Sketching Infant Grammars

Clifton Pye

The University of Kansas

Abstract

Hippolyte Taine and Charles Darwin initiated a tradition of publishing short sketches of children's language. The tradition was lost as children stopped acquiring indigenous languages in North America and psychologists took over the study of children's language. This paper looks to the tradition of publishing sketch grammars of children's language as an important addition to the documentation and maintenance of endangered languages. Children's language constitutes the essential link in the transmission of language from one generation to the next. The paper discusses how sketches grammars of child language can contribute to the linguistic training of a language community, language preservation and acquisition theory.

Keywords: language documentation, child language, community investigators

1 Introduction

Hippolyte Taine (1877) and Charles Darwin (1877) initiated a tradition of publishing short sketches of children's language. Anthropologists and linguists maintained this tradition through the middle of the last century (c.f. Burling 1959; Casagrande 1948; Chamberlain 1890; Dennis 1940; Kroeber 1916; Sapir 1929). The tradition was lost as children stopped acquiring indigenous languages in North America and psychologists began to dominate the study of children's language. There is presently a specialized literature on children's language that is dominated by theoretical debates on the acquisition of English (c.f. Ambridge & Lieven 2011; Ingram 1989).

This paper attempts to revive the traditional practice of documenting children's language as an important addition to the documentation and revitalization of endangered languages. Children's language constitutes the essential link in the transmission of language from one generation to the next. In many communities, this link is broken before the community is aware that their children are no longer acquiring their language. Older children are heard speaking the language outside the home while toddlers inside the home have stopped speaking the language.

Language revival programs have gone to great lengths to teach their language to children. An understanding of how children acquire language naturally offers important lessons about the optimum way to share language with children. The most important lesson from the documentation of children's language in many cultures is how parents and grandparents express their affection for children through language and gesture. A second essential lesson is that children acquiring a language naturally make mistakes that their caregivers ignore. This aspect of children's language shows that caretakers focus on dealing with the child rather than the child's language.

It will only be possible to document how children acquire most of the world's languages until the end of this century (Pye 2020). At present, we only have documentation of children's language for approximately 2% of the world's languages and most of these languages are spoken in Europe (Kidd & Garcia 2022; Lieven & Stoll 2010). Reviving the traditional child language sketch enables investigators who are not language acquisition specialists to help document the acquisition of the world's endangered languages. This paper outlines a procedure for incorporating the documentation of child language into language documentation and language revitalization projects. The sketch model of child language documentation adopts simple procedures that can be used without specialized training or equipment.

2 The place of child language in language documentation and revitalization

Children's language deserves to be documented like any language variety. Children produce a slower, abbreviated version of the adult language. As Yuen Ren Chao (1951) stated, children's language is "language in the making". Recordings of children's language allow us to see the parts of the adult language that children select to form their early utterances. The diversity across adult languages ensures a diversity of child languages. Elements like the passive voice or ergative marking that are rare in most languages may appear frequently in the speech of children acquiring a few languages. Rather than mechanically applying a universal bootstrap in the form of principles and parameters (Snyder 2021) or imitation (Behrens 2021), children in the words of Hippolyte Taine (1877) are original geniuses who create their own language varieties.

Languages disappear when parents no longer transmit their first language(s) to their children. The transmission chain is surprisingly fragile; it can be broken in a short period of time even in languages with many speakers due to the stigma attached to speaking minority languages (Campbell 2017). Two-year-olds and their mothers are the more endangered speakers of endangered languages. Faustino Montes (2022) reports an incident in which a young mother chastised her mother in Totonac because her one-year-old son embarrassed her by responding to a taxi driver in Totonac. Now, only a handful of Totonac-speaking children enter preschool in Mexico.

Documenting the cultural modes of language transmission warrants critical attention in any language documentation project. Societies have evolved different childcare practices to suit a myriad of economic and environmental contingencies. It is important to recognize the range of childcare practices as they establish the contexts for language transmission. In some societies mothers leave toddlers to play on the ground; in others, mothers carry toddlers on their backs, their sides or in front. In certain societies caretakers use purpose-built contraptions to push children along paved surfaces. Children are surrounded by different family members depending on the activities of their parents and family structure. In some cultures, the baby is the center of attention, in some the oldest child gets top billing, while in some the parent or grandparent takes the center stage. Documenting these practices is key to understanding the contexts of language transmission.

It is vitally important to preserve some record of language transmission where this is still possible. Central to the language transmission process is the forging of a language identity. Caretakers determine their children's language identity based on their vision of their children's future. Two-year-old children are language sponges who soak up the language(s) in their environment. They note the smallest nuances in the adult language(s), and are particularly sensitive to the manner in which older speakers address them. If their parents only use the endangered language when speaking to other adults and a majority language in speaking to two-year-olds then two-year-olds will infer that they are majority language speakers and not speakers of the endangered language. Two-year-olds adopt their parents' vision of themselves and use the language(s) addressed to them even when their parents do not use it with each other.

This is one place where a project that documents child language can play an important role in revitalizing endangered languages. A child language project works with a child's family as much as it documents a child's language. The investigator can give caretakers more confidence in using their first language and reassure nervous parents that their children can learn more than one language at a time. The use of recorders and computers adds prestige to the language by demonstrating that it can be written on computers, and that it has rules of grammar that children learn.

3 Recording and transcribing child language

It is essential to consider the end goals of a child language project before making the first recording with a child. It is necessary to clearly explain the goals of a child language project to the child's family as well as the language community in order to obtain their consent and support for the project. It is also important to provide children and their families with information about the purpose of language documentation, and how their identities will be protected. They should be informed about the significance of their contributions as well. The end goals of the project should be achievable within the means and resources that are available within the community. Ultimately, this means

severely restricting the project so that its results can be returned to the child's family and community within a short period of time. Putting the needs of the families and community first is the basis for a community-oriented child language project.

Seeking permission from families and the language community to record two-year-olds requires careful consideration. Requesting informed consent from all participants, including the children, is a delicate process that is difficult to do without a deep understanding of the community and its language situation. Many minority communities have good reason to be suspicious of outsiders because of their experiences with strangers taking their resources or their children (Nolan 2020). For this reason, a project documenting child language is best left to members from the child's community, who can identify the conditions for doing research in the community, and who can introduce the research to the community.

Recording hour-long samples of a two-year-old interacting with their family provides the most information about the transmission process with the least effort. Each recording session should probably not extend longer than an hour in order not to exhaust the child or the patience of the child's family. Recordings can be made on different days in order to sample a wider array of contexts and situations. It is best to prepare for long periods of silence as these are a natural part of the life of families and contribute important information about the rate of interactions between family members.

Two-year-olds make ideal subjects because they are more talkative than one-year-olds and demonstrate the beginnings of language acquisition more than three-year-olds. If no two-year-old speakers are available, it may be necessary to record any child no matter what their age is if they are using their parents' first language. Even four or five-year-old children still display incomplete control of their first language.

The need to finish the project in as timely a manner as possible means that the investigator should record a small sample of child language, somewhere between three to five hours. I estimate that each hour that is recorded requires up to two weeks to fully process. Each additional recording results in a delay in getting the information to the community. A five-hour language sample would require approximately ten weeks to process. For community-oriented projects, such delays can provoke frustration with the investigators.

The main difficulty in processing language samples is the time required to transcribe them. This is the main reason for recording a limited number of samples as it reduces the stress of the transcribers who invariably feel as though they are behind from the start. A five-hour language sample will give the transcribers more confidence that they can accomplish the transcription phase of the project in a timely manner, and not feel discouraged by the mountain of recordings that awaits their attention.

There is evidence that even a modest language sample will document many important features of child language. An extensive literature exists on sampling techniques that can instruct the small sample approach (c.f. Fujiki & Brinton 1985; Gun

& Eisenberg 2015). Tomasello & Stahl's (2004) research suggests that a five-hour sample of child language would register a child's use of some rare constructions. A three to five-hour sample of child language will document important features of parent-child interaction as well as offer basic information on child's vocabulary, syllable structures, phonology, and morphosyntax (Pye 2020).

There are many reasons why documenting child language makes an excellent goal for community-oriented language documentation projects. The first is that child language documentation provides solid training in language documentation practices for community members. Community members have essential knowledge of their community's organization and have the best idea of how a language documentation project of any kind can further their community's endeavors. Community members know best how to explain the goals of the project to a child's family members. They may also have good ideas about how the results from the project can benefit the child's family and their community.

Community members know the culturally appropriate ways to interact with two-year-olds. They may know of certain routines in which two-year-olds interact with strangers or family members. They may know the right way to enter a family's home and who to talk to about the goals of the project. They may be able to contact family members initially in public spaces such as markets or bus stops in order to determine if they would be interested in participating in the project. Community members will also understand best how to gain the informed consent of the family for the project.

The community members, who I refer to as community investigators, will have good ideas for setting up the recording equipment so that it is protected from the children, family pets, and the weather. Making the recordings with children provides community investigators with some motivation to see the project through to completion. They will have witnessed the humorous interactions that take place between the family members as well as unusual events which provide them with targets to note in their transcriptions.

It is best to assume that community investigators only have a limited amount of time that they can devote to transcription. Anything that can be done to make the transcription phase more efficient will help ensure that all of the recordings get transcribed. I suggest using a minimal transcription technique that deprecates the addition of morpheme breakdowns and morpheme glosses at the transcription phase. The addition of tiers for morpheme breakdowns and glosses adds to the total transcription time and increases the possibility for error, especially for languages with incomplete linguistic analyses. As the Mixe example in (1a) demonstrates, the addition of morpheme breakdowns and glosses requires a sophisticated understanding of the language and training in morpheme glossing conventions. The minimal coding transcription in (1b) is more efficient. It is always possible to add morpheme breakdowns and glosses at a later stage in a project.

(1) Mixe (Pérez Martínez 2020: 38)
    a.   Example with morpheme breakdowns
        mkakyëxänëtëpaj?
        m-kay-këëx-'an-në-të-p='aaj
        2SG.I-eat-stop-IRR-now-PL-INC.I=INTER
        "Are they going to stop eating now?"
    b.   Minimal coding transcription
        mkakyëxänëtëpaj?
        "Are they going to stop eating now?"

The minimal transcription technique supports community-oriented language projects by minimizing the linguistic training necessary for community investigators. Morpheme breakdowns and morpheme glosses are unnecessary for community members who speak the language. Minimizing the transcription also helps to define goals for the project that can be accomplished rapidly. It is important to guard against importing goals that are incompatible with the needs of the language community. I suggest more realistic goals below.

I suggest using the ELAN transcription program (ELAN 2021) because it can be downloaded for free from the Max Planck Institute for Psycholinguistics and because it links the recording to the transcription. I use three tiers in ELAN for the children's utterances. The first tier contains a transcription of the child's utterance in a practical orthography developed for the adult language. The second tier contains the adult equivalent for the child's utterance that attempts to stay close to the words in the child's utterance. The third tier contains a translation of the adult utterance into a national language.

The interpretation of children's utterances is an imperfect art (c.f. Bloom 1973). The surrounding context and an intimate knowledge of child culture often provide clues to a child's intentions. The adult equivalent for the child's utterance should be as complete as necessary without adding extra information. The adult equivalent should contain fully inflected word forms that adults use in casual conversation. If necessary, a comment tier can be used to provide additional information about the context for the child's utterance and explain obscure references. More information about these transcription techniques is available on my webpage (http://pyersqr.org/minimal).

4 Goals for community child language documentation projects

An important goal for community-oriented child language documentation projects is to add to the community's linguistic infrastructure by training community investigators to record and transcribe language samples. A child language project provides practical training in recording conversations in natural environments as well as in using a

computer in order to transcribe the language sample. Conversations with two-year-olds typically involve short utterances with many repetitions. This makes the transcriber's job all the easier. The transcriber will also have to determine which utterances are audible as well as interpretable. This experience provides community investigators with the basic skills necessary for processing language samples of any kind. The investigators can apply their linguistic skills to other language projects in the community such as recording community meetings or producing language materials for schools.

Community investigators can also serve as language ambassadors who can explain the history and structure of their language to the community. In the first place, they can show community members how they process recordings and write their language on computers. They can help translate any documents that community members have between languages. Community investigators can explain the significance of their language to families with young children, and encourage families to maintain their language.

Community investigators can work with their community on the best way to preserve the language materials for the community's benefit. The investigators can devise terms for accessing the material with the community. These terms can then be used to archive the language materials on a remote server like Google Drive where other community members can access the material, and where the material will be protected from local disasters such as hurricanes and fires. Storage limits on Google Drive provide another reason to start a documentation project with a small number of recordings.

One goal that is important to community-oriented language documentation work is to think carefully about products that are useful to the child's family and community. These products should be accessible to family members without encumbering them with printed matter that will be tossed aside after a few days. This topic deserves greater attention from linguists working with endangered language communities. One idea is to think in terms of products that family and community members can access on their cellphones. It is easy to put together a short picture book of the child's day that contains pictures from the child's home accompanied by selected portions of the transcriptions. The book can be made into a pdf document and shared via social media with protections for the identities of the family members.

Perhaps the most critical information that comes from a project documenting child language is the evidence it provides of children's communicative competence in their home environment. It is important that this information be shared with local school teachers who are willing to incorporate the information into their lesson plans. Community investigators can share some details of the children's language with teachers such as their vocabulary and grammar. They can work with teachers to encourage children to use their home language in school, perhaps by devoting some time to letting children talk about an event in their home language. Community

investigators can inform teachers about the structure and history of the children's home language and help to dispel any ideas that the children's language is inferior to any other language.

5 Two short-term projects

Assuming that the language investigator only has a limited amount of time to devote to producing a study of child language, I suggest focusing on two short-term projects. Having concrete projects already in mind when beginning a child language project eases the burden of explaining the project to the children's families and the language community. Concrete goals also reassure the families that the project will not take up too much of their time. For these reasons, I suggest two initial projects: 1. Setting up an archive of the recordings and transcriptions, and 2. Producing a record of the child's lexicon.

5.1 Archiving recordings and transcriptions

Without any doubt, the recordings and transcriptions constitute the fundamental contribution of any project documenting children's language. They deserve the utmost care and respect. The investigators, families, and language community should understand that the primary goal of documenting child language is to preserve a record that the community can draw upon for centuries to come. The documentation will need to serve both immediate and future requests. For these reasons, the investigators and the community will need to plan on ways to preserve the documentation. Electronic documents have an illusion of permanency that is shattered the moment that a computer ceases to operate or a thumb drive fails. Therefore, the first rule of language documentation is to store copies of the documents in multiple locations. It is best to store at least one copy of the documents on a remote drive in order to avoid the effects of human and natural disasters. The construction of a language archive on a remote drive should begin a discussion about who will be able to access the archive and what level of protections will be given to the families that contributed the information to the archive. The material will be most useful if it is accessible in the cloud, but the internet brings a host of problems that include hacking and possible misappropriation of the material. The safety of the children and the community has to be a paramount concern when archiving recordings.

       The easiest way to set up an archive on a remote drive is to use one of the public storage drives such as Google Drive. Google Drive allots 15GB of free storage space, which is more than sufficient for storing small language samples. Google Drive lets the owner of the archive restrict access to the materials as they wish. The main disadvantage of creating a private language archive is that the community will bear sole responsibility for its maintenance. Should any disruption to the community occur, the

archive would be lost. Therefore, it might be wise to place a copy of the material in an institutional language archive. The CHILDES archive (MacWhinney & Snow 1985) is well-known in language acquisition research, but many other language archives now exist such as The Archive of the Indigenous Languages of Latin America (https://ailla.utexas.org), and the Documenting Endangered Languages Programme (https://www.eldp.net) at the School of Oriental and African Language. The Language in Time and Space webpage supplies a list of language archives around the world (https://lucian.uchicago.edu/blogs/langtimespace/resources/resourcesarchives/). The Open Language Archives Community offers information on the best current practice for digital archiving of language resources (https://www.language-archives.org/).

The families and the community should understand the intrinsic value of the materials that they archive. Language materials have an intellectual value that cannot be translated into immediate economic advantages. The language materials deserve the same level of copyright protection as any other intellectual creation. In keeping with the Creative Commons license agreement (https://creativecommons.org), anyone who cites material in the archive should acknowledge the authors of the material. Such citations would acknowledge the existence and intellectual contributions of the community.

An archive of child language materials provides a potential resource for anyone from the community who pursues an advanced degree in anthropology, psychology, linguistics, or speech. The material would provide plenty of data for a master's thesis or a doctoral dissertation. The community would be able to bank this material for future scholars from the community.

5.2 A child lexicon

One of the easiest and yet most significant studies that can be done with a minimal transcription of child language is a report on the child's vocabulary. The lexicon has been a neglected area of research on language acquisition partly due to a lack of appreciation for the degree of lexical variation across languages. Typologists have long emphasized the differences between words across languages (Haspelmath 2011; Gijn & Zúñiga 2014). One source of variation is the difference between word units defined prosodically and syntactically (Bickel & Nichols 2007; Mithun 2014).

The grammatical classes of words differ between languages and therefore define a basic research question that concerns how children establish the grammatical classes for words in the adult language. Languages may have distinct noun classes for alienable and inalienable nouns, grammatical gender, or arbitrary noun classes. Languages may have distinct verb classes for transitive, intransitive, stative, existential, and derived verb classes or arbitrary classes for verbs (Grinevald 2003). Languages may or may not have distinct classes of adjectives, articles, pronouns, prepositions, or positional verbs. The particle classes never cease to puzzle linguists. Publishing a

simple record of children's first words is already a major step in documenting the acquisition of a language (c.f. Nelson 1973; 2014; Rescorla 1980).

The first step in lexical analysis is to produce a lexical concordance for the children's words. A lexical concordance provides the utterance contexts for a child's production of each word in the recording. It is best to use the adult targets as the concordance keys in order to group together a child's attempts at each adult word. I have posted step-by-step instructions for producing a lexical concordance on my webpage (https://pyersqr.org/minimal/Processing.pdf). The first publication for any investigation of children's language should be a lexical concordance for each recording.

Adding codes for the lexical categories to the lexical concordance makes it possible to sort children's words by their lexical category. The sorting can be accomplished rapidly and accurately in any spreadsheet program. The example in (2) contains a column for lexical categories that were added to the concordance by hand. I use the code LD for one of the four verb categories specific to Xi'iùy (Northern Pame) and the code CN for common nouns (Berthiaume 2012). I distinguish common nouns from proper nouns and pronouns because the categories have different characteristics across languages.

(2) A lexical concordance for a Xi'iùy child with lexical category codes

Word	Category	Begin Time	Child's Utterance	Adult Target	Translation
daʔtsəlʔ	LD	03:54.9	kiʔil	daʔtsəlʔ	It bit
daʔuaɲ	LD	20:56.8	ane	daʔuaɲ	It threw it
danãs	CN	12:44.2	nas	danãs	Orange
dapaj	CN	05:15.5	paj	dapaj	Tomato
		16:23.2	paj	dapaj	tomato
		16:30.6	paj	dapaj	tomato
		07:21.6	paj	dapaj	tomato

Surprisingly little information is available on children's lexical development for indigenous languages. The Xi'iùy example shows a child's production of a language-specific category of verbs. The examples in (3) show children speaking the Mayan language Mam using intransitive verbs in contexts where English or Spanish speakers would use transitive verbs.

(3) Mam children's use of intransitive motion verbs as substitutes for transitive verbs
    (Pye 20717)

a. WEN (2;0.2)
ku? pe tu?n?.
=ma   pa   0-kub'   kape   t-u?n-a.
  REC   already   ABS3-go.down   coffee ERG2-by-ENC
'Did you already pick the coffee?' (lit. 'Did the coffee already go down by you?)

b. JOS (2;7)
ma? aʃ x wonn.
=ma   xaw- ʂ   xal   w-u?n-a.
  REC   rise-away   CL.NH ERG1-by-ENC
'I lifted it up' (lit. 'It rose up by me')

WEN produced the oblique agent phrase using the relational noun *t-uʔn-a* 'by you' in (3a), and JOS produced the oblique agent phrase *w-uʔn-a* 'by me' in (3b). Relational nouns take the same possessive morphology as common nouns but are used to express syntactic relations. The use of relational nouns to express agents in oblique phrases enables Mayan languages to use intransitive verbs to express events that involve agents and patients. Grammatical words such as relational nouns are an important part of the lexical acquisition process and constitute their own domain of study. Grammatical words are also one of the primary sources of diversity across languages.

The lexicon provides a wealth of information about the structure of children's words and its relation to the structure of words in the adult language. Two-year-olds typically omit some of the syllables from polysyllabic words. They will say *nana* in place of the adult word *banana*. A study of children's syllable omissions provides an interesting topic of research especially in light of the wide variation in word lengths across adult languages (Demuth 1996). The two-year-old lexicon also reveals information on what sounds the children produce, omit or use as substitutes. Two-year-old phonologies vary due to the structure of sounds in adult phonologies. This is another topic that deserves further investigation (Pye et al. 2017).

6 Conclusion

Documenting the speech of children acquiring endangered languages is an essential element of language documentation. In many cases it may not be possible to record children speaking an endangered language because the language is no longer being transmitted to children or the community is reluctant to involve their children in the study (Woodbury 2003). These cases underline the importance of recording children whenever possible as another opportunity may never come again. As yet, no record exists for children speaking the majority of human languages.

The goal of this paper is to encourage non-specialists and particularly members of indigenous communities to preserve a record of child language while this is still possible. Members of indigenous communities are in the ideal position to document the language transmission process in their communities. No specialized training is necessary beyond the knowledge necessary to record conversations and transcribe the conversations on a computer. Any level of transcription, no matter its imperfections, is preferable to no transcription whatsoever.

A modest child language documentation project is ideal for learning how to use a computer for linguistic transcription. Once the computer is set up with a transcription program, conversations with two-year-olds provide an interesting topic of research. It is always fun to hear what a two-year-old says next. They produce simple utterances with a lot of repetitions that help novice investigators learn the transcription process.

The resulting recordings and transcriptions provide a permanent contribution to the intellectual history of the community. The materials should be valued as such and protected by the same copyright restrictions that apply to other types of intellectual property. The material should only be used by permission of the children's family and community and should acknowledge the community in which the material was originally recorded. An archive of child language contributes to the linguistic infrastructure and provides the community with insights that can be used in creating language materials for schools and social media.[1]

7 Abbreviations

Abbreviations used in this text are as follows: ABS 'absolute'; CL 'noun classifier'; CN 'common noun'; 1 'first person'; 2 'second person'; 3 'third person'; ENC 'enclitic'; ERG 'ergative'; I 'independent conjugation'; INC 'incompletive aspect'; INTER 'interrogative'; IRR 'irrealis'; LD 'Xi'iùy verb class'; NH 'nonhuman'; PL 'plural'; REC 'recent tense'; SG 'singular'.

---

8 References

Ambridge, Ben, and Elena V.M. Lieven. 2011. *Child acquisition: Contrasting theoretical approaches.* Cambridge: Cambridge University Press.

Behrens, Heike. 2021. Constructivist approaches to first language acquisition. *Journal of Child Language 48*. 959-983.

Berthiaume, Scott. 2012. *A phonological grammar of Northern Pame*. SIL ebooks, 37. Dallas, TX: Summer Institute of Linguistics.

Bickel, Balthasar, and Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.). *Language typology and syntactic description. Vol. 3: Grammatical categories and the lexicon,* 2nd edition, 169-240. Cambridge: Cambridge University Press.

Bloom, Lois. 1973. *Language development: Form and function in emerging grammars.* Cambridge, MA: MIT Press.

Burling, Robbins. 1959. Language development of a Garo and English speaking child. *Word 15*. 45-68.

Campbell, Lyle. 2017. On how and why languages become endangered. *Language 93*. e224-e233.

Casagrande, Joseph B. 1948. Comanche baby language. *International Journal of American Linguistics 14*. 27-44.

Chamberlain, Alexander F. 1890. Notes on Indian child-language. *American Anthropologist 3*. 237-241.

Chao, Yuen Ren. 1951. The Cantian idiolect: An analysis of the Chinese spoken by a twenty-eight-months-old child. *University of California Publications in Sinitic Philology 11*. 27-44.

Darwin, Charled. 1877. A biographical sketch of an infant. *Mind 2*. 285-294. [Reprinted in Aaron Bar-Adon and Werner F. Leopold (eds.). 1971. *Child language: A book of readings*, 26-28. Englewood Cliffs, NJ: Prentice-Hall.]

Dennis, Wayne. 1940. *The hope child*. New York: Appleton-Century.

Demuth, Katherine. 1996. The prosodic structure of early words. In James L. Morgan and Katherine Demuth (eds.). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, 171-184. Mahwah, NJ: Erlbaum.

ELAN computer software. 2018. Nijmegen: Max Planck Institute for Psycholinguistics. Retrieved from https://tla.mpi.nl/tools/tla-tools/elan/

Fujiki, Martin, and Bonne Brinton. 1985. Sampling reliability in spontaneous language sampling. Paper presented at the Annual Convention of the American Speech-Language-Hearing Association. Retrieved from https://eric.ed.gov/?id=ED267523

Grinevald, Colette. 2003. Classifier systems in the context of a typology of nominal classification. In Karen Emmorey (ed.). *Perspectives on classifier constructions in sign languages*, 91-110. Mahwah, NJ: Erlbaum.

Gun, Ling-Yu, and Sarita Eisenberg. 2015. Sample length affects the reliability of language sample measures in 3-year-olds: Evidence from parent-elicited conversational samples. *Language, Speech, and Hearing Services in Schools 46*. 141-153. Retrieved from: https://ncbi.nlm.nih.gov/pmc/articles/PMC4610271

Haspelmath, Matthew. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica 45*. 31-80.

Ingram, David. 1989. *First language acquisition: Method, description and explanation.* Cambridge: Cambridge University Press.

Kidd, Evan and Rowena Garcia. 2022. How diverse is child language acquisition research? *First Language* Advance online publication. https://doi.org/10.1177/01427237211066405

Kroeber, Alfred L. 1916. The speech of a Zuñi child. *The American Anthropologist 18*. 529-534.

Lieven, Elena, and Sabine Stoll. 2010. Language development. In Marc H. Bornstein (ed.). *The handbook of cross-cultural developmental science,* 143-160. New York: Psychology Press.

MacWhinney, Brian, and Catherine E. Snow. 1985. The child language data exchange system. *Journal of Child Language 12.* 271-296.

Mithun, Marianne. 2014. Morphology: What's in a word? In. Carol Genetti (ed.) *How languages work: An introduction to language and linguistics*, 71-99. Cambridge: Cambridge University Press.

Montes Castañeda, Faustino. 2022. *La adquisición prosódica y morfológica del verbo totonaco de Chumatlán.* Doctoral dissertation, Centro de Investigaciones y Estudios Superiores en Antropología social, México.

Nelson, Katherine. 1973. Structure and strategy in learning to talk. *Monographs of the Society of Research in Child Development 38*, no. 149.

Nelson, Katherine. 2014. A matter of meaning: Reflections on forty years of JCL. *Journal of Child Language 41*, Supplement 1. 93-104.

Nolan, Rachel. 2020. Language barrier: For indigenous people who cross the border, a translation crisis follows. *The New Yorker*, January 6, 2020 (pp. 26-31).

Perez Martínez, Yenny Aracely. 2020. La adquisición de la morfología verbal del mixe de Tamazulápam, Oaxaca. Master's thesis, Centro de Investigaciones y Estudios Superiores en Antropología social, México.

Pye, Clifton, 2017. *The comparative method of language acquisition research.* Chicago: The University of Chicago Press.

Pye, Clifton. 2020. Documenting the acquisition of indigenous languages. *Journal of Child Language 48.* 454-479. https://doi.org/10.1017/S0305000920000318

Pye, Clifton, Barbara Pfeiler, and Pedro Mateo Pedro. 2017. Mayan language acquisition. In Judith Aissen, Nora C. England and Roberto Zavala Maldonado (eds.). *The Mayan languages*, 19-42. London: Routledge.

Rescorla, Leslie. 1980. Overextensions in early language development. *Journal of Child Language 7*. 321-335.

Snyder, William. 2021. A parametric approach to the acquisition of syntax. *Journal of Child Language 48*. 862-887.

Taine, Hippolyte. 1877. Acquisition of language by children. *Mind 2*. 252-259. [Reprinted in Aaron Bar-Adon and Werner F. Leopolod (eds.). 1971. *Child language: A book of readings*, 20-26. Englewood Cliffs, NJ: Prentice Hall.]

Tomasello, Michael, and Daniel Stahl. 2004. Sampling children's spontaneous speech: How much is enough? *Journal of Child Language 31*. 101-121.

Van Gijn, Rik, and Fernando Zúñiga. 2014. Word and the Americanist perspective. *Morphology 24* (Special Issue on Word in the Languages of the Americas), 135-160.

Woodbury, Anthony. 2003. Defining documentary linguistics. In Peter K. Austin (ed.). *Language documentation and description*, Vol. I, 35-51. London: School of Oriental and African Studies.