Sketch Acquisition Manual (SAM), Part II:
The acquisition sketch

Rebecca Defina,[1,2] Shanley E. M. Allen,[3] Lucinda Davidson,[1,2] Birgit Hellwig,[4] Barbara F. Kelly,[1] Evan Kidd [2,5,6]

[1] University of Melbourne, [2] ARC Centre of Excellence for the Dynamics of Language,
[3] University of Kaiserslautern‑Landau, [4] University of Cologne,
[5] Max Planck Institute for Psycholinguistics, [6] Australian National University

Abstract

This paper presents the second part of a guide for documenting and describing child language, child-directed language and socialization patterns in diverse languages and cultures. The guide is intended for anyone interested in working across child language and language documentation, including, for example, field linguists and language documenters, community language workers, child language researchers or graduate students. We assume some basic familiarity with language documentation principles and methods, and, based on this, provide step-by-step suggestions for collecting, analyzing and presenting child data. This second part of the guide focuses on developing a child language acquisition sketch. It takes the sketch corpus as its basis (which was introduced in the first part of this guide), and presents a model for analyzing and describing the corpus data.

Introduction

Any theory of language, including how it is learned, must be built upon a typologically diverse representative sample of the world's languages. However, our current evidential base in language acquisition is unrepresentative and skewed towards English and other big European languages (Kidd & Garcia 2022). The Sketch Acquisition Manual addresses this bias by aiming to increase coverage within the domain of first language acquisition. In Part I of this manual, we outlined the pathways for collecting and processing child language and child-directed language data in under-studied languages. In Part II, we lay out some methodological guidelines for analyzing and presenting an 'acquisition sketch' of this data. The model we present here guides the reader through the writing of a sketch, from the introductory demographic and socially-situated material through to sections and subsections on key areas of linguistic analysis. We provide guidance for topics that we consider core and those we consider optional.

Before getting started on the details, it is necessary to highlight again the goals of the sketch format in order to arrive at a realistic assessment of what the sketches can and cannot contribute. The sketch format is designed to continue the descriptive focus of early studies in language acquisition, enriching it with a documentary perspective and adding the missing cross-linguistic perspective for a better understanding of universality and variability in the domain of child language and child-directed language. While the sketch corpora are in many instances too limited in size to directly contribute to some contemporary research questions, they will broaden our understanding of the variation space; an understanding that is indispensable to any theory of language learning.

When developing the sketch format, we have compared the kind of information contained in our sketch corpora with that contained in our larger child language corpora and/or our knowledge from previous studies of the acquisition of specific phenomena. The encouraging result is that the limited data present in the sketch corpora was enough to isolate and describe many of the features of both child language and child-directed language identified in more rigorous setups. While we could not detect every pattern (especially low-frequency phenomena were often not represented), we were able to detect the salient patterns. And even though the data was too limited to prove that a pattern can be generalized beyond the sketch corpus, its distribution in the sketch corpus matched the patterns found in our larger corpora closely enough to suggest that we can form valid hypotheses on the basis of the limited sketch data.

Nevertheless, we acknowledge that the data set is limited in size (five hours of data) and most likely constitutes an opportunistic and non-balanced sample. For our analysis, this means that the absence of a pattern does not necessarily mean that the child does not know it. It might simply not feature in our sample (see especially Tomasello & Stahl 2004). Conversely, the presence of a pattern does not mean that the child knows it and uses it productively. It might simply be rote-learned (i.e. the child does not analyze its underlying compositionality) or a repetition of what an interlocutor had produced earlier. Furthermore, there is considerable individual variation in children's development, making it problematic to generalize from one child to all the children of the same age. In this context, recall that we use "age" as a proxy for developmental stage (see Part I, Section 2.1.4), and that it is entirely possible that the focus children are developing at different rates.

In your analysis we therefore ask you to keep the limitations of the data set in mind and adhere to the following guiding principles.

First, the focus is on describing what children of different ages are doing and producing in the recordings. It is always tempting to speculate what children know at a certain age, but we must be aware that their productions attested in the sketch corpus may or may not fully reflect this knowledge. In some cases, individual productions will invite cautious inferences about children's acquisition trajectories and knowledge states (especially in the case of non-target-like forms; Bowerman 1982), but be wary of going too far beyond your data. For instance, in a morphologically complex language we are likely to find that a child will only use a subset of morphemes within the larger (adult) paradigm. In a sketch it would be appropriate to describe this subset and how it expands across development, for example, attaching to more and more nouns/verbs, occurring in paradigmatic opposition to (an)other morpheme(s), or being overgeneralized to infelicitous contexts (see Section 6.3 for more information). However, you should stop short of making inferences about children's full knowledge of the paradigm and whether the knowledge they display is fully productive.

Second, we recommend contextualizing your analyses by taking into account the preceding and following utterances. This will give you information on whether or not the

child is repeating something they have encountered before, as well as access to the adult's response to it (they may be expanding on a child's utterance, they may ignore it, they may correct it, etc.) (see especially Sections 5.1 and 5.2). Also consider the overall context of the recording. For instance, a context where a child accompanies their mother to the garden may trigger different language than a context that features a casual conversation in the evening before bedtime (see also Section 3.2).

Third, the sketches will strike a balance between qualitative description and quantitative analysis, where appropriate. We recommend making use of descriptive statistics wherever possible in order to show the distribution of a pattern in the corpus. Some features of the corpus may be amenable to statistical analysis if enough tokens are observed, using either simple non-parametric techniques or more modern approaches that can handle relatively small data sets (e.g. conditional inference trees and random forests; see Levshina 2021). We stress that these are not necessary. In general, be aware of the limited opportunities you have to adopt a quantitative approach, and always consider how you can use qualitative description to suggest developmental patterns.

Fourth, when preparing your sketch, keep in mind that the focus is to show the development of language across ages. We therefore recommend analyzing the data from the different ages separately, both the language used with children (as interlocutors are likely to engage differently with children of different ages) and the language used by children of different ages. In this context, recall that individual children may differ considerably, even if they are of the same age. If you notice such differences, it may become necessary to analyze individual children separately, but it may also be sufficient to simply report on any such observations in your acquisition sketch. This analytic approach will allow you to detect similarities and differences across individuals and ages. On this basis, your sketch should then focus on capturing developments. In some cases, it will be possible to offer a description of developments across all five age points. In other cases, you may only be able to describe developments between, say a younger age group and an older age group. In yet other cases, the data may be inconclusive and you will not be able to offer any generalizations. All such observations are valuable and should be noted in your sketch.

Finally, identifying patterns is easier if you know what you are looking for. We therefore recommend analyzing the data from the older children first, as the main structures are likely to be present at this age, giving you an idea of what the younger children are developing towards. Once you have a handle on the language of the older children, move your way backwards to the younger children. Comparing the language of older children to that of younger children, as well as comparing the language directed to older children with that directed to younger children, will likely help you to identify salient developments. The adult language may also provide a useful reference point. However, your sketch corpus is unlikely to contain enough language directed by adults to other adults for systematic comparison with the child data. While adult data from other sources may also be helpful, it is possible that data from other corpora may not be comparable (e.g. it may represent different genres and registers), and data from a reference grammar may focus on structure rather than language use.

The following sections describe the contents of the sketch, covering general information on the language (Section 1) and its typological profile (Section 2), the sketch data (Section 3), the learning environment and ethnotheories (Section 4), child-directed language (Section 5) and child language (Section 6). We also offer suggestions for developing community materials (Section 7). We are aware that we cover a lot of ground, and that it will not be possible for the sketches to address all topics in equal depth. Languages differ, and topics may thus be more or less relevant. Similarly, individual skills, backgrounds and interests

vary, and different researchers and communities may want to pursue specific topics in more detail, while backgrounding others. While we strive to ensure some form of comparability across the sketches, we consider it entirely legitimate for the specifics of a language and/or your interests to guide you in your decisions. In this manual, we attempt to strike a balance between comparability and variability by distinguishing between what we consider to be core and optional topics, as discussed below.

The manual is written with a view to publishing the resulting sketches in the peer-reviewed open access series *Child Language Documentation: The Acquisition Sketch Project* of *Language Documentation and Conservation*. Of course, this is not the only imaginable publisher and/or output. It is equally possible for you to base your approach only loosely on the sketch format, setting different priorities and aiming for different kinds of outputs. But in general, and especially for sketches within the series, we would like you to address all sections of the table of contents below as faithfully as possible. Within each section, we distinguish between core topics that should be addressed and optional topics (labeled "extensions") that can be addressed. Do not be intimidated by the wealth of information; we have attempted to cater for as many different interests as possible, but we do not expect every sketch to cover every topic. We highlight the core topics, and we leave it to you to decide if you want to pursue any of the other topics. Each section ends with a box that summarizes the discussion, listing core topics and possible extensions. For each topic that you describe, please feel free to decide on the level of detail. Depending on your interests, the structure of the language and/or the available data, you may either give a brief summary of salient issues or else pursue a more comprehensive study.

When preparing your acquisition sketch, please do not feel that you have to account for every single piece of data. Given that five hours of recording look like a manageable dataset, there is a strong temptation to analyze every utterance within it. Yet, these five hours contain a wealth of information, including many non-interpretable utterances, and there will inevitably be cases that will elude your analysis. Keep in mind that this is a *sketch* description. We anticipate that the sketches will typically be longer than an average journal article (e.g. 10,000 to 20,000 words), whereby a full sketch is likely to be closer to the upper limit, while a sketch focused on a smaller set of phenomena should be closer to the length of a journal article.

We also encourage you to consult the example sketches included in this volume, as they will provide you with ideas of how others have handled the challenge of balancing core and optional topics. Please feel free to get in touch with the contact persons to discuss any questions (see the introduction to this special publication for a list).

Below, we provide a table of contents. Note that we have organized the presentation of the two main sections (5. Child-directed language, 6. Child language) around the major subfields of linguistics (phonology and prosody, lexicon and semantics, morphology, syntax, gesture). The subsections within Sections 5 and 6 are not always parallel, though, since not all topics apply equally to child-directed language and child language. For example, scaffolding techniques such as routines (Section 5.1) and responses to non-target-like utterances (Section 5.2) are best analyzed within interactional contexts involving adults or older caregivers, and we thus discuss them in the child-directed language section only. Similarly, morphological and syntactic phenomena have received little attention in child-directed language so we discuss them together as 'morphosyntax' in Section 5.5, whereas there is extensive literature on both the morphology and the syntax of child language so we discuss them separately in Sections 6.3 and 6.4. Please take this table of contents as a suggestion only. Depending on the language and the development of your analysis, you

might consider a different structure more appropriate. We only ask you to attempt to cover all of the core topics and guiding questions highlighted throughout the following sections.

Table of contents:

Sections 1 to 6 of the acquisition sketch are essentially an academic report on language acquisition and socialization. In addition, we would like you to consider complementing this academic report with a description of any outcomes of the sketch project for the community. Below, we discuss some possibilities of how research findings may benefit the community (see Section 7), and we encourage you to think about them and/or pursue other possibilities. Please detail your approach in your acquisition sketch, in the hope that your reports will be sources of inspiration for others.

1   General information

At the outset of your sketch, please provide a short introduction to the language and its users. This should be kept very brief and is primarily for identification and to highlight any particular issues of relevance. This section should:

- Identify the language: Provide an ISO or Glottocode. If multiple names are in use, indicate the one you will use and provide a list of other names commonly in use. Indicate the language family, if known. Briefly discuss the dialect situation and identify the focus dialect.

- Locate the language: Where in the world is it typically used?

- Identify the people: Roughly how many people use the language, and do they form a distinct subset of the population?

- Discuss the linguistic situation across the lifespan: Are language users evenly balanced across age groups or is there a language shift occurring? Are language users becoming

bilingual in other regional languages at later ages; for instance, as a result of schooling or requirements of local adult life?

- Discuss any other important ethnographic information or linguistic setting information: Is the language one of several used within the community? Is the language restricted to particular domains of use (e.g. within the home)? What language(s) are used in school?

- Refer to any other studies of child language acquisition which may have been done in this language or community or among its neighbors.

2   Language typology overview

It is helpful to situate the acquisition sketch within the general grammatical properties of the language. However, it is not necessary to provide a full grammar of the adult language, and an acquisition sketch does not need to be delayed until the adult language has been described. In many cases, waiting until the adult language is well described risks missing the window of opportunity to document child language acquisition. Even when this is not the case, the acquisition sketch may be completed in tandem with or prior to a description of the adult language. As we detail below, descriptions of child language should be based on the observed productions of the children and thus do not require a judgement regarding mastery of adult-like grammar.

Regardless of the state of description of the adult language, it is useful to provide some information here to help the reader interpret the rest of the sketch, but this section should not develop into a sketch of the adult language. Information which would be good to provide includes:

- An overview of the basic phonology and how you represent it in your sketch. Especially any notable phonological properties, e.g., does the language make use of lexical tone?

- Word order; for example, is there a fixed word order (in clauses and noun phrases)? If so, what is it?

- Morphological type, for example is the language particularly isolating or synthetic? Any head- or dependent-marking properties? In morphologically rich languages, it may be useful to give a rough template of the structure of the predicate and/or noun.

This overview should be kept brief. Where relevant, more information should be given in Sections 5 and 6. For example, if the language has case, mention it here as a dependent-marking property, but postpone any details of the case categories or forms until Section 6.3 on the children's use of case marking. Also, refer to any published grammar or other descriptive work, if available.

3   Data

As discussed in Part I, Section 2.2.3, we assume that you will archive your sketch data. In the acquisition sketch, please provide the link to this deposit. Readers will then be able to browse the metadata information on participants and sessions, allowing a more comprehensive overview of your data set. In your sketch, it will be enough to briefly summarize the main information about the focus children (Section 3.1) and the recordings (Section 3.2). If your fieldsite is bi-/multilingual, please detail how you approached this topic in your sketch (Section 3.3).

3.1 Focus children

Please list the focus children and their ages. It is helpful to provide a table as in Tables 1 and 2 in Part I, Section 2.1.1. Make sure to use pseudonyms for all child and adult participants to preserve their identity.

Introduce each of the focus children: Are they older, younger, middle siblings or only children? Which family members do they live with? Which other languages (if any) are used in the home? Is there anything else of note? For example, do they attend formal education?

3.2 Recordings and other participants

Please provide the following information:

- A brief overview of each of the recordings used for the sketch. In what setting were they recorded (e.g. in the home, in the bush)? In what sorts of activities were the children principally engaged (e.g. playing outside with peers, reading with a caregiver, gathering food, painting, eating dinner).

- A list of other participants appearing in the recordings. If known, also list the relationship to the focus child. For children, provide (approximate) age.

- A table of data, such as that for the Pitjantjatjara sketch in Table 1.

Table 1. Data set for the Pitjantjatjara sketch. The table lists the age bracket, the focus child ID reference, the number of utterances produced by the focus child, the number of other children present at the recording, the number of utterances produced by those other children, the number of adults present, and the number of utterances produced by those adults.

Age	Focus child	Focus child utterances	No. of other children	Other child utterances	No. of adults	Adult utterances
2;0	ANT	176	3	409	3	283
	ANN	177	5	399	5	281
2;6	ANT	230	4	355	3	229
	FRE	194	1	139	1	217
3;0	ANT	219	2	283	3	329
	REN	298	1	5	4	567
3;6	ISY	202	5	395	2	162
	REN	171	3	363	2	127
4;0	ISY	212	3	297	8	470
	FRE	227	3	297	8	470

3.3 Multilingualism

Given the prevalence of bi- and multilingualism in today's world, your field setting may also be bi-/multilingual and the families in your study may be exposed to and/or learning more

than one language. This raises several issues that may be relevant for your sketch, and we ask you to briefly report on your approach in this section.

In one possible scenario, the children will primarily use the focal language and the other language will be less prevalent (e.g. Cree-English: Henke 2019; Q'anjob'al Maya-Spanish: Mateo Pedro 2021). Here, we suggest that the sketch be on the acquisition of the focal language. To facilitate that, try to collect data in situations where the language is the predominant one (e.g. with friends who primarily use that language, at home, in the bush). If both languages are present in the data, write the sketch using the utterances in the focal language. Utterances in the other language should remain in the data set, but will not be analyzed as part of the sketch. Instead, you should describe the extent of the use of the other language at the beginning of the sketch (see Section 1). For example, you may be able to identify patterns regarding which language is used in which types of situations, or with which types of interlocutors.

In another possible scenario, many/all members of the community will be highly bilingual and will regularly alternate between two or more languages within the same conversation or even within the same utterance. In this case, you should decide whether it makes most sense to focus the sketch on just one language, or on the mixed language use (e.g. Warlpiri-English-Kriol-Light Warlpiri: O'Shannessy 2008). Your choice will undoubtedly be informed by your knowledge of the sociolinguistics of the community.

In both types of scenarios, you will very likely encounter 'code-mixed' utterances that include two languages within one sentence – typically about 5% of utterances in studies of bilingual children (e.g. French-English: Genesee et al. 1995; Inuktitut-English: Allen et al. 2002). If your data set contains enough of these utterances, it would be useful to characterize the different types of mixing in your sketch. Muysken's (2000) typology of three patterns of mixing offers a helpful framework (see a further extension of this typology in Muysken 2013):

- Insertion mixing: inserting a content word of language A into a sentence otherwise of language B, where the inserted word typically follows the grammar of language B. This is the most common type of mixing, especially when the two languages have different typologies.

- Alternation mixing: one or more multi-word sequences in each language, each following the grammar of its own language. The switch from one language to the other is typically found at points where the grammatical structure and/or word order of the two language systems is similar or identical.

- Congruent lexicalization: words from each of languages A and B are interspersed throughout the sentence. This is the least common type of mixing, typically found where the grammatical structure of the two languages is very similar.

You may also notice effects of influence or transfer from one language to the other, typically most evident in vocabulary and structural aspects like word order. For example, children growing up bilingual in English and Italian often produce an overt pronoun subject in Italian where monolingual Italian-speaking children would omit the subject, arguably because English requires overt subjects (Serratrice et al. 2004; see also Serratrice 2013). If you find clear patterns like this in your data, it would be interesting to include them in the sketch. However, most such patterns require more data to substantiate than is likely to be available for your sketch. In this case, it would be more appropriate to mention the patterns briefly but not document them quantitatively.

Code mixing and cross-linguistic influence in situations of language contact can lead to long-term diachronic changes in one or both languages, or to the emergence of a new

language. In some cases, differences are evident between younger and older language users in synchronic time. A particularly clear example of this is Light Warlpiri spoken in northern Australia, which has emerged from contact between Warlpiri and English/Kriol (e.g. O'Shannessy 2013). If you find signs of this in your sketch data, it would be useful to comment on this or document it.

Studies focused specifically on bi- and multilingual acquisition often try to quantify how much of each language a child uses or is exposed to and in what contexts. A recent and comprehensive initiative in this direction is the Q-Bex project (Quantifying Bilingual Experience), which includes guidelines for what factors to consider as well as a questionnaire in several languages (De Cat et al. 2021). Bi-/multilingual acquisition studies often also compare the trajectory of acquisition of a particular linguistic phenomenon across the two languages of the child, or compare the trajectory of acquisition in bilingual versus monolingual children. However, the sketch data is not likely to be comprehensive enough to do this in a meaningful way unless the phenomenon is very frequent.

4   Learning environment and ethnotheories

As children learn language, they also acquire important cultural information which enables them to act as a member of their particular culture and society. This process of language socialization occurs in interactional contexts and can be both explicit and implicit in nature. The acquisition sketch should include information on identifiable language socialization practices and salient conversational sequences in which children and their interlocutors engage (e.g. prompting routines; see Section 5.1), as well as on local views and attitudes towards language development and language use. During the initial phase of corpus construction and transcription (see especially Part I, Sections 2.2.1, 2.2.3 and 3.2.3), you will have collected numerous anthropological and sociolinguistic notes that you can now draw on. In addition, the sketch corpus itself will contain relevant information.

Below, we provide a list of questions, structured around the topics of ethnolinguistic theories on language development, views on language use, and observations from interactional contexts. Box 1 summarizes the discussion and separates the various possibilities into core topics (that we would like you to address) and extensions (that are optional). If you are interested in pursuing anthropological approaches beyond the sketch format, we recommend that you consult the language socialization literature (e.g. starting with the classic volume by Schieffelin & Ochs 1986; see also the more recent overview by Brown & Gaskins 2014).

Ethnotheories on language development
Socialization practice evolves from and gives rise to different conceptualizations about children's language development and community attitudes toward children's language use. In framing the learning situation in the sketch, you might consider some of the following questions:

- Do people have ideas about how children acquire language? Do they assume that children can or should be taught language? For example, do adults play a role in children's language development?

- In a multilingual community, are there different attitudes about acquiring different languages? For example, people may consider different languages more or less difficult for children.

- Do people recognize any salient developmental stages? Are there any stages where children are supposed to master specific (linguistic or non-linguistic) skills? Which ages? Which skills?

- Are children's gestures, gaze and early vocalizations considered language?

- Do people have an opinion on when children start using language or what their first words are? When do people consider that children have acquired most of the features of their language?

- Do people assess children's language? How? For example, people may consider some children to have 'better' language skills than others because they master some specific vocabulary or some specific structural feature.

Views on language use

People may have specific views on the kind of language used by and with children. Please use the following questions as a guide:

- Do people recognize a special register used with children ('child-directed language'; see Section 5)? Do they reflect on why they are using (or not using) such a register? Are people aware of any specific lexical or structural properties of the register?

- If there is a child-directed language register, when is it used? How common is it? Note that the sketch corpus is likely to be biased towards child-directed interaction and thus cannot be used to assess the amount of child-directed language vis-à-vis other forms of input arising from multiparty or peer interaction. However, you can include any relevant anthropological observations in your sketch. If you are interested in pursuing this topic further, consider one of the following options: day-long recordings (as done by Casillas et al. 2020), or systematic observations (as done by Cristia et al. 2019).

- Do people focus differently on children's comprehension of language vs. their production of language? If so, does this have any effect on their attitude about children as conversational partners or about children's language skills? For example, people may report not talking to children until they are able to respond verbally.

- Are there any features of child language that strike their caregivers as salient? Any common non-target-like forms (i.e. differences from typical adult language), speech acts, routines?

Interactional contexts

Children participate in different types of communicative interactions, in particular child-directed interactions (between a caregiver and a child), multiparty interactions (with several adults and children) and/or peer interactions (children amongst themselves). The following set of questions serve as a guide for identifying and describing typical communicative interactions:

- In what types of communicative interactions do children typically engage and in what language(s)?

- What activities are associated with the different types of communicative interaction? For example, peer interaction may take place when children are sent to the bush to look for firewood.

- What characterizes children's behavior in these interactions? For example, do children produce language and/or gestures to demonstrate their active participation in multiparty

interactions? What social actions do they appear to be pursuing through them? How do the children's interlocutors - of different ages - respond to these productions? Do they facilitate the children's contributions? Do they discourage the children? Correct them? Ignore them?

- Do you observe any age- or gender-related differences in communicative interaction? For example, adults may preferably communicate with older children and/or communicate with younger children through older children. Or mothers and fathers may engage in different types of interactions with their children.

- What does an average day look like for a child? How are the different types of communicative interactions typically distributed over the day?

Box 1. Key areas of focus in learning environment and ethnotheories.

Core	
(i)	What ideas do people have about how children acquire language?
(ii)	What salient developmental stages of language acquisition do people recognize (if any)?
(iii)	What do people think about using a special register when communicating with children ('child-directed language'; see Section 5)?
(iv)	In what types of communicative interactions do children typically engage and in what language(s)?
Extension	
(v)	Expand on (i): investigate whether adults think that language is a developmental skill that needs to be taught.
(vi)	Expand on (iii): investigate when child-directed language is used.
(vii)	Investigate whether people focus differently on children's comprehension of language vs. their production of language.
(viii)	Investigate whether adults recognize any salient features of child language.
(ix)	Expand on (iv): investigate the structure and contexts of different types of communicative interactions.

5   Child-directed language

Language learning takes place in the interaction between children and their interlocutors, and researching language acquisition thus includes studying the children's input. In many societies, we observe the existence of a special register (labeled child-directed language) that is used when communicating with young children and that tends to be characterized by the structural and lexical features summarized in Table 2 (see Newport 1977 and other early contributions in Snow & Ferguson 1977 and Gallaway & Richards 1994; see Pizer et al. 2011 for child-directed signing). The following sections discuss important aspects of this register in more detail: routines, speech acts and drawing attention (Section 5.1), responses to children's utterances (Section 5.2), phonological and prosodic features (Section 5.3), lexical features (Section 5.4), morphosyntactic features (Section 5.5), and gesture (Section 5.6). Note that this section focuses on the structural and lexical properties of child-directed language (if

it exists). Community views on child-directed language and the distribution of child-directed interaction vis-à-vis other types of interactions are discussed in Section 4.

Table 2. Overview: Typical features of child-directed language.

short (but complete) utterances
few hesitations and errors
exaggerated pitch contours, larger and more clearly articulated signs
high F0
long duration and pauses
restricted vocabulary with reference to the present time and location
nursery vocabulary
many questions, imperatives and prompts
many full and partial repetitions, variation sets

In the early literature, the focus of research was on the identification and description of this register. Around the same time, anthropological studies questioned the universality of the register and its features (cf. especially studies from within the language socialization paradigm, e.g. Schieffelin & Ochs 1986), and we know of counterexamples to at least some of the proposed features (e.g. Pye 1986: 88 reports a low F0 in the speech of K'iche' Mayan mothers). The number of such studies remains small, though, and for most languages we know little about the existence of such a register or its characteristic features. We nevertheless recommend that you take the features summarized in Table 2 as your starting point, as they are attested widely across many languages. Be aware, though, that the language may not have such a register (see also Section 4) and/or that not all features may be present and/or that other features may be characteristic. Throughout your analysis, keep in mind that there may be age-related differences – in particular, differences based on the age of the child (i.e. younger vs. older focus children). If there is enough data available, you may also want to take into account possible differences based on the age of the interlocutor (i.e. adults vs. older children).

On the basis of the sketch corpus, you will be able to develop a good idea about the (non-) existence of such a register and its characteristic features. Drawing on your knowledge of the adult language, you will be able to form hypotheses about differences between this register and adult-directed language. A systematic comparison of child-directed and adult-directed language is unlikely to be possible, though, as the sketch corpus will not contain enough adult-to-adult language. If you are interested in investigating this topic systematically, consider a more controlled study (e.g. as done by Frye 2022). Within the scope of the sketch format, you can pursue this topic to some extent by paying attention to differences in the language directed at different age groups. Many studies have found that the characteristic features of child-directed language decrease as the child's age increases, so a comparison of the language directed at the youngest versus oldest children in your sketch corpus may well give you insights into the existence and properties of such a register.

The section on child-directed language in your acquisition sketch will contribute to our knowledge of universality and variability in the domain of child-directed language. This knowledge has implications for current debates on the role of child-directed language in

language learning: the facilitating role of specific features as well as the role of child-directed language versus other forms of input (e.g. overheard language) (cf. Saint-Georges et al. 2013 for a systematic review; cf. Shneidman & Woodward 2016 on different forms of input). While the sketch corpora are too limited to directly engage with this line of research, they will contribute indirectly by mapping out the variation – a variation that needs to be taken into account for any theory of learning. At the same time, your research will contribute to issues in language documentation. Several studies have used child-directed language to shed light onto the structure of the adult language (Hellwig & Jung 2020; Henke & Brittain 2022). Equally, documenting Indigenous language learning contexts can feed into language maintenance and revitalization efforts (Child Language Research and Revitalization Working Group 2017).

5.1   Scaffolding techniques: Routines, speech acts, and drawing attention

This section covers a number of scaffolding techniques that are used by adults to engage children in interaction, in particular, their use of routines, speech acts and ways of drawing attention. Given the asymmetry between the interlocutors, we subsume all techniques within the section on child-directed language, even though children are active participants contributing to the exchange. In this section, you should aim for a characterization of salient techniques attested in your sketch corpus. The focus should be on their use by adults (and older children) towards the younger children, as well as on young children's responses and contributions to the interaction. In addition, we include some pointers in case you would also like to pursue this line of research further and investigate routines, speech acts and attention direction among the young children. In this case, we recommend that you include any discussion within this section, as it will tie in seamlessly with your description of the techniques employed in child-directed language.

The following paragraphs are structured around topics (under the headings of 'routines', 'speech acts', 'drawing attention', and 'children's responses'), covering both core and extensional topics; Box 2 then explicitly distinguishes between core and extensions.

Routines

- Are there any salient routines when talking to children? For example, anthropological research often reports on practices that can be considered functional equivalents of child-directed language, e.g. the *a:la:ma* (also known as *elema*) 'say like that' routine of the Kaluli where adults model the child's utterance and prompt the child to repeat it (Ochs & Schieffelin 1996: 86-87; see also example 1). If you observe any such routines, do children engage in them as well and use the same kinds of routines in their interaction with others?

- Do adults prompt children as a means of teaching appropriate pragmatic behavior such as by modeling "say thank you" or by framing utterances telling a child what to say, such as "say, 'I don't like it'".

- Are there any other interactional routines? For example, what do people do to soothe or comfort children? To distract them? To entertain them? To tease them? To prevent them from getting hurt? Are there differences in routines with strangers vs. with familiar people?

- Are there any songs/chants/lullabies that are commonly used? Any games?

One example is the prompting routines Murrinhpatha-speaking caregivers often engage in with young children. These routines tend to focus on important cultural information, such as children's connections to particular areas of land or entities within that land (totems). In example (1),[1] Bernadette prompts her daughter Tabitha (3;0) to state her main *ngakumarl* 'totem'. She does so by providing her daughter with the information, followed by the directive *thama* 'you say it'. The mother then issues a display question (line iii) to test or consolidate Tabitha's knowledge of this content.

(1)   Murrinhpatha: Bernadette (mother) and Tabitha (3;0)
        (LAMP_20140313_LD_01_00:06:33.245; Davidson 2018: 257)

    i.   B.   *ngakumarl*   *ngay-ka*   *ku*   *tek*
            totem   1SG-TOP   NC:ANIM   red_tailed_black_cockatoo

            thama
            2SG.SAY/DO.FUT

            'Say "my totem is the red tailed black cockatoo"'

    ii.   T.   *ngay-ka*   *ngay-ka*   *ngakumarl*   *ngay-ka*
            1SG-TOP   1SG-TOP   totem   1SG-TOP

            ku   *tek*
            NC:ANIM   red_tailed_black_cockatoo

            'My... my... my totem is the red tailed black cockatoo'

    iii.   B.   *thangku*   *ngakumarl*   *nhinhi-yu*
            what   totem   2SG-CTC

            'What's your totem?'

    iv.   T.   *ku tek*
            'The red tailed black cockatoo'

Salient speech acts and their distribution

- Pay special attention to imperatives, questions, and labeling things since they are particularly important speech acts in child-directed language. In which contexts do you observe them? Which forms/structures are used? For example, there might be a preference for polar or tag questions over content questions, or a preference for 'where' questions over 'what' questions. What is their relative proportion in the sketch corpus (bearing in mind that the distribution in the sketch corpus may not be representative)? Note that some speech acts (such as labeling things) are often accompanied by pointing and other gestures (see Section 5.6).

- In which speech acts do children engage? For example, if children produce questions, how do they ask them? Specifically:

  - What structures do children use when asking questions?

  - To whom do they direct their questions?

  - What is the pragmatic function of the questions?

---

[1] Our glossing follows the Leipzig Glossing Rules. In addition, we use the following abbreviations: ANIM – animate; AUG – augmentative; CTG – contingent mood; DM – discourse marker; MOD – modalis case; NC – noun class; PAR – participial mood; POL – politeness morpheme (used to 'soften' the imperative).

- Does the form or pragmatic function of questions change according to the child's addressee?

- How do addressees respond?

• What asymmetries occur in the speech acts between adults and children? For example, do children give commands to adults or other children?

Drawing attention

• What do people do to capture children's attention? What do people draw children's attention to?

• What do children do to draw attention to themselves and towards third referents? How do their interlocutors respond? What do they respond to?

• What do people do if they do not like what children are doing/attending to?

Children's responses

• How do children respond to the above routines, speech acts and ways of capturing and directing attention? For example, do they use negation as a frequent response to adult questions or commands? Do they imitate or repeat (all or part of) their interlocutors' utterances? If yes, what do they repeat (e.g. the end of utterances)? Does this differ across ages?

Box 2. Key areas of focus in child-directed language: Routines, speech acts, drawing attention.

Core
(i) What routines and speech acts are used regularly between adults and children?
(ii) How do adults capture children's attention and garner a response? How do adults respond to children's attempts to get their attention?
(iii) How do children respond to routines, speech acts and ways of drawing attention?
Extension
(iv) Is there any evidence of different routines and speech acts used between children and peers compared to adults?
(v) Are there any asymmetries occurring in the speech acts between adults and children?

5.2   Responses to non-target-like utterances

One fruitful area of research is responses to non-target-like aspects of children's utterances including their articulations, lexical choices and morphosyntactic structures. Such responses tell us something about the contexts of learning and the existence of feedback that facilitates learning (e.g. children may be teased or corrected). They also tell us something about how adults and older children perceive the language of young children: which sounds, signs and structures they consider difficult, or which non-target-like forms they consider salient enough to react to (possibly even taking them over into child-directed language). Last but

not least, they add to our knowledge of the adult language, providing us with negative evidence from naturalistic data.

Research in several societies shows that explicit corrections are rare. Instead, interlocutors tend to resort to a variety of implicit strategies where they repeat a child's utterance, correcting the non-target-like element and thereby modeling the target form to the child (e.g. Chouinard & Clark 2003; Clark 1987; Saxton 1997, 2000). Such corrections are usually not intended to teach language, but instead form part of communicative exchanges, where, for instance, an adult seeks confirmation that they have understood the child correctly. Example (2) from Qaqet illustrates such an exchange. The young child's utterance is incomprehensible, and the mother responds first with a series of clarification questions ('who', 'huh?'), then deduces from the context that the child probably pointed out a bird, and produces the target form using the question intonation to ask for confirmation of her interpretation ('a bird now?'); when the child confirms, she expands on the initial utterance ('the bird is gone').

(2)   Qaqet: BLN (mother) and ZDL (1;6) (LongYJL20150805_1 1331.957 1345.906)

i.	ZDL	*papapa* [points up]	'??'
ii.	BLN	*nema?*	'who?' [restricted clarification question]
iii.	BLN	*nema?*	'who?' [restricted clarification question]
iv.	ZDL	*pata* [points up]	'??'
v.	BLN	*ah?*	'huh?' [open clarification question]
vi.	ZDL	*papaita* [points up]	'??'
vii.	BLN	*baluskia?*	'(is it) a bird now?' [offering target form]
viii.	ZDL	*paita*	'??'
ix.	BLN	*balus kiamit*	'the bird is gone' [expansion]

In the Qaqet corpus, we identified all non-target-like utterances and annotated on a separate tier for a) the non-target-like aspect of the utterance (pronunciation, lexicon, etc.), b) the response (ignore, clarification request, etc.), and c) the child's reaction (ignore, repeat, etc.). We want to re-iterate, though, that it is not necessary to systematically annotate your corpus: you can always choose to do a manual search, eyeballing the data.

The following questions serve as guidelines for identifying and describing the relevant phenomena:

• Do people respond to non-target-like utterances? Or do they tend to ignore them?

• If people respond to non-target-like utterances, how do they do so? Note whether you observe any of the following:

  - Adults or older children imitating and mimicking the utterances of younger children? Teasing them about their errors?

  - Explicit corrections? For example, interlocutors may tell the child explicitly that they produced something wrong, and model how to produce it instead.

  - Implicit corrections? For instance, interlocutors may repeat the child's utterance and expand on it by adding missing words or morphemes, replacing non-target-like forms, or giving additional information not provided by the child.

- If you observe any of the above, make a note of the context (if possible). For example, ignoring may be common in multiparty interaction, but rare in dyadic interaction. Equally, explicit correction may be found in staged play contexts, but may be rare in other contexts. Interlocutors may also react to non-target-like utterances differently depending on the child's age.

- What kinds of non-target-like utterances do interlocutors react to? For instance, they may react to articulations, lexical choices, grammatical structures, pragmatics, etc.

- What is the form of the response? Interlocutors may imitate the entire utterance or may integrate (part of) the imitation into a clause, for example. Or they may expand by adding specific items (e.g. subject indexes, or object noun phrases).

- How do the children respond? Do interlocutors try to get children to repeat the target form?

Box 3. Key areas of focus in child-directed language: Responses to non-target-like utterances.

Core
(i)   Do interlocutors respond to non-target-like utterances of children?
(ii)   What kinds of non-target-like utterances do interlocutors react to?
(iii)   What is the form of the responses?
Extension
(iv)   How do children respond?

5.3   Phonology and prosody

Child-directed language is often characterized by specific prosodic features. In many cultures adults tend to articulate more slowly and clearly, in what can be perceived as an exaggerated form of adult language. When describing child-directed language, we therefore recommend noting any conspicuous features in the following areas:

- Pitch and suprasegmentals: Do you observe a high F0? Exaggerated pitch contours? Exaggerated signing?

- Speech and sign rate: Do you observe a long duration of utterances? Of (vowel) phonemes? Long pauses between utterances?

- Hesitations and disfluencies: Do you observe fewer hesitations than in language addressed to older children or adults (if you have comparison data)? What are the characteristics of hesitations – for example, are there hesitation particles (such as *uh*), do they occur with a specific pitch contour (such as a level pitch), etc.?

- Articulation: Do you see adults communicating with more opened mouths or larger signs or seeming to articulate words more carefully?

A comprehensive articulatory or acoustic analysis is beyond the scope of the sketch format, but consider the following pragmatic approach. On the basis of your knowledge of the adult language, you will probably be able to identify the most conspicuous differences impressionistically. Once you have identified candidate features, conduct spot-checks comparing the language addressed to the youngest versus the oldest children in the sketch

corpus. If you have data from the same adult interlocutor addressing both younger and older children, choose this adult. Otherwise, try to choose interlocutors who are as similar as possible (e.g. same gender, similar age) and utterances of the same structural type such as directives or questions.

For example, we conducted such spot-checks on polar questions in the Qaqet sketch corpus. We annotated the corpus for speech acts (on a separate tier labeled "speech acts"), extracted intonation units that contained polar questions, and inspected their pitch contours in Praat (Boersma & Weenink 2022). Regardless of addressee, all Qaqet polar questions are characterized by a distinctive rise-fall contour, but those units addressed to children (up to age 3;6) exhibit exaggerated pitch movements (illustrated in Figure 1a) compared to polar questions addressed to older children and adults (illustrated in Figure 1b).

Figure 1a. Polar question addressed by AMT to YDS (2;0): *kua nyinarli?* 'do you hear?' (LongYDS20150506_1   1167.752 1168.561).
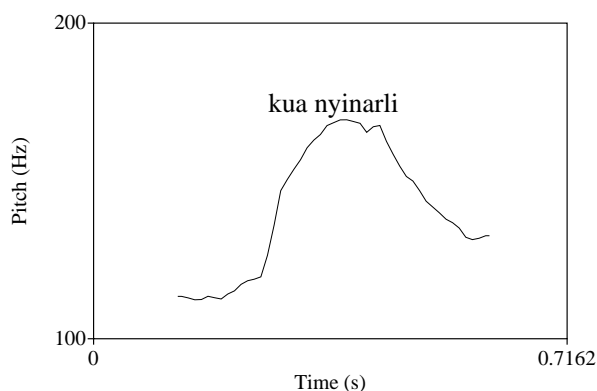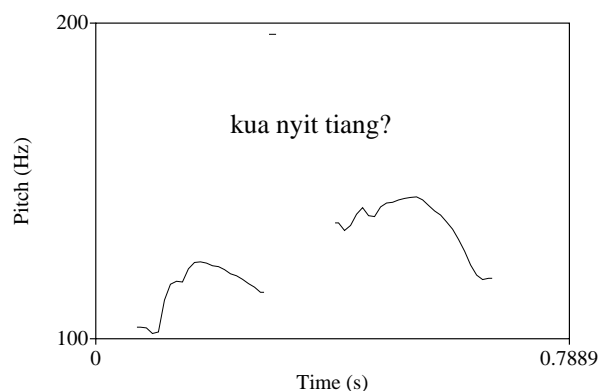
Figure 1b. Polar question addressed by AMT to ZVI (adult): *kua nyit tiang?* 'do you get others?' (LongYDS20151204_1 1001.330 1002.025).

Possible measurements include (see Frye 2022 for details):

- Measuring F0, maximum pitch and minimum pitch for individual utterances, and on this basis calculating mean F0 and frequency range. Praat or other software designed for acoustic analysis will be helpful here.

- Relating the number of words (or morphemes or syllables; see Section 6.3) to the length of speech/sign (excluding pauses), resulting in measures such as X number of words per second of speech/sign.

- Relating the number of hesitation pauses to the number of words, resulting in measures such as X number of hesitation pauses per X number of words.

Treat the above remarks as recommendations only. Depending on your background and research interests, you can expand on them (and go beyond spot-checks and conduct a systematic comparison across all ages) or restrict the investigation to a qualitative description.

Aside from a description of prosodic features, your sketch should also include a description of any divergences from adult-directed phonology and syllable structure. During the initial interviews and observations, adults may have reported on some phonemes being

'difficult' for children. If this is the case, pay special attention to their use of these phonemes, for example, they may substitute them. In addition:

- Do you observe interlocutors substituting phonemes? (e.g. *twain* for *train*)

- Do you observe the omission or addition of phonemes? (e.g. *dolly* for *doll*)

- Do you observe a reduction of consonant clusters? (e.g. *tummy* for *stomach*)

- Do you observe that adults produce underlying forms and suspend morphophonological changes?

Box 4. Key areas of focus in child-directed language: Phonology and prosody.

Core	
(i)	Are there any conspicuous prosodic features (pitch and suprasegmentals, speech or sign rate, hesitations, disfluencies)? Please report your impressions (based, for example, on your knowledge of the adult language and spot-checks of the sketch data).
(ii)	Are there any divergences from adult-directed phonology and syllable structure?
Extension	
(iii)	Expand on (i): systematically measure selected prosodic features.

5.4   Lexicon and semantics

In some languages, there is a distinct nursery vocabulary, such as *beddy-byes* (for going to bed) or *doo-doo* (for poo) in English. If there is such a nursery vocabulary, it is very likely that adults are aware of it and are able to supply typical lexemes during your initial interviews. Please take both the interview data and the sketch corpus data into account when answering the following questions:

- Is there such a nursery vocabulary?

- If yes, what are its structural properties? For example, are the words derived from standard lexemes? If yes, how (e.g. via shortening, reduplication, phonological changes, the addition of diminutives)? Are they onomatopoetic words such as *woof-woof* (for dog)? Or are they distinct and apparently underived lexemes?

- What sort of words have replacement nursery lexemes? Include a list of nursery lexemes attested in the sketch corpus, across semantic fields (see Section 6.2 for more detail on semantic fields). If possible, also include known or salient nursery lexemes not represented in the sketch corpus.

- To what extent and in what contexts is nursery vocabulary used by adults vs. children? In given interactional sequences, to what extent is its use initiated by adults vs. children? To what extent is its use reciprocal vs. non-reciprocal across conversational turns, e.g. with the child using an adult form followed by the adult using a nursery form or vice versa?

For example, Pitjantjatjara child-directed language makes extensive use of a nursery lexicon. Most words are derived via phonological substitutions which are seen as typical of children's language. Some words are apparently underived, though – particularly those used

frequently with very young children. Some examples of both types are in Table 3 (underlined consonants indicate retroflex place of articulation, so *r* is a retroflex approximant and *n* is a retroflex nasal). Note that nursery words formed via phonological substitution are lexically specific; for example, the nursery form of *kuṯa* is always *tjutja*, not *tjutju* or *utja*. All words in Table 3 occur within the Pitjantjatjara sketch corpus except for *ngunytji* 'untrue', though it does occur in the wider Pitjantjatjara acquisition corpus.

Table 3. Examples of Pitjantjatjara nursery lexicon.

Adult form	Nursery form	English	Phonological substitution
nyara	*nyaya*	there	r̲/r -> y
waṉi	*wanyi*	throw	retroflex -> palatal
ngunti	*ngunytji*	false/untrue	alveolar -> palatal
kami	*ami*	grandmother	initial consonant elision
kangkuṟu	*kaku*	older sister	cluster reduction & final syllable elision
kuṯa	*tjutja*	older brother	retroflex -> palatal & consonant harmony
kapi	*apa*	water	vowel harmony & initial consonant elision
mai 'vegetable food'/ *kuka* 'meat food'	*taa*	food	
kunkunpa	*paipai*	sleeping	
ipi	*ama*	breastmilk	

Irrespective of the existence of a nursery vocabulary, we recommend characterizing the semantic content of child-directed language:

- What do adults and older children talk about to younger children? Are there any recurring topics? For example, locating referents, or talking about kinship relations?

- What is the distribution of semantic fields in the input? We recommend that you categorize the vocabulary according to general semantic fields (see Section 6.2 for possibilities).

- Do you observe a preference for reference to current time and location (vs. past/future or distant locations)? Pay special attention to any other references, for instance narrations about past events.

Box 5. Key areas of focus in child-directed language: Lexicon and semantics.

Core	
(i)	Is there a nursery vocabulary? If yes, what are its structural properties? Which words have replacement nursery lexemes?
(ii)	What do adults and older children talk about to younger children?
Extension	
(iii)	What is the distribution of semantic fields?
(iv)	Is there a preference for talking about the present time and location?

5.5 Morphosyntax

In many societies, child-directed utterances tend to be short but complete and correct, becoming longer and more complex with the increasing age of the child. We suggest that you address the following questions:

- Are there any differences in the length of utterances addressed to younger vs. older children? The length is typically measured in terms of words or morphemes or syllables per utterance (the mean length of utterance (MLU); see Section 6.3 for a discussion of possibilities and recommendations). Calculate the MLU in child-directed language for at least part of your corpus, ideally allowing you to compare the MLU of the same interlocutor addressing younger vs. older children.

- Do you observe any morphosyntactic differences to adult-directed language? For example:

  - Are child-directed utterances essentially correct and complete? Do you observe structures that are ungrammatical in adult-directed language? Any obvious morphosyntactic simplifications, such as the omission of inflections or articles?

  - What are the typical morphosyntactic structures? For example, do you observe a preference for particular constituent orders (e.g. AVO in free-order languages), or for specific TAM morphology or speech acts (e.g. imperative morphology)?

In the Inuktitut sketch, for example, the MLU of child-directed utterances increased from a mean of 2.43 morphemes per utterance for children aged 1;4 to a mean of 4.40 for children aged 3;4. Typical examples at each age are in (3). We expected that mothers might simplify utterances by omitting verbal inflections given the complexity of the verbal inflection paradigm in Inuktitut, but we did not find any evidence of this. However, we did find that the majority of inflections at 1;4 were imperatives, and that inflections for more conceptually complex verbal moods such as conditional or dubitative did not appear until 2;4 and were used rarely.

(3)   Inuktitut: a. Mother of Jini (1;4); b. Mother of Lizzie (3;2); c. Mother of Louisa (3;3)

　　a. *qai-git*　　　　　// *taku-guk*　　　　　　　 // *ijukka-tuq*
　　　　come-IMP.2SG.SBJ　// see-IMP.2SG.SBJ>3SG.OBJ　// fall-PAR.3SG.SBJ
　　　　'Come here.'　　　 // 'Look at it.'　　　　　　// 'It fell.'

　　b. *aanni-tau-tsarua-ravit.*
　　　　hurt-PASS-might-CTG.2SG.SBJ
　　　　'You might get hurt.'

c. *qausi-alu-nnik   atu-qatta-qit?*
    wet-AUG-MOD.DU   wear-HAB-INT.2SG.SBJ
    'Did you wear wet (boots)?'

Child-directed language is furthermore often characterized through various forms of repetitions, in particular, partial or varied repetitions, where interlocutors repeat their message but introduce variations of the form, as in example (4) from English. Küntay & Slobin (1996) coined the term "variation set" for this phenomenon. These sets serve a communicative function, and interlocutors produce them to help focus the child's attention and to encourage the child to respond. From a learning perspective, they are known to facilitate language learning, since they allow the learner to compare across near-identical utterances. From a descriptive perspective, they give us insights into the permutation possibilities of a language.

(4)   Küntay & Slobin (1996: 267):

       Who did we see when we went out shopping today?

       Who did we see?

       Who did we see in the store?

       Who did we see today?

       When we went out shopping, who did we see?

Variation sets have been found to be very frequent in the input to young children in many languages (making up around 20% or more of child-directed language), and become less frequent as the child grows older. In some languages, however, there is reported variation in their structures, their functions, and their frequencies. In one study of variation sets in polysynthetic languages, for example, the variation sets become more (not less) frequent in the input to older children (Lester et al. 2022).

In your acquisition sketch, we suggest that you aim to identify the various types of (exact and varied) repetitions, addressing the following questions:

• For all forms of exact and varied repetitions: How common are they? Do you observe any age-related differences? Both could be discussed with respect to the age of the interlocutor (e.g. adult vs. older children) and the age of the child (younger vs. older children).

• For variation sets:

  - What are the functions of variation sets? For example, Küntay & Slobin (2002) distinguish between control-oriented sets (asking a child to do something), information-querying sets (asking a question that requires a verbal reaction) and ideational sets (imparting information, e.g. labeling something for the child).

  - Who produces them (e.g. adults, older children)? Are they produced by a single interlocutor, or do several participants contribute and repeat different parts?

  - What is the role of the child? Are they expected to respond? Are variation sets terminated when the child responds? What happens when the child does not respond?

  - What are the structural properties of the sets? Do you observe any of the following: addition and deletion of material (e.g. adding or deleting a morpheme), substitution of material (e.g. pronouns for nouns), re-ordering of material (e.g. changes in constituent order)? What kind of material is typically added, deleted, substituted, re-ordered (e.g. adpositional phrases, deictic morphology etc.)?

- Do you observe formulaic utterances? These are similar to variation sets, but they differ in that the message is not kept constant (e.g. *has mommy gone to work? has daddy gone to work?*)?

- Do you observe exact repetitions? In which contexts?

- Do you observe interlocutors repeating (all or part of) a child's utterance? In which contexts? What is repeated?

Box 6. Key areas of focus in child-directed language: Morphosyntax.

Core	
(i)	Do you observe any differences in the mean length of utterance? If yes, in what way?
(ii)	Do you observe any morphosyntactic differences to adult-directed language? If yes, which ones?
(iii)	Do you observe any repetitions and variations? If yes, which types (e.g. variation sets, formulaic utterances etc.)?
Extension	
(iv)	Expand on (iii): What are the structural properties of repetition and variation?
(v)	Expand on (iii): What are the functions of repetition and variation? Who produces them? How does the child respond to them?

5.6   Gestures

Language is not unimodal and we encourage you to maintain awareness of gesture during your sketch analysis. In terms of child-directed language, there is a high functional load on pointing and gaze in establishing and negotiating joint attention, a critical foundation of social and language learning (e.g. Baldwin 1995; Goldin-Meadow 2007; Rowe 2000). Gesture analysis is exceptionally time consuming, however, and so a full analysis goes well beyond the scope of the sketch format. In most cases we do not recommend detailed gesture coding (one exception is if the language contains an alternate sign system). If you decide to code for gesture, it will often be enough to simply describe the gesture type and its function on a separate tier in your transcription (e.g. POINT: imperative, EMBLEMATIC: head nod). From these codes a taxonomy of gesture types can be described, giving an overview of non-verbal communication in the language. The following Figures show an example of adult pointing (Figure 2a) and child pointing (Figure 2b); both from the Murrinhpatha context. Note that in 2a the adult's fingers are not fully retracted and she faces away from the object of joint attention but this is still a clear pointing gesture. In 2b the child's gaze is congruent with the direction of pointing.

Figure 2a. Adult pointing gesture

Figure 2b. Child pointing gesture

Two aspects of gesture in child-directed language are particularly worth paying attention to:

- Deictic gesture: Deictic gestures, which for adults involve pointing, are key components of joint attention modulation (alongside gaze). We encourage you to report on the following topics:

  - What sorts of things are children's conversational partners pointing at? For example, objects, people, actions, etc.

  - What functions does pointing have in the interaction (see Section 6.5)? Is it being used to direct attention or control behavior (i.e. used imperatively), in discussing an object of shared attention, or in negotiating an activity (i.e. used declaratively)? (See e.g. Salo et al. 2019.)

  - Do interlocutors respond to children's pointing gestures? How?

  - How are they pointing? With an extended index finger, middle finger, whole hand, head tilt, lip point, or something else?

  - Do you notice any changes in pointing behavior relating to the age of the child?

  Note that all of this is likely to be impressionistic based on what you notice as you transcribe and code the data. If you choose not to code pointing directed to children, you could include some discussion of pointing with a general discussion of who talks to children and how (see Section 4).

- Conventionalized gestures: These are recognizable symbolic gestures with a lexical-like meaning (e.g. Matsumoto & Hwang 2013). For instance, a head nod to mean 'yes', a finger wagging to mean 'no', or an index finger in front of the lips to mean 'quiet'. Conventionalized gestures are likely to be highly frequent with children and, as with the verbal lexicon, there may be specialized forms for use with children. We encourage you to notice the range and frequency of such gestural communications and include them in your consideration of child-directed lexicon (see Section 5.4). They are likely to be more or less important in different communities. For instance, in the Pitjantjatjara context, the general community makes wide-scale use of a highly developed alternate sign language (Ellis et al. 2019). It would then be a key component of the acquisition sketch to describe the extent to which these conventionalized signs are used with children.

Box 7. Key areas of focus in child-directed language: Gestures.

Core	
(i)	What types of deictic gestures do caregivers use?
(i)	What conventionalized gestures do caregivers use (including any alternate sign systems)?
Extension (see Section 6.5 for details)	
(iii)	Do caregivers use representational gestures, and if so, how?
(iv)	For each type of gesture, how do caregivers use them to support communicative development in combination with speech/sign?
(v)	For each type of gesture, how does their use by caregivers change across development of the child?

6   Child language

In this section, we turn to child language, covering phonology and prosody (Section 6.1), lexicon and semantics (Section 6.2), morphology (Section 6.3), and syntax (Section 6.4). We also include some pointers to the description of gesture (Section 6.5).

You should aim to give an overview of what children produce at each of the five different age points, with a focus on the observed developments across the ages. For example, you might observe that the children regularly substitute a certain phoneme at age 2;0, but start producing it at age 2;6, and seem to have target-like command over it at age 3;0. Depending on the available data and the phenomenon, it will not always be possible (or useful) to identify and describe developments across every six-month interval. For example, the development of a phenomenon may only be of interest for the two younger age groups, or your data may only allow for a comparison between a younger and an older age group.[2] In other words, you do not need to describe each phenomenon for each age group separately, but should feel free to adapt your description accordingly.

When analyzing the data, please distinguish between utterances that children produce spontaneously and those where they repeat an interlocutor's previous utterance. On the one hand, repetitions should be treated with caution, as children may simply be repeating them as a whole without analyzing their component parts. On the other hand, describing what the children repeat in a more or less target-like way can give you valuable insights. Thus, patterns of repetition may allow you to form reasonable hypotheses about the children's developing knowledge.

Furthermore, we suggest that you watch out for the following kinds of phenomena, as they may be indicative of children's developing knowledge:

- The expansion of an element to other contexts; for example, a morpheme that is initially only used with a few nouns attaches to more and more nouns.

- The emergence of small paradigms; for instance, the use of the same verb root with different tense/aspect morphemes.

---

[2] Again, please recall that we use "age" as a proxy for developmental stage. You may well decide not to compare between a younger versus older age group, but between a low versus high MLU group (see Part I, Section 2.1.4).

- Overgeneralizations and systematic errors (e.g. extending the use of the English past tense suffix to produce novel forms like *runned* instead of *ran*). If they are persistent, they may be a systematic generalization, and shed light on children's understanding. However, be careful to not read too much into them. If they are rare, they could simply be an articulation error made in that specific moment.

As always, keep in mind the limitations of the data set and be conservative about any generalizations you make.

6.1   Phonology and prosody

Children develop a sensitivity towards the phonological system of their language long before age 1;0, learning to distinguish the relevant phonemic contrasts (e.g. Werker et al. 2012), and producing articulations that become increasingly similar to the phonemes of their language (e.g. de Boysson-Bardies & Vihman 1991). This development continues well into the period covered by the sketch corpus, with children mastering the basics of phonology at about age 2;6, and continuing to acquire more complex aspects of phonology, phonotactics and prosody over the following years (e.g. Johnson & Reimers 2010). The sketch corpus will feature a large number of both target- and non-target-like realizations of words, thus enabling you to describe salient patterns and put forward hypotheses about developments in this area. We recommend paying special attention to contexts where children repeat their interlocutors' utterances: their repetitions may be more or less target-like, providing important clues to their developing phonological knowledge. Describing phonological development will provide you with crucial information for evaluating questions in lexical and morphological development, such as helping you to decide if a non-target-like form is a morphological error or a regular phonological substitution.

A useful way to approach this topic is to work with lists of the language's phonemes, allophones and syllable structures, ticking off the productions you observe for children of each age bracket. We suggest that you describe the following issues and their development over the different age points:

- Phoneme inventory:

  - Which phonemes are the children using? If possible and applicable, include any observations on differences in production to that of the adults.

  - Do you observe any phonemes that are regularly substituted by other phonemes? Or any regular omission of a particular phoneme? For instance, Pitjantjatjara-speaking children often substitute a palatal approximant *j* for an alveolar trill *r*.

  - Which phonemes are the children not using? (Keep in mind, though, that non-attested phonemes may simply constitute gaps in the data.)

- Phonological processes:

  - Do the children produce allophones and phonological processes in the appropriate contexts? Which ones? Or do they produce underlying forms?

  - Do you observe any non-target-like assimilations? For example, do they produce a consonant harmony where all the consonants within a word are at the same place of articulation?

- Phonotactics:

  - What kinds of syllable structures do the children produce?

- What do they do with complex syllable types? For example, do they simplify consonant clusters by omitting or adding sounds, or do they omit coda consonants? For instance, Pitjantjatjara often features non-homorganic nasal-stop clusters. Children often produce these as homorganic clusters. For example, *punka<u>n</u>u* 'fell' was produced by A at 2;6 as *puntunu* and by D at 2;8 and 3;3 as *pungkanu*.

If you are working with a tonal language, you can do some initial work on charting use of tones by the children you are recording. However, discussion of complex lexical tone and tone processes such as tone sandhi is beyond the scope of this sketch work. We suggest that you focus on the following features of tone and their development over the different ages:

- At what age or stage of the phonological acquisition process do children use tone and show evidence of having acquired it?

- What differences in tone use do you note across different ages?

- What does the acquisition of tone look like in relation to other phonological contrasts such as the acquisition of consonants and vowels?

See Singh & Fu (2016) for more information about the acquisition of lexical tone.

An important topic that spans across different levels of analysis (phonology, lexicon and morphology) is the structure of words. We introduce this topic here in the phonology section, but depending on the language and the focus of your description, you may want to discuss it under lexicon or morphology instead. The issue is that very young children often truncate longer words. Research shows that languages differ in this respect, and that children tend to reduce complex target words to a single morpheme in an agglutinative language (e.g. Inuktitut), but to a single stressed syllable in a more fusional language (i.e. often only part of a morpheme, or a string that goes across morpheme boundaries; e.g. K'iche' Mayan, Mohawk). In the Inuktitut example in (5), the child is kissing a character on the TV. The sister comments on this using various affixes appended to the verb root, while the child keeps responding with simply the verb root.

(5)   Inuktitut: Child (Jini 1;4) and her sister

Sister: *maa-li-ruk.*
           kiss-POL-IMP.2SG.SBJ>3SG.OBJ
           'Kiss her.'

Child: *maa.*
           kiss
           'Kiss.'

Sister: *maa-pait.*
           kiss-PAR.2SG.SBJ>3SG.OBJ
           'You kissed her.'

Child: *maa.*
           kiss
           'Kiss.'

Sister: *maa-tau-laur-langa.*
           kiss-PASS-POL-IMP.1SG.SBJ
           'Kiss me.' [lit: let me be kissed]

> Child: *maa.*
>   kiss
>   'Kiss.'

In contrast, children learning K'iche' Mayan tend to produce the final syllable of the target verbal word regardless of whether it is a morpheme or not. In example (6), *loh* contains the final consonant from the verb *il* 'see' and the termination morpheme *oh*. Similarly, *lik* contains the final consonant from the verb *ar* 'sleep' and the termination morpheme *ik*.

(6)   K'iche' Mayan (Pye 1983: 587): A1 Tiya:n (2;2)

> Adult: *la: utz kawiloh*
>   'Do you like it?'

> Child: *jah, loh (= jah, kinwiloh)*
>   'Huh? I like it.'

> Adult: *kawarik*
>   'He's sleeping.'

> Child: *lik (= kawarik)*
>   'He's sleeping.'

For further information on this topic, see Tzakosta & van de Weijer (2006) for a discussion of some of the phonetic, morphological and semantic factors involved in child truncations; see Demuth (1996) and Fikkert (1994) for prosodic theories of development. For a good example of investigating this topic with small amounts of data in Murrinhpatha, see Forshaw (2021).

Given the importance of this topic, we recommend describing any patterns you observe in the truncation of words. In this context, it is useful to pay special attention to stress (if the language has stress) and/or edge position (i.e. initial or final syllables), for example:

* Do children only utter the stressed syllable of longer words?

* Do they predominantly produce words that exhibit a specific stress pattern in the adult language?

* Do they generalize a (dominant) stress pattern of the adult language, and always place stress on the same syllable?

You may also want to describe other prosodic aspects of the children's production. It may be the case that the prosodic system of the adult language is not well understood yet, making it difficult to compare and detect non-target-like productions. But you can observe what the children are producing, and may well be able to address some of the following issues:

* Do you observe any salient prosodic patterns? For example, do children use recurring pitch contours? Do they produce a specific question intonation, imperative intonation, etc.?

* When an utterance contains two or more words, do children integrate them prosodically? Or do they break them up into two or more prosodic units?

Box 8. Key areas of focus in child language: Phonology and prosody.

Core

(i)   What is the distribution of phonemes, tonemes and syllable structures across development?

(ii)   Do you observe any (target- or non-target-like) phonological processes? Which ones and at which ages?

(iii)   Do you observe any patterns in the truncation of words? Which ones and at which ages?

Extension

(iv)   Expand on (iii): investigate children's use of stress.

(v)   Do you observe any salient prosodic patterns in child language? Which ones?

6.2   Lexicon and semantics

Children typically produce their first word at around twelve months of age (e.g. Casillas et al. 2020; Fenson et al. 2007), so it is reasonable to expect that even the youngest focus children will produce single-word utterances. Thus, a primary and achievable goal of the sketch will be to document both the content of the emerging lexicon and how it changes across the ages of the focus children. Describing the developing lexicon provides important clues about the nature of development across the entire linguistic system (see also Section 6.1 on word structure). In older children, the emergence of morphosyntactic knowledge has been linked to the acquisition of a 'critical mass' of lexical items (Marchman & Bates 1994). Therefore, it is likely that a researcher will observe growth in grammar as the lexicon increases in size.

In this vein, a sketch will most optimally provide information about the different tokens and types produced by the focus children, which should minimally be coded for the following information: (i) syntactic category (e.g. nouns, verbs, other prominent parts of speech, making special note of language-specific grammatical categories)[3] and (ii) semantic field. The semantic fields will no doubt have both a universal and context-specific flavor. Some childhood experiences are universal (e.g. kin terms, foodstuffs, basic human activities like eating and going to the toilet), while others are culturally-specific. We recommend attempting to categorize the children's vocabulary according to general classes such as: kin terms, deixis (see also Section 6.5), body parts, social conventions (e.g. greetings), natural kinds (e.g. animals, plants), inanimate technologies (e.g. vehicles, digital technologies), and terms for culturally-relevant belief systems (e.g. religious entities or mythical creatures). A possible extension would be a more systematic investigation into specific semantic fields. A good candidate could be the domain of plants or animals, e.g. you could take children on "plant walks" to elicit their semantic knowledge about plants (as done by Stross 1969; see also Si 2020 on possible relationships between language loss and loss of traditional ecological knowledge).

One useful starting point for a richer sketch of the lexicon is the set of semantic fields used in the MacArthur-Bates Communicative Development Inventories (MB-CDI; Fenson et al. 2007). An example comes from Inuktitut, where we created a version of the MB-CDI using the data from an acquisition study (Allen et al. 2017). We started with the 21 semantic

---

[3] Please use the syntactic categories of the adult language. These may or may not correspond to the children's categories, but it will likely not be possible to decide one way or other.

categories in the original English MB-CDI, and then sorted the vocabulary from the Inuktitut acquisition data into those categories. Not surprisingly, many words in the English version were also relevant for Inuktitut (e.g. doll, dog, soup, dress, tooth, blanket, window, cloud, mother, yes, tomorrow, me, what, here, walk, eat, small). However, the vocabulary from our acquisition data led us to adapt the words in several sections to reflect Inuit culture (e.g. animals, vehicles, food/drink, clothing) and to reflect aspects of the Inuktitut semantic system that are more elaborated than in English (e.g. baby words, kinship terms, locatives). Some examples are in (7). Although this was done on the basis of some 20 hours of data rather than the 5 hours planned for the acquisition sketch, the same principle could easily be applied for the sketch data.

(7)   Inuktitut:

Animals: *aiviq* 'walrus', *umimmak* 'muskox', *uviluq* 'mussel', *tuktuk* 'caribou'

Vehicles: *anartauti* 'septic truck', *haanta* 'all-terrain vehicle', *qajariaq* 'canoe'

Food/Drink: *panirtitaq* 'bannock', *puijiviniq* 'seal meat', *misiraq* 'whale oil'

Clothing: *amauti* 'parka for carrying baby', *atigi* 'parka', *pualuq* 'mitten'

Baby Words: *aataaq* 'hurt', *amaama* 'bottle, suckle', *ammu* 'sleep', *vuvu* 'vehicle'

People: *ajak* 'maternal aunt', *akkak* 'paternal uncle', *najak* 'sister of boy'

Locations: *avani* 'there', *maani* 'here', *paani* 'up there', *kanani* 'down here'

Another common question in children's lexical development focuses on the relative distribution of nouns and verbs in children's early lexicon. While early work from English claimed that nouns strongly outweigh verbs in children's early lexicon, later work from typologically different languages was instrumental in refining that claim, see Tardif (1996) and Waxman et al. (2013). For an excellent example of how a categorization of lexical development can be done with a small amount of data, see Taverna & Waxman (2020).

One obstacle in working with children's data that will be evident at the word level is that children do not always produce adult-like forms. Pye (2021) recommends producing a lexical concordance for child utterances and adult targets and makes further recommendations for how that could be used to study other typologically interesting features of the language. We highly recommend the paper, which provides helpful direction in how a researcher could extend the use of their data looking at the lexicon in both the sketch and in work that might be conducted with their data beyond the sketch. For instance, one interesting topic concerns the study of word structure across development (see Section 6.1).

A possible further extension concerns semantic development. This is a fascinating but surprisingly less-studied feature of child language, likely because inferring what children mean is a difficult task. Some features of semantic development, such as children's early inferences about category boundaries and how they change across development, require detailed observations of individual children that will be beyond the scope of the sketch (see Bowerman 1980). This also applies to common topics in the study of semantics, such as children's acquisition of quantifiers and evidentiality. However, other features of semantic development may be more tractable and worthy of mention. For instance, the study of adpositions lends itself well to the study of semantic development, and there may be enough data to describe their use across development (Clark & Carpenter 1989; Johnston & Slobin 1979; Tomasello 1987). Sometimes the language itself will make inferences about children's semantic category development more tractable, such as when a language has a classifier system that makes distinctions based on semantic criteria (e.g. the elaborate classifier systems found in some Australian languages).

Box 9. Key areas of focus in child language: Lexicon and semantics.

Core	
(i)	What is the distribution of word classes and semantic fields across development?
(ii)	What is the relative distribution of nouns to verbs across development?
(iii)	When do children begin to use other word classes?
(iv)	Do you observe any patterns in the truncation of words? Which ones and at which ages? (See Section 6.1 on word structure)
Extension	
(v)	What additional observations can be made about children's semantic development, given the language and the data?

6.3   Morphology

Languages differ widely in their morphological complexity (e.g. isolating vs. polysynthetic, agglutinative vs. fusional), and each type of language presents different challenges for learning morphology. Thus, you should begin by thinking carefully about the morphology of the language and what challenges it is likely to present. The morphological system is also very likely to affect patterns of acquisition. For example, children are faster and show fewer non-target-like forms learning inflectional systems that are systematic (e.g. Turkish, Inuktitut) than systems that are not systematic (i.e. with many empty cells like English, or with lots of syncretism like German), regardless of the relative complexity of the system. In addition, very young children tend to reduce morphologically complex target words (see Section 6.1 on word structure).

Regardless of the language typology, important morphological developments will take place during the ages covered by the sketch corpus. For example, the first instances of inflectional morphology will likely start to appear around age 2;0 and become more frequent and regularized over the following age points. The main challenges will be to identify the morphemes of each type at each stage, and then to hypothesize about possible developments between stages. Depending on what is relevant for the language, we suggest focusing on:

- Nominal morphology: case, number, gender, noun class, etc.

- Verbal morphology: agreement, tense, aspect, evidentiality, etc.

- Grammatical morphology: passive, antipassive, causative, negation, etc.

It will likely make most sense to focus on large-scale developments and on identifying indicators of potential stages in acquisition. In the Pitjantjatjara sketch corpus, for example, Wighton (2021) was able to distinguish three stages for verbal inflection (children whose mean length of utterance was under 2 vs. 2-3 vs. over 3). This does not mean that there are no smaller-scale stages in between, but the sketch data will likely be too limited to detect them. When trying to identify stages, it is useful to pay attention to the following kinds of evidence:

- an increase in the number of categories that children mark (e.g. different cases, number categories, tense/aspect categories, subject or object indexes)

- an increase in the number of nouns/verbs that these morphemes attach to

- an increase in the complexity of the forms

For example, children learning morphological causatives in Inuktitut go through three main stages (Allen 1998). During the first stage (ca. 1;4-2;6), they produce no morphological causatives; all causative notions are expressed with lexical causatives, usually with an imperative inflection (as in example 8). In the second stage (ca. 2;0-2;10), their morphological causatives are restricted to one of three forms: verb root and causative morpheme with no inflection (9a), causative morpheme with imperative inflection but no verb root (9b), or causative with both verb root and imperative inflection but always the same fixed form (9c). Structures with no verb root are sometimes produced where an adult would use a morphological causative, but often where an adult would use a lexical causative. Thus, it seems that the children are using the rootless morphological causative as a strategy since they are not yet sure which verb root takes which causative form. Structures with the same fixed form are likely used as one unanalyzed unit. In the third stage (ca. 2;4-3;6), morphological causatives are used for the most part with the proper verb roots and inflections, and with a wider range of inflections including declaratives (10a) and interrogatives (10b). Any increase in number and complexity is thus likely to be indicative of children starting to analyze a particular morpheme, extending it to other nouns/verbs, and gradually integrating it into mini paradigms contrasting different person or number categories.

(8)   Inuktitut: a. Tumasi (1;9); b. Jini (1;4)

   a.   *ukkui-*
         open.door
         'Open the door.'

   b.   *qai-lau-ruk*
         come-POL-IMP.2SG.SBJ>3SG.OBJ
         'Give it to me.'

(9)   Inuktitut: a. Sarah (2;0); b. Lizzie (2;6); c. Elijah (2;0)

   a.   *tii-tur-ti-lau.*
         tea-consume-CAUS-POL
         'Let (her) have tea.'

   b.   ti-lau-ruk   //   ti-lau-nnga
         CAUS-POL-IMP.2SG.SBJ>3SG.OBJ   //   CAUS-POL-IMP-2SG.SBJ>1SG.OBJ
         'Make it do X.'   //   'Make (let) me do X.'

   c.   *taku-ti-lau-nnga*   //   *sikituur-ti-lau-nnga*   //   *piir-ti-lau-nnga*
         see-CAUS-POL-IMP.2SG.SBJ>1SG.OBJ   //   ride.skidoo-...   //   remove-...
         'Let me see   //   ride the skidoo   //   get off.'

(10)   Inuktitut: a. Paul (3;3); b. Elijah (2;5)

   a.   *panik   itsiva-ti-tait*
         daughter   sit-CAUS-PAR.2SG.SBJ>3SG.OBJ
         'Daughter, you made it sit.'

   b.   *... akki-mik   taku-ti-laa-raminga?*
         ... fish.hook-ABS.SG   see-CAUS-FUT-CTG.4SG.SBJ>1SG.OBJ
         '... will he let me see the fish hook?'

An important but difficult question is whether a given morpheme is productive for the child or whether it is (part of) an unanalyzed, rote-learned, or repeated form. A good way to approach this issue is to provide ratios, for example the percentage of obligatory contexts in which a morpheme is used, and/or the percentage of forms of the adult paradigm used by the

child. Furthermore, several tests can help to discern productivity, including: (i) use of the morpheme in conjunction with more than one other morpheme (e.g. *-ed* in *walked*, *shouted*, *cleaned*), (ii) use of other morphemes in the same slot (e.g. *-ed* in *walked* vs. *-s* in *walks* vs. *-ing* in *walking*), and (iii) overgeneralization of the morpheme (e.g. *-ed* in *goed*, *runned*, *eated*). Allen's (1996: 165-171) assessment of the productivity of noun incorporation in Inuktitut offers a good example of how to apply these tests. However, there probably will not be enough data in your corpus to decide definitively if most morphemes are productive. Thus, we suggest that you mention patterns that are strongly suggestive one way or the other (e.g. if a given morpheme only ever appears in one combination, or if a given morpheme clearly passes all the tests just mentioned), but that you otherwise remain neutral on this question.

Another concern is unambiguously identifying a particular morpheme (see also Part I, Section 3.3). The initial and final points in development are usually less problematic: an early absence of the morpheme in question, and a later target-like realization of the morpheme. In between, it is very likely that children will produce non-target-like realizations including partial morphemes (e.g. only a vowel) and considerable variations in morpheme articulation (e.g. pronouncing a consonant in different ways). Very young children may also produce 'filler syllables' – single syllables that appear in the position where a particular class of morpheme would normally appear (e.g. article, subject index), and thus show early knowledge of the class. You may not be able to resolve this analytical uncertainty, but you can address it by focusing on the context of the utterance. For example, a child may have produced an unidentifiable vowel morpheme in a structural position and in a context where adults would use an accusative suffix. This does not necessarily mean that the child intended to produce the accusative suffix, but you can now compare all accusative contexts in your data and analyze what the child produced in these contexts (and possibly compare it to all nominative or all dative contexts).

The average number of morphemes per utterance is also widely used to compare children across stages of linguistic development, using the measure known as MLU (mean length of utterance). Calculating the MLU is straightforward: total the number of morphemes in a sample of 100 consecutive utterances, and divide by 100. In English, for which it was originally developed (Brown 1973), MLU typically increases with age and language ability until at least age 4;0, following the logic that each additional morpheme represents a new piece of grammatical knowledge. Although MLU is also used in many other languages, it may be calculated differently depending on the typology of the language. For languages with clear morpheme boundaries and few portmanteau morphemes (e.g. Turkish), calculating MLU by morphemes is very sensible. However, calculating MLU by words makes more sense for isolating languages (e.g. Mandarin), languages with highly fusional or portmanteau morphology (e.g. Dene), or languages where new meaning units are added by changes in articulation rather than adding morphemes (e.g. Irish). Calculating MLU by syllable is also possible – a useful option for languages with regular syllables that are easy to hear in speech (e.g. Inuktitut). Allen & Dench (2015) provide an excellent summary of the relevant literature, as well as a comparison of the three options using Inuktitut as a test case.

A set of guidelines for how to count morphemes is provided in Brown (1973: 54) and summarized in Parker & Brorson (2005: 374). Several types of words and utterances are typically excluded from the analysis because they do not index the child's linguistic ability. These include filled pauses (*um*), utterances that are not fully intelligible, routinized utterances (songs, nursery rhymes), and immediate partial or full repetitions (of self or interlocutor). Language-specific rules indicate what to count as one versus two morphemes, largely on the basis of what unit is most likely to be productive for the child. In English, for example, only one morpheme is typically assigned to compound words (*birthday*), reflexives

(*herself*), diminutives (*doggie*), irregular forms (*went*), proper names (*Mickey Mouse*), and contractions (*don't*) and concatenations (*gonna*) that are likely not productive. In contrast, two morphemes are assigned to overgeneralizations (*goed, runned*) and to contractions that are likely productive (*she's*). A similar set of guidelines would need to be developed for your sketch language. Despite its drawbacks, MLU is currently the best general measure of development in the field. We suggest that you calculate it for 100 utterances of each session in your sketch corpus. For corpora prepared in CHAT format, MLU can be calculated automatically using the relevant program in CLAN (MLU based on morphemes requires morphological annotation, whereas MLU based on words can be calculated without that).

Box 10. Key areas of focus in child language: Morphology.

<div style="border:1px solid black">

Core

(i)   What types of morphemes occur at each stage? What kinds of nominal morphology? Verbal morphology? Grammatical morphology?
- nominal morphology (e.g. case, number, gender, class)
- verbal morphology (e.g. agreement, tense, aspect, evidentiality)
- grammatical morphology (e.g. passive, antipassive, causative, negation)

(ii)   Is there any evidence for stages of development for a given type of morpheme? Do you observe an increase in the number of categories marked? In the number of items the morpheme attaches to? In the complexity of forms?

(iii)   What, if any, evidence can be used to show that children have a productive understanding of the morphology (rather than reproducing memorized forms)?

(iv)   What is the mean length of utterance for each session? Is this best measured by morpheme, word, or syllable?

Extension

(v)   For one or two morpheme types that are particularly salient in the language (and where there is enough data available), do a more systematic analysis.

</div>

6.4   Syntax

Languages also differ widely in their syntactic structure (e.g. strict vs. flexible word order, syntactic operations signaled through word order vs. morphology). Thus, each language will present different challenges for the acquisition of syntax. Regardless of these differences, however, you will see an increasing complexity in the syntactic structures across the different ages covered in the sketch corpus. In the youngest age group, one-word utterances are likely to predominate (i.e. no syntax at all). Two-word and/or two-morpheme utterances will probably start to appear around 2;0, and utterances will become longer on average over the next two years of age. More complex structures will likely start to appear by age 2;6 or 3;0, depending on the child.

   The structure of the language may well influence how the earliest stage of syntax is characterized. In English, this is commonly referred to as the 'two-word stage' (because children produce two separate words in their earliest utterances containing syntactic relations) and the 'telegraphic stage' (because children omit function words and content words likely interpretable by the interlocutor so the resulting utterance sounds like a telegram, e.g. *want cookie* vs. *I want a cookie*). If the language has very little in the way of

morphology, then this stage will likely also contain two-word utterances. However, if the language is morphologically 'rich', then it will more likely be characterized by two-morpheme utterances since the syntactic advances will be signaled in morphology (e.g. verbal inflection rather than independent subject). This is illustrated in (11): an utterance that would constitute four morphemes in the target is rendered as a two-morpheme utterance by a Murrinhpatha-speaking child. Similarly, if the language has many function words and requires subjects, then children's earliest utterances with syntactic relations will probably also look telegraphic. However, if the language realizes function words in other ways (e.g. through affixes) and allows subject omission, then this stage will probably not look too different from typical adult utterances.

(11)   Murrinhpatha: Child (Acacia 2;7)
       (LAMP_20130524_WF_01_V1 00:20:33; Forshaw 2016: 356)

       Child   ˈipirt=nga?

       Target   nguˈngu-**pirt**-nu=**nga**
                 1SG.SBJ.REMOVE(32).FUT-take.off-FUT=DM
                 'I will take it off?' [Acacia is fiddling with the buckle on the microphone backpack she is wearing]

Early word combinations will be one main focus of this section. Here we suggest that you first characterize which types of words are combined (again, please use the categories of the adult language; see Footnote 3). For example, children may combine two words from open classes – typically noun+noun or noun+verb, but maybe also noun+adjective or verb+adverb. Or they may combine a word from an open class with a function word (e.g. with negator *no*, or additive particle *too*). Observe typical combinations and describe the order of elements. For example, the function word may always appear in a specific position, or the two content words may always be combined in a certain order. It is likely that early word and morpheme combinations will have a lexically-specific flavor (e.g. *want + X, see + X, where's + X?*, see Lieven et al. 1997), which are taken to indicate that early syntax is built around an inventory of concrete lexically-based constructions (see Ambridge & Lieven 2011). This may mean that the construction is a useful unit of analysis for the sketch. In general, the acquisition literature claims that word order differences between child and adult language are rare, but data from a wider variety of languages may call this assumption into question. For languages with flexible word order, it would be useful to characterize which word orders are most frequent for the children at different ages, at least for a couple of common combinations. While the sketch data is likely to be too sparse to allow clear hypotheses regarding exactly how and why children are varying word order, looking at variations within individual contexts, particularly with the same verb, can provide some clues. For instance, in example (12) from Pitjantjatjara-speaking Rachel (3;6), one possibility is that she is fronting objects on initial mentions. This would then be a hypothesis to test in subsequent work.

(12)   Pitjantjatjara: Rachel (3;6) talking to a younger (Iti) and older cousin (Murphy).
       (PITJACQ_Rachel_20170507_00:03:48)

       i.   *Iti,*   *puli*   *katu*   *nya-nganyi*
            baby   hill   top   see-PRS
                      O   V
            'Baby, see the hill top'

ii. *Murphy,   nya-nganyi   puli   katu   mula*
     Murphy   see-PRS   hill   top   true
     A   V   O
     'Murphy, you can see the hill top, right'

iii. *puli   katu   xxx   ngara-nyi-n*
     hill   top   stand-PRS-?
     'standing xxx on the hill top'

iv. *Murphy   nya-nganyi   puli   katu=nta*
     Murphy   see-PRS   hill   top=and.then
     A   V   O
     'Murphy can you see the hill top?'

v.   *nyuntu   nya-nganyi=n?*
     2SG.NOM   see-PRS=2SG.NOM
     A   V
     'Can you see (it)?'

vi. *mamu   ngayu   nya-nganyi*
     monster   1SG.NOM   see-PRS
     O   A   V
     'I see a monster'

Characterizing the functions of early word combinations is also important. Typical early functions include location (agent plus location, theme plus location), possession (possessor plus possessed), and verb plus argument (verb plus agent, verb plus patient/theme). Be aware, though, that the interpretation of early utterances is strongly dependent on the context, given that children typically use one or two salient words to characterize an entire situation. As a result, a combination such as *baby blanket* could be interpreted as a possessive structure ('baby's blanket'), an action ('baby plays with blanket'), a location ('baby is on blanket'), or a request ('give me my blanket'). Or *granny* in a combination such as *granny go* could be an agent ('granny went'), but also a goal ('to granny (we) go'). Thus, it is very helpful to pay close attention to the context of the utterance, and also to accept that some utterances will not be able to be interpreted unambiguously.

The way in which these early words are combined is also worth noting, especially in terms of morphology and prosody. Early word combinations may or may not have the relevant morphology (see Section 6.3) and may or may not show prosodic integration (see Section 6.1). For example, children may produce words with two primary accents or separate pitch contours, and/or with a pause between them. Make a note of any morphological and prosodic peculiarities of such early combinations, as well as the appearance of more target-like patterns.

Argument realization has received a lot of attention in the acquisition literature. The key question is the form in which arguments are realized – as a lexical noun phrase, demonstrative, pronoun, or omitted (e.g. Allen et al. 2015; Hyams 2011). Children are sensitive early on to the tendency for arguments that are accessible to the interlocutor to be omitted or realized as pronouns, and for less accessible arguments to be realized as full NPs. Following Preferred Argument Structure, higher ranking arguments (e.g. referents taking the A role in transitive sentences)[4] tend to be more accessible than lower ranking, dependent

---

[4] Some researchers prefer to approach argument structure through syntactic functions (subject, object etc.), and others, through semantic roles (agent, patient etc.). Please feel free to adhere to your preferred approach. For

arguments (e.g. referents taking an O or S role). We suggest that you document the proportion of referents realized as A, S and O in each form at each age in your sketch data, and to comment on any developmental patterns. If you have more resources, you could also determine whether the choice of form correlates with accessibility of the referent, for example in terms of new versus given. Another possibility is to look at the relationship between gestures and argument realization, especially if gestures are salient in the language (cf. Özçalişkan & Goldin-Meadow 2005; see Section 6.5). At early ages, children may combine speech and gesture to produce a complete utterance. For example, they may use a point or other deictic gesture to realize an argument, or may use an iconic gesture to realize an argument (e.g. *milk* = hand pretending to hold a bottle) or a verb (e.g. *eat* = hand pretending to put food in mouth).

Another main focus of this section should be on the appearance and development of different syntactic constructions. Depending on the language, this may include several of the following:

- nominal structures: possession, noun modification

- verb types: intransitive, transitive, ditransitive

- syntactic functions: subject, object, indirect object

- valency alternation: causative, passive, antipassive, noun incorporation

- early constructions: negation, imperatives, questions, coordination

- later constructions: subordinate clauses, relative clauses

We suggest that you start by identifying the constructions used by the oldest age groups in your corpus, and describing their properties (e.g., word order, relevant morphology). Then analyze the distribution of these constructions among the younger age groups. Wherever possible, note the age when a construction first appears in your corpus as well as any salient developments. For example, in the Inuktitut sketch corpus, there is very little noun incorporation in the youngest age group (restricted to the copula), but its use increased in both frequency and variety by 2;6 (to include direction and possession), and was more prevalent by 3;0.

---

the Sketch Manual, we use the labels A (more agent-like argument in a transitive clause), S (only argument in an intransitive clause) and O (more patient-like argument in a transitive clause).

Box 11. Key areas of focus in child language: Syntax.

Core	
(i)	Is the two-unit stage best characterized as two-word or two-morpheme?
(ii)	What types of word combinations are used at different ages? Is there any evidence for stages of development? Are there any patterns in the types of words combined? In the order of words in word combinations? In the function of the combinations?
(iii)	What word orders are used for expressing the verb and its arguments – e.g. SV/VS, OV/VO, AOV/AVO/VAO? What orders are most frequent? Does this change with age?
(iv)	What syntactic constructions are used? Is there evidence for development or stages? Select a few structures that are most relevant for the language and/or that show developmental patterns, for example nominal structures, verb types, syntactic functions, valency alternation, early constructions, later constructions.
Extension	
(v)	For one or two structures that are particularly salient in the language (and where there is enough data available), do a more systematic analysis.
(vi)	How are arguments realized (lexical NP, demonstrative, pronoun, null form)? Does this differ for subject vs. object? Does this change with age?
(vii)	Is the form of argument realization (lexical, demonstrative, pronoun, null) associated with accessibility of the referent?
(viii)	Is there a relationship between argument realization and production of gestures?

6.5   Gestures

Non-verbal communication forms an important means by which children break into language. Accordingly, much of infants' early communicative acts either involve non-verbal gesture or combine early acquired words with gestures, most typically deictic pointing gestures (e.g. Bates et al. 1979). Thus, gesture and speech/sign are closely coupled in acquisition. As children develop, their dependence on gesture decreases. However, gestures still form an important additional channel of communication, even in the adult system. An interesting feature of communicative gesture is that there appear to be quasi-universal and culturally-specific modes of gesture, which makes this a particularly interesting phenomenon to address in a sketch (Lieven & Stoll 2013).

Deictic (or pointing; see also Section 5.6) gestures are the first manual gestures to emerge, typically in pre-linguistic infants during the first year of life, expressing functions such as *giving*, *showing*, *requesting*, and *attention-directing*. The form of these will differ depending on the function. For instance, requesting may be expressed as the hand extended with a flat palm, while showing will involve either an 'in-hand' display of an object or a point to an object in space. Declarative pointing, which emerges around twelve months, has been argued to have special significance for children's cognitive development (Tomasello et al. 2007), heralding a step-change in children's understanding of and interest in others' mind states. Thus, when an infant directs the attention of another to an object using a declarative point, they arguably demonstrate an emerging understanding of others as individual agents whose contents of mind can be manipulated. Declarative pointing has been argued to be a universal

in both form and function (Liszkowski et al. 2012). While a sketch beginning at two years is unlikely to capture the emergence of deictic gestures, identifying their presence or absence is recommended. This may be done in concert with deictic vocabulary (see Section 6.2). It would also be important to note if the language has a mode of pointing that is outside of the typically-used extended index-finger (e.g. lip-pointing), and if there are culturally-specific contexts in which pointing is used (see Cooperrider et al. 2018; Wilkins 2003).

Conventionalized gestures (e.g. waving hello/goodbye, gestures of negation) will also form a significant component of young children's early communicative repertoire (and of child-directed language; see Section 5.6). They are interesting because of their cultural specificity; for instance, Inuit raise their eyebrows to denote 'yes', and scrunch their nose to indicate 'no'. Thus, they will no doubt be noticeable and should be relatively straightforward to code and describe. Some cultures may be more gesture-rich than others (see Marentette et al. 2016), which will likely be most observable in this category, including prominent alternate sign systems (Ellis et al. 2019).

Two other categories of gestures may appear but may be given less attention. *Representational* or *iconic* gestures are imagistic gestures that have semantic content. These can include hand or body movements, or facial expressions. For instance, a child raising their hand to their mouth and simulating eating would refer to EATING; pretending to stir an imaginary spoon in a pot to COOKING. Thus, representational gestures are frequently action-based, embodied expressions with specific meaning, which is relatively stable across different uses. Representational gestures are fairly rare in child-directed language (Iverson et al. 1999), but may form a crucial iconic gateway for children into spoken or signed language (Morin-Lessard et al. 2021). They are also not always easy to spot. We do not recommend spending large amounts of time coding and describing them, if you choose to code them at all. A simple description of when they appear (i.e. are they more common in younger children?), what they look like (e.g. *child mimes chewing something to denote* EATING) and in what contexts (e.g. child requests food) is enough.

The final gesture category, *beat* gestures, are staccato hand gestures that closely align with the prosodic characteristics of speech or sign but which have no semantic content. While they are no doubt interesting in their own right, they are beyond the scope of a sketch.

Box 12. Key areas of focus in child language: Gestures.

Core	
(i)	What types of deictic gestures do children use?
(i)	What conventionalized gestures do children use (including any alternate sign systems)?
Extension	
(iii)	Do children use representational gestures, and if so, how?
(iv)	For each type of gesture, how do they support communicative development in combination with speech/sign?
(v)	For each type of gesture, how does their use change across development?

7   Community report

In the preceding sections we have focused on the academic report of how children acquire the community language. However, it is just as important to communicate the findings of an acquisition sketch study to the local community. By definition, a community which is the focus of an acquisition sketch study will have limited research-based information regarding how children learn their language. In many cases, they will be interacting with language testing tools and education materials designed for another language. In all cases, what you can contribute from the sketch project will be valuable. There is, however, still a need to be careful not to generalize too far from the sketch findings and it is likely that more care will be needed around this in communicating your findings to the community than in communicating your findings to academia.

The nature of exactly what sort of information is most relevant, how it should be presented, and what resources are available for creating community-oriented outputs will vary widely. For instance, a community which is literate and teaching the focus language in school will likely be interested in literacy and education-focused materials, such as readers for different ages/levels and reports for teachers. This includes materials for early childhood education, cf. e.g. the "language nest" approach employed in the revitalization of languages such as Māori or Hawai'ian (e.g. Chambers & Saddleman 2020; see Koller et al., forthcoming for a general introduction). Furthermore, a sketch will show recurring patterns and routines that are frequent enough to emerge even in a small data set. It is likely that these routines include patterns that can function as carrier phrases for the introduction of new words. Such carrier phrases, in turn, play an important role in language teaching, and can be utilized in teaching materials. You may be able to work with community educators to guide the development of such resources throughout the sketch project. In contrast, a community with no general literacy in the focus language, and which may be under pressure from a dominant regional language, may be more interested in information regarding the extent to which children are using and learning the focus language versus the other community languages. Alternatively, the community may be interested in materials for parents that explain about developmental stages. For example, the Inuit-run school board in arctic Quebec, *Kativik Ilisarniliriniq*, has created a host of educational resources for parents and teachers based at least partly on previous studies of language development in Inuit communities (https://www.kativik.qc.ca/en/educational-resources/). You and the community will be in the best position to determine what community-oriented outputs are most valuable. Discuss options with the community and determine a plan for developing these alongside your more academic-focused sketch. It is vital for the knowledge gained through the sketch writing process to be made accessible to the community. Below is an incomplete and not mutually exclusive list of potential ideas. We encourage you to include a summary of your activities in the acquisition sketch, in the hope that it will serve as a source of inspiration for others.

- Possible ways of presenting the information:
  - Formal written document

    This is most similar to the academic sketch but would still need extensive adjustment to tailor to a more general audience.

  - Videos

    Videos can provide a good alternative to written reports, particularly (but not only) in communities with low literacy rates. A good example of this can be seen in O'Shannessy et al. (2020) where researchers and educators reflect on the process of using videos, discussions and workshops to translate research findings into educator

professional development outcomes at Yuendumu school in Australia (available at https://carmeloshannessy.net/#alw).

- Informal oral presentations or discussions

    You may find it best to communicate findings through oral presentations in community meetings and discussions. This could be especially valuable when there is already an established community meeting format for information sharing.

- Suggestions for what information to include:
  - Everything

    One possibility is to write a full alternative sketch covering all information provided in the academic sketch, but targeted for a community or professional audience.

  - Highlights or showcases of language socialization strategies and/or child linguistic competence

    It could be particularly beneficial to highlight features which clearly demonstrate competence or skill. This is particularly important in cases of language endangerment or where the language and its community is viewed as a minority or inferior in some way, where a focus on strengths can serve to showcase highlights of children's linguistic competence as well as caregiver's nuanced practices of socialization and language teaching. For instance, in some Australian Indigenous communities, teachers sometimes say that children come to school with 'no language at all' because they don't speak English nor do they speak the traditional language 'properly'. In such cases, it is very valuable to be able to showcase the linguistic skills of preschool aged children to demonstrate the rich knowledge and skills they are in fact bringing with them when they start school. Likewise, highlighting culturally nuanced and appropriate socialization techniques employed by adults can be helpful in combatting the deficit mentality in relation to caregiving and parenting.

  - Vocabulary checklist

    A community and language relevant vocabulary checklist is something which communities and education or speech pathology professionals often request. This is especially the case if they are already engaging with language development testing, potentially using materials which are not tailored to the community language(s) such as a MacArthur-Bates Communicative Development Inventory in a dominant regional language. However, developing or adapting such a test to your focus language is a time intensive endeavor. For guides and suggestions from previous successful checklist developments, see Jones et al. (2020) for Kriol and Reese et al. (2015) for Samoan and Tongan; for a general screening tool, see Dench et al. (2011) for Inuktitut; for a morphosyntactic assessment, see Allen et al. (2019) for Inuktitut.

    An alternative approach is to create wordlists from your sketch data. This can be an initial start on a larger checklist development project. The idea here would be to extend the lexicon component of your sketch and make it accessible to a wider audience. The list(s) should include information on the ages children are observed using each word and which words are more or less frequent in your recordings. This information can then be used by others to develop readers or initial language development testing materials. For instance, Rebecca Defina has collaborated with speech pathologist Michelle Harvey working in the Pitjantjatjara-speaking communities. Defina provided lists of particularly frequent and infrequent words and Harvey utilized this to develop pilot testing materials for use in schools.

- Language use

  In cases of potential language shift, some community members may be particularly concerned regarding the extent to which the children are fully acquiring the language. Acquisition sketch projects provide an opportunity to more neutrally and measurably check the extent to which children in the community are acquiring the language or shifting to a new language. Community members may be surprised to learn that children are in fact using a lot of the traditional language, although they are also using some highly salient words from the majority languages they encounter in their community.

We strongly encourage you to consider, discuss, and plan community outputs from the outset of your project. Work with the community to determine what information they are most interested in and how best to provide it. Throughout the project you and the community may adjust these plans, but this can best be done as part of ongoing continued discussions. Do not let community outputs be an afterthought which can all too easily be left out when time or funding runs out.

8   Outlook

And thus concludes our sketch manual. In writing the manual we aimed to integrate key elements of language documentation and child language acquisition, in an attempt to build a renewed enthusiasm for collecting crosslinguistic data in a way that responds to the large problem of language endangerment and loss. Our solution to the problem is one that borrows ideas from both fields and is thus, potentially, of interest to many. Our hope is that researchers, community language workers, and any other interested parties take up the challenge of collecting data and writing a sketch. The papers that follow in this special section are the first sketches, many of which are written by us and/or our students. These will also help future sketch writers to write their own sketches. The success of the project will be determined by the number of people that take up the challenge – we hope to have convinced you that it is a challenge that is worth the effort. Happy sketching!

References

Allen, Shanley E.M. 1996. *Aspects of argument structure acquisition in Inuktitut.* Amsterdam: John Benjamins. https://doi.org/10.1075/lald.13

Allen, Shanley E.M. 1998. Categories within the verb category: Learning the causative in Inuktitut. *Linguistics* 36(4). 633–677. https://doi.org/10.1515/ling.1998.36.4.633

Allen, Shanley E.M., Mary Cain, Catherine B. Dench, Catherine Genest & Natacha Trudeau. 2017. Inuktitut adaptation of the MacArthur-Bates Communicative Development Inventory. Poster at *14th International Congress for the Study of Child Language*, July, Lyon, France.

Allen, Shanley E.M. & Catherine B. Dench. 2015. Calculating mean length of utterance for Eastern Canadian Inuktitut. *First Language* 35(4-5). 377–406. https://doi.org/10.1177/0142723715596648

Allen, Shanley E.M., Catherine B. Dench & Kerry Isakson. 2019. InuLARSP: An adaptation of the Language Assessment Remediation and Screening Procedure for Inuktitut. In Martin J. Ball, Paul Fletcher & David Crystal (eds.), *Grammatical profiles: Further languages of LARSP*, 267–293. Bristol, UK: Multilingual Matters. https://doi.org/10.21832/9781788924399-014

Allen, Shanley E.M., Fred H. Genesee, Sarah A. Fish & Martha B. Crago. 2002. Patterns of code mixing in English-Inuktitut bilinguals. In Mary Andronis, Christopher Ball, Heidi Elston & Sylvain Neuvel (eds.), *Proceedings of the 37th Annual Meeting of the Chicago Linguistic Society* (vol. 2), 171–188. Chicago, IL: Chicago Linguistic Society.

Allen, Shanley E.M., Mary E. Hughes & Barbora Skarabela. 2015. The role of cognitive accessibility in children's referential choice. In Ludovica Serratrice & Shanley E.M. Allen (eds.), *The acquisition of reference*, 123–154. Amsterdam: John Benjamins. https://doi.org/10.1075/tilar.15.06all

Ambridge, Ben & Elena V. M. Lieven. 2011. *Child language acquisition.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511975073

Baldwin, Dare A. 1995. Understanding the link between joint attention and language. In Chris Moore & Philip J. Dunham (eds.), *Joint attention: Its origins and role in development*, 131–158. Hillsdale, NJ: Lawrence Erlbaum.

Bates, Elizabeth, Laura Benigni, Inge Bretherton, Luigia Camaioni & Virginia Volterra. 1979. *The emergence of symbols: Cognition and communication in infancy.* New York: Academic Press.

Boersma, Paul & David Weenink. 2022. Praat: Doing phonetics by computer [Computer program]. Version 6.2.09, retrieved 15 February 2022 from http://www.praat.org/

Bowerman, Melissa. 1980. The structure and origin of semantic categories in the language learning child. In Mary L. Foster & Stanley H. Brandes (eds.), *Symbol as sense*, 277–299. New York: Academic Press.

Bowerman, Melissa. 1982. Starting to talk worse: Clues to language acquisition from children's late speech errors. In Sidney Strauss (ed.), *U-shaped behavioral growth*, 101–145. New York: Academic Press. https://doi.org/10.1016/B978-0-12-673020-3.50012-4

Brown, Penelope & Suzanne Gaskins. 2014. Language acquisition and language socialization. In N. J. Enfield, Paul Kockelman & Jack Sidnell (eds.), *The Cambridge handbook of linguistic anthropology*, 187–226. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139342872.010

Brown, Roger. 1973. *A first language: The early stages.* Cambridge, MA: Harvard University Press. https://doi.org/10.4159/harvard.9780674732469

Casillas, Marisa, Penelope Brown & Stephen C. Levinson. 2020. Early language experience in a Tseltal Mayan village. *Child Development* 91(5). 1819–1835. https://doi.org/10.1111/cdev.13349

Chambers, Natalie A. & Danielle Saddleman. 2020. Moving towards a language nest: Stories and insights from nkmalqs. *First Peoples Child & Family Review* 15(1). 27–43. Retrieved from https://fpcfr.com/index.php/FPCFR/article/view/399

Child Language Research and Revitalization Working Group. 2017. *Language documentation, revitalization, and reclamation: Supporting young learners and their communities.* Waltham, MA: EDC.

Chouinard, Michelle M. & Eve V. Clark. 2003. Adult reformulations of child errors as negative evidence. *Journal of Child Language* 30(3). 637–669. https://doi.org/10.1017/S0305000903005701

Clark, Eve V. 1987. The principle of contrast: A constraint on language acquisition. In Brian MacWhinney (ed.), *Mechanisms of language acquisition*, 1–33. Hillsdale, NJ: Lawrence Erlbaum. https://doi.org/10.4324/9781315798721-7

Clark, Eve V. & Kathie L. Carpenter. 1989. On children's uses of *from*, *by* and *with* in oblique noun phrases. *Journal of Child Language* 16(2). 349–364. https://doi.org/10.1017/S030500090001045X

Cooperrider, Kensy, James Slotta & Rafael Núñez. 2018. The preference for pointing with the hand is not universal. *Cognitive Science* 42(4). 1375–1390. https://doi.org/10.1111/cogs.12585

Cristia, Alejandrina, Emmanuel Dupoux, Michael Gurven & Jonathan Stieglitz. 2019. Child-directed speech is infrequent in a forager-farmer population: A time allocation study. *Child Development* 90(3). 759–773. https://doi.org/10.1111/cdev.12974

Davidson, Lucinda K. 2018. *Allies and adversaries: Categories in Murrinhpatha speaking children's talk.* PhD Thesis, The University of Melbourne.

de Boysson-Bardies, Bénédicte & Marilyn M. Vihman. 1991. Adaptation to language: Evidence from babbling and first words in four languages. *Language* 67(2). 297–319. https://doi.org/10.2307/415108

De Cat, Cecile, Draško Kašćelan, Philippe Prevost, Ludovica Serratrice, Laurie Tuller & Sharon Unsworth. 2021. Delphi consensus survey on how to document bilingual experience. https://doi.org/10.31219/osf.io/ebh3c

Demuth, Katherine. 1996. Stages in the acquisition of prosodic structure. In Eve V. Clark (ed.), *Proceedings of the 27th Child Language Research Forum*, 39–48. Stanford: Stanford University Press.

Dench, Catherine B., Patricia L. Cleave, Jane Tagak & Janice Beddard. 2011. The development of an Inuktitut and English language screening tool in Nunavut. *Canadian Journal of Speech-Language Pathology and Audiology* 35(2). 168–177.

Ellis, Elizabeth M., Jennifer Green, Inge Kral & Lauren W. Reed. 2019. 'Mara yurriku': Western Desert sign languages. *Australian Aboriginal Studies* 2019(2). 89–111. http://hdl.handle.net/1885/216033

Fenson, Larry, Virginia A. Marchman & Donna Thal. 2007. *MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual.* Towsen: Paul H. Brookes.

Fikkert, J. Paula M. 1994. *On the acquisition of prosodic structure.* PhD Thesis, Leiden University.

Forshaw, William. 2016. *Little kids, big verbs: Acquisition of Murrinhpatha bipartite stem verbs.* PhD Thesis, University of Melbourne.

Forshaw, William. 2021. *The acquisition of complex morphology: Insights from Murrinhpatha.* Amsterdam: John Benjamins. https://doi.org/10.1075/tilar.30

Frye, Henrike. 2022. *Child-directed speech in Qaqet: A language of East New Britain, Papua New Guinea.* Canberra: ANU Press.

Gallaway, Clare & Brian J. Richards (eds.). 1994. *Input and interaction in language acquisition.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620690

Genesee, Fred H., Elena Nicoladis & Johanne Paradis. 1995. Language differentiation in early bilingual development. *Journal of Child Language* 22(3). 611–631. https://doi.org/10.1017/S0305000900009971

Goldin-Meadow, Susan J. 2007. Pointing sets the stage for learning language - and creating language. *Child Development* 78(3). 741–745. https://doi.org/10.1111/j.1467-8624.2007.01029.x

Hellwig, Birgit & Dagmar Jung. 2020. Child-directed language – and how it informs the documentation and description of the adult language. *Language Documentation and Conservation* 14. 188–214. http://hdl.handle.net/10125/24920

Henke, Ryan. 2019. The development of possession in the L1 acquisition of Northern East Cree. *Journal of Child Language* 46(5). 980–997. https://doi.org/10.1017/S0305000919000217

Henke, Ryan & Julie Brittain. 2022. Obviative demonstratives in Northern East Cree: Insights from child-directed speech. *International Journal of American Linguistics* 88(1). 53–93. https://doi.org/10.1086/717057

Hyams, Nina. 2011. Missing subjects in early child language. In Jill de Villiers & Tom Roeper (eds.), *Handbook of generative approaches to language acquisition*, 13–52. Dordrecht: Springer. https://doi.org/10.1007/978-94-007-1688-9_2

Iverson, Jana M., Olga Capirci, Emiddia Longobardi & M. Cristina Caselli. 1999. Gesturing in mother–child interactions. *Cognitive Development* 14(1). 57–75. https://doi.org/10.1016/S0885-2014(99)80018-5

Johnson, Wyn & Paula Reimers. 2010. *Patterns in child phonology*. Edinburgh: Edinburgh University Press. https://doi.org/10.1515/9780748642489

Johnston, Judith R. & Dan I. Slobin. 1979. The development of locative expressions in English, Italian, Serbo-Croatian and Turkish. *Journal of Child Language* 6(3). 529–545. https://doi.org/10.1017/S030500090000252X

Jones, Caroline, Eugenie Collyer, Jaidine Fejo, Chantelle Khamchuang, Anita Painter, Lee Rosas, Karen Mattock, Alicia Dunajcik, Paola Escudero & Anne Dwyer. 2020. Developing a parent vocabulary checklist for young Indigenous children growing up multilingual in the Katherine region of Australia's Northern Territory. *International Journal of Speech-Language Pathology* 22(5). 583–590. https://doi.org/10.1080/17549507.2020.1718209

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

Koller, Eve, Gabriela Amaller & Kenedi Cooper. Forthcoming. Language nests. *Oxford Bibliographies.* Oxford: Oxford University Press.

Küntay, Aylın & Dan I. Slobin. 1996. Listening to a Turkish mother: Some puzzles for acquisition. In Dan I. Slobin, Julie Gerhardt, Amy Kyratzis & Jiansheng Guo (eds.), *Social interaction, social context, and language: Essays in honor of Susan Ervin-Tripp*, 265–286. Hillsdale, NJ: Lawrence Erlbaum.

Küntay, Aylın & Dan I. Slobin. 2002. Putting interaction back into child language: Examples from Turkish. *Psychology of Language and Communication* 6. 5–14.

Lester, Nicholas A., Steven Moran, Aylin Küntay, Shanley E. M. Allen, Barbara Pfeiler & Sabine Stoll. 2022. Detecting structured repetition in child-surrounding speech: Evidence from maximally diverse languages. *Cognition* 221. 104986. https://doi.org/10.1016/j.cognition.2021.104986.

Levshina, Natalia. 2021. Conditional inference trees and random forests. In Magali Paquot & Stefan T. Gries (eds.), *A practical handbook of corpus linguistics*, 611–643. New York: Springer. https://doi.org/10.1007/978-3-030-46216-1_25

Lieven, Elena V. M., Julian M. Pine & Gillian Baldwin. 1997. Lexically-based learning and early grammatical development. *Journal of Child Language* 24(1). 187–219. https://doi.org/10.1017/S0305000996002930

Lieven, Elena V. M. & Sabine Stoll. 2013. Early communicative development in two cultures: A comparison of the communicative environments of children from two cultures. *Human Development* 56. 178–206. https://doi.org/10.1159/000351073

Liszkowski, Ulf, Penny Brown, Tara Callaghan, Akira Takada & Conny de Vos. 2012. A prelinguistic gestural universal of human communication. *Cognitive Science* 36(4). 698–713. https://doi.org/10.1111/j.1551-6709.2011.01228.x

Marchman, Virginia A. & Elizabeth Bates. 1994. Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language* 21(2). 339–366. https://doi.org/10.1017/S0305000900009302

Marentette, Paula, Paola Pettenati, Arianna Bello & Virginia Volterra. 2016. Gesture and symbolic representation in Italian and English-speaking Canadian 2-year-olds. *Child Development* 87(3). 944–961. https://doi.org/10.1111/cdev.12523

Mateo Pedro, Pedro. 2021. The acquisition of causatives in Q'anjob'al Maya. *First Language* 41(4). 451–477. https://doi.org/10.1177/0142723720966827

Matsumoto, David & Hyisung S. Hwang. 2013. Emblematic gestures (emblems). In Kenneth D. Keith (ed.), *The encyclopedia of cross-cultural psychology* (2nd edition), 464-466. Chichester: Wiley-Blackwell. https://doi.org/10.1002/9781118339893.wbeccp188

Morin-Lessard, Elizabeth, Rochelle F. Hentges, Suzanne C. Tough & Susan A. Graham. 2021. Developmental pathways between infant gestures and symbolic actions, and children's communicative skills at age 5: Findings from the all our families pregnancy cohort. *Child Development* 92(3). 799–810. https://doi.org/10.1111/cdev.13567

Muysken, Pieter C. 2000. *Bilingual speech: A typology of code-mixing.* Cambridge: Cambridge University Press.

Muysken, Pieter C. 2013. Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition* 16(4). 709–730. https://doi.org/10.1017/S1366728912000727

Newport, Elissa L. 1977. Motherese: The speech of mothers to young children. In N. John Castellan, David B. Pisoni & George R. Potts (eds.), *Cognitive theory* (vol. 2), 177–217. Hillsdale, NJ: Lawrence Erlbaum.

Ochs, Elinor & Bambi Schieffelin. 1996. The impact of language socialization on grammatical development. In Paul Fletcher & Brian MacWhinney (eds.), *The handbook of child language*, 73–94. Cambridge, MA: Basil Blackwell. https://doi.org/10.1111/b.9780631203124.1996.00004.x

O'Shannessy, Carmel. 2008. Children's production of their heritage language and a new mixed language. In Jane Simpson & Gillian Wigglesworth (eds.), *Children's language and multilingualism: Indigenous language use at home and school*, 261–282. London: Continuum Publishing. https://doi.org/ 10.5040/9781474212144.ch-012

O'Shannessy, Carmel. 2013.The role of multiple sources in the formation of an innovative auxiliary category in Light Warlpiri, a new Australian mixed language. *Language* 89(2). 328–353. https://doi.org/10.1353/lan.2013.0025

O'Shannessy, Carmel, Samantha Disbray, Barbara Martin & Gretel MacDonald. 2020. (Re)turning research into pedagogical practice: A case study of translational language research in Warlpiri. *Language Documentation & Conservation*, Special Publication 18. 139–151. http://hdl.handle.net/10125/24881/

Özçalişkan, Seyda & Susan Goldin-Meadow. 2005. Gesture is at the cutting edge of early language development. *Cognition* 96(3). B101–B113. https://doi.org/10.1016/j.cognition.2005.01.001

Parker, Matthew A. & Kent Brorson. 2005. A comparative study between mean length of utterance in morphemes (MLUm) and mean length of utterance in words (MLUw). *First Language* 25(3). 365–376. https://doi.org/10.1177/0142723705059114

Pizer, Ginger, Richard P. Meier & Kathleen Shaw Points. 2011. Child-directed signing as a linguistic register. In Rachel Channon & Harry van der Hulst (eds.), *Formational units in sign languages*, 65–84. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9781614510680.65

Pye, Clifton. 1983. Mayan telegraphese: Intonational determinants of inflectional development in Quiche Mayan. *Language* 59(3). 583–604. https://doi.org/10.2307/413905

Pye, Clifton. 1986. Quiché Mayan speech to children. *Journal of Child Language* 13(1). 85–100. https://doi.org/10.1017/S0305000900000313

Pye, Clifton. 2021. Documenting the acquisition of indigenous languages. *Journal of Child Language* 48(3). 454–479. https://doi.org/10.1017/S0305000920000318

Reese, Elaine, Elaine Ballard, Mele Taumoepeau, Melenaite Taumoefolau, Susan B. Morton, Cameron Grant, Polly Atatoa-Carr, Stuart McNaughton, Johanna Schmidt, Jatender Mohal & Lana Perese. 2015. Estimating language skills in Samoan-and Tongan-speaking children growing up in New Zealand. *First Language* 35(4-5). 407–427. https://doi.org/10.1177/0142723715596099

Rowe, Meredith L. 2000. Pointing and talk by low-income mothers and their 14-month-old children. *First Language* 20(60). 305–330. https://doi.org/10.1177/014272370002006005

Saint-Georges, Catherine, Mohamed Chetouani, Raquel Cassel, Fabio Apicella, Ammar Mahdhaoui, Filippo Muratori, Marie-Christine Laznik & David Cohen. 2013. Motherese in interaction: At the cross-road of emotion and cognition? (A systematic review). *PLoS ONE* 8(10). e78103. https://doi.org/10.1371/journal.pone.0078103

Salo, Virginia C., Bethany Reeb-Sutherland, Tahl I. Frenkel, Lindsay C. Bowman & Meredith L. Rowe. 2019. Does intention matter? Relations between parent pointing, infant pointing, and developing language ability. *Journal of Cognition and Development* 20(5). 635–655. https://doi.org/10.1080/15248372.2019.1648266

Saxton, Matthew. 1997. The Contrast Theory of negative input. *Journal of Child Language* 24(1). 139–161. https://doi.org/10.1017/s030500099600298x

Saxton, Matthew. 2000. Negative evidence and negative feedback: Immediate effects on the grammaticality of child speech. *First Language* 20(60). 221–252. https://doi.org/10.1177/014272370002006001

Schieffelin, Bambi & Elinor Ochs (eds.). 1986. *Language socialization across cultures.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511620898

Serratrice, Ludoviva. 2013. Crosslinguistic influence in bilingual development: Determinants and mechanisms. *Linguistic Approaches to Bilingualism* 3(1). 3–25. https://doi.org/10.1075/lab.3.1.01ser

Serratrice, Ludovica, Antonella Sorace & Sandra Paoli. 2004. Crosslinguistic influence at the syntax–pragmatics interface: Subjects and objects in English–Italian bilingual and monolingual acquisition. *Bilingualism: Language and Cognition* 7(3). 183–205. https://doi.org/10.1017/S1366728904001610

Shneidman, Laura & Amanda L. Woodward. 2016. Are child-directed interactions the cradle of social learning? *Psychological Bulletin* 142(1). 1–17. https://doi.org/10.1037/bul0000023

Si, Aung. 2020. Patterns in the transmission of traditional ecological knowledge: A case study from Arnhem Land, Australia. *Journal of Ethnobiology and Ethnomedicine* 16. 52. https://doi.org/10.1186/s13002-020-00403-2

Singh, Leher & Charlene S. L. Fu. 2016. A new view of language development: The acquisition of lexical tone. *Child Development* 87(3). 834–854. https://doi.org/10.1111/cdev.12512

Snow, Catherine E. & Charles A. Ferguson (eds.). 1977. *Talking to children: Language input and acquisition.* Cambridge: Cambridge University Press.

Stross, Brian. 1969. Language acquisition by Tenejapa Tzeltal children. *Working Paper* 20. Berkeley: Language Behavior Research Laboratory, University of California.

Tardif, Twila. 1996. Nouns are not always learned before verbs: Evidence from Mandarin speakers' early vocabularies. *Developmental Psychology 32*(3). 492–504. https://doi.org/10.1037/0012-1649.32.3.492

Taverna, Andrea S. & Sandra R. Waxman. 2020. Early lexical acquisition in the Wichi language. *Journal of Child Language* 47(5). 1052–1072. https://doi.org/10.1017/S0305000919000898

Tomasello, Michael. 1987. Learning to use prepositions: A case study. *Journal of Child Language* 14(1). 79–98. https://doi.org/10.1017/S0305000900012745

Tomasello, Michael, Malinda Carpenter & Ulf Liszkowski. 2007. A new look at infant pointing. *Child Development* 78(3). 705–722. https://doi.org/10.1111/j.1467-8624.2007.01025.x

Tomasello, Michael & Daniel Stahl. 2004. Sampling children's spontaneous speech: How much is enough? *Journal of Child Language* 31(1). 101–121. https://doi.org/10.1017/S0305000903005944

Tzakosta, Marina & Jeroen van de Weijer. 2006. On the role of phonological, morphological and semantic headedness in acquisition. In Piotr Banski, Beata Lukaszewicz & Monika Opalinska (eds.), *Studies in constraint-based phonology*, 223–244. Warsaw: University of Warsao.

Waxman, Sandra R., Xiaolan Fu, Sudha Arunachalam, Erin Leddon, Kathleen Geraghty & Hyun-joo Song. 2013. Are nouns learned before verbs? Infants provide insight into a long-standing debate. *Child Development Perspectives* 7(3). 155–159. https://doi.org/10.1111/cdep.12032

Werker, Janet F., H. Henny Yeung & Katherine A. Yoshida 2012. How do infants become experts at native-speech perception? *Current Directions in Psychological Science* 21(4). 221–226. https://doi.org/10.1177/0963721412449459

Wighton, Wanyima. 2021. *Children's verbal inflection development in Pitjantjatjara: An acquisition sketch.* Graduate Diploma in Arts (Advanced) thesis, University of Melbourne. http://hdl.handle.net/11343/290224

Wilkins, David. 2003. Why pointing with the index finger is not a universal (in sociocultural and semiotic terms). In Sotaro Kita (ed.), *Pointing: where language, culture, and cognition meet*, 171–215. Hillsdale, NJ: Lawrence Erlbaum.