of 42 child subjects. Two adults were excluded for failure to address the primary experimenter using the informal 2nd person *tú*, leaving a total of 23 adult subjects.

5. Results

Adult and child responses are reported in Table 4 (highlighted cells represent target answers). To test whether comprehension is conditioned by the phonological differences between agreement and clitics or by the contrast between 3rd person and the other two persons, we ran 2 x 5 repeated-measures ANOVAs on the proportion of target answer produced by children and by adults, with form (agreement, clitic) and features (1st singular, 2nd singular, 3rd singular, 1st plural, 3rd plural) as within-subjects factors. We then followed up with paired-sample t-tests comparing the proportion of target responses across agreement markers and clitics with the same person and number features, as well as paired-sample comparisons between each 3rd person agreement marker or clitic and the corresponding 1st and 2nd person agreement and clitic forms with the same number feature.

Table 3. adult responses (N = 23; target responses in shaded cells)

response	agreement					clitics				
	1Sg	1Pl	2Sg	3Sg	3Pl	1Sg	1Pl	2Sg	3Sg	3Pl
investigator	56	0	1	14	1	67	0	0	19	1
child & investigator	7	66	3	5	26	1	68	6	0	14
child	2	0	64	17	0	1	0	63	5	0
other female-Sg	4	0	0	30	3	0	0	0	44	0
other female-Pl	0	3	1	3	38	0	1	0	1	54
other	0	0	0	0	1	0	0	0	0	0
no answer	0	0	0	0	0	0	0	0	0	0
proportion target	0.91	0.96	0.97	0.48	0.55	0.99	0.99	1.00	0.65	0.78
(SD)	(0.22)	(0.15)	(0.13)	(0.37)	(0.27)	(0.06)	(0.06)	(0)	(0.35)	(0.31)

Table 4. child responses (N = 42; target responses in shaded cells)

response	agreement					clitics				
	1Sg	1Pl	2Sg	3Sg	3Pl	1Sg	1Pl	2Sg	3Sg	3Pl
investigator	57	9	2	15	12	88	11	5	29	6
child & investigator	54	86	31	33	44	23	90	39	9	40
child	6	14	88	39	34	4	14	77	28	22
other female-Sg	3	1	2	20	13	3	4	0	48	8
other female-Pl	5	15	3	17	21	8	6	5	11	49
other	1	0	0	0	2	0	0	0	1	1
no answer	0	1	0	2	0	0	1	0	0	0
proportion target	0.88	0.69	0.94	0.30	0.17	0.88	0.72	0.92	0.47	0.39
(SD)	(0.19)	(0.30)	(0.12)	(0.34)	(0.22)	(0.19)	(0.31)	(0.17)	(0.37)	(0.38)

For adults, there was a significant main effect of form ($F(1,22) = 20.63$, $p < 0.001$) and features ($F(4,88) = 31.24$, $p < 0.001$) as well as a significant interaction ($F(4,88) = 2.65$, $p = 0.039$). Paired-sample t-tests comparing individual agreement markers to clitics revealed that the proportion of target answers was significantly higher for 3rd person clitics relative to 3rd person agreement markers (singulars: $t(22) = 2.15$, $p = 0.04$; plurals: $t(22) = 3.27$, $p < 0.01$), but this contrast was not significant in either the 1st or 2nd person conditions (all $t < 1.74$, all $p > 0.09$). Paired-sample t-tests comparing 3rd person clitics and agreement to 1st and 2nd person clitics and agreement revealed that accuracy was significantly higher for 1st and 2nd person singular relative to 3rd person singular in both the agreement block (1S/3S: $t(22) = 4.83$, $p < 0.001$; 2S/3S: $t(22) = 6.55$, $p < 0.001$) and the clitic block (1S/3S: $t(22) = 4.41$, $p < 0.001$; 2S/3S: $t(22) = 4.70$, $p < 0.001$), and likewise for 1st person plural relative to 3rd person plural in both blocks (agreement: $t(22) = 6.14$, $p < 0.001$; clitics: $t(22) = 3.10$, $p < 0.01$).

For children, the same repeated-measures ANOVA revealed the same main effects of form ($F(1,41) = 9.03$, $p < 0.01$) and features ($F(4,164) = 97.41$, $p < 0.001$), plus an interaction ($F(4,164) = 4.84$, $p < 0.01$). As with adults, paired-sample t-tests revealed significantly higher accuracy for 3rd person clitics relative to 3rd person agreement markers (singulars: $t(41) = 2.49$, $p = 0.02$; plurals: $t(41) = 3.94$, $p < 0.001$) and no significant differences within the 1st or 2nd persons (all $t < 0.91$, all $p > 0.37$). Likewise, accuracy was significantly higher for 1st and 2nd person singular relative to 3rd person singular in both the agreement block (1S/3S: $t(41) = 9.63$, $p < 0.001$; 2S/3S: $t(41) = 11.47$, $p < 0.001$) and the

clitic block (1S/3S: $t(41) = 6.44$, $p < 0.001$; 2S/3S: $t(41) = 7.55$, $p < 0.001$), as well as for

1st person plural relative to 3rd person plural in both agreement ($t(41) = 9.86$, $p < 0.001$) and

clitic ($t(41) = 4.62$, $p < 0.001$) blocks.

In sum, for both children and adults, performance in the 1st and 2nd person agreement

conditions patterns with performance in the 1st and 2nd person clitic conditions, rather than

3rd person agreement conditions. That is, consistent with the semantic hypothesis,

comprehension of 1st and 2nd person agreement and accusative clitics is significantly higher

relative to the 3rd person, for both singular and plural members of these paradigms.

Meanwhile, we find only partial support for the phonological hypothesis. Accuracy in the

1st and 2nd person agreement conditions is not significantly different from accuracy in the

corresponding clitic conditions. Only in the 3rd person is there a significant difference

between clitics and agreement. The next two sections take a closer look at children's

comprehension of 1st and 2nd person agreement and the nature of children and adults' errors

in 3rd person agreement and clitic conditions.

5.1. Developmental analysis of 1st and 2nd person agreement

As a group, children in our sample appear to comprehend 1st person plural agreement

marker /-mos/, 1st person singular /-o/, and 2nd person singular /-s/. To further explore the

development of this ability, we used three separate linear regression models to test the

effect of age in months as a predictor of accuracy scores in each of these three conditions.

Next, we divided children into five age groups (2;3-2;11, $N = 7$; 3;0-3;9, $N = 10$; 4;1-4;11,

N = 13; 5;0-5;11, N = 7; 6;0-6;7, N = 5) and used single-sample t-tests to compare each group's mean accuracy score to chance in each condition. In the 1st plural agreement condition, the regression model revealed no significant effect of age month, and t-tests revealed significantly greater-than-than chance accuracy (20%) within every age group (all M > 61%, all t > 2.273, all p < 0.034). In the 1st and 2nd person singular agreement conditions, where chance is equal to 40%, the regression models again revealed no significant effect of age month, and t-tests again revealed greater-than-chance accuracy within every age group (all M > 80%, all t > 6.081, all p < 0.001). Thus, even in a supposedly demanding picture-selection task, we find evidence that children comprehend 1st and 2nd person agreement markers as early as 27-35 months—much younger than even the earliest reported comprehension for Spanish 3rd person agreement (41-50 months, González-Gómez 2017).

5.2. Error analysis of 3rd person agreement and clitics

To further explore children's comprehension of 3rd person agreement and clitics, we must first understand why adults produced so many non-target answers in this condition. The vast majority of adults' non-target responses involved a violation of the person feature (agreement block: 62 out of 67 total errors; clitic block: 38 out of 39 errors). This was also true for children's errors: 177 out of 190 total errors in the agreement block and 134 out of 143 total errors in the clitic block involved a person violation. Why would these

participants allow a 3rd person clitic or agreement marker to refer to a picture of themselves or the experimenter?

As noted in the introduction, 3rd person pronouns do not necessarily exclude reference to the speaker or hearer, and listeners interpret these pronouns by attending not only to which available referents match their formal features, but also by attending to which referents are most prominent in the physical and/or linguistic context. If prominence is a crucial factor in this experiment, then it may be that participants are calculating prominence, not within the display of each individual experimental item (where each referent is equally visually prominent) but across experimental items, where the most prominent referent is the one that was most recently mentioned, i.e., the person or persons depicted in the photo chosen in the immediately preceding filler trial. Participants may be willing to allow 3rd person agreement and clitics to pick out this highly prominent referent even if it includes the speaker and/or addressee—as long as it has the right number and gender properties. If participants calculate prominence in this way, we would expect them to repeat their response from the preceding filler trial whenever that response is compatible in number and/or gender with the agreement or clitic form being tested. In the rest of this section we test this hypothesis, first for adults and then for children.

Figure 1 shows the proportion of responses that match the response provided during the immediately preceding filler trial[5]. For the 3rd person condition, responses are subdivided

---

[5]This data is taken from the 14 experimental trials that were immediately preceded by a filler and excludes the 14 trials preceded by a distractor and the 2 trials appearing at the beginning of each block. (See section 4.2.)

depending on whether or not the picture chosen during the immediately preceding filler trial was compatible in number and/or gender with the 3rd person agreement marker or clitic being tested (black: compatible in number (and gender, if applicable); white: incompatible with either number or gender or both). For the agreement block, only number features are relevant because agreement markers do not carry gender; however, for the clitic block both number and gender are relevant. For example, if the preceding filler response was the photo of the experimenter and the participant together, and if the form being tested was the 3rd plural feminine clitic *las*, then this photo would provide a prominent, number-compatible referent for the clitic, but it would only be compatible in gender if the participant herself was female. (Recall that all experimenters were female, so only the gender of the participant him/herself could influence gender compatibility.) As can be seen from the figure, adults never repeated their response from the preceding filler trial if that response clashed with either the number or gender features of the 3rd person expression being tested.

Figure 1 also includes the proportion of matching responses in 1st and 2nd person conditions (gray), which serves as a baseline. Because 1st and 2nd person forms do not need to rely on the preceding discourse to select a referent, such repeat responses should be due to sheer coincidence. Indeed, one-sample t-tests show that the proportion of responses matching the preceding filler response in 1st and 2nd person conditions was not different from chance (1/5, or 20%) in either block, for either adults or children (2-sided, all $p > 0.08$). Therefore, we use participants' proportion of matching responses in 1st and 2nd

person conditions as a baseline against which to compare their proportion of matching responses in the 3rd person conditions.
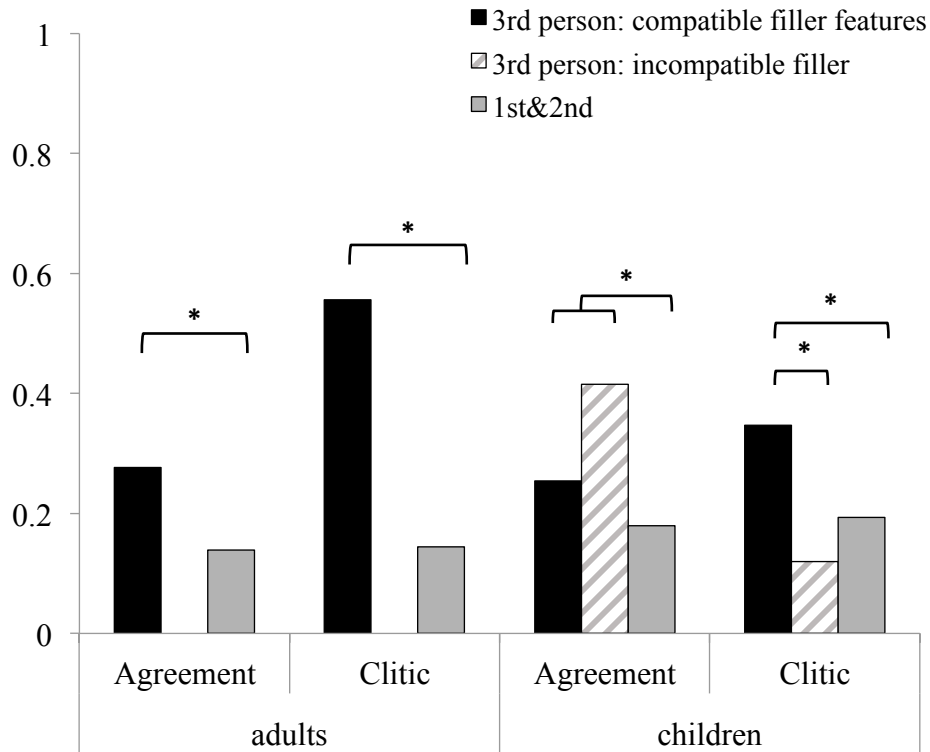
Figure 1. Proportion of responses that simply repeat the response given in the immediately preceding filler trial. (i) Black bars represent 3rd person conditions in which the preceding response happened to match the number and/or gender features of the clitic or agreement marker being tested. (ii) White bars represent 3rd person conditions in which this response happened to clash with the number and/or gender features. (iii) Gray bars 1st & 2nd person conditions, which serve as a baseline.

If adults use both grammatical features and discourse prominence to locate the antecedent of 3rd person null subjects and object clitics, then we would expect them to repeat their immediately preceding response more often than baseline (i.e., more often than

occurred in 1st and 2nd person conditions), but only when that response is compatible with the number and/or gender features of the clitic or agreement form being tested. To test this, we ran two analyses. First, we used one-sided chi-squared tests of proportion to compare the frequency of repeated responses in 3rd person conditions with number and gender-compatible preceding responses (black bars) to the frequency of repeated responses in 1st and 2nd person conditions (baseline: gray bars). This revealed a significant difference in both the agreement block ($M1 = 0.28$, $M2 = 0.14$, $\chi(1) = 2.77$, $p = 0.048$) and the clitic block ($M1 = 0.56$, $M2 = 0.15$, $\chi(1) = 18.83$, $p < 0.001$). Second, we compared the frequency of repeated responses when the preceding response was number- and gender-compatible (black bars) to cases when the preceding response happened to be incompatible (white bars). This difference was highly significant (both $\chi > 6.9$, both $p < 0.01$), because in the latter case adults never once repeated that response.

Next, we ran the same two analyses on child responses to test whether they too rely on both grammatical features and discourse prominence. In the clitic block, children showed the same pattern as adults: when the response from the preceding filler was compatible in number and gender with the 3rd person form being tested, then children chose this photo significantly more often relative to baseline ($M1 = 0.40$, $M2 = 0.19$, $\chi(1) = 6.44$, $p < 0.01$), and significantly more often relative to when it was incompatible with those number and/or gender features ($M1 = 0.40$, $M2 = 0.11$, $\chi(1) = 9.77$, $p < 0.001$). Thus, in the clitic block, we have evidence that children, like adults, are sensitive to both grammatical features and discourse salience when locating the referent of a 3rd person accusative clitic.

In the agreement block, however, children showed only part of the adult pattern: the effect of feature compatibility disappeared. Here, children were no more likely to repeat the preceding filler response when it was compatible in number with the form being tested than when it was incompatible ($M1 = 0.25$, $M2 = 0.42$, $\chi(1) = 2.82$, $p = 0.95$). Collapsing across these two cases, we found that children were more likely *overall* to repeat the preceding filler response in 3rd person agreement conditions, relative to baseline ($M1 = 0.33$, $M2 = 0.18$, $\chi(1) = 6.82$, $p < 0.01$). Thus, in the agreement block we have evidence that children rely on discourse prominence but not number marking when locating the referent of a null subject accompanied by 3rd person agreement.

In sum, adults tend to choose the referent of a 3rd person clitic or agreement marker by looking to the most recently selected photo compatible with its number and/or gender features. Children also look to the most recently selected photo for a referent, but they are more willing than adults to accept an antecedent with incompatible number and/or gender properties when it comes to interpreting 3rd person agreement markers.

6. Discussion

Early investigations into the comprehension of subject-verb agreement suggested that preschoolers do not comprehend the agreement markers that they themselves produce (Gxilishe et al. 2009, Johnson et al. 2005, Pérez-Leroux 2005, Rastegar et al., 2010). Our results show that this is actually not the case for all agreement markers: Mexican children as young as 2;3-2;11 reliably link 1st singular /-o/, 2nd singular /-s/ and 1st plural /-mos/ to a

photo depicting the speaker, themselves, and both people, respectively. What's more, their overall level of accuracy is statistically indistinguishable from the level of accuracy they achieve when interpreting accusative clitics. Whatever the reason for children's apparent inability to comprehend 3rd person singular and plural agreement markers in those original picture-selection studies, it is not necessarily due to their status as agreement markers *per se*.

What could be the reason for children's difficulty interpreting 3rd person agreement, then? This study offers a new perspective on the two explanations that researchers have offered up to now. First, some researchers have pointed out that picture selection imposes greater processing demands than spontaneous production (Brandt-Kobele and Höhle 2010, Legendre et al. 2014, Verhagen & Blom 2014, González-Gómez et al. 2017). What this study shows is that even in an unusually demanding version of the picture selection task—where the choice is between not two but five different pictures—children still demonstrate early comprehension of 1st and 2nd person agreement, suggesting that the choice of experimental task is not the sole factor driving children's performance. Low accuracy scores for 3rd person agreement cannot be entirely blamed on the picture selection paradigm.

Second, other researchers have pointed out that the comprehension delay does not apply to all languages equally (Legendre et al. 2010a, 2014) and have suggested that morphophonological properties like salience could be crucial to the early comprehension of agreement. This study does not find strong evidence that salience plays a role in children's

comprehension of Spanish agreement markers. What it does suggest, however, is the importance of the preceding discourse to the interpretation of both 3rd person agreement and clitics. Similar to previously reported picture-selection studies, children in this study failed to reliably link 3rd person singular and plural agreement markers to a photo of the target referent. However, they also had trouble linking 3rd person clitics to the target referent, and even more importantly, adults did not always link 3rd person agreement and clitics to the target referent. Instead, both adults and children allowed these 3rd person expressions to pick out referents from the preceding discourse—that is, the referent that had been selected in the immediately preceding filler trial—even if that referent was the speaker or addressee. This shows that children are aware that 3rd person expressions are discourse-dependent in a way that 1st and 2nd person are not, and it raises the possibility that children's low accuracy scores reported in previous picture selection tasks may have something to do with uncertainty over the discourse prominence of the pictured referents. However, since not all of these studies report their fillers and none report adult responses, this suggestion remains to be tested. Future work can test this by explicitly manipulating the content of the preceding discourse and by consistently including adult participants.

Children's interpretation of 3rd person expressions did differ from that of adults in one respect: unlike adults, children allowed 3rd person singular and plural agreement markers to pick out a recently-mentioned discourse referent even if that referent had the wrong number properties. There are two possible interpretations of this result. First, it may be that children really do not comprehend the number features of 3rd person singular and plural agreement

markers, as suggested by the original picture-selection studies. Alternately, however, it may be that children simply prioritize the need for a discourse-recent referent over the need for a number-compatible referent. One reason they may do so is that highly reduced referring expressions (such as null subjects) exhibit an even stronger association with discourse-prominent referents than less reduced expressions (such as clitics) across many languages (Ariel 1988, Gundel et al. 1993).

If this second alternative is correct, it would certainly not be the first time that pragmatic considerations have affected children's deployment of a grammatical constraint. Children have widely been observed to violate Principle B in their interpretations of object pronouns (Chien & Wexler 1990, Thornton & Wexler 1999; see Elbourne 2005 for review). Yet when pragmatic conditions are altered such that the grammatically illicit referent is either less discourse-prominent or no longer part of the QUD, performance becomes adult-like (Conroy et al. 2009, Spenader, Smits & Hendricks 2009).

Summarizing, this paper expands the body of available evidence on children's acquisition of agreement by examining their comprehension of agreement in all three persons, systematically comparing this to their comprehension of clitics. The picture that emerges is one of early, accurate comprehension of 1st and 2nd person agreement and an awareness that 3rd person agreement, like other 3rd person expressions, is sensitive to the properties of the surrounding discourse. This does not definitively solve the mystery of why preschoolers fail to use 3rd person singular and plural agreement markers in some comprehension tasks. But it does suggest that the morphophonological properties of

agreement and the cognitive demands of picture selection are not the whole story. Rather, the problem may lie in the discourse setup of these tasks, which allows children and adults to oscillate between choosing the referent of a 3rd person referring expression from the visual display of the current experimental trial and choosing from the context of the experiment as a whole. If that is the case, future work will have to tease apart the conditions under which children and adults choose one source of information over the other.

Whatever the experimental method, future investigations into the acquisition of agreement morphology must take into account the complexity of interpreting 3rd person referring expressions. As more and more evidence emerges that young children are sensitive to a range of discourse cues (Song & Fisher 2005, Pyykkonen et al. 2010, Hughes & Allen 2013, Hartshorne et al. 2015, and others), researchers interested in children's comprehension of the formal properties of agreement must be aware that children may also be sensitive to the discourse properties of these same expressions.

References

Alexiadou, A. and Anagnostopoulou, E. (1998). Parametrizing Agr: Word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory*, 16:491–539.

Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition*, *76*(1), B13-B26.

Ariel, M.(1988). Referring and accessibility. *Journal of Linguistics*, 24(1):65– 87.

Barbosa, M. D. P. P. (2009). Two kinds of subject pro. *Studia Linguistica*, 63(1):2–58.

Benki, J.R.(2003). Analysis of English nonsense syllable recognition in noise. *Phonetica*, 60(2):129–157.

Benveniste, E. (1971). Problems in general linguistics, volume 8. University of Miami Press.

Brandt-Kobele, O.-C. and Höhle, B. (2010). What asymmetries within comprehension reveal about asymmetries between comprehension and production: The case of verb inflection in language acquisition. *Lingua*, 120(8):1910– 1925.

Brener, R.(1983). Learning the deictic meaning of personal pronouns. *Journal of Psycholinguistics*, 12(3):235–262.

Chien, Y.-C. and Wexler, K. (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition*, 1(2):225–295.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Conroy, A., Takahashi, E., Lidz, J., and Phillips, C. (2009). Equal treatment for all antecedents: How children succeed with Principle B. *Linguistic Inquiry*, 40(3):446–486.

Corbett, G. (2006). *Agreement*, volume 109. Cambridge University Press.

Davies, B., Rattanasone, N. X., and Demuth, K. (2016). Two-year-olds' sensitivity to inflectional plural morphology: Allomorphic effects. *Language Learning and Development*. 13(1): 38-53.

De Cat, C. (2002). *French Dislocation*. PhD thesis, University of York.

Elbourne, P. (2005). On the acquisition of Principle B. *Linguistic Inquiry*, 36(3):333–365.

González-Gómez, N., Hsin, L., Barriere, I., Nazzi, T., and Legendre, G. (2017). Agarra, agarran: Evidence of early comprehension of subject-verb agreement in Spanish. *Journal of Experimental Child Psychology*, 160:33–49.

Gundel, J., Hedberg, N., and Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69(2):274–307.

Gxilishe,S., Smouse, M.R., Xhalisa,T., and de Villiers, J. (2009). Children's insensitivity to information from the target of agreement: The case of Xhosa. In Crawford, J., Otaki, K., and Takahashi, M., editors, Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008), pages 46–53, Somerville, MA. Cascadilla Proceedings Project.

Harley, H. and Ritter, E. (2002). Person and number in pronouns: A feature- geometric analysis. *Language*, 78(3):482–526.

Hartshorne, J. K., Nappa, R., and Snedeker, J. (2015). Development of the first-mention bias. *Journal of Child Language*, 42(2):423–446.

Heim, I. (1991). Artikel und definitheit. In von Stechow, A. and Wunderlich, D., editors, *Semantik: Ein internationales Handbuch der zeitgen Aussischen Forschung*, pages 487–535. de Gruyter, Berlin.

Hughes, M. and Allen, S. (2013). The effect of individual discourse- pragmatic features on referential choice in child English. *Journal of Pragmatics*, 56:15–30.

Johnson, V., de Villiers, J., and Seymore, H. (2005). Agreement without understanding? The case of third person singular /s/. *First Language*, 25(3):317–330.

Legendre, G., Barriere, I., Goyet, L., and Nazzi, T. (2010a). Comprehension of infrequent subject-verb agreement forms: Evidence from French-learning children. *Child Development*, 81(6):1859–1875.

Legendre, G., Culbertson, J., Barriere, I., Nazzi, T., and Goyet, L. (2010b). Experimental and empirical evidence for the status and acquisition of subject clitics and agreement marking in adult and child spoken French. In Torrens, V., Escobar, L., Gavarró, A., and Gutiérrez, J., editors, *Movement and Clitics: Adult and Child Grammar*, chapter 15, pages 333–360. Cambridge Scholars Publishing, Newcastle.

Legendre, G., Culbertson, J., Zaroukian, E., Hsin, L., Barrière, I., and Nazzi, T. (2014). Is children's comprehension of subject-verb agreement universally late? Comparative evidence from French, English, and Spanish. *Lingua*. http://dx.doi.org/10.1016/j.lingua.2013.05.004.

Legendre, G. and Smolensky, P. (2012). On the asymmetrical difficulty of acquiring person reference in French: production versus comprehension. *Journal of Logic, Language and Information*, 21(1):7–30.

Miller, K. and Schmitt, C. (2014). Spanish-speaking children's use of verbal inflection in comprehension. *Lingua*, 144:40–57.

Nevins, A. (2007). The representation of third person and its consequences for person-case effects. *Natural Language & Linguistic Theory*, 25(2):273–313.

Pérez-Leroux, A. T. (2005). Number problems in children. In Gurski, C., editor, *Proceedings of the 2005 Canadian Linguistics Association Annual Conference*, pages 1–12.

Phillips, C. (1995). Syntax at age two: Cross-linguistic differences. *MIT working papers in linguistics*, 26:325–382.

Polišenska, D. (2010). *Dutch children's Acquisition of verbal and adjectival inflection*. PhD thesis, University of Amsterdam, Center for Language and Communication.

Pyykkönen, P., Matthews, D., and Järvikivi, J. (2010). Three-year-olds are sensitive to semantic prominence during online language comprehension: A visual world study of pronoun resolution. *Language and Cognitive Processes*, 25(1):115–129.

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rastegar, Z., Shirazi, H., and Sadighi, F. (2012). An amazing conundrum in children's comprehension and production of verb inflection. *World Applied Sciences Journal*, 18(8):1095–1101.

Redford, M. A. and Diehl, R. L. (1999). The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *The Journal of the Acoustical Society of America*, 106(3):1555–1565.

Roberts, I. G. (2010). *Agreement and head movement: Clitics, incorporation, and defective goals*, volume 59. MIT Press.

Sauerland, U. (2003). On the semantic markedness of phi-features. In Harbour, D., Adger, D., and Bejar, S., editors, *Phi-Theory: Phi-Features Across Modules and Interfaces*, chapter 3, pages 57–82.

Song, H.-J. and Fisher, C. (2005). Who's "she?": Discourse prominence influences preschooler's comprehension of pronouns. *Journal of Memory and Language*, 52(1):29–57.

Spenader, J. , Smits, E.-J., and Hendriks, P. (2009). Coherent discourse solves the pronoun interpretation problem. *Journal of Child Language*, 36:23–52.

Thornton, R. and Wexler, K. (1999). *Principle B, VP ellipsis, and interpretation in child grammar*. MIT press.

Verhagen, J. and Blom, E. (2014). Asymmetries in the acquisition of subject-verb agreement in Dutch? Evidence from comprehension and production. *First Language*, 34(4):315–335.