Early consonant production in Tseltal and Yélî Dnye

Bram Peute
Radboud University Nijmegen
Houtlaan 4, 6525 XZ Nijmegen
Leiden University
Rapenburg 70, 2311 EZ Leiden

aakvs@xs4all.nl

Marisa Casillas
Max Planck Institute for Psycholinguistics
Wundtlaan 1, 6525 XD Nijmegen
University of Chicago
Rosenwald Hall, 3rd Floor, 1101 E. 58th St., Chicago, IL, 60637 USA
mcasillas@uchicago.edu

Word count: 7794

Abstract

Recent evidence shows that children reach expected basic linguistic milestones in two rural Indigenous communities, Tseltal and Yélî Dnye, despite infrequent exposure to child-directed speech. However, those results were partly based on vocal maturity measures that are fairly robust to environmental variation, e.g. the onset of babbling. Directed speech input has been traditionally linked to lexical development, which is by contrast environmentally sensitive. We investigate the relation between child-directed speech and early phonological development in these two communities, focussing on a phonological benchmark that links children's pre-lexical and early lexical development: the production of consonants. We find that, while Tseltal and Yélî children's canonical babble onset align with previously attested patterns, their early consonant acquisition shows some divergence from prior expectations. These preliminary results suggest that early consonant production may demonstrate greater environmental sensitivity than canonical babble, possibly via similar mechanisms that link linguistic input and lexical development.

Keywords: language acquisition; consonant production; vocal motor schemes; canonical babble; child directed speech

1. Introduction

Children's acquisition of spoken language, from early vocalizations to their first words, offers insight into the evolutionary origins and environmental sensitivity of the human language learning system. The trajectory of early vocal development is well documented across a diverse set of language communities: infants start producing protophones soon after birth, then around 6 to 8 months of age they begin to babble, and then some time around the first birthday their first words appear (Oller 1980; Oller et al. 1998; Oller 2000; Lee et al. 2017; Cychosz et al. in press.). Canonical babbling (henceforth CB) serves as an important step along the road to first words. CB is composed of well-formed syllables, typically reduplicated consonant-vowel structures, like "mama" or "da" (Lee et al. 2017). This kind of early babble is a milestone in the child's motor development that also marks the beginning of their ability to reliably and distinguishably produce the phonemes of their home language(s) (McGillion et al. 2017).

The robustness of this overall developmental trajectory is underscored in contexts where children are infrequently directly addressed; despite the infrequent use of child-directed speech (henceforth CDS) in some communities—as is the case in the unrelated Tseltal and Yélî Dnye speech communities—children's lexical vocalizations showed a similar onset as compared to children in more CDS-rich settings (Casillas et al. 2020a; b). This result appears, at face value, to run counter to other work clearly showing that CDS is strongly associated with faster-growing receptive and productive vocabularies (e.g., Hart & Risley 1995; Hoff 2003; Shneidman & Goldin-Meadow 2012; Weisleder & Fernald 2013). Critically, however, the investigations by Casillas and colleagues were not focused on vocabulary development. The authors instead focused on few broad categories of vocalization type, several of which are known to be fairly robust to environmental variation (e.g., Oller et al. 1995; Lee et al. 2017; Cychosz et al. in press). A stronger test of the impact of CDS on early vocal development would therefore be to look at patterns in production that link children's pre-lexical development to their eventual lexical development. If there were still evidence for non-delayed development with this stronger test, it would suggest that these children indeed have some means of gathering sufficient relevant language information on the same timescale as children in CDS-rich environments, but without encountering frequent CDS. Furthermore, while past work on early vocal development has been linguistically diverse (de Boysson-Bardies & Vihman 1991; Kunnari 2003; Fikkert & Levelt 2008; Lee et al. 2010; Shneidman & Goldin-Meadow 2012; Weisleder & Fernald 2013; Lee et al. 2017) links between early phonological development, caregiving practices, and everyday language use have yet to be drawn for any rural, Indigenous context. Especially considering that many languages spoken in rural and traditional or Indigenous contexts have typological features underrepresented in developmental research, exploring the predictive relationship between productions at the pre-lexical and early lexical stages of development in these communities is a key next step for research on language development.

We here investigate children's "vocal motor schemes" (henceforth VMS), which are pre-existing word-like motor-plans developed through babbling. VMSs provide a bridge between pre-lexical babble and early productive vocabulary. They provide the child with the means of producing "auditory approximations" to the target words produced by adults, without the child having to deduce the exact phonemes involved (Vihman 1993). This capacity for consistent phonetic patterning is theorized to help train the child in producing consonants, described as "a generalized action plan that generates consistent phonetic forms... a formalized pattern of motor activity that does not require heavy cognitive resources to enact" (McCune & Vihman 2001: 152). McCune and Vihman (2001) operationalize VMS as a measure of continuous phonological training and oral motor control as follows: If a child is able to produce 10 realizations of a given consonant in 3 or 4 contiguous monthly sessions, the child is said to have acquired the VMS for that consonant. The English-learning children in McCune and Vihman's (2001) study acquired two VMS consonants between the age of 9 and 14 months. This transition in the child's phonological capabilities, in addition to its strong association with the onset of babbling (DePaolis et al. 2011; Majorano et al. 2013; McGillion et al. 2017), also predicts individual differences in the onset of children's first recognizable words (McGillion et al. 2017). We hypothesized that, much like children's early vocabularies (Hart & Risley 1995; Hoff 2003; Shneidman & Goldin-Meadow 2012; Weisleder & Fernald 2013), children's VMS development might show sensitivity to differences in children's language environments. If so, we might expect that children in language communities with infrequent CDS, such as Tseltal and Yélî children, show overall later VMS development than children in environments with more frequent CDS (e.g., the English acquiring children in McCune & Vihman (2001)).

1.1 The present study

The present study focuses on the acquisition of consonants (both in onset and coda position) by children born into two speech communities: Tenejapan Tseltal and Yélî Dnye. Tseltal is a Mayan language spoken in the highlands of Chiapas in Southern Mexico. This language, which has five main dialects, is in vigorous use with estimates of over 400,000 to 500,000 speakers who are bilinguals in Spanish, and an estimated 40,000 to 50,000 monolingual speakers (Polian 2013; Eberhard et al. 2020). Yélî Dnye is an isolate spoken on Rossel Island, in the Milne Bay Province of Papua New Guinea. The language, which has two main dialects, is in vigorous use as well, with a loosely estimated 5000–7000 native speakers, most of whom are at least partly bilingual in English, Tok Pisin, and/or other languages of Papua New Guinea (Levinson in press; Eberhard et al. 2020). Many aspects of daily life, as well as the overall rate of CDS, are similar between these two patrilocally organized, horticulturalist communities (Casillas et al. 2020a; b). Before age 3;0, children in both communities are directly addressed around 3–3.5 minutes per hour during waking days at home (Casillas et al. 2020a; b). While the quantity of directed linguistic input is relatively similar, the consonant inventories of the two languages are highly distinct from each other. Yélî Dnye has 60 contrastive consonants compared to Tseltal's 20, with the additional consonants including both single- and doubly-articulated stops with nasal, labial, and/or palatal releases and pre-nasalization (Levinson in press; Polian 2013). The fact that these two language communities use similar rates of directed input but have highly divergent phonological inventories makes the comparison of children's early consonant development particularly interesting.

If Tseltal and Yélî Dnye-acquiring children show similar VMS development to same-aged English acquiring children, it would suggest that even lexically-related early phonological development is cross-culturally robust to variation in CDS use, and that these children may instead use other cues from their linguistic environment to spur their early linguistic growth. For example, many scholars focusing on language development in non-urban contexts have highlighted the role of observable third-party behavior for learning (e.g., Rogoff et al. 1993; Chavajay & Rogoff 1999; Gaskins & Paradise 2010). In addition to VMS, we also attempt to replicate the finding that CB onset shows no delayed development in these two communities, this time using a larger sample and independent annotations from the original study data (Casillas et al. 2020a; b).

1.2 Predictions

Based on prior findings, we hypothesize that the onset of CB occurs between 6 and 8 months in both datasets (Oller 1980; Oller et al. 1998; Oller 2000; Lee et al. 2017; Cychosz et al. in press). We also predict that VMS acquisition by Yélî and Tseltal children will be comparable to that of Western children; that is, at least 2 VMS consonants will emerge between 9 and 14 months (McCune & Vihman 2001). We make this prediction in line with the prior finding that the onset of single- and multi-word utterances is non delayed in these communities (Casillas et a. 2020a; b), and despite the fact that VMS is related to early lexical development, which is known to be sensitive to ambient CDS rates. Finally, given the differences in phonological inventories between Tseltal and Yélî Dnye, we predict that Yélî children acquire their VMS consonants slightly later than Tseltal children because the Yélî consonant inventory is large, complex, and acoustically packed (Levinson in press; see also Cristia & Casillas under review). That said, differences based on inventory effects may not emerge with children this young (Jakobson 1968; Vihman & de Boysson-Bardies 1994; Monnin & Lœvenbruck 2010).

2. Methods

2.1 The communities

We analyzed a total of 15 hours and 45 minutes of Tseltal and 4 hours and 30 minutes of Yélî Dnye audio recordings for the phonological content of children's spontaneous vocalizations. The recordings were collected in 2015 (Tseltal) and 2016 (Yélî Dnye), and can be accessed via the Casillas HomeBank corpus (Casillas et al. 2017). Participant consent processes and data collection were conducted in accordance with ethical guidelines approved by the Radboud University Social Sciences Ethics Committee (Casillas 2020b).

The Tseltal speaking children come from a farming community in the highlands of Chiapas in Southern Mexico, where they are typically raised in patrilineally organized, multigenerational households. During the day, infants are carried on their mother's back while she goes about her business, or they are left at home with other family members while the mother works elsewhere (e.g., in the field). The majority of children in this community grow up monolingually until they go to school (Casillas et al. 2020b), and their linguistic environments have been characterized as non-child-centered and non-object-centered (Brown 1998; 2011; 2014). CDS from adults is typically limited until the infants themselves start to seek out interaction (Brown 2014) and continues to be relatively infrequent through 3;0 (Casillas et al. 2020b). The Tseltal data used in the current study include 20 children (M = 10 months; median = 9; range = 5–19), including four children whose recordings were used in the Casillas et al. (2020b) study. We focused on children between 5 and 20 months because children are expected to begin CB production around 6 months and sometimes even earlier, and by 20 months nearly all children would be expected to have started producing recognizable words (Oller et al. 1998; McGillion et al. 2017; Casillas et al. 2020b).

The Yélî Dnye speaking children live in a collection of small settlements on the north-eastern shore of Rossel Island, which is located 250 nautical miles off the south coast of mainland Papua New Guinea. Children grow up in hamlets with patrilocally organized household clusters, where there is often a shared open space between households. During the day, children are carried in their caregivers' arms, and they are frequently passed around between community members—even those from far outside the natal hamlet—who return the child to the mother for feedings (Casillas et al. 2020a; Brown & Casillas in press). Yélî children mostly grow up speaking Yélî Dnye at home, although English, Tok Pisin, and other regional languages are often spoken by adults and school-aged children (Brown & Casillas in press; Casillas et al. 2020a). Children begin to learn English once they start school. The linguistic environment of Yélî children can be characterized as child-centered (Brown & Casillas in press; Ochs & Schieffelin 1984); children are considered a shared responsibility, as well as a source of joy and entertainment for caregivers, and, as such, interaction with infants and young children on Rossel Island is initiated by women, men, girls, and boys alike (Casillas, Brown, & Levinson 2020a). The Yélî data we use in the current study include all 12 children in the 5–20-month age range in the 2016 Yélî Dnye Casillas HomeBank corpus (M = 12.4 months; median = 12.5; range = 8–17), including four children whose recordings were used in the Casillas et al. (2020a) study.

Table 1. Descriptive statistics of the age of the children in the present dataset.

Language	N	Mean	Median	Range	Std. Deviation
Tseltal	20	10	9	5-19	4.025
Yélî Dnye	12	12.4	12.5	8-17	2.596

2.2 The data

For both datasets, the recordings were made using an audio recorder (Olympus WS-832 or WS-853) and photo camera (Narrative Clip 1) strapped to the child's chest during most of a waking day at home. For young infants and very small children, the primary caregiver wore the photo camera (see Casillas et al., 2020b for details). The recordings give a natural picture of language use over the course of multiple home activity contexts. We do not use the image data in this study—only the audio recordings.

We base our analyses on a subsample of each child's spontaneous vocalizations from the day. We selected short random clips from each recording, following the random sampling procedure used for the eight children whose recordings were previously annotated for Casillas et al. (2020a) and (2020b). That is, we focused our analyses on spontaneous vocalizations by the target child within nine randomly selected and non-overlapping clips for each recording; clips were 5 minutes long for the Tseltal data and 2.5 minutes long for the Yélî Dnye data. The reason for this disparity in clip duration is that the number of speakers and amount of background noise in the Yélî dataset makes it particularly time consuming to annotate, limiting what can be feasibly accomplished during researcher visits to the island (Casillas et al. 2020a). We take this disparity into account in the analysis, as explained below. Each target child vocalization in each clip was classified and broadly phonetically transcribed. For the eight recordings with existing annotations, phonetic transcriptions were added to the existing speech annotations (without reference to their previously added vocal maturity classification). For all other recordings, the child's vocalizations within each clip were first diarized, and subsequently classified and phonetically transcribed. Vocalization annotation followed the scheme shown in Table 2.

Table 2. Vocalization annotation scheme.

Annotation	**Meaning**
hamuwa	example of CB
N	non-canonical vocalizations
L	laughing
Y	crying

We transcribed children's canonical babbling using the International Phonetic Alphabet (International Phonetic Association, 2020) on the basis of the phoneme inventory of each language, as it is spoken by adults. Instances in which children produced phonemes that are not present in the adult language were transcribed as perceived by the first author.

The phonemes used in transcribing the Tseltal data are based on Polian's (2013) grammar, the inventory for which is shown in Tables 3 and 4.

Table 3. Tseltal consonants (non-bold, no line), with the additional, non-native phones we found in Tseltal children's spontaneous vocalizations (bold, underlined).

	Bilabial	Alveodental	Palatoalveolar	Velar	Glottal
Plosives	p b p'	t **<u>d</u>** t'		k k'	ʔ
Affricates		ts ts'	tʃ **<u>dʒ</u>** tʃ'		
Fricatives		s	ʃ **<u>ʒ</u>**	x	h
Nasals	m	n			
Laterals		l			
Rhotics		r			
Approximants	β		j	w	

Table 4. Tseltal vowels (non-bold, no line), with the additional, non-native phones we found in Tseltal children's spontaneous vocalizations (bold, underlined).

	Front	Central	Back
Close	i		u
Close-mid	e	**<u>ə</u>**	o
Open		a	

The phonemes used in transcribing the Yélî Dnye data are based on Levinson's (in press) grammar, the inventory for which is shown in Tables 5 and 6. The majority of the phones were not attested in the child vocalizations (e.g., we found no doubly articulated consonants).

Table 5. Yélî Dnye consonants (non-bold and no line) with additional, non-native phones found in Yélî children's spontaneous vocalizations (bold, underlined). SBC stands for 'simultaneous bilabial closure'.

	Bilababial				Alveolar		Alv.+SBC		Post-Alv.		Post-Alv.+SBC		Velar		
		Pal.	Lab.	Both		Pal.		Pal.		Pal.		Pal.		Pal.	Lab.
Plosives	p	pʲ	pʷ	pʲʷ	t	tʃ	tp	tpʲ	ṭ	ṭʲ	ṭp	ṭpʲ	k **g**	kʲ	kʷ
Prenasalized plosives	mb	mbʲ	mbʷ	mbʲʷ	nt	ndʒ	nmdb	nmdbʲ	ṇḍ	ṇḍʲ	ṇmḍb	ṇmḍbʲ	ŋg		ŋgʷ
Nasally-released plosives						ʃ			ṭṇ	ṭṇʲ	ṭpṇm	ṭpṇmʲ	kŋ		kŋʷ
Nasals	m	mʲ	mʷ	mʲʷ	n	nʲ	nm	nmʲ	ṇ	ṇʲ	ṇm	ṇmʲ	ŋ		ŋʷ
Rhotics					**r**										
Continuants	w~β	βʲ							j l	lʲ	lβʲ				

	Velar+SBC		Glottal
		Pal.	
	kp	kpʲ	**ʔ**
	ŋmgb		
	kpŋm		
			h
	ŋm		
	ɣ **x**		

Table 6. Yélî Dnye vowels.

	Front		Central		Back	
	Oral	Nasal	Oral	Nasal	Oral	Nasal
Close	i	ĩ	ɨ		u	ũ
Near-close	e		ə	ə̃		
Open-mid	ɛ	ɛ̃			ɔ	ɔ̃
(Near-)open	æ		ɐ	ã	ɑ	ɑ̃

2.3 The framework

In order to compare the onset and development of canonical babble produced by Tseltal and Yélî children to each other and also to other children discussed in the literature (e.g., Lee et al. 2017; Cychosz et al. in press.), we calculated their canonical proportion (CP; see formula). To make this calculation, we looked per child at what percentage of their vocalizations consisted of CB. Lee et al. (2017) used a similar measure to quantify the amount of CB, but looked at canonical productions by syllables, a measurement they called the canonical babbling ratio (CBR). In contrast, we looked at CB by whole vocalization, in line with the approach of Cychosz et al. (in press), hence the difference in terminology. To calculate the CP, in the present study any vocalization containing one or more canonical syllables was counted as an instance of canonical babble, and we subsequently divided the number of canonical babble vocalizations by the total number of vocalizations, while excluding vocalizations comprised solely of laughing or crying (Table 2).

Formula for calculating CP:

$$\frac{\#voc\text{'}s\ with\ canonical\ babbling}{\#voc\text{'}s\ with\ canonical\ babbling + \#voc\text{'}s\ without\ canonical\ babbling}$$

We deployed VMS as a measure of consonant development, though we make some adjustments to McCune and Vihman's (2001) original measure in order to accommodate our cross-sectional, daylong recording dataset. Our lack of longitudinal data removes our ability to observe individual children's development. However, the daylong recording collection affords us more data per individual and the production setting is highly natural. We define VMS as follows: If a Tseltal child produced 10+ realizations of a consonant within their total 45 randomly sampled minutes, or if a Yélî child produced 5+ realizations of a consonant within their total 22.5 randomly sampled minutes, then we consider the child to have acquired the VMS for that consonant. A visual aid of this process using hypothetical data is shown in Figure 1. On the basis of vocalization rates in the data from Casillas et al. (2020a) and (2020b) we had anticipated that our random clip sampling would typically yield 100+ vocalizations per child, a lower estimate for what is needed to secure stable VMS scores (Vihman et al. 1985; Vihman et al. 1994; Vihman personal communication). As noted below, there were a few children for whom we found fewer than 100 vocalizations in their random clips, namely 1 Yélî Dnye and 3 Tseltal-acquiring children.

We analyzed the number of VMS consonants acquired per child per language. Following McCune and Vihman (2001), we collapsed voiced and voiceless variants of produced consonants, as this distinction is often not yet mastered by young children (Eilers et al. 1984) and is also not contrastive in either language. We also only counted supraglottal consonants toward children's total VMS score, following prior work: Glottals and glides already occur frequently in the early part of children's first year while supraglottal consonants only begin to occur around 6-8 months (McCune & Vihman 2001).

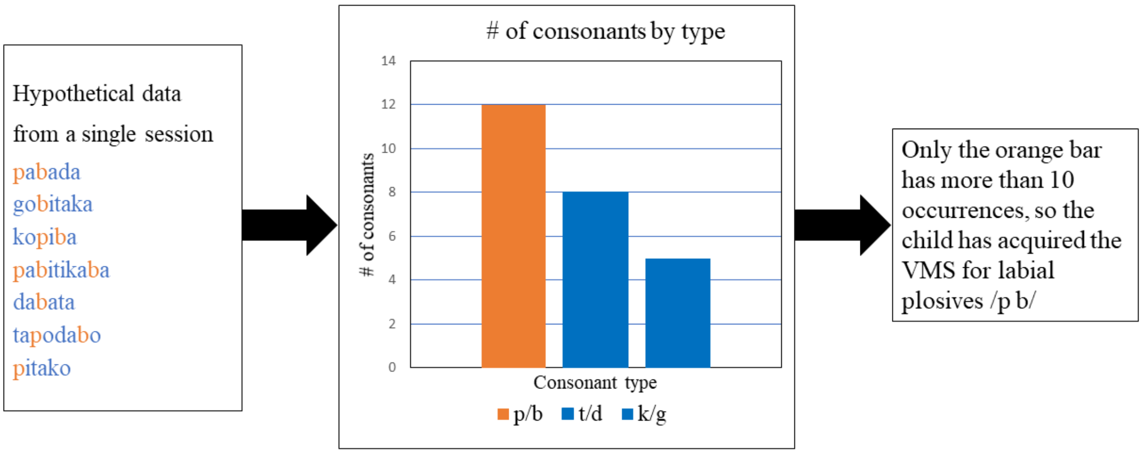

Figure 1. Illustration of how VMSs are derived from the data. Note again that, because of the difference in the amount of annotated data between corpora, we define 5 occurrences of a consonant as a VMS in the case of Yélî Dnye, instead of 10 occurrences as we do for Tseltal.

3. Results

We modeled our two dependent measures—canonical proportion and the number of VMSs produced by each child—using a linear regression with fixed effects of age (numeric, in months), language (Tseltal/Yélî Dnye), and their interaction. Because we only had one datapoint per child (e.g., the number of VMSs that child displayed), we were unable to include a random effect of child. The analysis was conducted in R (R Core Team, 2018) using lme4 (Bates et al. 2015), and all plots were generated with ggplot2 (Wickham, 2009). We first review findings regarding canonical babble, then findings for VMS acquisition.

3.1 Canonical babble

All Tseltal and Yélî speaking children 8 months and older used canonical babbles, consistent with previous findings (Oller 1980; Oller et al. 1998; Oller 2000; Lee et al. 2017). Furthermore, and also consistent with other cross-linguistic data, both populations of children had a CP greater than 0.15 after age 0;10 (Lee et al. 2017; Cychosz et al. in press); CP less than 0.15 might indicate delay in development (Oller et al. 1995). Differences in CP between the youngest and oldest children in each corpus were apparent (Figure 2), with a more gradual increase apparent among the Tseltal children. Note, however, that the youngest children in our Yélî sample are older than the youngest children in the Tseltal sample, and as such already produce vocalizations with canonical babble quite frequently.

A linear regression of CP revealed a significant positive effect of age ($p<.01$), no significant effect of language ($p=.15$) and no age-by-language interaction ($p=.16$). In other words, while the proportional amount of canonical babble increased with age overall, there is no evidence for differing developmental rates per language group, neither overall nor specifically for age-related increase between languages.

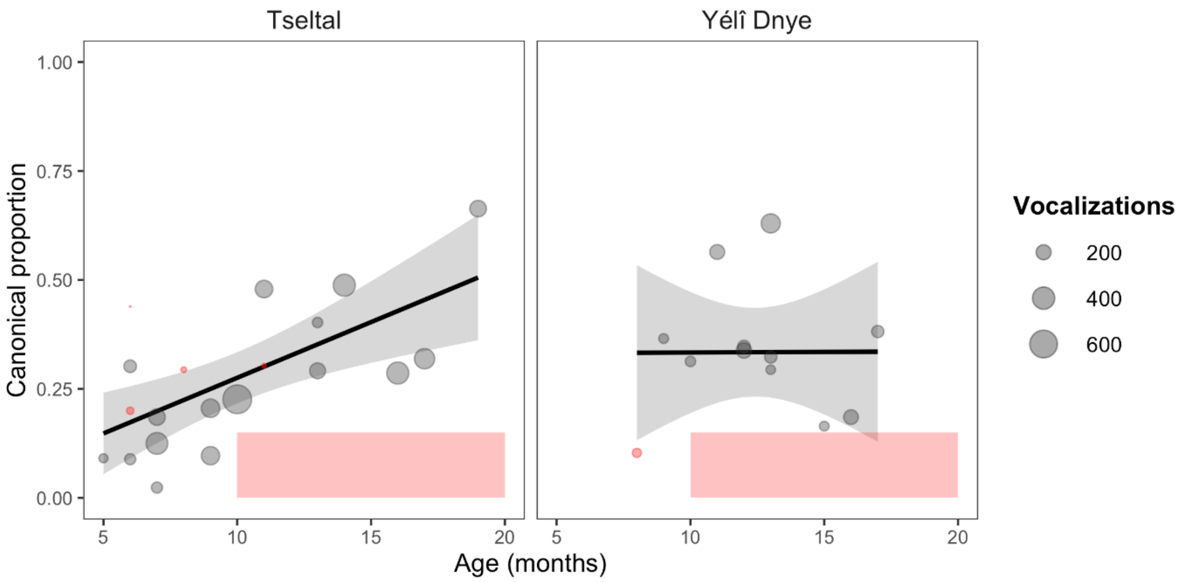

Figure 2. CP of Tseltal and Yélî children. Scores falling in the shaded red region would be unexpected, based on the benchmark of 0.15 CP at age 0;10 (Oller et al., 1995; Lee et al., 2017; Cychosz et al., in press). Points in red indicate fewer than 100 vocalizations found.

Table 7. Output of the CP regression analysis.

Coefficients:				
Term	**Estimate**	**Std. Error**	**Statistic**	**p-value**
(Intercept)	0.02	0.082	0.243	0.81
Age in months	0.026	0.026	3.339	<.01
Language	0.311	0.211	1.474	0.152
Age in months:Language	-0.025	0.017	-1.478	0.151

3.2 VMS consonants

Prior work suggests that Western, English-acquiring children typically acquire 2 VMS consonants between 9 and 14 months (McCune & Vihman 2001); VMS scores not expected on this benchmark (i.e., < 2 VMS from age 9 months and beyond) would fall within the shaded red region indicated in Figure 2. We see that, while most Tseltal and Yélî Dnye showed VMS scores aligning with results based on English, this was not the case for all children: 3 Tseltal-learning children and 5 Yélî Dnye-learning produced fewer VMS consonants than has been attested in their English-acquiring same-aged counterparts. The data used for the graphs in Figure 3 are shown in Appendix Table 1 and Appendix Table 2.

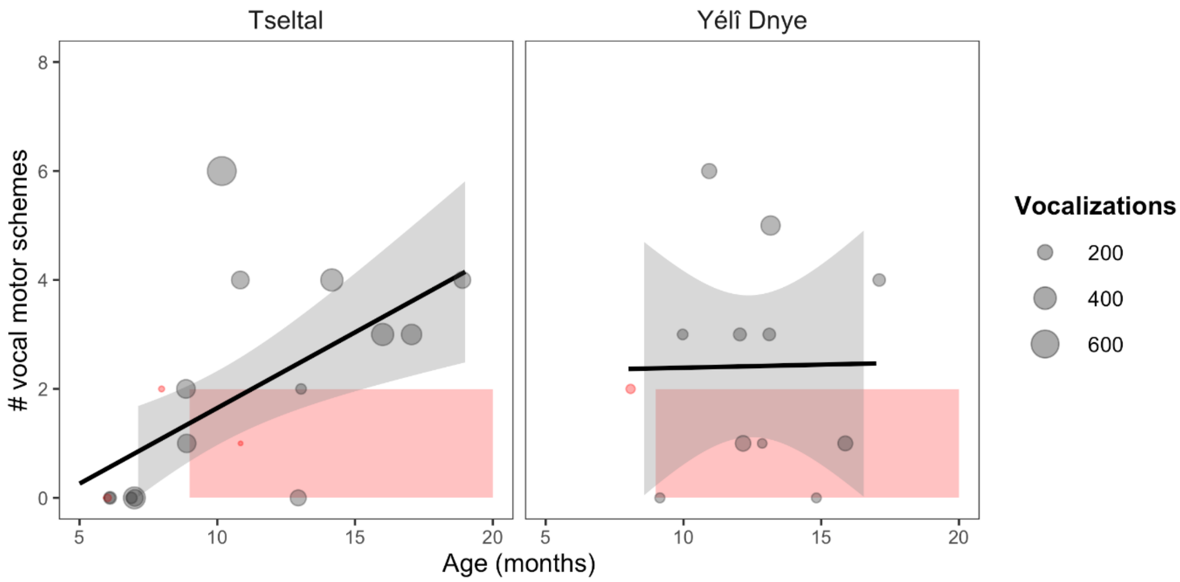

Figure 3. VMS count of Tseltal and Yélî children; the red area indicates <2 VMSs beyond 9 months. Points in red indicate fewer than 100 vocalizations found.

The most common VMS consonants produced by children in both samples were [t/d] and [m] (Table 8). The difference in distribution of the plosives in the two languages shows an interesting pattern: While in both cases [t/d] is the most prevalent type of plosive (in line with the hypothesis that the alveolar place of articulation is universally less marked (Paradis & Prunett 1991; Tsuji et al. 2015), in Tseltal the second most prevalent is [p/b] with few instances of [k/g], while in Yélî Dnye [k/g] is the second most prevalent plosive type and there are few instances of [p/b]. The graphs for the total number of phones used across tokens analyzed in each language are shown in Appendix Figure 4.

Table 8. Distribution of VMS consonants acquired by language sample.

	p/b	t/d	k/g	m	n	l
Tseltal	7	8	2	8	5	3
Yélî Dnye	3	7	6	7	2	4

A linear regression of VMS counts revealed a significant positive effect of age (p<.01), no significant effect of language (p=.20) and no age-by-language interaction (p=.27). In other words, while the number of VMSs acquired increased with age overall, there is no evidence for different patterning by language group, neither overall nor specifically for age-related increase between languages.

Table 9. Output of the VMS regression analysis.

Coefficients:				
Term	**Estimate**	**Std. Error**	**Statistic**	**p-value**
(Intercept)	-1.128	1.002	-1.125	0.27
Age in months	0.278	0.093	2.987	<.01
Language	3.404	2.565	1.327	0.195
Age in months:Language	-0.266	0.208	-1.281	0.211

4. Discussion

Children living in the Tseltal and Yélî communities are infrequently directly spoken to, yet prior work shows no apparent delay in their early linguistic development (Casillas et al. 2020a; b). This finding seems counterintuitive, as CDS has previously been strongly associated with with faster-growing receptive and productive vocabularies (Hart & Risley 1995; Hoff 2003; Shneidman & Goldin-Meadow 2012; Weisleder & Fernald 2013). This apparent discrepancy may lie in the fact that prior work on Tseltal and Yélî Dnye used vocal maturity measures that are, in fact, fairly robust to environmental variation. In the current study, we investigated whether there was also no evidence of delay on a measure of early phonological development that has been shown to relate to early lexical development: Vocal motor schemes. We also tested whether the initial finding of non-delayed CB onset would hold up with a larger sample of children than was studied previously.

We predicted that, replicating prior work, CB development would show no delays in either corpus (Casillas et al. 2020a; b; Cychozs et al. in press). Following prior results showing no delays in the emergence of single- and multi-word utterances (Casillas et al. 2020a; b), we further hypothesized that VMS acquisition by Yélî and Tseltal children would be on par with previous results from English-acquiring children: at least 2 VMS consonants emerging between 9 and 14 months. We also tentatively predicted that Yélî children would acquire their VMS consonants slightly later than Tseltal children because the Yélî consonantal inventory is large, complex, and acoustically packed (Levinson in press).

4.1 Canonical babbling

We looked at the development of canonical babbling produced by Tseltal and Yélî children by calculating the proportion of canonical babbling in their total vocalizations. Prior work has shown that most children start using canonical babbling after 8 months, with canonical babble making up at least 0.15 of their total syllables by 0;10 (Oller 1980; Oller et al. 1998; Oller 2000; Lee et al. 2017; Cychosz et al. in press). We found that both Tseltal and Yélî children have indeed already surpassed this benchmark of a CP of 0.15 at 7 months, in line with the cross-linguistic findings of Cychosz et al. (in press); this is significantly earlier than the 10 months reported by Lee et al. (2017), which might be because CP is a less fine-grained measure than CBR. Then, in line with prior work on these populations (Casillas et al. 2020a; b) and with our own predictions, the onset and use of canonical babbling in these two communities have a similar developmental path to that found in children from other language communities, despite the relatively infrequent rate of directed speech in the communities of study.

4.2 Vocal motor schemes

In comparison to what has been attested for English-acquiring children (McCune & Vihman 2001), we found that some Tseltal and Yélî Dnye-learning children showed fewer VMSs than expected. That is, of the children 9 months and older in each corpus, 3 out of 11 Tseltal children (ages 0;9, 0;11, 1;11) and 5 out of 11 Yélî children (ages 0;9, 1;0, 1;1, 1;3, 1;4) produced fewer than 2 VMS consonants. While relatively more Yélî children fell into this unexpected scoring range, and despite the differences in phonological inventory between the languages, we found no evidence for an overall significant effect of language on VMS consonant production, nor any interaction effect of language and age on VMS count. In fact, the distribution of VMS consonants was similar between the two languages.

4.3 Infrequent CDS and phonological development: CB vs VMS

Keeping in mind that our results are preliminary, that some of our VMS scores are based on fewer vocalizations than hoped for, and that 2 VMS by age 0;9 is a less robustly attested benchmark than the benchmark we use for CB (particularly with respect to daylong recording data), we now tentatively discuss the possibility that VMS is indeed sensitive to environmental variation and that, therefore, some children are likely to show slower VMS development in these two communities where CDS is relatively infrequent.

Prior research suggests that CB onset is more a meaningful benchmark for motor development than it is for language development (Vihman et al. 2009). Under this view, CB is a hallmark of the development of rhythmic motor tools that develop in the first year of life (Iverson, Hall, Nickel & Wozniak 2007), and develops as a result of both proprioceptive and auditory experience (Westermann & Miranda 2004; Guenther & Vladusic 2012) with syllable production, but doesn't itself require stable articulatory or phonological representations. This perspective—that CB demonstrates a motor skill that helps prepare the child for language, rather than an early step in language development—aligns with findings showing its cross-cultural and cross-linguistic developmental robustness (Oller 1980; Oller et al. 1998; Oller 2000; Lee et al. 2017; Cychosz et al. in press).

In contrast, VMS gives us insight into early, stable consonant productions that approximate consonant categories in the ambient language (Vihman 1993), and so may be expected to link more tightly to early linguistic representations and, thereby, early productive vocabulary. Over the first year of life, we know that children's ambient language environment comes to shape how they perceive speech (Werker & Tees 1984; Monnin & Lœvenbruck 2010). In fact, CDS (and not just ambient speech in general) has been proposed to facilitate early phone discrimination (Kuhl et al. 2003; Kuhl 2007), and so may thereby help children acquire phonological categories sooner. In production, we can also imagine that the elicitation of child-produced speech by caregivers engaging in interactive CDS may result in children getting more frequent practice in targeting more adult-like phonological forms (Kuhl & Meltzoff 1996; Kuhl 2007). If we understand VMS as reflecting something about the initial stabilization of phonological categories rather than a simple practiced motor skill—e.g., via the coupling of articulatory parameters and auditory perception (Westermann & Miranda 2004)—we can predict that it is sensitive to the child's exposure to CDS and, thereby, linked to early lexical development (Vihman 1993; McCune & Vihman 2001; McGillion et al. 2017), which is also sensitive to CDS.

Total linguistic input in these communities is not at all sparse: On the basis of past work by Casillas and colleagues (2020a; b), we can also say that there is a great deal of other ambient speech present in these two communities for children to learn from beyond CDS, much of it directed to other children within earshot of the target children. If future work consistently finds larger VMS counts in linguistic communities where children experience not just a lot of linguistic input, but a lot of directed linguistic input, it would stand as stronger evidence that speech directed to the target child, and not exposure to speech in general (nor CDS in general), shapes phonological development in this transitional period between babble and first words. Until we gain greater clarity of the relation between VMS and language environment from further work, the current findings should only be taken as preliminary.

4.4 Further research

The current study has a number of important limitations. While our sample size is an improvement over previous work done in these communities (Casillas et al. 2020a; b), it is still quite limited considering our cross-sectional design and broad age span (0;5-1;8). We were also unable to achieve at least 100 spontaneous vocalizations for every child with our random sampling technique, which means some individual estimates may not be as stable as hoped. Because we had no measure of CDS input rate for each child in the study, we were also only able to compare individual productions patterns to community-wide input patterns. Individual input rate estimates would allow us to investigate whether VMS production relates to directed input *within* each community, regardless of the benchmark for English.

Finally, while phonetically trained, our transcriber did not speak the target languages and so did not benefit from the contextual and lexical information that could have rendered the transcriptions closer to what would be heard by a native annotator. Future work linking VMS to individualized input rate estimates and vocabulary size in these and comparably designed corpora from other languages and cultural communities are need to clarify the results presented in the current study.

5. Conclusion

We find that, while Tseltal and Yélî children have a similar developmental trajectory to English-acquiring children in their production of canonical babble, they are more variable than expected in their production of VMS consonants. This phenomenon could be influenced by the fact that CDS is infrequent in Tseltal and Yélî communities, but future work should more closely examine individual differences within the community and the cross-linguistic applicability of VMS counts as a benchmark before coming to a strong conclusion that, like lexical development, early consonant development is impacted by CDS rate.

Acknowledgements

We are deeply grateful to the participating communities, families, and children represented in these datasets. We thank Rebeca Guzmán López, Humbertina Gómez Pérez, Juan Méndez Girón, Taakêmê Ńamono, Ndapw:éé Yidika, and Y:aaw:aa Pikuwa for the ir help in corpus creation and initial annotation. We also thank the National Research Institutes of PNG and Milne Bay Province administration for their support. This work was funded by an NWO Veni Innovational grant to MC (275-89-033). Last but not least, we thank Dr. Paula Fikkert for her supervision and support

Appendix

Table 1. Overview of VMS consonant use in the Tseltal dataset.

Aclew ID	Age in months	p/b	t/d	k/g	m	f	n	l	r	s/z	ʃ/ʒ	ŋ	x/ɣ	Total VMS	Total consonants	Total vocalizations	tʃ/dʒ	w	j	ʔ	h
6964	5	-	10	-	-	-	-	-	-	-	-	-	-	1	28	100	-	-	-	-	-
4935	6	-	-	-	-	-	-	-	-	-	-	-	-	0	33	54	-	-	-	-	19
6107	6	-	-	-	-	-	-	-	-	-	-	-	-	0	92	152	-	11	-	23	53
1991	6	-	-	-	-	-	-	-	-	-	-	-	-	0	34	81	-	-	-	-	-
8179	6	-	-	-	-	-	-	-	-	-	-	-	-	0	19	127	-	-	-	-	-
1188	7	-	-	-	-	-	-	-	-	-	-	-	-	0	61	233	-	15	-	-	39
456	7	-	-	-	-	-	-	-	-	-	-	-	-	0	30	381	-	-	-	-	-
2109	7	-	-	-	-	-	-	-	-	-	-	-	-	0	6	128	-	-	-	-	-
3214	8	-	-	-	10	-	10	-	-	-	-	-	-	2	34	68	-	-	-	-	-
7273	9	14	-	-	-	-	-	-	-	-	-	-	-	1	44	279	-	-	-	-	14
7439	9	-	10	-	11	-	-	-	-	-	-	-	-	2	80	284	-	-	-	-	14
3591	10	15	31	40	12	-	21	11	-	-	-	-	-	6	223	643	-	13	20	-	52
909	11	10	-	-	-	-	-	-	-	-	-	-	-	1	46	61	-	10	-	-	14
8787	11	64	107	14	16	-	-	-	-	-	-	-	-	4	260	260	-	15	13	-	10
684	13	21	-	-	14	-	-	-	-	-	-	-	-	2	85	116	-	-	-	-	23
9415	13	-	-	-	-	-	-	-	-	-	-	-	-	0	82	218	-	-	-	14	51
7326	14	17	120	-	-	-	86	52	-	-	-	-	-	4	374	392	-	15	23	14	26
6028	16	-	15	-	74	-	11	-	-	-	-	-	-	3	183	391	-	10	-	10	26
5147	17	-	24	-	18	-	-	15	-	-	-	-	-	3	198	335	-	-	16	-	89
9592	19	32	107	-	35	-	21	-	-	-	-	-	-	4	266	235	-	11	12	-	20

Table 2. Overview of VMS consonant use in the Yélî Dnye dataset.

Aclew ID	Age in months	p/b	t/d	k/g	m	f	n	l	r	s/z	ʃ/ʒ	ŋ	x/ɣ	Total VMS	Total consonants	Total vocalizations	tʃ/dʒ	w	j	ʔ	h
1143	8	-	6	-	-	-	-	13	-	-	-	-	-	2	24	99	-	-	-	-	-
8018	9	-	-	-	-	-	-	-	-	-	-	-	-	0	43	109	-	-	6	-	21
8274	10	-	15	-	22	-	-	19	-	-	-	-	-	3	81	118	-	-	-	-	11
5623	11	22	28	5	54	-	20	11	-	-	-	-	-	6	209	197	-	23	21	5	17
3446	12	-	-	-	50	-	-	-	-	-	-	-	-	1	91	208	-	-	-	-	29
4624	12	-	16	6	-	-	5	-	-	-	-	-	-	3	72	153	-	-	8	-	31
2977	13	-	10	10	5	-	-	-	-	-	-	-	-	3	55	148	-	-	6	-	14
5139	13	27	40	11	8	-	-	12	-	-	-	-	-	5	384	303	-	83	23	12	165
7035	13	-	-	-	9	-	-	-	-	-	-	-	-	1	33	104	-	-	-	-	14
1504	15	-	-	-	-	-	-	-	-	-	-	-	-	0	15	107	-	-	-	-	12
9562	16	-	-	7	-	-	-	-	-	-	-	-	-	1	43	193	-	10	-	-	13
6934	17	21	19	9	14	-	-	-	-	-	-	-	-	4	114	146	-	16	-	15	13

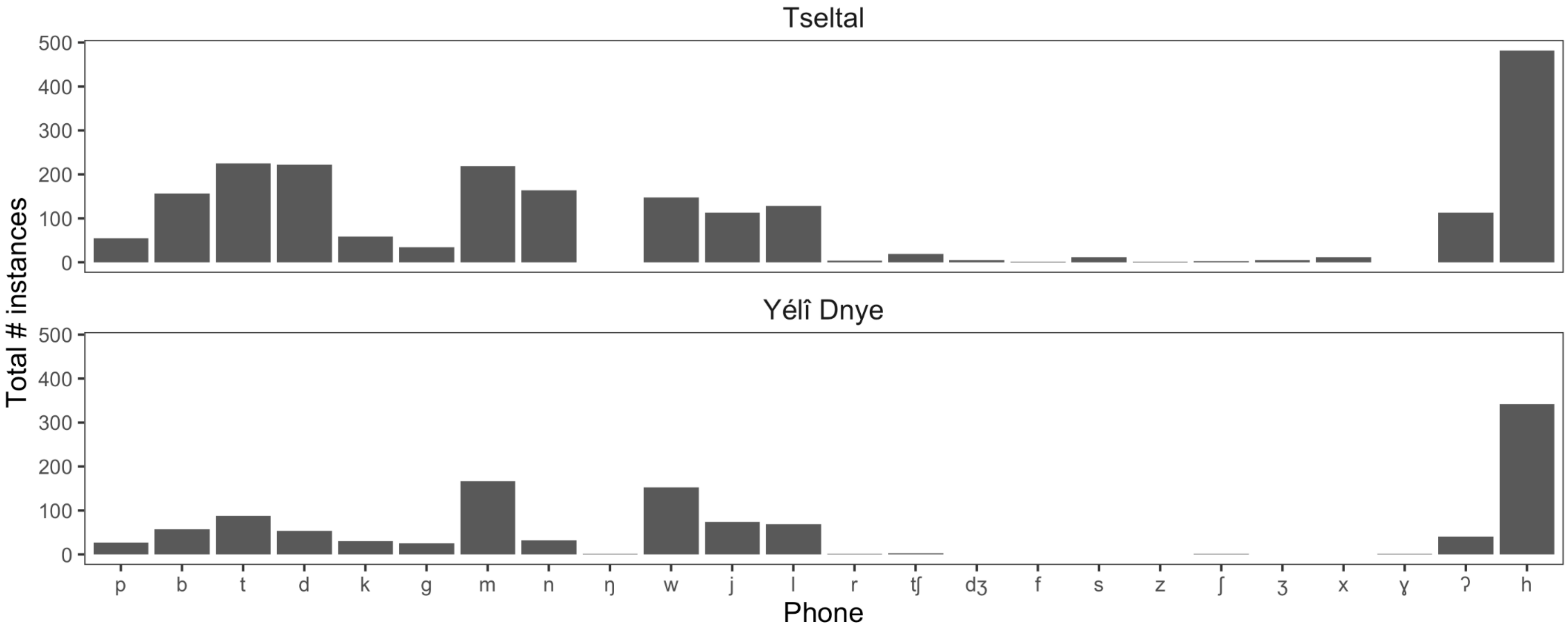

Figure 1. Total number of phones used across tokens analyzed in each language. While the phone frequencies are overall similar, Yélî Dnye appears to have an overrepresentation of /m/ and /w/, and underrepresentation of /ʔ/ compared to Tseltal.

References

Bates, Douglas, Martin Mächler, Ben Bolker, & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1). 1–48. doi:10.18637/jss.v067.i01.

Brown, Penelope. 1998. Conversational structure and language acquisition: The role of repetition in Tzeltal adult and child speech. *Journal of Linguistic Anthropology*. 8(2). 197–221.

Brown, Penelope. 2011. The cultural organization of attention. In Alessandro Duranti, Elinor Ochs & Bambi B. Schieffelin (eds.), *Handbook of language socialization*, 29–55. Malden, MA: Wiley-Blackwell.

Brown, Penelope 2014. The interactional context of language learning in Tzeltal. In Inbal Arnon, Marisa Casillas, Chigusa Kurumada & Bruno Estigarribia (eds.), *Language in interaction: Studies in honor of Eve V. Clark*, 51–82. Amsterdam, NL: John Benjamins.

Brown, Penelope & Marisa Casillas. In press. Childrearing through social interaction on Rossel Island, PNG. In Alicia J. Fentiman & Mary Goody (eds.), *Esther Goody revisited: Exploring the legacy of an original inter-disciplinarian*. New York, NY: Berghahn.

Casillas, Marisa, Penelope Brown & Stephen C. Levinson. 2020a. Early language experience in a Papuan village. *Journal of Child Language* 1–23.

Casillas, Marisa, Penelope Brown & Stephen C. Levinson. 2020b. Early language experience in a Tseltal Mayan village. *Child Development* 91(5). 1819–1835.

Casillas, Marisa, Penelope Brown & Stephen C. Levinson. 2017. *Casillas HomeBank corpus*.

Chavajay, Pablo & Barbara Rogoff. 1999. Cultural variation in management of attention by children and their caregivers. *Developmental Psychology* 35(4). 1079–1090. doi:10.1037/0012-1649.35.4.1079.

Cristia, Alejandrina, Marisa Casillas. Under review. Non-word repetition in children learning Yélî Dnye.

Cychosz, Margaret, Alejandrina Cristia, Elika Bergelson, Marisa Casillas, Gladys Baudet, Anne S. Warlaumont, Camila Scaff, Lisa Yankowitz & Amanda Seidl. In press. Vocal development in a large-scale crosslinguistic corpus. *Developmental Science*, XX.

de Boysson-Bardies, Bénédicte & Marilyn May Vihman. 1991. Adaptation to language: Evidence from babbling and first words in four languages. *Language*. 67(2). 297–319.

DePaolis, Rory A., Marilyn May Vihman & Tamar Keren-Portnoy. 2011. Do production patterns influence the processing of speech in prelinguistic infants?. *Infant Behavior and Development* 34(4). 590–601.

Eberhard, David. M., Gary. F. Simons & Charles D. Fennig (eds.). 2020. *Ethnologue: Languages of the World.* Twenty-third edition. Dallas, Texas: SIL International. Online version: www.ethnologue.com.ru.idm.oclc.org.

Eilers, Rebecca E., D. Kimbrough Oller & Carmen R. Benito-Garcia. 1984. The acquisition of voicing contrasts in Spanish and English learning infants and children: A longitudinal study. *Journal of Child Language* 11(2). 313–336.

Fikkert, Paula & Claartje Levelt. 2008. How does Place fall into Place?. *Contrast in phonology: Theory, perception, acquisition* 231–268.

Gaskins, Susan & Ruth Paradise. 2010. Learning through observation in daily life. In David F. Lancey, John Bock & Susan Gaskins (eds.), *The anthropology of learning in childhood*, 85–117. Walnut Creek, CA: Rowman AltaMira Press.

Guenther, Frank H. & Tony Vladusich. 2012. A neural theory of speech acquisition and production. *Journal of neurolinguistics* 25(5). 408–422.

Hart, Betty & Todd R. Risley. 1995. *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Paul H. Brookes.

Hoff, Erika. 2003. The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development* 74(5) 1368–1378

International Phonetic Association. 2020. doi: www.internationalphoneticassociation.org/content/full-ipa-chart

Iverson, Jana M., Amanda J. Hall, Lindsay Nickel. & Robert H. Wozniak. 2007. The relationship between reduplicated babble onset and laterality biases in infant rhythmic arm movements. *Brain and Language* 101(3). 198–207.

Jakobson, Roman. 1968. Child language, aphasia and phono-logical universals. The Hague, The Netherlands: Mouton.

Kuhl, Patricia K. & Andrew N. Meltzoff. 1996. Infant vocalizations in response to speech: Vocal imitation and developmental change. *The Journal of the Acoustical Society of America* 100(4). 2425–2438.

Kuhl, Patricia K., Feng-Ming Tsao & Huei-Mei Liu. 2003. Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences* 100(15). 9096–9101.

Kunnari, Sari. 2003. Consonant inventories: a longitudinal study of Finnish-speaking children. *Journal of multilingual communication disorders* 1(2) 124–131.

Lee, Chia-Cheng, Yuna Jhang, George Relyea, Li-mei Chen & D. Kimbrough Oller. 2017. Babbling development as seen in canonical babbling ratios: A naturalistic evaluation of all-day recordings. *Infant Behavior and Development* 50. 140-153.

Lee, Sue Ann S., Barbara Davis & Peter MacNeilage. 2010. Universal production patterns and ambient language influences in babbling: A cross-linguistic study of Korean-and English-learning infants. *Journal of Child Language* 37(2). 293–318.

Levinson, Stephen C. 2007. A grammar of Yélı Dnye, the Papuan language of Rossel Island. In press.

Majorano, Marinella, Marilyn May Vihman & Rory A. DePaolis. 2014. The relationship between infants' production experience and their processing of speech. *Language Learning and Development* 10(2). 179–204.

McCune, Lorriane & Marilyn May Vihman. 2001. Early phonetic and lexical development. *Journal of Speech, Language, and Hearing Research* 44. 670–684.

McGillion, Michelle, Jane S. Herbert, Julian Pine, Marilyn May Vihman, Rory A. DePaolis,
Tamar Keren-Portnoy & Danielle Matthews. 2017. What paves the way to conventional language? The predictive value of babble, pointing, and socioeconomic status. *Child Development* 88(1) 156–166.

Monnin, Julia & Hélène Lœvenbruck. 2010. Language-specific influence on phoneme ` development: French and Drehu data. *Proceedings of Interspeech 2010* 1882-1885. Makuhari, Japan, September 26–30, 2010.

Ochs, Elinor & Bambi B. Schieffelin. 1984. Language acquisition and socialization: Three developmental stories and their implications. *Culture theory: Essays on mind, self, and emotion* 276–322. Cambridge University Press.

Oller, D. Kimbrough. 1980. The emergence of the sounds of speech in infancy. *Child Phonology* 1. 93–112.

Oller, D. Kimbrough. 2000. *The emergence of the speech capacity.* Lawrence Erlbaum Associates.

Oller, D. Kimbrough, Sharyse Levine, Alan B. Cobo-Lewis, Rebecca E. Eilers & Barbara Zurer Pearson. 1998. Vocal precursors to linguistic communication: How babbling is connected to meaningful speech. *Exploring the speech-language connection* 8. 1–25.

Oller, D. Kimbrough, Rebecca E. Eilers, Devorah Basinger, Michelle L. Steffens & Richard Urbano. 1995. Extreme poverty and the development of precursors to the speech capacity. *First Language* 15(44). 167–187.
Paradis, Carole, & Jean-François Prunet. 1991. The special status of coronals: Internal and external evidence. *Phonetics and Phonology* 2. San Diego: Academic Press.
Polian, Gilles. 2013. *Gramática del tseltal de Oxchuc*. Centro de Investigaciones y Estudios Superiores en Antropología Social, México.
R Core Team. 2018. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/
Rogoff, Barbara, Jayanthi Mistry, Artin Göncü & Christine Mosier. 1993. Guided participation in cultural activity by toddlers and caregivers. *Monographs of the Society for Research in Child Development*, 58(8). I–179. doi:10.2307/1166109.
Shneidman, Laura A. & Susan Goldin-Meadow. 2012. Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science* 15(5). 659–673.
Tsuji, Sho, Reiko Mazuka, Alejandrina Cristia & Paula Fikkert. 2015. Even at 4 months, a labial is a good enough coronal, but not vice versa. *Cognition* 134. 252–256.
Vihman, Marily May. 1993. Vocal motor schemes, variation and the production-perception link. *Journal of Phonetics* 21. 163–169.
Vihman, Marilyn May & Bénédicte de Boysson-Bardies. 1994. The nature and origins of ambient language influence on infant vocal production and early words. *Phonetica*, 51(1–3). 159–169.
Vihman, Marilyn May, Rory A. DePaolis & Tamar Keren-Portnoy. 2009. A dynamic systems approach to babbling and words. *British Journal of Psychology* 108(1). 1–27.
Vihman, Marilyn May, Marlys A. Macken, Ruth Miller, Hazel Simmons & Jim Miller. 1985. From babbling to speech: A re-assessment of the continuity issue. *Language* 61(2). 397–445.
Vihman, Marily May, Edwin Kay., Bénédictie de Boysson-Bardies, Catherine Durand, & Ulla Sundberg. 1994. External sources of individual differences? A cross-linguistic analysis of the phonetics of mothers' speech to 1-yr-old children. *Developmental Psychology*, 30(5), 651–662.
Wickham, Hadley. 2009. *Ggplot2: Elegant graphics for data analysis.* New York, NY: Springer-Verlag. Retrieved from http://ggplot2.org
Werker, Janet F. & Richard C. Trees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7(11). 49–63.
Westermann, Gert & Eduardo Reck Miranda. 2004. A new model of sensorimotor coupling in the development of speech. *Brain and language* 89(2). 393–400.
Weisleder, Adriana & Anne Fernald. 2013. Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science* 24(11). 2143–2152.