Learning speaker- and addressee-centered demonstratives in Ticuna

Amalia Horan SKILTON

1University of Texas at Austin, USA

Address for Correspondence: amalia.skilton@austin.utexas.edu
Department of Linguistics
University of Texas at Austin
305 E. 23rd Street STOP B5100
Austin, TX 78712

Acknowledgments: I completed some of this research as a guest of Caroline Rowland at the Language Development Department at the Max Planck Institute for Psycholinguistics and the Baby & Child Research Center, Radboud University Nijmegen. I thank Marisa Casillas for comments on a previous version of this manuscript.

Competing interests: The author(s) declare none.

Keywords: Demonstratives; Deixis; Language of Space; Language Diversity.
Title:
Learning speaker- and addressee-centered demonstratives in Ticuna

Abstract:
English-acquiring children produce demonstrative words (e.g., this/that, here/there) very early in development – but do not display adult-like production or comprehension of the items until very late. We ask whether this pattern of early emergence, but late mastery, also holds in Ticuna, an Amazonian language with an unusually complex demonstrative system. It does not. In a cross-sectional, observational study of 45 children aged 1;0 to 4;11, Ticuna children displayed early emergence only for speaker-centered demonstratives (those which relate the demonstrative referent to the speaker), not for addressee-centered ones (which relate the referent to the addressee). The late emergence of addressee-centered demonstratives cannot arise from input frequency; rather, the Ticuna addressee-centered demonstratives emerge late because children, due to their bias toward egocentric spatial cognition, struggle to comprehend them. This result suggests that children’s cognitive biases, such as the bias toward egocentrism, can outweigh frequency effects in demonstrative acquisition.

Keywords: Demonstratives; Deixis; Language of Space; Language Diversity

Word Count: 9,985 (excludes abstract and references)
1. Introduction

Demonstratives, such as *this*/*that* and *here*/*there*, behave paradoxically in language acquisition – displaying early emergence, but late mastery (Clark & Sengul 1978; Tanz 1980; Küntay & Özyürek 2006, a.o.). English-acquiring children often produce demonstratives during the one-word stage, when other function words are barely present (Clark 1978; Gonzalez-Peña et al. 2020). At this stage, the items are extraordinarily frequent, representing up to 7% of all word tokens (Diessel 2006: 483; Gonzalez-Peña et al. 2020: 5). Somewhat later in development, combinations of demonstratives and deictic gestures pave the way for two-word utterances (Iverson & Goldin-Meadow 2005), and demonstrative-noun combinations represent a prominent two-word utterance type (Capirci et al. 1996: 663).

But children’s early production is not adult-like. Rather, young children often use proximal (*here*) and distal (*there*) demonstratives interchangeably (Clark 1978: 102-104; Clark & Sengul 1978: 471). And challenges with demonstratives persist far into development: English-learning children may not display adult-like comprehension of the items until six to seven years (Tanz 1980; Webb & Abrahamson 1976; Clark & Sengul 1978). This same pattern of early emergence, but late mastery, is documented not only for the demonstratives of English, but also for Spanish (Rodrigo et al. 2004), Turkish (Küntay & Özyürek 2006), Mandarin (Chu & Minai 2018), and Yucatec Maya (Espinosa Ochoa 2009), and in diary data for Bulgarian, Dutch, German, and Japanese (Clark 1978: 95).

To explain the cross-linguistic early emergence of demonstratives, authors invoke *frequency*. Children produce demonstratives early because the items are both exceptionally frequent in caregivers’ speech (Gonzalez-Peña et al. 2020: 7) and centrally involved in joint attention (Diessel 2006: 472). To explain late mastery, on the other hand, authors invoke a conflict between demonstrative semantics and children’s cognitive bias toward spatial *egocentrism* (Clark 1978: 89; Küntay & Özyürek 2006: 318). Semantically, demonstratives are *deictic*, picking out different referents when produced by different speakers: for example, if A stands at the north end of a room and B at the south, A’s tokens of *here* index the north end, but B’s the south. Thus, adult-like production and comprehension of demonstratives require the ability to construe demonstrative referents from others’ spatial perspectives. Because of children’s bias toward egocentric spatial cognition, they struggle with this perspective-shifting operation – tending to construe spatial arrays only from their own perspective (Clark & Sengul 1978). This difficulty with perspective-shifting, authors argue, directly causes non-adult-like comprehension of demonstratives, and indirectly causes non-adult-like production.

Thus, the dominant hypothesis is that frequency drives the early emergence of demonstratives, while the items’ deictic semantics (because of egocentrism) cause late mastery. Both high frequency and deictic semantics are robust properties of demonstratives across languages (Diessel 2006; Hanks 2011; Levinson 2004). As a consequence, this account suggests that the early emergence-late mastery pattern will be close to universal across languages, appearing for any demonstrative that is both very frequent and semantically deictic.

Against this background, we investigated the L1 acquisition of demonstratives in Ticuna, an Indigenous Amazonian language with four deictic demonstratives. In a cross-sectional study of
Learning speaker- and addressee-centered demonstratives in Ticuna

45 Ticuna-acquiring children (1;0-4;11), we discovered that only some of the language’s demonstratives followed the early emergence-late mastery pattern seen in English, Mandarin, and Turkish. Specifically, the subset of Ticuna demonstratives which relate referents to the speaker only – equivalent to English this near me – displayed the expected early emergence-late mastery pattern. Children began to produce these demonstratives during the one-word stage, but did not attain adult-like use of the items within the age range of the study. On the other hand, Ticuna demonstratives which relate referents to the addressee – equivalent to that near you – displayed late emergence, appearing only after 3;0, as well as late mastery. This finding indicates that children’s cognitive egocentrism can affect when demonstratives emerge, as well as when they are mastered.

Before turning to the specifics of this study, we first present background information about cross-linguistic diversity in demonstrative meaning and acquisition, as well as about the demonstrative system of Ticuna.

2. Background on Demonstratives

2.1. Diversity in demonstrative meaning

To describe demonstrative semantics, we employ the concepts of deictic content and origo. The deictic content of a demonstrative is the information which it conveys about the referent; the origo of a demonstrative is the discourse participant(s) to whom it relates the referent (Bühler 1934). For example, on a traditional analysis of the English demonstratives, that conveys that the demonstrative referent is far from the speaker. Thus, the deictic content of that is spatial, conveying ‘far from origo,’ and the origo is the speaker.

The deictic content and origo of demonstratives vary across languages. For example, the origo of a demonstrative may be the speaker, the addressee, or the dyad of both discourse participants (Levinson 2004: 109; Hanks 2011: 331-334). Likewise, the deictic content of demonstratives can concern many different properties of the referent, such as location in space, perceptual accessibility, or status as an object of joint attention (Levinson 2004: 117; Hanks 2011: 329).

The demonstrative system of Turkish (Küntay & Özyürek 2006: 307) provides examples of variation, in deictic content and origo, among the demonstratives of a single language. Turkish has three demonstratives: bu, o, and şu. Bu conveys that the referent is close to the speaker, while o conveys that the referent is far from the speaker. Thus, the deictic content of both bu and o is spatial, and their origo is the speaker. In contrast, the third demonstrative, şu, conveys that the speaker is calling new joint attention to the referent – that is, that the addressee is currently disattending the referent. Thus, the deictic content of şu is attentional (conveying ‘origo not attending’) and the origo is the addressee.

Because şu relates referents to the addressee, speakers must monitor their addressees’ attention states in order to use it appropriately (Küntay & Özyürek 2006: 304). This illustrates a larger generalization about the relationship between egocentrism and variation in the demonstrative origo. Appropriate use of speaker-centered demonstratives (those which have the speaker as their
(origo) requires monitoring only ego’s relation to the demonstrative referent. In contrast, appropriate use of addressee-centered demonstratives (those with addressee as the origo) requires monitoring the addressee’s relation as well.

2.2. Egocentrism and demonstrative acquisition

Egocentrism impacts children’s production and comprehension of demonstratives in two ways. The first impact of egocentrism relates to the comprehension of speaker-centered demonstratives. It is apparent in contexts where the child and the speaker have different spatial relations to the demonstrative referent: for example, where the referent is far from the child, but close to the speaker. In contexts of this type, English- and Mandarin-acquiring children aged 3 to 6 years tend to construe the speaker’s proximal demonstratives as indexing referents that are near the child, but far from the speaker. Conversely, children of this age construe distal demonstratives in these contexts as indexing referents that are far from the child, but close to the speaker (Clark & Sengul 1978, Chu & Minai 2018). Thus, children aged 3 to 6 years tend to represent themselves, rather than other speakers, as the origo of all speaker-centered demonstratives. According to Clark & Sengul (1978) and Chu & Minai (2018), it is this egocentric tendency which drives children’s late mastery of demonstratives in comprehension.

The second impact of egocentrism relates to the production of addressee-centered demonstratives. Turkish-acquiring children aged 4 to 6 years produce addressee-centered şu much less frequently than adults (Küntay and Özyürek 2006: 313). Furthermore, while Turkish adults’ use of şu is entirely addressee-centered – displaying effects of the addressee’s visual attention, but no effects of the referent’s location relative to the speaker – children’s use of the item is not; they instead use şu as a speaker-proximal demonstrative equivalent to bu (Küntay & Özyürek 2006: 315). Both of these forms of late mastery are attributed to children’s difficulty with monitoring the addressee’s attention state, a form of (not strictly spatial) egocentrism (Küntay & Özyürek 2006: 318).

While the studies discussed here provide extensive evidence that egocentrism inhibits children’s mastery of demonstratives, they focused exclusively on children aged 3;0 and above. Thus, it is still unknown whether egocentrism also affects the emergence of demonstratives in children under 3;0.

3. Language background

3.1. Location, classification, and speaker population

Ticuna is a language isolate spoken in northwestern Amazonia. Most speakers live along the main course of the Amazon/Solimões river in northeastern Peru, southern Colombia, and western Brazil. Between 38,860 and 69,000 people speak Ticuna as a first language; it is the most widely spoken Indigenous language of Brazil (Lewis et al. 2014; ISA 2017). The language is still acquired by children in most communities in Peru and Brazil, but not in Colombia (Santos 2004).
3.2. Field site

Data discussed in this paper comes from our own fieldwork in the town of Cushillococha, located in Mariscal Ramon Castilla district, Loreto region, Peru. The author has conducted research with Ticuna speakers in Cushillococha since 2015, with a total of ~13 months of fieldwork.

Cushillococha is a land-titled Indigenous community with ~5,000 residents, most of whom work as independent farmers. Most households also participate in the cash economy, either by marketing produce in the nearby provincial capital (population ~12,000; located 8km away), or by working for wages in the government, construction and transportation sectors.

At this time, effectively everyone in Cushillococha speaks Ticuna natively. Most residents use Ticuna as their dominant language in all domains of life, including official domains like church and school. Education from preschool to high school is offered in the town and is under local control. Classes are taught monolingually in Ticuna for children aged 3 to 8 years; afterward, Spanish is taught as a second language. Most adults and children aged over 8 years know some Spanish, with abilities ranging from limited passive knowledge to L2 fluency.

3.3. Typological characteristics of Ticuna

Two typological properties of Ticuna are relevant to the data in this paper. First, the language displays lexical and grammatical tone. The variety of Ticuna spoken in Cushillococha has eight lexical tones on monosyllables (Anderson 1959, cf. Montes Rodriguez 1995). Transcriptions use IPA (v. 2020) and represent tones with raised numerals; 1 is the lowest tone.

Second, Ticuna nouns are exhaustively divided into five noun classes, primarily based on semantic principles such as animacy. Most noun phrase constituents, including demonstratives, agree in noun class with the head of their phrase.

3.4. Demonstrative system of adult Ticuna

Adult Ticuna has six nominal demonstratives, which are equivalent to English this/that, and six locative demonstratives, which are equivalent to here/there. Of the six demonstratives in each syntactic category, four – the same four in both paradigms – have productive deictic uses. This article discusses only the four productive, deictic items.

Nominal demonstratives (this/that) appear as arguments and modifiers. They display noun class agreement. Locative demonstratives (here/there) appear only as adjuncts. They do not display noun class agreement, but do display case marking, conveyed by tone. We call these items ‘locative demonstratives,’ rather than using the more common term ‘demonstrative adverbs,’ because they pattern syntactically with nouns, not adverbs.

Table 1 shows the forms of the four deictic demonstratives, with a gloss and English paraphrase for each item. Nominal demonstratives appear in the upper portion of the table, locative demonstratives in the lower portion. Glosses and paraphrases in Table 1 reflect the analysis of
the adult demonstrative system in [Redacted for anonymous review], which is based on controlled production tasks with adults and recordings of informal conversation.

### Table 1. Forms of the Ticuna nominal and locative demonstratives.

<table>
<thead>
<tr>
<th>Nominal Demonstrative</th>
<th>Noun Class Agreement Forms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gloss</strong></td>
<td><strong>Paraphrase</strong></td>
</tr>
<tr>
<td>Speaker-Proximal</td>
<td>‘this near me’</td>
</tr>
<tr>
<td>Speaker-Distal</td>
<td>‘that far from me’</td>
</tr>
<tr>
<td>Dyad-Proximal¹</td>
<td>‘this between us’</td>
</tr>
<tr>
<td>Addressee-Proximal</td>
<td>‘that near you’</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Locative Demonstrative</th>
<th>Allative Case</th>
<th>Locative Case</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gloss</strong></td>
<td><strong>Paraphrase</strong></td>
<td><strong>nu²a²</strong></td>
</tr>
<tr>
<td>Speaker-Proximal</td>
<td>‘here near me’</td>
<td>je²a²</td>
</tr>
<tr>
<td>Speaker-Distal</td>
<td>‘there far from me’</td>
<td>je³a²</td>
</tr>
<tr>
<td>Dyad-Proximal</td>
<td>‘here between us’</td>
<td>je³a²</td>
</tr>
<tr>
<td>Addressee-Proximal</td>
<td>‘there near you’</td>
<td>je³ma²</td>
</tr>
</tbody>
</table>

Of the demonstratives in Table 1, only the Addressee-Proximal items (that/there near you) can also be used as anaphors, analogous to English anaphoric that and there. The other deictic demonstratives do not allow anaphoric uses.

### 3.5. Frequency of demonstratives

Between 2017 and 2018, we collected a corpus of informal conversation between adults in Ticuna, following the methods outlined in Rossi et al. (2020). This corpus contains 1h49m of transcribed conversation drawn from 8 interactions (mean time per interaction = 13m 38s, SD = 7m 37s). [Redacted], an L1 Ticuna speaker, approved the transcription of the corpus and provided a Spanish translation of every turn.

The adult conversational corpus contains 2,360 turns of adult-directed speech (ADS), defined as talk produced by adults and addressed to other adults. 2,224 turns (94.1%) were at least partially intelligible and contained at least one Ticuna word. Table 2 reports the frequency of each demonstrative in these 2,224 turns of conversational ADS. Because of the extremely high frequency of all demonstratives, frequencies are reported per 100 words. Means and standard deviations of frequency are calculated by interaction, not by speaker.

---

¹ The Dyad-Proximal and Addressee-Proximal are phonologically similar, but behave morphologically as monomorphemic – affixes cannot intervene between the first and second syllables. Thus, synchronically they are not derived from a shared root.
Looking first to the nominal demonstratives (upper portion of Table 2), we see that the Speaker-Proximal (*this near me*) and Addressee-Proximal (*that near you*) are both exceptionally frequent in ADS; on average, each represents **2.06-2.07% of all ADS word tokens**. Their frequencies are statistically indistinguishable ($W = 36, p = 0.7209$). The Speaker-Distal demonstrative (*that far from me*) is significantly less frequent than the Speaker-Proximal ($W = 64, p = 1.55e^{-4}$) or Addressee-Proximal ($W = 63, p = 3.11e^{-4}$); in turn, the Dyad-Proximal (*this between us*) is significantly less frequent than the Speaker-Distal ($W = 53, p = 0.031$).

Turning to the locative demonstratives (lower portion of Table 2), the Speaker-Proximal (*here near me*), Addressee-Proximal (*there near you*), and Speaker-Distal (*there far from me*) terms are all indistinguishable in ADS frequency (in an ANOVA, $F(2,21) = 0.83, p = 0.45$). On average, each of these demonstratives represents **1.08-1.58% of ADS word tokens**. However, the Dyad-Proximal locative demonstrative (*here between us*) is significantly less frequent than any of the other locative types (compared to the Speaker-Proximal, $W = 61.5, p = 0.0023$).

### 4. Predictions

In combination with the language-specific facts just discussed, the theories discussed in §2 yield several predictions about Ticuna children’s acquisition of demonstratives.

To the extent that cognitive egocentrism is the central factor in Ticuna-learning children’s acquisition of demonstratives, we predict that they will acquire the language’s egocentric demonstratives before the non-egocentric ones. Thus, the Speaker-Proximal and Speaker-Distal will be learned before the Addressee-Proximal and Dyad-Proximal. Additionally, to the extent that Ticuna children share the proximity bias documented for children acquiring other languages – such as English (Tanz 1980), Spanish (Rodrigo et al. 2004), and Yucatec Maya (Espinosa...
Ochoa 2009: 100-104) – we predict that they will acquire Speaker-Proximal demonstratives before Speaker-Distal demonstratives. Thus, children are expected to acquire demonstratives in the order Speaker-Proximal > Speaker-Distal > {Addressee-Proximal, Dyad-Proximal}. Theories that foreground cognitive biases treat demonstrative semantics as controlling the order of acquisition. The nominal and locative demonstratives have near-identical semantics; therefore, our predicted order of acquisition is identical for nominal and locative demonstratives.

In contrast, if cognitive biases do not substantially influence the order of acquisition of demonstratives, we expect that the items – like other function words (Ambridge et al. 2015: 243-248; Lieven 2010; Rowland et al. 2003) – will be acquired roughly in order of word frequency: children will produce and master more frequent demonstratives before less frequent ones. Because the ADS frequency ranking of demonstratives is different between syntactic categories (§3.5), frequency yields different predictions for nominal and locative demonstratives. To the extent that ADS and CDS frequency are the same, children are predicted to learn nominal demonstratives in the order {Speaker-Proximal, Addressee-Proximal} > Speaker-Distal > Dyad-Proximal, and locative demonstratives in the order {Speaker-Proximal, Addressee-Proximal, Speaker-Distal} > Dyad-Proximal.

Where these competing sets of predictions differ is in the predicted behavior of Addressee-Proximal demonstratives (paraphrased above as that/there near you). The Ticuna Addressee-Proximals are very frequent – in ADS, as frequent as Speaker-Proximals – but using them correctly requires non-egocentric construals of the demonstrative referent. Therefore, theories that foreground cognitive egocentrism predict that Addressee-Proximal demonstratives will be learned later than comparable items (i.e., after speaker-centered demonstratives), while theories that foreground frequency predict that they will be learned earlier (i.e., at the same time as speaker-centered demonstratives).

Our predictions from cognitive biases are based on studies which examined late mastery (Clark & Sengul 1978, Tanz 1980, Küntay & Özyürek 2006, a.o.) rather than early emergence. Consequently, the predicted orders given above concern the order of mastery, not of first appearance. However, as we are not aware of any studies on the order of emergence of demonstratives, we assume the same predictions for order of emergence as for order of mastery.

Concerning the absolute age of emergence and mastery, children acquiring languages with two or three demonstrative terms, such as English or Spanish, typically produce all of the terms by 3;0 (Gonzalez-Peña et al. 2020). For age of emergence, we therefore predict that Ticuna children will produce all four demonstrative types before 3;0. For age of mastery, we predict that Ticuna-acquiring children will pattern like English-, Mandarin- and Turkish-acquiring children (§1): they will not display adult-like use of any demonstrative until ~6 years.

Qualitatively, we refer to demonstratives which emerge before 3;0 or are mastered before 6;0 as displaying ‘early’ emergence or mastery, and demonstratives which emerge after 3;0 or are mastered after 6;0 as displaying ‘late’ emergence or mastery. ‘Early’ emergence and ‘late’ mastery are predicted by the literature; neither late emergence nor early mastery is predicted for any demonstrative type.
5. Methods

5.1. Participants

We recruited 46 child participants, born to 40 mothers, from families residing in Cushillococha. Participants were aged between 1;0 and 4;11 (mean age = 2;10.12, SD age = 14 months 28 days). We recruited only children whose primary caregivers spoke Ticuna as an L1; who were acquiring Ticuna as an L1; and who appeared to be typically developing. 12 participants were siblings. 10 further participants lived in the same household but were not siblings.

By caregiver report, 14 of the 46 participants were or had been regularly exposed to Spanish at home. We observed 2 more participants speaking in Spanish on the study recordings or during home visits. The other 29 participants appeared to be monolingual in Ticuna. With this high prevalence of monolingualism, the participants were relatively similar to other Latin American Indigenous children represented in the acquisition literature (e.g., the Q’anjob’al- and Tseltal-acquiring children studied by Mateo Pedro [2010] and Casillas et al. [2020]), and relatively dissimilar to North American Indigenous children (e.g., the Navajo-English bilingual children studied by Chee [2017]).

The first author, who is a proficient L2 speaker of both Ticuna and Spanish, conducted recruitment and all study procedures. The researcher communicated with families monolingually in Ticuna, using Spanish only with non-Ticuna family members. Study procedures were approved by the Institutional Review Board of the University of Texas at Austin.

5.2. Procedures

Children completed three procedures: a daylong audio recording, an object play session, and a free play session. Only object play and free play data are analyzed in this study. Procedures were completed in the participants’ homes in 2019, taking place on three different days within a 10-day period. Families were paid for participation.

5.2.1. Object Play

Participants were video recorded for 30 minutes of object play with one primary caregiver. The 6 sibling pairs were recorded together. Thus, there were 40 recordings with 40 unique caregivers. 34 recordings have two participants (one child and one caregiver) and six have three participants (two children and one caregiver).

The stimulus for the object play was a locally acquired set of 50 marbles. Playing marbles is a common activity for children of the participants’ age in Cushillococha. Participants did not play with the marbles following any conventional set of rules. Thus, the task represents object play, not a structured game.

Object play was recorded with two high-definition video cameras (Sony PJR540 and Canon XA30) at opposing angles. Additionally, audio tracks were recorded via body-worn audio recorders (Olympus VP10); a stand-mounted microphone (Rode NT4); and the internal
microphones of the cameras. Multiple redundant audio tracks were necessary because participants’ homes – which were all located 50 to 200m from a busy road – were typically extremely noisy.

Our aim in object play was to gather maximally comparable data for each participant. Therefore, the first author was present to operate the video cameras and discourage non-target household members from entering the scene.

Some research on child development in Indigenous Latin American (Brown 2011: 37) and Australian (Kelly et al. 2015: 296-297) settings has found that participants are uncomfortable with dyadic interactions in general and dyadic child-caregiver play in particular. In contrast, dyadic child-caregiver interactions are relatively common in Cushillococha, and participants in this study appeared very comfortable with the object play task.

5.2.2. *Free Play*

Participants were video recorded for 60 minutes of free play with one or two primary caregivers. One child was withdrawn from the study before completing this procedure due to family travel. She was excluded from all analyses, leaving 45 complete participants.

The 6 sibling pairs, and the 5 pairs of non-sibling participants who lived together, were each recorded together. This yielded a total of 34 free play recordings (23 with one child, 11 with two children). 52 unique caregivers participated in the free play recordings. Thus, recordings contain between two and four target participants (one to two children and one to two caregivers). Recordings were made with the same equipment as for object play.

Our goal in free play was to record maximally naturalistic interaction between children and caregivers. Thus, during free play sessions, we told participants that they could do whatever they wanted, as long as they remained in the room with the cameras. We also explained that other family members were welcome to enter the scene, which they often did. To facilitate natural interaction, the researcher left the room once the recording began.

5.3. *Sampling and Annotation*

5.3.1. *Sampling*

In object play recordings (total time: 23 hours 20 minutes), participants’ turns were distributed relatively evenly across time. Thus, for transcription we sampled the first 10 minutes of talk in each session, measuring from the first turn after participants received the objects. In recordings with two child participants, we sampled the first 20m rather than 10m.

In free play recordings (total time: 35 hours 18 minutes), turns occurred in short bursts taking place at unpredictable timepoints. Thus, to maximize the quantity of speech included in the sample, we sampled free play recordings based on child volubility (cf. Casillas et al. 2020: 1824). Specifically, we identified the 10m of each free play recording which contained the highest proportion of child and child-directed vocalizations per time. To identify the high-volubility samples, we viewed all free play videos in their entirety. In 10 recordings (with 16
participants), the location of the highest volubility 10m segment was apparent on viewing the recording. In the other 24 free play recordings (with 29 participants), the highest volubility segment was identified automatically via pitch criteria. Specifically, we used the ‘Silent vs. Sounding’ function of Praat to automatically label all intervals in the free play audio with F0 > 300Hz and duration > 100ms. A custom-written R script was then used to identify the 10m segment with the highest proportion of time meeting the pitch criteria.

In recordings with two child participants, we sampled two 10m segments rather than one, following the same procedure.

5.3.2. Annotation

All 45 complete participants had 10m of usable data from object play, and 44 of 45 participants had 10m of usable data from free play, for a total of 89 10-minute samples. One participant did not have usable data from free play because she was on camera for only 3m of her high-volubility segment. Calculations for this child were performed using her productions from the 10m high-volubility segment defined for her co-participant brother.

In each 10m sample, we phonetically transcribed all vocalizations by the target participants, as well as vocalizations by other speakers directed to them. If the 10m timeframe ended during a vocalization, we continued transcribing until the end of the vocalization. 17 of the 89 10-minute samples were transcribed by the first author in collaboration with Angel Bitancourt Serra, an L1 Ticuna speaker, in the Cushillococha area. The remaining 72 samples were transcribed by the first author only. Speech in the samples was translated into Spanish (for data transcribed with Bitancourt Serra) or English (for the remaining data).

Following transcription and translation, the first author performed addressee coding for all linguistic turns in the free play samples. Coding distinguished between speech directed to adults, to target children, and to non-target children. Since object play samples generally included only target participants, we did not code them for addressee. Instead, we assume that all caregiver turns in object play are directed to the target children and vice versa.

All annotation was performed in ELAN (Wittenburg et al. 2006) using the Transcription Mode interface (Dingemanse et al. 2012). The total transcribed time of the 89 samples was 15h 12m (7h 44m object play, 7h 28m free play). Together, they contained 24,491 turns at talk. 13,217 (54.0%) turns were produced by target caregivers, 8,480 (34.6%) by target children, 1,450 (5.9%) by non-target children, and 1,344 (5.5%) by non-target adults.

5.4. Post-Processing

We exported the transcripts from ELAN as CSV files. Using R (v. 4.0.1; R Core Team 2020) and regular expressions implemented with the stringr package (v. 1.4.0; Wickham 2019), we automatically coded each turn at talk for the number of tokens of each demonstrative lexical item which it contained. This coding collapsed demonstratives across the non-deictic features of noun class, number, and case. Following the automatic coding, we manually checked the output and added coding for tokens of demonstratives which were not identified by the search.
5.5. Reliability

Most of the data in the sample was transcribed by the first author, who is not an L1 speaker of Ticuna. In order to assess the reliability of the author’s transcriptions against L1 speaker transcriptions, the first author blindly re-transcribed 40 minutes (23.5%) of the data that was originally transcribed with Angel Bitancourt Serra. The author’s blinded transcriptions and Bitancourt Serra’s transcriptions displayed substantial inter-rater agreement on the number and identity of demonstratives per turn. Raw agreement was 82.4% (Cohen’s $k = 0.69$) for object play data and 86.8% (Cohen’s $k = 0.67$) for free play data.\(^2\)

The R scripts used for sampling, post-processing, and reliability calculations are available, along with the raw and post-processed data, at https://osf.io/x7p2f/?view_only=d03141ccf7494d61b6dd163f70f10339.

6. Results

Analyses were conducted in R (v. 4.0.1; R Core Team 2020) using tidyverse (v. 1.3.0; Wickham et al. 2019) packages. Analysis scripts are available at https://osf.io/x7p2f/?view_only=d03141ccf7494d61b6dd163f70f10339.

We collapsed data from free play and object play, as a series of eight ANOVAs (one for each of the four demonstrative types and two syntactic categories) did not identify significant differences in the production of demonstratives by session type. Below, we report results first for caregivers’ production, then for the production of the target children. Following these results, we analyze correlations between child and caregiver production.

6.1. Caregiver Production

Caregivers produced 13,217 total turns in the corpus. We analyzed turns that (a) contained at least one intelligible word of Ticuna and (b) were directed to the target children (including turns directed to a target child plus another addressee). 10,467 turns, containing 29,545 total word tokens, met these criteria.

Figure 1 [next page] displays caregivers’ frequency of production of each demonstrative per 100 words. Frequencies were calculated independently for each transcript (i.e., recording session) and, in multi-caregiver transcripts, for each caregiver.

\(^2\) In the plurality of cases where the transcriptions disagreed, they disagreed because the author transcribed zero demonstratives and Bitancourt Serra transcribed one token of a Speaker-Proximal demonstrative. Note that underestimation of the frequency of the Speaker-Proximal demonstratives would not affect any of our conclusions: in the data reported in §6, the Speaker-Proximals are already the most frequent type for all age groups.
Figure 1. Frequencies of nominal and locative demonstratives in caregivers’ target child-directed speech (TCDS).

As Figure 1 shows, in target child-directed speech (TCDS) caregivers most often used Speaker-Proximal demonstratives. Following the Speaker-Proximal, they used Addressee-Proximal and Speaker-Distal demonstratives with similar frequency. Caregivers used Dyad-Proximal demonstratives least often. This frequency ranking, which held for both nominal and locative demonstratives, can be schematized as Speaker-Proximal > {Addressee-Proximal, Speaker-Distal} > Dyad-Proximal.

To compare the frequency of each demonstrative in ADS vs. TCDS, we carried out a series of eight Wilcoxon rank sum tests, one for each of the four demonstrative lexical items and two syntactic categories. Each Wilcoxon test compared the frequency of a given demonstrative in adult TCDS vs. the conversational ADS corpus (§3.5). We found that the Speaker-Proximal locative demonstrative (*here near me*) was significantly more frequent in TCDS than ADS (mean tokens per 100 words in TCDS = 5.27, in ADS = 1.58, \(W = 81\), \(p = 5.71 \times 10^{-04}\)). Furthermore, the Addressee-Proximal nominal demonstrative (*that near you*) was significantly less frequent in TCDS than ADS (mean tokens per 100 words in TCDS = 1.28, in ADS = 2.07, \(W = 468.5\), \(p = 0.025\)). The other six demonstratives did not display significantly different frequency in ADS vs. TCDS (all \(p\)’s > 0.15).
Given this new information about CDS frequency, we amended our frequency-based predictions from §4 to correspond to the rank order of demonstratives in TCDS. Based on TCDS frequencies, we now predict that children will acquire both nominal and locative demonstratives in the order Speaker-Proximal > {Addressee-Proximal, Speaker-Distal} > Dyad-Proximal.

Despite the differences between ADS and TCDS in demonstrative frequency, the new order of demonstrative acquisition predicted by TCDS frequency still differs from the order predicted by theories based on cognitive biases. TCDS frequency predicts that the Addressee-Proximal and Speaker-Distal will appear at the same time, while cognitive biases predict that the Addressee-Proximal will appear after the Speaker-Distal.

### 6.2. Child Production

Target children produced 8,480 total turns in the corpus. 5,729 turns (67.6%) contained at least one intelligible word of Ticuna. Most of the remaining turns were made up of fussing and noncanonical babbling (1321 turns), laughter (606 turns), or crying (370 turns). We analyzed all 5,729 turns with intelligible lexical content, regardless of addressee; these turns contained 11,422 total word tokens.

To give a sense of the content of this data, (1) and (2) provide two pseudo-random examples of child turns from the corpus. (1) contains two tokens of demonstratives, given in bold.

1. Child 40 (3;11) and Child 41 (1;6) Object Play, 6:22-6:34
   (Child 40 is across the room from caregiver chasing a marble, while Child 41 sits in caregiver’s lap)
   Child40. ma³ ku³na³tʃe³ʔ [standing up and walking toward caregiver]
   ku³= na³= tʃe³ʔ
   mother.VOC 2SG= 3(IV)OBJ= chop
   Mom, you chop (= flick) it!

   Child40. ku³⁴ da³⁴ʔe²²ma⁴ e⁴ri⁴ ti⁴⁴ʔi⁴⁴, tʃe⁴ʔ (7.7s silence)
   ku³⁴ri⁴ da³⁴ʔe⁴ =ā⁴ma⁴ e³⁴ri⁴ ti⁴⁴ʔi⁴⁴ tʃe³ʔ
   2SG.POSS DEM.NOM:SpkrProx(I) =other INTJ 3(I)OBJ:CI COP chop
   It’s yours, here’s the other one, chop (= flick) it!

   Child40. m³ [sitting down]
   Mm.

   Child40. ma³ nu³ʔa² na³tʃe³ʔ [flicking the marble]
   ma³ nu³ʔa² na³= tʃe³ʔ
   mother.VOC DEM.LOC:SpkrProx 3(IV)OBJ= chop
   Mom, chop (= flick) it from here!

---

3 Glosses use Leipzig abbreviations and the following additional abbreviations: INFER = inferential (evidential), INTJ = interjection, PERF = perfect, SP = Spanish loan word, VOC = vocative. Roman numerals in glosses denote noun class agreement.
(2) Child 6 (4;6) Free Play, 24:38-25:00  
(Child 6 and a neighbor child are with caregiver; neighbor child’s father has just told caregiver not to give him any milk)  

Caregiver.  *ku³¹na¹ta⁴ʔa¹ʔɪ³rɪ³ma³rɪ³* [to neighbor child]  

<table>
<thead>
<tr>
<th>2SG</th>
<th>DAT</th>
<th>3(I)SBJ=</th>
<th>be.stingy</th>
<th>INFER</th>
<th>PERF</th>
</tr>
</thead>
</table>

He (=neighbor’s father) is already being stingy with you, I guess.

Child6.  *ku³¹na¹ta⁴ʔa¹ʔɪ³rɪ³*  

<table>
<thead>
<tr>
<th>2SG</th>
<th>DAT</th>
<th>3(I)SBJ=</th>
<th>be.stingy</th>
<th>INTJ</th>
<th>=again</th>
</tr>
</thead>
</table>

He's being stingy with you again.

Child6.  *e³ʔɪ'ka⁵ma³mi⁵*  

<table>
<thead>
<tr>
<th>INTJ:let’s.see</th>
<th>mother.VOC.SP</th>
</tr>
</thead>
</table>

Right, mom?

Child6.  *ku³¹na¹ta⁴ʔa¹ʔɪ³*  

<table>
<thead>
<tr>
<th>2SG</th>
<th>DAT</th>
<th>3(I)SBJ=</th>
<th>be.stingy</th>
</tr>
</thead>
</table>

He's being stingy with you.

Our cross-sectional analysis divided the participants into four one-year age groups. Table 3 shows participant and data characteristics for each age group. In Table 3, MLU was calculated in terms of words, as the corpus is not labeled for morpheme boundaries. Since Ticuna is morphologically polysynthetic, these MLU values are lower than would be expected for children acquiring an analytic language (see Allen 1996: 43-46 on MLU in polysynthetic languages).

| Table 3. Characteristics of participants, by age group. Mean word token, word type, and MLU values are calculated by session. Standard deviations are presented in parentheses. |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Age Group       | 1;0-1;11   | 2;0-2;11   | 3;0-3;11   | 4;0-4;11   | Total  |
| Participants    | 14        | 10         | 11          | 10          | 45     |
| Females         | 10        | 4          | 6           | 6           | 26     |
| Males           | 4         | 6          | 5           | 4           | 19     |
| Mean Age        | 1;5.10    | 2;5.23     | 3;7.20      | 4;6.12      | -      |
| Age Range       | 1;0.1 - 1;10.21 | 2;1.3 - 2;10.3 | 3;1.13 - 3;11.22 | 4;1.7 - 4;11.8 | -      |
| Samples Analyzed| 28        | 20         | 21          | 20          | 89     |
| Mean Word Tokens| 40.1 (41.1)| 107.1 (124.0) | 190.7 (152.4) | 203.6 (152.2) | --     |
| Mean Word Types | 26.2 (24.9)| 61.9 (67.7)  | 96.9 (68.7)  | 110.1 (65.7) | --     |
| Mean MLU        | 1.4 (0.4) | 1.6 (0.2)  | 1.9 (0.5)   | 2.2 (0.6)   | --     |
6.2.1. Type Analysis

We begin by analyzing participants’ inventory of demonstrative types across age groups. Figure 2 [next page] presents the type data, showing the proportion of participants in each age group who ever produced a given demonstrative type. In the adult portion of the type analysis, only TCDS was analyzed. Additionally, if the same child appeared with different caregivers in the two recording sessions, the two caregivers were treated as one individual in the adult type analysis.
Figure 2. Participants producing each demonstrative type, by age group and syntactic category.
As Figure 2 shows, all age groups – including one- and two-year-olds – produced the Speaker-Proximal nominal demonstrative type (this near me) robustly. One- and two-year-olds were less likely to produce the Speaker-Proximal locative demonstrative type (here near me) than its nominal counterpart; however, prevalence of the Speaker-Proximal locative type approached ceiling for all other age groups. Similarly, the Speaker-Distal nominal and locative demonstrative types (that/there far from me) were present for significant numbers of participants in all age groups, again including one-year-olds. Thus, both of the two speaker-centered demonstrative types – Speaker-Proximal and Speaker-Distal – were produced across all age groups, including the one-year-old group.

In contrast to the speaker-centered demonstratives, the proportion of participants who produced the Addressee-Proximal nominal demonstrative type (that near you) varied substantially between age groups. None of the one-year-olds produced this nominal demonstrative type, and only two of 10 two-year-olds produced it. Subsequently, following the two-year-old age group, there was a sharp rise in the proportion of participants producing the Addressee-Proximal nominal type. After the rise of this type in the three-year-old age group, the Addressee-Proximal nominal demonstrative became comparable to the speaker-centered nominal items in prevalence.

In comparison to the Addressee-Proximal nominal type, the Addressee-Proximal locative demonstrative (there near you) did not display a sharp rise in prevalence between the two- and three-year-old age groups. Though the prevalence of this type increased with time, it was never produced by more than 50% of children in any age group. This low prevalence of the Addressee-Proximal locative type is not adult-like, since the type was produced by 41 (97.6%) of 42 caregivers (i.e., was comparable in adult prevalence to the Addressee-Proximal nominal type).

The Dyad-Proximal nominal demonstrative type (this between us) followed the same pattern as the Addressee-Proximal nominal type, with prevalence rising between the two- and three-year-old age groups. However, the Dyad-Proximal’s rise in prevalence was much less sharp than the Addressee-Proximal’s. Additionally, while 83.3–95.2% of caregivers produced each of the other three demonstrative types at least once, only 27 of 42 caregivers (64.3%) produced the Dyad-Proximal demonstrative at least once. The Dyad-Proximal type’s lower prevalence among adults, as well as its shallower rise between the two- and three-year-old age groups, likely reflects its low frequency in general (recall from §3.5 that this is the least frequent demonstrative in ADS). The Dyad-Proximal locative type (here between us) displayed similar trends in prevalence to its nominal counterpart.

Interim Discussion

The type data represented in Figure 2 support a Speaker-Proximal > Speaker-Distal > {Addressee-Proximal, Dyad-Proximal} order of emergence of demonstrative types. Between 1;0 and 3;0, Ticuna children learn the language’s two speaker-centered demonstratives: first the Speaker-Proximal, then the Speaker-Distal. Later, at ~3;0, the Addressee-Proximal and Dyad-Proximal demonstratives appear. Thus, in type terms, children display early emergence (between 1;0 and 3;0) only for the speaker-centered – Speaker-Proximal and Speaker-Distal –
demonstratives. Non-speaker-centered demonstratives, especially the Addressee-Proximal, display **late** emergence, after 3;0.

### 6.2.2. Nominal Demonstratives: Token Analysis

As well as participants’ inventory of demonstrative types, we also analyze the token frequency of each demonstrative type in their production. Children displayed order of magnitude differences in token frequency between nominal and locative demonstratives. Thus, we report these analyses separately for the two syntactic categories.

Figure 3 [next page] visualizes the token data for nominal demonstratives, showing the token count of each type per 100 words in each age group of participants. The adult group includes data from TCDS only. Note that to facilitate comparison across age groups, each demonstrative occupies its own panel in Figure 3; scales vary between panels.
Figure 3. Token frequency of nominal demonstratives, by type and age group. Outliers (1.0% of the data) are suppressed.
As Figure 3 shows, participants across all age groups produced the Speaker-Proximal nominal demonstrative (this near me) with very high token frequency. In children’s speech, the noun class I form of the Speaker-Proximal nominal demonstrative was the single most frequent word in the entire corpus, and the noun class IV form was the 13th most frequent word. Similarly, in caregivers’ speech, the class I form of the Speaker-Proximal was the 10th most frequent word overall, and the class IV form was the 12th most frequent. In contrast, none of the other nominal demonstratives (except the Addressee-Proximal, see below) were among the 20 most frequent words in either children’s or caregivers’ speech.

Children in all age groups produced the Speaker-Proximal nominal demonstrative numerically much more than adults. To evaluate the significance of this difference, we conducted a series of four pairwise Wilcoxon rank-sum tests comparing the frequency of the Speaker-Proximal for each child age group to its frequency for adults. Two-year-olds, three-year-olds, and four-year-olds produced the Speaker-Proximal more than adults (two-year-olds: $W = 1317, p = 3.6e-7$; three-year-olds: $W = 1268.5, p = 0.0010$; four-year-olds: $W = 1311, p = 5.7e-6$). One-year-olds did not produce the Speaker-Proximal significantly more than adults ($W = 1309, p = 0.15$).

The token frequency of the Speaker-Distal nominal demonstrative (that far from me) varied more between age groups. Recall from §6.2.1 that most one-year-olds did not produce this demonstrative type. The one-year-olds who did produce the Speaker-Distal, however, used it very frequently; consequently, the Speaker-Distal made up 4.20% of word tokens in the average one-year-old’s speech. Among children aged two and over, the Speaker-Distal was less frequent and displayed relatively stable average frequency, ranging from 0.82% to 2.28% of all word tokens. Pairwise Wilcoxon rank-sum tests showed that one-year-olds produced the Speaker-Distal significantly more per word than adults ($W = 637, p = 0.0005$). However, two-, three- and four-year-olds did not display significantly different frequencies for the Speaker-Distal than adults (all $p$’s > 0.1).

Age groups varied more in the token frequency of the Addressee-Proximal nominal demonstrative (that near you). As described in §6.2.1, children under 3;0 almost never produced this type. Moreover, among children over 3;0, the token frequency of the Addressee-Proximal remained much lower than adults. It made up on average only 0.57% of word tokens for three-year-olds and 0.43% for four-year-olds, compared to 1.3% of word tokens produced by adults. Pairwise Wilcoxon tests indicated that all age groups of children used the Addressee-Proximal less than adults (one-year-olds: $W = 224, p = 9.4e-11$; two-year-olds: $W = 202, p = 5.1e-7$; three-year-olds: $W = 554.5, p = 0.0090$; four-year-olds: $W = 412, p = 8.8e-4$).

The Dyad-Proximal nominal demonstrative (this between us) presents difficulties for comparison because of its low frequency and prevalence among all age groups. Statistically, however, pairwise Wilcoxon comparisons indicated that one-year-olds and two-year-olds used the Dyad-Proximal less per word than adults (one-year-olds: $W = 699, p = 0.0005$, two-year-olds: $W = 559, p = 0.047$). Three- and four-year-olds’ frequencies of the item, in contrast, were not significantly different from adults’ (both $p$’s > 0.1).

Interim Discussion
Children’s tokens of nominal demonstratives were distributed very asymmetrically across types. Compared to adults, children across all age groups produced many more tokens of the Speaker-Proximal nominal demonstrative type, and many fewer tokens of the Addressee-Proximal nominal demonstrative type. Importantly, children’s lower-than-adult use of the Addressee-Proximal type persisted even in the age groups (three-year-olds and four-year-olds) where the majority of children used the type at least once. That is, even after children’s inventory of nominal demonstrative types increased at ~3;0 to include the Addressee-Proximal, their distribution of tokens still did not include an adult-like proportion of this type.

While children overused the Speaker-Proximal demonstrative and underused the Addressee-Proximal in comparison to adults, their frequencies of other demonstratives were relatively adult-like, especially for older age groups. Children aged 2;0 and above displayed frequencies of the Speaker-Distal demonstrative which are statistically indistinguishable from adults’: children 3;0 and above displayed adult-like frequencies of the Dyad-Proximal demonstrative as well.

In sum, token distributions indicate that children displayed late mastery – i.e., did not achieve adult-like production by 4;11 – for only some demonstratives, the Speaker-Proximal and Addressee-Proximal. By contrast, children in this study did display adult-like frequencies for the Speaker-Distal demonstrative by 2;0 and for the Dyad-Proximal demonstrative by 3;0.

6.2.3. Locative Demonstratives: Token Analysis

Children produced ~10 times fewer locative demonstrative tokens than nominal demonstrative tokens. Figure 4 [next page] visualizes the token data for locative demonstratives in each age group. The adult group again includes TCDS only.
Figure 4. Token frequency of locative demonstratives, by type and age group. Outliers (0.4% of the data) are suppressed.
As visual comparison of the adult and child age groups in Figure 4 makes clear, children produced all locative demonstrative types much less frequently than adults. However, the magnitude of differences between children and adults’ token frequencies varied between types.

The **Speaker-Proximal** locative demonstrative (*here near me*) was used much less by children than adults. On average, this item made up 5.3% of all adult word tokens, and two forms of this demonstrative were respectively the **third** and **ninth most frequent** words in caregivers’ speech. In contrast, among children the Speaker-Proximal made up on average only 0.43% (for one-year-olds) to 2.6% (for three-year-olds) of word tokens. Despite these lower frequencies, a form of the Speaker-Proximal locative item was still the **19th most frequent** word in children’s speech.

Pairwise Wilcoxon tests indicated that all age groups of children used the Speaker-Proximal locative item less frequently than adults (one-year-olds: $W = 163, p = 1.3e-11$; two-year-olds: $W = 160, p = 1.1e-7$; three-year-olds: $W = 448, p = 5.4e-4$; four-year-olds: $W = 317, p = 3.8e-5$).

The **Speaker-Distal** locative demonstrative (*there far from me*) displayed less divergence between child and adult frequencies. This item made up on average 1.34% of adult word tokens. For children, frequencies of the item ranged from 0.37% (for four-year-olds) to 0.95% (for two-year-olds) of all word tokens. A Wilcoxon test did not identify differences between two-year-olds’ frequencies of the Speaker-Distal locative item and adults’ ($W = 609.5, p = 0.20$). However, pairwise tests showed that all other age groups used the Speaker-Distal significantly less than adults (one-year-olds: $W = 557, p = 7.0e-5$; three-year-olds: $W = 557, p = 0.0098$; four-year-olds: $W = 389, p = 4.1e-4$).

The **Addressee-Proximal** locative demonstrative (*there near you*) – like its nominal counterpart – displayed much lower token frequency for children than adults. For adults, the Addressee-Proximal locative demonstrative represented on average 1.15% of all word tokens; for children, its average frequencies ranged from 0% (for one-year-olds) to 0.40% (for four-year-olds) of all word tokens. As for the Speaker-Proximal, pairwise Wilcoxon tests indicated that all age groups of children produced the Addressee-Proximal locative demonstrative significantly less than adults (one-year-olds: $W = 210, p = 5.3e-11$; two-year-olds: $W = 203.5, p = 5.7e-7$; three-year-olds: $W = 295.5, p = 1.6e-6$; four-year-olds: $W = 401.5, p = 6.3e-4$).

Last, the **Dyad-Proximal** locative demonstrative (*here between us*) displayed the least divergence between child and adult token frequency. For adults, the Dyad-Proximal locative item made up on average 0.43% of all word tokens. For children, average frequencies of the Dyad-Proximal ranged from 0% (for one-year-olds) to 0.39% (for three-year-olds) of all word tokens. Pairwise Wilcoxon tests showed that one- and two-year-olds produced the Dyad-Proximal locative item significantly less often than adults (one-year-olds: $W = 532, p = 3.4e-6$; two-year-olds: $W = 397, p = 4.4e-4$). However, three-year-olds and four-year-olds’ frequencies of the Dyad-Proximal were not significantly different from adults’ (three-year-olds: $W = 658, p = 0.059$; four-year-olds: $W = 599, p = 0.073$).

**Interim Discussion**

Though children displayed locative demonstrative types early in acquisition, they used the locative items much less frequently than adults. This pattern held across all age groups and all of
the three most frequent demonstrative types, with the (presumably chance) exception of two-year-olds’ productions of the Speaker-Distal locative demonstratives. The pattern did not hold for the Dyad-Proximal locative demonstrative, which three- and four-year-olds produced with statistically indistinguishable frequency from adults. However, the diverging behavior of the Dyad-Proximal may be an artefact of its relatively low token frequency in all age groups.

Since children displayed non-adult-like low frequencies of the Speaker-Proximal, Speaker-Distal, and Addressee-Proximal locative demonstrative types throughout the timeframe of this study, our findings support that Ticuna locative demonstratives, as a group, display late mastery. However, children’s non-adult-like use of locative demonstratives is different from their non-adult-like use of nominal demonstratives. Among nominal demonstratives, children overused some types (i.e., the Speaker-Proximal) and underused others (i.e., the Addressee-Proximal). Among locative demonstratives, on the other hand, children underused all items regardless of type.

6.3. Relationships between Child and Caregiver Speech

In order to assess whether differences in children’s demonstrative production were due to differences in the lexical composition of TCDS, we analyzed correlations between the frequency of each demonstrative type in children’s speech vs. in their caregivers’ speech. In this analysis, children who were recorded with two caregivers simultaneously were compared to both caregivers (i.e., child’s frequency vs. caregiver 1’s frequency and child’s frequency vs. caregiver 2’s frequency were treated as two independent observations). Child and caregiver frequencies of the Speaker-Proximal nominal demonstrative showed a significant positive association (Spearman’s $\rho = 0.26$, $p = 0.0085$). For each other nominal and locative demonstrative type, the association between child and caregiver frequencies was not significant (all $p$’s $> 0.5$, all $|\rho| < 0.18$).

Additionally, to evaluate whether differences between age groups in demonstrative production were related to differences in the composition of TCDS directed to younger vs. older children, we analyzed correlations between the frequency of each demonstrative in caregivers’ speech and children’s age (calculated in days). We observed a significant positive association between child age and caregiver frequency of the Dyad-Proximal nominal demonstrative (Spearman’s $\rho = 0.22$, $p = 0.029$). There was also a significant positive association between child age and caregiver frequency of the Dyad-Proximal locative demonstrative (Spearman’s $\rho = 0.28$, $p = 0.0045$). For each other nominal and locative demonstrative type, the association between child age and caregiver frequency was not significant (all $p$’s $> 0.05$, all $|\rho| < 0.19$).

7. General Discussion

We conducted a cross-sectional, observational study of Ticuna-acquiring children’s production of demonstratives. 45 children aged 1;0 to 4;11 were recorded interacting with their caregivers, once in object play (30 minutes) and once in a free play session (60 minutes). We analyzed 20 minutes of recording time per child – 10 minutes from object play, 10 from free play. Our analysis replicated the finding, from studies of English and other global languages, that
demonstratives display early emergence but late mastery in L1 acquisition. However, we replicated this finding in Ticuna only for speaker-centered demonstratives. Addressee-centered demonstratives did not display early emergence, but instead late emergence and even later mastery. We also observed the novel and unpredicted finding that children were much more likely to produce nominal demonstratives (equivalent to this/that) than locative ones (equivalent to here/there).

### 7.1. Early Emergence

As discussed in §4, theories which foreground cognitive biases predict that children – because of their bias toward egocentrism – will produce speaker-centered terms before non-speaker-centered ones, and proximal terms before distal ones. Thus, the predicted order of emergence of the Ticuna demonstratives is Speaker-Proximal > Speaker-Distal > {Addressee-Proximal, Dyad-Proximal}. In contrast, if the emergence of demonstratives was controlled only by input frequency, per §6.1 we expect the order Speaker-Proximal > {Speaker-Distal, Addressee-Proximal} > Dyad-Proximal.

The results of this study clearly support the predictions based on cognitive bias. Among the children that we sampled, nominal and locative demonstrative types emerged in the order Speaker-Proximal > Speaker-Distal > {Addressee-Proximal, Dyad-Proximal}. We did not observe evidence of the Addressee-Proximal appearing simultaneously with the Speaker-Distal, as would be predicted by frequency alone. More specifically, one-year-olds robustly produced the Speaker-Proximal nominal demonstrative, but no other demonstrative type. Likewise, two-year-olds robustly produced the Speaker-Proximal and Speaker-Distal types, but still effectively did not produce any non-speaker-centered terms. In other words, children displayed early (before 3;0) emergence only for speaker-centered demonstratives. Speaker-centered demonstrative types appeared in the one-word stage; non-speaker-centered types, in contrast, mostly did not appear until at least 3;0. This result fails to support our prediction that all four demonstrative types would emerge before 3;0 (cf. Rodrigo et al. 2004, Gonzalez-Peña et al. 2020).

Importantly, frequency in CDS cannot account for the late emergence of non-speaker-centered demonstratives in Ticuna. We show in §6.1 that children heard Speaker-Distal and Addressee-Proximal demonstrative types equally often in CDS; furthermore, per §6.3, children’s age did not correlate with the CDS frequency of these items. Yet despite the equal CDS frequency of the Speaker-Distal and Addressee-Proximal, children began producing the Speaker-Distal demonstrative substantially earlier than the Addressee-Proximal. Thus, our results for the Addressee-Proximal demonstrative suggest that children’s cognitive bias toward egocentrism can outweigh frequency in the acquisition of demonstratives.

This result raises several questions for future research on the L1 acquisition of Ticuna and other languages with addressee-centered demonstratives. One immediate question is whether addressee-centered demonstratives in other languages which have them – for example, Turkish or Finnish (Nahkola et al. 2020: 250) – display the same late emergence as the Ticuna Addressee-Proximal. On the analysis that the late emergence of the Addressee-Proximal reflects
an innate bias toward egocentric spatial cognition, we predict that addressee-centered demonstratives will always emerge late.

Our results about late emergence also lead to questions for our own analysis. One of these questions relates to overextension: if a Ticuna child does not produce the Addressee-Proximal demonstrative, then how do they index referents that are near the addressee? We take up this issue in the following section.

7.2. Late Mastery

Cross-linguistic literature on the acquisition of demonstratives indicates that children are unlikely to attain adult-like production or comprehension of the items until ~6;0. Consequently, this literature predicts that children in this study – the oldest of whom was 4;11 – will not display adult-like use of any demonstrative.

For nominal demonstratives, this prediction was in part correct. Across all age groups, children displayed significantly different frequencies than adults for the Speaker-Proximal and Addressee-Proximal nominal demonstratives. However, the way that children differed from adults varied between these types. Children of all age groups produced the Speaker-Proximal type much more frequently than adults, but the Addressee-Proximal much less frequently. The low frequency of the Addressee-Proximal was not simply due to one- and two-year-olds’ lack of this type. Even in age groups where the Addressee-Proximal type was present for most children (i.e., three- and four-year-olds), it remained much less token-frequent in children’s production than in CDS.

Since children displayed much higher frequencies of the Speaker-Proximal demonstrative than adults, they likely overextended this item. More specifically, since children in all age groups paired overuse of the Speaker-Proximal nominal demonstrative with underuse of its Addresssee-Proximal counterpart, we suggest that children are overextending the Speaker-Proximal to contexts where adults would use the Addressee-Proximal. As this conclusion is based only on the overall frequency of each item, it is necessarily tentative. To provide more support for overextension of the Speaker-Proximal into Addressee-Proximal space, analysis of the contexts where children vs. adults produced Speaker-Proximal demonstratives (as in Künstay & Özyürek 2006) would be necessary.

In contrast to the Speaker-Proximal and Addressee-Proximal demonstrative types, the Speaker-Distal and Dyad-Proximal demonstrative types did not display consistent differences between child and adult frequency. For the Dyad-Proximal demonstrative type, this lack of a significant difference is difficult to interpret, as both children and adults used the item very infrequently. For the Speaker-Distal, on the other hand, the lack of significant differences in frequency between children (other than one-year-olds) and adults suggests that children may attain adult-like production relatively early, between 2;0 and 3;0.

Ticuna children’s early achievement of adult-like frequency for the Speaker-Distal contrasts with the behavior of children learning Spanish (Rodrigo et al. 2004), who underuse Speaker-Distals until at least 3;0, and children learning Yucatec Maya, who underuse Speaker-Distals through
5;5 (Espinosa Ochoa 2009: 113-114). We suggest that Ticuna children’s earlier mastery of the Speaker-Distal nominal demonstrative reflects that this term is exclusively speaker-centered (Redacted for anonymous review). In contrast, at least one of the putative ‘Speaker-Distal’ demonstratives of Yucatec may be better analyzed as dyad-distal, conveying that the referent is far from both speaker and addressee (Hanks 1990: 436-448); similarly, Jungbluth (2003) argues that the Spanish ‘Speaker-Distal’ demonstrative is dyad-distal as well.

Thus, as predicted in the English-based literature, children displayed late mastery for some demonstrative types. Across all age groups, they overused Speaker-Proximal (nominal) demonstratives in comparison to adults, and underused Addressee-Proximal forms. In other words, children’s late mastery of demonstratives included both a type-level bias against producing the Addressee-Proximal at all before 3;0, and a token-level bias against producing tokens of this form after 3;0. By contrast, our data presented no evidence for late mastery of the Speaker-Distal or Dyad-Proximal demonstratives. Thus, the bias which drives late mastery appears to be specific to the contrast between Speaker-Proximal and Addressee-Proximal forms.

### 7.3. Lateness of Locative Demonstratives

We predicted in §4 that children would acquire nominal and locative demonstratives on similar timeframes. In Ticuna, nominal and locative demonstratives have similar semantics (§3.4) and similar ADS and CDS frequency (§§3.5, 6.1). Thus, both cognitive biases and input frequency predict that the two syntactic categories of demonstratives will emerge simultaneously.

Our data supports this prediction at the type level – children’s inventories of nominal and locative demonstrative types increased at similar rates across time (§6.2.1). However, our findings do not support this prediction at the token level. Between 1;0 and 3;0, children produced ~10 tokens of nominal demonstratives (§6.2.2) for every 1 token of a locative demonstrative (§6.2.3). After 3;0, children’s token frequency of locative demonstratives increased, but still did not match their frequency of nominal demonstratives. This finding is inconsistent not only with theoretical predictions, but also with the literature on early L1 acquisition of demonstratives in other languages. English-acquiring children aged 1;0 to 3;0 produce locative demonstratives more frequently than nominal ones; Spanish-learning children in this age group produce locative and nominal demonstratives with similar frequency (Gonzalez-Peña et al. 2020: 11).

It is not possible to account for the lateness of locative demonstratives in Ticuna via overall frequency, morphosyntactic complexity, or frequency in isolation. First, frequency does not explain the late emergence of locatives, since they are overall equal in CDS frequency to nominal demonstratives (§6.1). Second, in the domain of morphosyntactic complexity, one could argue that locative demonstratives are acquired later in Ticuna because they are morphosyntactically more complex. But this is not true. At least in terms of allomorphy, Ticuna nominal demonstratives have greater morphosyntactic complexity than locative ones, since the nominal items display noun class agreement (§3.4) and the locative items do not.

Having eliminated frequency and morphosyntactic complexity, we also considered whether frequency in isolation could account for the lateness of locative demonstratives. Gonzalez-Peña
et al. (2020: 11) suggest that, for the English locative demonstrative *there*, frequency in isolation drives early emergence. Thus, Ticuna children might learn nominal demonstratives earlier because they are produced more often in isolation. We conducted an exploratory analysis to test whether this was true in the caregivers’ CDS in this study. It was false. Caregivers were as likely to produce locative demonstratives in isolation (mean proportion of tokens in isolation = 7.7%, SD = 12.5%) as to produce nominal demonstratives in isolation (mean tokens in isolation = 9.0%, SD = 8.9%) ($W = 2592, p = 0.093$). An additional analysis examined whether nominal demonstratives occurred in overall shorter utterances than locative demonstratives. They did not: utterances with nominal demonstratives (mean words = 3.97, SD = 2.21) were actually longer than utterances with locative ones (mean words = 3.67, SD = 1.93) ($W = 1892998, p = 7.47e-5$). Explorations of locative items’ lateness based on utterance length therefore fail as well.

Against this background, the strongest remaining explanation for Ticuna children’s preference for nominal over locative demonstratives is their role in presentative constructions. Across languages – English (Gonzalez-Peña 2020), Spanish (Rodrigo et al. 2004), and Yucatec Maya (Espinosa Ochoa 2009: 138) – presentative constructions, equivalent to English *here/there it is*, emerge as a very early use of demonstratives and make up a large proportion of children’s early demonstrative utterances. While English uses locative demonstratives as the predicate in presentative constructions (*here it is*), Ticuna uses nominal demonstratives (equivalent to *this it is*; e.g., example 1, line 2). Thus, children’s disproportionately high use of nominal vs. locative demonstratives may simply reflect that they (a) produce disproportionately many of their demonstrative tokens in presentative contexts and (b) use adult-like syntax, with a nominal demonstrative as the predicate, in their presentatives.

8. Conclusion

This study investigated young children’s acquisition of the demonstrative system of Ticuna, focusing on the contrast between acquisition of speaker-centered vs. addressee-centered demonstratives. For speaker-centered demonstratives, our results supported the same pattern of early emergence, but late mastery, seen in the L1 acquisition of demonstratives in well-studied languages such as English. Children began to produce Speaker-Proximal and Speaker-Distal demonstratives very early in development – often during the one-word stage. However, they did not display adult-like use of the items at any point in the timeframe of our study (1;0 – 4;11).

In contrast, our results for Addressee-Proximal demonstratives departed from the early emergence-late mastery pattern. While children produced speaker-centered demonstratives well before 2;0, they did not produce Addressee-Proximal demonstratives until after 3;0. Because children heard Addressee-Proximal demonstratives as frequently as some (early-appearing) speaker-centered demonstratives, the late emergence of addressee-centered demonstratives cannot be due to frequency effects. Instead, it likely reflects that children struggle to comprehend addressee-centered demonstratives, and therefore also to produce them, because of their bias toward egocentric spatial cognition. Moreover, since the Addressee-Proximal emerges late despite its high frequency in CDS, this result illustrates that children’s egocentrism can outweigh input frequency in the acquisition of spatial language. These observations highlight the importance of studying the L1 acquisition of typologically diverse languages, which offer new opportunities to test the predictions of classic literature.
References


Learning speaker- and addressee-centered demonstratives in Ticuna


