

Child-directed language – and how it informs the documentation and description of the adult language

Birgit Hellwig

University of Cologne

Dagmar Jung

University of Zurich

Language documentation efforts are most often concerned with the adult language and usually do not include the language used by and with children. Essential parts of the natural linguistic behaviour of communities thus remain undocumented, and a growing body of literature explores what language documentation, language maintenance, and language revitalization have to gain by including child language and child-directed language.

This paper adds a methodological perspective to the discussion, arguing that child language and child-directed language constitute data types that can inform our understanding of the adult language. For reasons of feasibility, the paper focuses on child-directed language only. Presenting data from two on-going language acquisition projects (Qaqet from Papua New Guinea and Dëne Sųłné from Canada), we illustrate how this data type provides insights into the metalinguistic knowledge of adult speakers. After an introduction to child-directed language, three case studies on the topics of variation sets, clarification processes, and discourse context are exemplified from both languages and related to our understanding of the adult language. Focusing on the potential of this data type, this paper argues in favour of extending our documentation efforts to events involving children.

1. Introduction: The scope of the paper Over the past few years, we have seen an increased interest in including child language and child-directed language within the language documentation paradigm. This interest is based on a number of good reasons (see below), and this paper explores one additional – methodological – reason that is rarely discussed in the literature: how child-directed language¹ can inform our understanding of the adult language and the metalinguistic knowledge of adults. A similar reason can be given for child language, too, but for reasons discussed later on, the paper focuses on child-directed language only. The paper is situated partly in the

¹This paper uses the term *child-directed language* in place of the more established term “child-directed speech”, as it is neutral with respect to the modality, covering both spoken and signed languages. As will be discussed in this paper, there is little information on child-directed speech in endangered languages; and there is even less information on child-directed sign. Our own expertise is on spoken languages, and we refer the interested reader to on-going research by Victoria Nyst on language socialization in deaf families in Africa.

literature that discusses extending the language documentation paradigm to include children, and partly in the literature that discusses methodological issues and data types. This first section gives a brief overview of relevant aspects of both literatures, setting the scene for the more detailed discussion of child-directed language (§2) and its potential for understanding the adult language (§3).

From the perspective of language documentation, the literature highlights two major reasons for documenting both child language and child-directed language (Eisenbeiß 2005; Anand et al. 2010; Kelly & Nordlinger 2014; Kelly et al. 2015; Hellwig 2019b). First, the language used by and with children constitutes observable linguistic behavior in any community, i.e., its documentation fits squarely within the scope of the language documentation paradigm. In the light of striving for comprehensiveness and representativeness in our documentary efforts (see e.g., Seifart 2008), an extension to include children seems inevitable.

And second, over and above this truism, the documentation of inter-generational language use is a resource for the development of future community-based language strengthening programs, contributing to two essential areas in language maintenance and revitalization: a better understanding of processes which lead to language shift, as well as an understanding of communicative practices that typically facilitate language learning in this kind of setting. Given that languages become endangered when they are no longer being transmitted to the next generation (Grenoble & Whaley 2006), research into language acquisition and socialization in such contexts becomes an important prerequisite for understanding the processes that underlie language endangerment and loss. Such knowledge, in turn, is needed for developing strategies to support communities in their language maintenance and revitalization efforts, and to inform educational policies – for detailed examples of the possibilities, see especially the white paper by the Child Language Research & Revitalization Working Group (2017) (see also Kelly et al. 2015; O’Grady & Hattori 2016). It is true that, in many communities, the language has become endangered to such an extent that it is no longer being acquired, and a documentation of its acquisition is thus no longer possible. But it is equally true that, for many other communities, acquisition and socialization in the community language (or languages) is still taking place – possibly in a multilingual context, or even evolving together with a language of wider communication (see e.g., McConvell & Meakins 2005; O’Shannessy 2012 for the emergence of mixed languages in this context).

Both reasons sketched out above motivate the increasing interest in child language and child-directed language within the language documentation framework. This interest from within language documentation is met by a complementary interest from within another discipline, psycholinguistics. Psycholinguistics is interested in studying the acquisition of a wide variety of languages, albeit for a different reason: to enable crosslinguistic comparisons and discover universal aspects of language learning and processing. It is by now widely recognized that our generalizations about human language are based on a biased sample of languages from the so-called WEIRD societies (i.e., Western, Educated, Industrialized, Rich, Democratic, a term coined by Henrich et al. 2010; see also Evans & Levinson 2009). The extent of this bias differs

across the different psycholinguistic disciplines, whereby the study of language acquisition is the only one that can lay a claim to a cross-linguistic perspective. Its cross-linguistic outlook is due to a small number of large-scale initiatives such as the classic series *The crosslinguistic study of language acquisition* (Slobin 1985–1997), the *Frog Story* project (Berman & Slobin 1994; Strömquist & Verhoeven 2004), the creation of CHILDES (Child Language Data Exchange System; MacWhinney 2000), or the language socialization paradigm within anthropology (Schieffelin & Ochs 1986). But despite this cross-linguistic tradition, Lieven & Stoll (2009:144) are forced to remark that “[i]f we take all the acquisition studies together (experiments and longitudinal studies), we know something about the acquisition of approximately 70 to 80 languages (i.e., approximately 1% of all the languages spoken today). This 1% of languages also includes languages for which only one acquisition study of a single feature exists [...]” (see also Kelly & Nordlinger 2014). In the other psycholinguistic disciplines, the cross-linguistic picture is even worse (see Anand et al. 2010 for a survey; see Norcliffe et al. 2015 for language processing). This bias is particularly untenable in the light of the central aim of all psycholinguistic research: to understand universal aspects of language processing and learning, and to distinguish them from language-specific aspects. Psycholinguistics thus has an intrinsic interest in extending its research to more languages from across the world.

Given the above considerations, it is not unsurprising that there is a general push both from language documentation and from psycholinguistics towards documenting language use by and with children – depending on the discipline, the goal is either to achieve a more comprehensive documentation that is not restricted to the adult language, and/or to better understand the mechanisms of language shift and language loss, and/or to broaden the database for our psycholinguistic generalizations. Despite a general consensus, though, relevant documentation projects are still extremely few and far between. The largest and most comprehensive such project is set within the *Chintang Language Research Program*, which includes a strong language acquisition component led by Sabine Stoll and Elena Lieven (e.g., Stoll et al. 2012; Stoll & Bickel 2013). Within the context of Sabine Stoll’s *ACQDIV (Acquisition processes in maximally diverse languages: min(d)ing the ambient language)* project (Moran et al. 2016), Dagmar Jung is currently constructing a child corpus of Dëne Şuñné, a Dene (Athapaskan) language, which also includes documentation components. We are familiar with a number of recent projects on Australian and Papuan languages that integrate documentation and acquisition perspectives: Murrinhpatha (e.g., Forshaw et al. 2017), Pitjantjatjara (led by Rebecca Defina), Ku Waru (e.g., Rumsey 2017), Nungon (led by Hannah Sarvasy), and Qaqet (e.g., Hellwig *to appear*). And we are aware of a handful of researchers starting or intending to start such research with South American and African communities. Outside of the language documentation paradigm, a number of large-scale projects investigate the acquisition of languages from non-WEIRD societies, many of them having originated in the context of Dan Slobin’s (1985–1997) *crosslinguistic study of language acquisition*. For overviews of these projects and their impact, we refer the interested reader to Bowerman (2011), Lieven & Stoll (2009), Slobin & Bowerman (2007), and Stoll (2009).

It is very likely that this fairly small number of existing projects is due to numerous methodological and ethical challenges (Kelly et al. 2015; Hellwig 2019b): to comprehensively investigate the acquisition of any language, large amounts of data need to be collected and processed, ideally with a longitudinal setup and ideally with a good number of children (for discussions of sampling issues, see Tomasello & Stahl 2004). The studies mentioned in the above paragraph predominantly follow longitudinal approaches and research several children, and more such studies are needed to reduce the bias in our databases. However, it is also clear that the number of such comprehensive projects will continue to remain small. Given the overall scarcity of resources (in terms of budget, personnel, and time), there is of necessity a trade-off, as documentation projects and communities have to weigh up interests and set priorities. And while there are very good reasons to document child language, there are equally good reasons to invest the scarce resources into different, equally important, aspects of documentation.

But while the comprehensive study of language acquisition may be beyond the possibilities of most documentation projects, it may still be possible to document child language and child-directed language on a smaller scale. As outlined in the summaries above, there are a number of good reasons for doing so. This paper, too, argues in favour of including at least the documentation of child-directed language in our documentation projects. More specifically, it does so by exploring one additional argument in its favour – an argument that has not featured prominently in our discussions so far: the potential of child-directed language to enhance our understanding of the adult language. In a similar way, Eisenbeiß (2005) makes a good case for including child language. The reason for our focus on child-directed language is feasibility: although research into children’s developing knowledge of a language will give insights into the adult language as well, its documentation tends to be much more involved, entailing a long-term commitment and a team of researchers focusing almost exclusively on this goal. The documentation of child-directed language, by contrast, can be realistically integrated into many documentation projects, capitalizing on language documentation’s expertise with documenting observable linguistic behavior, including different registers – with child-directed language being one of the registers that we can realistically include in a documentation.

More generally, the focus of this paper ties into the methodological debate on the role of different data types within language documentation. In his programmatic article, Himmelmann (1998:166) sketched out the scope of language documentation as being centered around the linguistic practices and traditions of speech communities, recognizing that “[they] are manifest in two ways: (1) the observable *linguistic behavior*, manifest in everyday interaction between members of the speech community, and (2) the native speakers’ *metalinguistic knowledge*, manifest in their ability to provide interpretations and systematizations for linguistic units and events” [emphasis original]. While the focus of documentary practice has perhaps been more of an elaboration of the first manifestation, the second manifestation has not been neglected either, as evidenced by the many discussions on the place of staged events and elicitation sessions within the context of language documentation (e.g., Foley 2003; Lehmann

2004; Seifart 2008; Lüpke 2009; Hellwig 2010; Majid 2011; Himmelmann 2012; San Roque et al. 2012; Silva & AnderBois 2016; Lahaussois & Vuillermet 2018). This paper situates itself within this overall methodological debate, and takes its specific inspiration from an intriguing comment found in a footnote in Himmelmann (2012:197):

One anonymous reviewer raises the issue of whether and how the practices covered by the broad definition of metalinguistic proposed here can be distinguished from the practices speakers engage in when interacting with small children acquiring a language, which surely should be considered part of (many) speakers' usual linguistic repertoire. This is an intriguing issue in need of further consideration. While there are certainly similarities and overlaps between these two kinds of practices [...], there are also clear differences with regard to intensity, reflective stance, objectification of linguistic units, and participant structure (adult-child, typically in kin relation, vs. adults who interact primarily in order to document linguistic structures and practices). It is highly likely that at least in some cultures and societies, metalinguistic skills displayed in linguistic elicitation and experimentation build on everyday practices in adult-child interaction. But I still believe that the differences are significant enough not to include them in the same category.

The paper takes the above comment as its starting point and explores whether and how the documentation of a specific “observable linguistic behavior” (child-directed language) can shed light on the “metalinguistic knowledge” of adult speakers. As Himmelmann (2012) sketched out above, we also do not assume that the metalinguistic skills displayed in child-directed language and those displayed in linguistic elicitation are identical. However, we are interested in the possibility of metalinguistic knowledge revealing itself in child-directed language, and thus in the potential of child-directed language as constituting yet another valuable resource for investigating such knowledge and thereby further enriching our documentary corpora and descriptions.

This paper arose in the context of two longitudinal projects on documenting the language used by and with children among the Qaqet of Papua New Guinea and the Dëne Słı́né of Canada. Despite a clear focus on language used by and around children, both projects led to insights that significantly shaped our understanding of the adult language – as an almost accidental by-product and much to our own surprise.

The paper is structured as follows: §2 gives an introduction to child-directed language, §3 presents the case studies, and §4 concludes this paper.

2. Speaking with children An important aspect of language acquisition research is studying the input that children receive. This importance is recognized across different theoretical frameworks. In a nativist model, it is assumed that the input is too defective for children to acquire their language, thus necessitating the existence of

an innate universal grammar, a “language acquisition device” – an assumption that can only be substantiated by studying the input. In an interactionist model, it is assumed that language learning takes place in the interaction between the child and her environment, with the child drawing on her innate ability to cooperate and to communicate cooperatively – and input studies directly investigate these mechanisms of learning.

Given its theoretical importance, the role of input has been extensively studied in the acquisition literature. Early studies date back to the 1970s, focusing on the description and characterization of a specific register: child-directed language (labeled “motherese” or “baby talk” in the early literature) (see especially the by now classical volumes of Snow & Ferguson 1977 and Gallaway & Richards 1994, and for a more recent systematic review Saint-Georges et al. 2013). This register is tied to the function of communicating with young children (i.e., with interlocutors who are linguistically and cognitively immature), and it is characterized through specific structural features. These features and their functions are extensively explored in the literature. Compared to adult-directed language, child-directed language tends to be characterized by all or many of the following features: a predominance of short utterances, which tend to be correct and complete, and with only a small proportion of hesitations and errors; the presence of specific prosodic characteristics such as an exaggeratedly large pitch range, a higher F_0 , a longer duration, and more pauses; the use of a restricted vocabulary with reference to the here and now; the presence of a nursery vocabulary; a high proportion of questions and imperatives; and – especially – repetitions and variations in various disguises. It seems likely that such features facilitate the acquisition process, with, e.g., intonational phenomena helping to identify word and phrase boundaries; a restriction to the here and now helping to learn words meanings; or variations helping to learn grammar. Accordingly, the contemporary literature is centrally concerned with investigating possible correlations between the features of child-directed language and the acquisition of these features by children.

While there are extensive discussions of child-directed language in the literature, existing studies are biased towards Western languages, and – increasingly – the larger Asian languages. As a result, we still know little about the universality of this register and its features – an issue already addressed in the early literature (e.g., Lieven 1994; see also Ferguson 1978). Nevertheless, studies that shed light on questions of universality do exist. Such studies not only come from within psycholinguistics (see the next paragraph for some studies), but also – and very importantly – from anthropological research on language socialization (e.g., Ochs & Schieffelin 1984; Schieffelin 1985; Schieffelin & Ochs 1986; Ochs 1988; Kulick 1992; Crago et al. 1993): they show that there are considerable differences across cultures in the ways that adults interact and talk with children, reflecting often very different ideologies about the development of children and the role of adults in this process. These studies come from a research tradition that differs in many respects from that of contemporary language acquisition research: a strong focus on qualitative (rather than quantitative) methods, including techniques of participant observation, and an overall interest in exploring cultural ideologies and beliefs. As a result, reports about the non-universality of child-

directed language are sometimes dismissed as “only” reflecting belief systems, but not actual practice – arguing that some form of child-directed language is inevitable, since “it is difficult to imagine how communication could be at all successful without resorting to at least some of the characteristic features of Child Directed Speech” (Saxton 2009:80).

On the one hand, the above statement is probably true on one level: all anthropological studies reporting on the absence of child-directed language simultaneously report on the presence of other practices that are used in the interaction with young children and that can be considered functional equivalents, e.g., the Kaluli in Papua New Guinea employ a prompting routine (starting with the verb *elema* ‘say like that’) that models utterances for the child (Ochs & Schieffelin 1996:86–87). But on the other hand, it severely underestimates the extent of the attested variation. Given that the number of cross-linguistic and cross-cultural studies still remains small, it is impossible to comprehensively map this variation. Nevertheless, even the existing small sample already provides counter-examples to most of the features proposed to be characteristic of child-directed language. For example, Pye (1986:88) and Pye et al. (2017:21) show that K’iche’ Mayan mothers continue to speak at their normal – and sometimes even faster – speech rate and with a low F_0 , which is sometimes even reduced to a whisper. More generally, there exists evidence for considerable variation in the learning environments (see, e.g., Crago & Allen 1997; Mastin & Vogt 2016; O’Shannessy 2015; Vogt et al. 2015; Shneidman & Woodward 2016; Cristia et al. 2019): in many cultures, overheard language plays a much more prominent role than child-directed language; triadic joint engagement (where child and caregiver jointly interact over an object) is often less important than other types of engagement (such as engagement over persons and the social world); interaction is frequently not with a single principle caregiver (e.g., the mother), but with multiple caregivers, including especially older children; dyadic interaction is less frequent than multi-party interaction; and acquisition and socialization tend to take place in a multilingual society, and only rarely in contexts of monolingualism or bilingualism within the family.

The effects of these considerable differences on language learning, and hence on our theories of learning, remain to be seen. For now, what language documentation can realistically contribute to this debate is a more solid foundation that maps the extent of the variation: documenting how different cultures interact with their children; investigating the presence or absence of a child-directed language register; describing its structural features (if present); and more generally researching how adults and other children talk to, and interact with, young children. Such an effort would not only be of relevance to psycholinguistics (contributing to our understanding of universality and variation), but also to language documentation (documenting an observable linguistic behavior) and to language maintenance (documenting appropriate communicative practices that facilitate language learning). And, as we hope to show in the next section, such an effort would also enhance our understanding of the adult language.

3. Case studies This section presents a number of case studies with an aim to illustrate the potential of child-directed language to shed light on our understanding of the adult language. These case studies emerged in the context of two projects documenting language acquisition and socialization. One project is set among the Qaqet in Papua New Guinea. Qaqet is a Papuan language that is spoken by 15,000 speakers in the Gazelle Peninsula of East New Britain Province. In the remote interior mountains, the language continues to be strong, and children acquire it as their first and dominant language. In the more accessible coastal regions, by contrast, the language is rapidly giving way to the national *lingua franca* Tok Pisin, and children acquire both languages simultaneously, with Tok Pisin being arguably the dominant language. Our project follows a number of children aged 2 to 4 years from both regions, creating a longitudinal corpus that allows us insights into the development of individual children, and that enables a comparison of acquisition and socialization within each sociocultural context and across the two contexts.

The other project is based in two Dëne Sųłné communities in Northwestern Saskatchewan (Canada) where Dene is still the dominant language of daily interaction. There are about 3,000 speakers in the two communities. Until recently all children entering school would speak Dene, and bilingual language competence (Dene and English) is still the norm. In the village school, school children now prefer English. At the reserve school, Dene language use is at the tipping point towards a preference of English, although the implementation of a transitional immersion program had a stabilizing effect on language use (Jung et al. 2018). The Dëne Sųłné Language Acquisition Study (DESLAS) is a longitudinal project that has families record normal interactions at home including children age 2 to 4. This corpus has been designed according to the experiences with the Chintang language documentation project.

In a Western context, we can take it for granted that any project on language development will be carried out against the background of a comprehensive knowledge of the structure and use of the adult language, and that it will include native speaker linguists. In a language documentation context, by contrast, either one is unlikely to be the case (see Kelly & Nordlinger 2014 for a discussion of this issue in the context of language documentation; see Demuth & Ellis 2009 for an appraisal of how their extensive acquisition corpus of Sesotho shaped their understanding of the adult language): our understanding of the adult language is bound to be limited, and native speaker linguists are not always available. Furthermore, in the context of language endangerment and shift, it is not always clear which language or which variety of a language the children are actually acquiring.

This is also true of our projects: although both projects have a clear focus on children, it was impossible to not develop insights into the adult language in parallel. This happened on innumerable occasions – some of them shedding light on an isolated word or morpheme only, but many having a more profound impact on our overall understanding of the adult language. To give an idea of the possibilities, we have chosen a number of illustrative examples, organized with respect to three properties that are considered typical of child-directed language: insights emerging from the prevalence of repetition and especially of so-called variation sets in child-directed

language (§3.1), from the existence of clarification processes (§3.2), and from the discourse context (§3.3). While the universality of child-directed language remains an empirical question, some or all of these three properties are strong contenders for having a wider cross-linguistic distribution.

3.1 Variation sets It has long been noted that the language from caregivers to young children is characterized through repetitions, especially partial repetitions. Küntay & Slobin (1996) have coined the term *variation set* for this phenomenon, defining it as “partial repetitions of maternal utterances, with changes in lexical items, grammatical morphology, and/or word order, maintaining a constant communicative intent” (Küntay & Slobin 1996:267). Such sets serve a communicative function, as mothers and other caregivers produce them in order to “attract and hold the child’s attention until some kind of desired response is produced – either an action or a verbalization” (Küntay & Slobin 2002:8). There are indications that variation sets are widespread across the world’s languages (Slobin et al. 2010), and there is growing evidence that they make up a large proportion of the input that children hear on a daily basis – corpus studies show that such sets account for minimally 20% of a young child’s input (Newport et al. 1977; Küntay & Slobin 1996; Onnis et al. 2008), becoming less frequent as the child grows older (Gleitman et al. 1984; Wirén et al. 2016).

While the variation sets are produced for communicative reasons, they also (as a non-intended by-product) facilitate language development: keeping the larger part of an utterance constant enables the (child) learner to isolate lexical items and grammatical structures through comparison and contrast across the entire set. The focus of current research is on such developmental issues, showing quantitatively for selected phenomena that variation sets in the input are, indeed, predictive of a (child) learner’s later production.

At the same time, these structural properties make them of interest from the perspective of studying and understanding the grammatical properties of the adult language. The caregivers choose to express a single meaning in a series of utterances that differ formally from each other. And this means that variation sets allow us to study the permutation possibilities of a language in a natural, non-elicited, context: removing or adding a word or phrase gives us information about the combinatorial possibilities and about the optionality of constituents; substituting one morpheme, word, or phrase for another gives us information about their equivalence; and changing their ordering gives us information about variation in word and constituent order.

Such variation sets are very common in both our longitudinal corpora, and they proved invaluable for our analysis of the adult language. We have selected three examples of such sets in Qaqet and in Dëne Sųłné to give a small impression of the possibilities. Example (1) from Qaqet is directed from a mother at her young daughter (aged 2;1), intending her to blow on a smoldering fire to re-ignite the flames. After the first utterance, her daughter moves towards the fire, starts to blow on it but gets distracted, and the mother repeats her instructions another three times until she is satisfied that her daughter carries out the action. While the mother’s intention thus remains constant across her four utterances, the utterances themselves differ on a

The understanding of Qaqet argument structure, as sketched out above, arose partly through classic fieldwork techniques – analyzing the distribution of verbs in our adult text corpus, and exploring the extent of the combinatorial and permutation possibilities during elicitation sessions. But in addition, and to a not inconsiderable degree, it benefitted from taking variation sets into account. Example (1) presents just one of very many variation sets in our child corpus that feature argument structure alternations, thus giving us access to large numbers of naturally-occurring utterances which are formally different, but semantically similar.

Example (2) from Dëne Sųłné shows an interaction between a mother talking with her son (2;4) and nephew about a puppy dog. We have chosen this example to show a different kind of variation: here, the mother expands on the incomplete utterances of her son, repeating and varying the verb forms of ‘to be gone’ (*húlě*) in the first set, and the local specifications (*wě/juwě* ‘over there’) in the second set. Verb stems in Dene most often end in a vowel and contract with vocalic particles that follow the verb sentence-finally. The patterns are very complex, making the identification of phonological processes and the recovery of underlying forms an analytic challenge for the linguist. The variation sets in the child corpus now provide us with many relevant examples, constituting a valuable data source for our analysis. In this example, the mother varies between the statement *húlě* ‘it is gone’ and the question *húla* ‘(what) is gone?’, a contraction of the verb form *húlě* and the content-question marker *ʔa*. The child’s first mentions of the short forms *lo* and *húlo* at the beginning could be instantiations of another verb/particle combination, the well-formed contracted yes/no-question, *húlě=o*. After the exchange containing the variations of *húlě*, the mother continues with a variation on the directional adverb (*ju*)*wě* ‘over there’ in combination with the allative marking postposition *-ts’én*.

(2) S: *mom lo* [child shows empty hands]

mom húlě
mom 3SBJ-gone
 ‘mom, it is gone!’

M: *haǰ*
 ‘what are you saying?’

S: *húlo*
húlě
 ‘it is gone’

M: *dlát’i húlě*
what 3SBJ-gone
 ‘what, gone?’

S: *puppy*

M: *puppy?* *dási húla?*
 puppy dlási húlë=a
 puppy where 3SBJ-gone=Q
 ‘the puppy? Where is it gone to?’

S: *wé odén* [pointing to the kitchen]
wé ho-ts’én
 there ARE-ALL
 ‘towards there’

M: *wéts’én*
wé-ts’én
 there-ALL
 ‘towards there?’

S: *mmb*

M: *juwé n̩hɔɪ, juwéts’én*
juwé n-ɪh-ɔɪ, juwé=ts’én
 over.there LEX-2SBJ.IPFV-look over.there=ALL
 ‘you take a look over there, towards over there’

(deslas-BCR-2015-06-09-BCD)

The next example in (3) stems from the same situation a little bit later in the conversation, with the mother talking to her son about the puppy. The son is avoiding the puppy in the living room. It illustrates yet another type of variation: formulaic utterances, where part of the utterance is kept constant, but its referent changes (varying between 2SG and 3SG in this example). Here, the caregiver produces a number of utterances featuring the same verb *l.ge* ‘crawl’ with different inflections according to person (second person subject and third person subject), grammatical tense/aspect (perfective, imperfective plus future particle, and imperfective), and goal phrases (distal location, human location, and proximal location).

(3) 1. *kosí bëyaë n̩hgedë*
ko-sí bë-yaë n̩-l.ge=dé
 DIST-towards 3O-underneath PRVB.2SG.SBJ.PFV-crawl=when
 ‘when you crawled underneath that way’

2. *nq n̩hgesq*
n̩-ghq ní-l.ge ha=sq
 2SG.OBJ-for PRVB.3SG.SBJ.IPFV-crawl FUT=DCL
 ‘he’s going to crawl to you’

3. *dēsí n̩je ɪghá bëch’azé*
dēsí ní-l.ge ɪghá bë-ch’azé
 PROX-towards PRVB-2SG.SBJ.IPFV-crawl hurry 3OBJ-from
 ‘hurry, you go this way away from him’

(deslas-BCR-2015-06-09-BCD)

The examples in this section illustrate a well-known phenomenon in child-directed language: the extensive use of variation. Such variations come in different types, and we have chosen our examples accordingly: sometimes, variations are initiated by the caregiver (as in example 1) and sometimes, they result from expansions of child utterances (as in example 2); they can maintain a constant intention (as in examples 1 and 2), or they can be part of formulaic sequences that introduce, e.g., changes in referents (as in example 3). This list is by no means exhaustive, but we hope that the examples suffice to illustrate their potential for our understanding of the adult language: corpora of child-directed language are likely to give us access to large numbers of variations, thus allowing us to study the permutation possibilities in a natural, non-elicited, context.

3.2 Simplification and clarification processes Research into child-directed language has found many processes that could be described as simplifying the adult language, “as in replacing difficult consonants with easy ones or eliminating inflections or replacing pronouns with proper names”, or as clarifying it, “as in speaking slowly, clearly and with many repetitions” (Brown 1977:4). The existence of such processes in the Western context is well documented, although there is some controversy over their interpretation: it is by no means clear what simplification means, how it can be measured, and whether child-directed language can really be characterized as a simplified register. But from the perspective of this paper, the main point of interest is the existence of such processes and what they reveal about the adult language. In this context, Ferguson (1977:212–213) provides the following characterization:

The processes that derive simplified registers from adult speech [...] are not always simplifying in nature. Some processes are clarifying in that they modify in the direction of greater redundancy, often by adding material to the model. Sentences may be pronounced more slowly and articulated more carefully; vowels normally reduced or elided may be supplied; words, phrases or whole sentences may be repeated; ambiguous words or near homophones may be replaced by synonyms; everything may be uttered more loudly; ambiguous constructions may be paraphrased. Such clarifying processes are part of the competence of all users of language and occur outside simplified registers. Details vary across languages and indeed across registers in the same language. For the linguist they are of special interest in the cues they offer for basic ‘underlying’ forms and they raise the old issue of ‘clarity norm’ versus ‘frequency norm’ [...].

Ferguson’s observation speaks directly to the issue raised in the introduction of this paper: these ‘simplifying’ and ‘clarifying’ skills displayed by adults when speaking with young children are, indeed, reminiscent of the skills displayed by native speakers when explaining the structure of their language to the linguist researcher.

The universality of such ‘simplifying’ and ‘clarifying’ processes in child-directed language is unclear, but, where they occur, they provide insights into the adult language. We illustrate this phenomenon with two examples from our corpora.

The first example comes from Qaqet phonology. In Qaqet, the three voiceless stop phonemes /p/, /t/, and /k/ are lenited to [β] (written <v> in the orthography), [ɾ] (written <r>), and [ɣ] ~ [j] (written <q>) in intervocalic environments. For example, the name *Qaqet* is phonologically /kakət/: the second /k/ always occurs intervocalically and therefore inevitably lenites to [ɣ]; while the first /k/ is usually preceded by a vowel-final article, and therefore also usually lenites to [ɣ], e.g., *aqaqet* [aɣaɣət] ‘the Qaqet people’.

Lenition is widespread and automatic, and it is only blocked in the case of geminated plosives. Nevertheless, there are exceptions. This concerns a good number of individual lexemes, including all known loanwords, e.g., *apusi* ‘cat’, *atapiuk* ‘cassava’, or *akap* ‘cup’ (from Tok Pisin *pusi*, *tapiok*, and *kap*). It also concerns the verb lexicon where the opposition between the stem-initial voiceless stop and its lenited counterpart conveys an aspectual distinction between a continuous and a non-continuous meaning, e.g., *kapileng* ‘he woke him up (CONT)’ vs. *kavileng* ‘he woke him up (NCONT)’, *katal* ‘he carried it (CONT)’ vs. *karal* ‘he carried it (NCONT)’, or *kakarlu* ‘he shouted (CONT)’ vs. *kaqarlu* ‘he shouted (NCONT)’. That is, some of the lenited consonants have to be analyzed as allophones of the voiceless plosives, while others have to be analyzed as phonemes in their own right – and it is not always straightforward to determine the phonemic status of a lenited sound. The main challenge here is posed by the root-initial consonant of nouns. Nouns are always preceded by a vowel-final article, even in their citation form, and their root-initial consonant thus always occurs intervocalically, i.e., in an environment where a voiceless plosive would surface lenited. Table 1 illustrates this problem with the help of the article *a* ‘noun marker’. In the case of the root-initial consonant being a voiced plosive, nasal, or liquid (as in the first three rows), the structure of the underlying form is straightforward. But in the case of the citation form containing a lenited consonant (as in the last three rows), it is impossible to be certain about its phonemic status: is it an allophone of the voiceless plosive, or is it a phoneme?

Table 1. Qaqet nouns

Gloss	Citation form	Underlying form
Dog	a ⁿ daŋ	a= ⁿ daŋ
Betelnut	aməɾik	a=məɾik
Coconut	alams	a=laməs
Breadfruit	aβas	a=pas? a=βas?
Taro	arim	a=tim? a=rim?
Frog	aɣəlmin	a=kəlmin? a=ɣəlmin?

In the adult language, there are very few contexts where the underlying consonant would surface. The main such context is the vocative, as vocative forms can (but do not have to) occur without a preceding article. Example (4) features the vocative *kəlmin* ‘Frog!’, thus revealing that [aɣəlmin] ‘frog’ in Table 1 above is underlyingly /a=kəlmin/.

- (4) *kəlmin, niɣuariⁿ di?*
kəlmin, ni=kuariⁿ di?
 frog 2SG=where
 ‘Frog, where are you?’

(R12ADNFrog 9.535 11.265)

Even though the vocative context reveals the underlying form, it turned out to be impossible to use it as a diagnostic in elicitation sessions: speakers were only prepared to use it with animate (preferably human) referents, and they were understandably very reluctant to use or accept it with inanimate referents – i.e., with the vast majority of the nominal lexicon.

Fortunately, the underlying consonants are also revealed in a second context: in child-directed language. This is the only context where adults regularly omit articles, thus allowing the initial consonant of noun roots to occur at the beginning of words. This is illustrated with example (5) by means of the word *aquukuka* ‘a sweet potato’. The adult pronunciation would be [aɣuukuka] (from /a=kuukuk-ka/ ‘NM=sweet.potato-SG.M’): the first /k/ always lenites to [ɣ] in the intervocalic environment following an article, the second /k/ unexpectedly never lenites (it is one of the lexical exceptions to the lenition rule), and the last /k/ is underlyingly geminated, and thus does not lenite either. In her first utterance, YDS (aged 2;0) has just spotted a sweet potato and points it out to her mother AMT. AMT acknowledges it, and models the word consisting of the root plus the appropriate noun class suffix – but she omits the article, which would normally be obligatory in this context. As a result, the underlying initial *k* surfaces. YDS repeats her pronunciation of the word, and then her older brother YRA (aged 3;2) chimes in, producing a morphologically correct form (including an article), but not leniting the initial *k* – an adult, even in child-directed language, would have lenited *k* in this context. Indeed, the acquisition of lenition is quite protracted in Qaqet, and even older children show considerable differences from the adult pattern.

- (5) YDS: **kaakak*
kuukuk
 sweet.potato
 ‘sweet potato’
- AMT: *ee, *kuukuka*
ee, kuukuk-ka
 yes sweet.potato-SG.M
 ‘yes, a sweet potato’
- YDS: **kaakak*
kuukuk
 sweet.potato
 ‘sweet potato’

YRA: **akuukuka*
a=kuukuk-ka
 NM=sweet.potato-SG.M
 ‘a sweet potato’

AMT: *ee*
ee
 yes
 ‘yes’

(LONGYDS20150516_I 297.285 301.465)

While the omission of articles is common in Qaqet child-directed language, it is not inevitable: there are many contexts where adults do produce articles. Nevertheless, it is the only known register where the omission of articles occurs with any frequency, thus constituting a natural context for underlying forms to surface. It even became possible to capitalize on this knowledge when eliciting nouns: asking speakers to imagine using a noun when talking to a child worked much better than asking speakers to imagine using it as a vocative.

The next example is from Dëne Sùhné where a mother is talking to her child: a usually fused enclitic (expressing negation) is separated from the verb to clarify the stem vs. particle parts. In this instance, the combined expression verb plus negative enclitic is used by the mother first (*narít'èla*), but is then immediately repeated as separate words (*narít'a ʔilë*), allowing the negation particle (*(h)ilë/(ʔ)ilë* ‘not, it is not’ to surface in its underlying form – a form that has undergone considerable morphophonological changes in the fused form. This example nicely illustrates the complexities of Dëne Sùhné morphology and phonology: a negated classificatory verb denoting a single compact object (something that bothers her daughter and that is stuck on her clothing) is used in the reversative aspect (*na-*) causing a middle voice derivation (*t-*), a common association in the grammar of Dene languages. When a stem-final vowel (in this case /-a/) and the enclitic-initial high-toned vowel /i-/ occur together in typical conversational speech, the vowels assimilate resulting in one high-toned vowel “shared” by verb stem and negative enclitic, resulting in this case in [é] (a half-closed unrounded front vowel).² The mother uses this fused construction first. When the child does not understand, she repeats it by producing the verb and the negation as separate words, also adding a declarative particle. When the daughter still does not react and continues with her action, the mother just says ‘no’ (in the sense of ‘this doesn’t work’).

- (6) 1. *narít'èla*
 na-dë-í-t-'a=ilë
 REV-QUL-3.PFV-MID-CLV.SC=NEG
 ‘it doesn’t come off’

²The vowel written as <ë> varies phonetically between an epsilon and a schwa, and is the vowel commonly used in affixes. <ë> in this variety of Dene tends towards a pronunciation as long vowel or diphthong [ei], preserving its history as heavy (dual) vowel in assimilative positions.

- (7) *belungera* *naqua* *amaguleng?*
be=lu-nget-a *ne=kua* *ama=guleng?*
 CONJ=DEM-N-DIST from/with=where ART=malay.apple
 ‘and those malay apples are from where?’
 (NIIAAGSiriniLobster2 12.234 13.522)

Our child corpus, by contrast, contains a large number of discontinuous noun phrases, thus allowing us to investigate their distribution in more detail. One salient context for their occurrence is exemplified in (8) (a continuation from example 1 above). The mother instructs her daughter (aged 2;1) to put some more logs onto the fire. In her first utterance, she does not overtly mention the ‘logs’, and instead resorts to an indefinite pronoun marked for neuter noun class. Qaqet has eight different noun classes, and it is a common discourse strategy among adults to not introduce participants by means of lexical nouns, but by means of pronominal forms that are marked for noun class. In the context of fire-making, the use of the neuter noun class is enough for a competent adult speaker to pick out the intended referent (the logs). But children are not yet competent speakers, and the daughter cannot pick out the referent and has to check back. This then is one important context where discontinuity arises: the mother adds the lexical noun to the right of the predicate, thus creating a discontinuous noun phrase consisting of an indefinite determiner (in Qaqet, indefinite expressions function both as determiners and heads) and a head noun.

- (8) AMT: *iang* *maqamuk*
 ia-nget *maqa-muk*
 other-N PROX-across
 ‘there are others over here’
 YDS: **ia?*
 ia
 other
 ‘others?’
 AMT: *iang* *maqamuq* *amaltany*
 ia-nget *maqa-muk* *ama=ltany*
 other-N PROX-across ART=ember
 ‘there are other logs over here’
 (LONGYDS20150705_I 361.510 365.205)

On the basis of the child corpus, it thus became possible to investigate the distribution of discontinuous noun phrases and to identify a number of discourse contexts that triggered their use, e.g., the clarification context illustrated in (8) above occurs frequently in our child corpus, but only rarely in our adult corpus. Without access to this interactional data, we would only have known about the existence of discontinuous phrases, but not about their distribution.

Similarly, our analysis of Dëne Sųhné benefited from the interactional nature of the child data. One intriguing area of the categorization of events in this language is the classificatory verb system (which the whole Dene language family is renowned for). Dëne Sųhné distinguishes nine different verb stems that combine with a wealth of derivational morphology to distinguish the handling and positioning of different types of classes of objects according to their animacy, shape, consistency, and number (such as long, compact, fabric-like, mushy, etc.; Kasyon 1997). These distinctions are well established, and adult speakers of the language are very aware of them. A second distinction is less obvious: controlled vs. uncontrolled handling that also entails polite vs. impolite and related secondary meanings (see Davidson et al. 1963; Rice 1989 for the related language Dene (Slave)). These uncontrolled classificatory verbs are less prominent, and less easy to elicit, even with appropriate contexts. Correspondingly there are controlled and uncontrolled manners of motion. Oftentimes they seem to the linguist to be lexicalized due to the infrequent occurrence of these meanings in words such as ‘slide’. But, in our natural discourse contexts, they are used more often, and luckily for the linguist the appropriate situations are right there in the documented video recordings. In the case of the child-directed language in (9) below, for example, the mother is holding the beam of the swing set, while the children are on the swing. The verb stem *-kéth* belongs to the set of verbs that expresses uncontrolled events.

- (9) *hustónë* *yékéth* *chá*
 u-was-tón-ë yé-Ø-kéth ch’á
 DIR-OPT.1SG.SBJ-hold-OPT up.over-IPFV.3SBJ-CLV.ST.UNCNTR so.that.not
 ‘I’ll hold this (beam of swing set) so then it doesn’t tip over’
 (deslas-BCR-2016-07-10-C)

The phenomena discussed in this section differ from those discussed in §3.1 and §3.2. The previous two sections have illustrated cases where child-directed language allowed us to study phenomena in a natural context that we normally study by means of elicitation. The phenomena in this section, by contrast, do not easily lend themselves to elicitation. To some extent, we would have been able to discover them in adult corpora, too, provided that they feature interactional genres. But there are two reasons why child corpora are specifically suited for such investigations. On the one hand, child corpora are necessarily interactive, i.e., any child corpus is bound to contain a good amount of interactive data. Adult corpora, by contrast, contain a mixture of genres and are frequently biased towards monological genres. On the other hand, communication with immature interlocutors relies heavily on real-world contexts, potentially bringing out discourse strategies more clearly. Specifically, in the case of example (8), situations where interlocutors demonstrate their lack of understanding are very frequent in our child corpus, but almost absent in our adult corpus. And in the case of example (9), situational contexts of certain types, such as e.g., uncontrolled handling and concomitant meaning nuances, feature more prominently in the recorded interaction with children than what can be found in a more typical interactive adult situation.

4. Conclusions As discussed in the introduction to this paper, both the fields of language documentation and of psycholinguistics have a strong interest in documenting child language and child-directed language in communities around the world. But despite this common interest, there are still very few language documentation projects that purposely include language used by and with children. Given the overall scarcity of resources, coupled with the methodological challenges involved in conducting such a project, it cannot come as a surprise that many documentation projects decide to focus on other, equally important, aspects of language use.

This paper argues for making the effort of including at least the documentation of child-directed language into our projects, showing with the help of three case studies that this register constitutes a rich data source with considerable potential for giving insights into the adult language, thus enhancing the quality of our documentations and descriptions. While the paper focuses on one benefit only (the better understanding of the adult language), this is by no means the only benefit. As summarized in the introduction, other benefits for language documentation include a more comprehensive documentation (that includes an up to now under-documented register) and the creation of an invaluable resource, which can support language maintenance and revitalization efforts subsequently. The benefits for psycholinguistics include a better understanding of the extent and limits of variation in child-directed language. Eventually, the documentation of this register in different socio-cultural contexts around the world is likely to have implications for theories of language learning, but this can only be the second step, necessitating dedicated studies later on. A first step would be to map cross-linguistic variation. When research into child-directed language in the Western world commenced, Garnica (1977:64) phrased the challenge as follows:

There are two questions relating to the verbal input to the child and his language development. First, what are the linguistic characteristics of speech addressed to the child? This is a purely descriptive question but necessarily preliminary to the second question: what features of the verbal environment are critical for learning language? Certainly, the latter question is more relevant to a theory of language development and, more broadly, to the problem of cognitive development. Its answer, however, depends significantly on the answer to the first question.

Four decades later, language documentation is in a very strong position to continue and expand these efforts, with results that benefit both disciplines, language documentation and psycholinguistics. Different from forty years ago, the language documentation framework will ensure that this register and its cross-linguistic variation is not just described, but documented.

Acknowledgements We would like to thank the Clearwater River Dene School and Nation, as well as the community of La Loche for their patient help and support in our work. This research has been partially funded by the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 615988 (PI Sabine Stoll). In Papua New Guinea,

our sincere thanks go to the communities of Raunsepna and Lamarain for their generosity and their help. This research has been funded by the Volkswagen Foundation's Lichtenberg program (Az 87 100), the Endangered Languages Documentation Programme (SG0110), and the Australian Research Council (FT0991412). We have presented some of the ideas in this paper at the 3rd Vielfaltslinguistikkonferenz 2019 in Cologne and in talks at the University of Leuven, the University of Toronto, the University of Zurich and La Trobe University, and we want to thank the participants for their comments and discussions. Many thanks also to two anonymous reviewers for their helpful feedback on earlier versions of this paper.

Abbreviations used in glosses ALL: allative, ARE: areal, ART: article, CLV: classificatory verb, CONJ: conjunction, DCL: declarative, DEM: demonstrative, DIR: directional, DIST: distal, FUT: future, IPFV: imperfective, LEX: lexicalized, M: masculine, MID: middle, N: neuter, NCONT: non-continuous (aspect), NEG: negation, NM: noun marker, NPST: non-past, OBJ: object, OPT: optative, PFV: perfective, PROX: proximate, PRVB: preverb, PURP: purposive, Q: question, QUL: qualifier, REV: reversative, SBJ: subject, SC: single compact (object), SG: singular, ST: stick-like (object), UNCNTR: uncontrolled

References

- Anand, Pranav, Sandra Chung, & Matthew Wagers. 2011. Widening the net: Challenges for gathering linguistic data in the digital age. In *NSF SBE 2020: Rebuilding the mosaic: Future research in the social, behavioral and economic sciences at the National Science Foundation in the next decade*.
- Berman, Ruth & Dan I. Slobin. 1994. *Relating events in narrative: A cross-linguistic developmental study*. Mahwah, NJ: Erlbaum.
- Bowerman, Melissa. 2011. Linguistic typology and first language acquisition. In Song, Jae Jung (ed.), *The Oxford handbook of linguistic typology*, 591–617. Oxford: Oxford University Press.
- Brown, Roger. 1977. Introduction. In Snow, Catherine E. & Charles A. Ferguson (eds.), *Talking to children: Language input and acquisition*, 1–27. Cambridge: Cambridge University Press.
- Child Language Research and Revitalization Working Group. 2017. *Language documentation, revitalization, and reclamation: Supporting young learners and their communities*. Waltham, MA: EDC.
- Crago, Martha B. & Shanley E.M. Allen. 1997. Linguistic and cultural aspects of simplicity and complexity in Inuktitut child directed speech. In Hughes, Elizabeth, Mary Hughes, & Annabel Greenhill (eds.), *Proceedings of the 21st Annual Boston University Conference on Language Development*, 91–102. Somerville, MA: Cascadilla Press.
- Crago, Martha B., Betsy Annahatak, & Lizzie Ningiuruvik. 1993. Changing patterns of language socialization in Inuit homes. *Anthropology & Education Quarterly* 24(3). 205–223. <https://www.jstor.org/stable/3195642>.

- 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. doi: 10.1111/cdev.12974.
- Davidson, William, Leon.W. Elford, & Harry Hoijer. 1963. Athapaskan classificatory verbs. In Harry Hoijer et al., *Studies in the Athapaskan languages*, 30–41. Berkeley: University of California Press.
- Demuth, Katherine & David Ellis. 2009. Revisiting the acquisition of Sesotho noun class prefixes. In Lieven, Elena, Jiansheng Guo, Nancy Budwig, Susan Ervin-Tripp, Keiko Nakamura, & Seyda Ozcaliskan (eds.), *Crosslinguistic approaches to the psychology of language*, 93–104. New York: Psychology Press.
- Eisenbeiß, Sonja. 2005. Documenting child language. In Austin, Peter K. (ed.), *Language documentation & description*, vol. 3, 106–140. London: SOAS. http://www.elpublishing.org/docs/1/03/ldd03_08.pdf.
- Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral & Brain Sciences* 32. 429–492. doi: 10.1017/S0140525X0999094X.
- Ferguson, Charles A. 1977. Baby talk as a simplified register. In Snow, Catherine E. & Charles A. Ferguson (eds.), *Talking to children: Language input and acquisition*, 209–235. Cambridge: Cambridge University Press.
- Ferguson, Charles A. 1978. Talking to children: A search for universals. In Greenberg, Joseph H. (ed.), *Universals of human language*, vol. 1, 203–224. Stanford, CA: Stanford University Press.
- Foley, William A. 2003. Genre, register and language documentation in literate and preliterate communities. In Austin, Peter K. (ed.), *Language documentation & description*, vol.1, 85–98. London: SOAS. http://www.elpublishing.org/docs/1/01/ldd01_08.pdf.
- Forshaw, William, Lucinda Davidson, Barbara Kelly, Rachel Nordlinger, Gillian Wigglesworth, & Joe Blythe. 2017. The acquisition of Murrinhpatha (Northern Australia). In Fortescue, Michael, Marianne Mithun, & Nicholas Evans (eds.), *The Oxford handbook of polysynthesis*, 473–494. Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199683208.013.27.
- Gallaway, Clare & Brian J. Richards (eds.). 1994. *Input and interaction in language acquisition*. Cambridge: Cambridge University Press.
- Garnica, Olga K. 1977. Some prosodic and paralinguistic features of speech to young children. In Snow, Catherine E. & Charles A. Ferguson (eds.), *Talking to children: Language input and acquisition*, 63–88. Cambridge: Cambridge University Press.
- Gleitman, Lila, Elissa Newport, & Henry Gleitman. 1984. The current status of the motherese hypothesis. *Journal of Child Language* 11. 43–79. doi: 10.1017/S0305000900005584.
- Grenoble, Lenore A. & Lindsay J. Whaley. 2006. *Saving languages: An introduction to language revitalization*. New York: Cambridge University Press.
- Hellwig, Birgit. 2010. Meaning and translation in linguistic fieldwork. *Studies in Language* 34(4). 802–831. doi: 10.1075/sl.34.4.02hel.

- Hellwig, Birgit. To appear. Child language documentation: A pilot project in Papua New Guinea. In Penfield, Susan (ed.), *Interdisciplinary language documentation in practice. Language Documentation & Conservation* [Special Publication]. Honolulu, HI: Hawai'i University Press.
- Hellwig, Birgit. 2019a. *A grammar of Qaqet*. Berlin: De Gruyter Mouton.
- Hellwig, Birgit. 2019b. Linguistic diversity, language documentation and psycholinguistics: The role of stimuli. In Lahaussais, Aimée & Marine Vuillermet (eds.), *Methodological tools for linguistic description and typology. Language Documentation & Conservation* [Special Publication 16]. Honolulu, HI: Hawai'i University Press. <http://hdl.handle.net/10125/24855>.
- Henrich, Joseph, Steven J. Heine, & Ara Norenzayan. 2010. The weirdest people in the world? *Behavioral & Brain Sciences* 33(2-3). 61–83. doi:10.1017/S0140525X0999152X.
- Himmelman, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36(1). 161–195. doi: 10.1515/ling.1998.36.1.161.
- Himmelman, Nikolaus P. 2012. Linguistic data types and the interface between language documentation and description. *Language Documentation & Conservation* 6. 187–207. <http://hdl.handle.net/10125/4503>.
- Jung, Dagmar, Mark Klein, & Sabine Stoll. 2018. Language transition(s): School responses to recent changes in language choice in a Northern Dene community (Canada). In Wigglesworth, Gillian, Jane Simpson, & Jill Vaughan (eds.), *Language practices of Indigenous children and youth: The transition from home to school*, 49–68. Palgrave Macmillan.
- Kasyon, Mary Jane. 1997. *Dene (Chipewyan) Language Classification Verbs*. Prince Albert, SK: Northern Canada Mission Distributors.
- Kelly, Barbara, William Forshaw, Rachel Nordlinger, & Gillian Wigglesworth. 2015. Linguistic diversity in first language acquisition research: Moving beyond the challenges. *First Language* 35(4-5). 286–304. doi: 10.1177%2F0142723715602350.
- Kelly, Barbara & Rachel Nordlinger. 2014. Fieldwork and first language acquisition. In Gawne, Lauren & Jill Vaughan (eds.), *Selected papers from the 44th Conference of the Australian Linguistic Society, 2013*, 178–192. Melbourne: University of Melbourne. <http://hdl.handle.net/11343/217034>.
- Kulick, Don. 1992. *Language shift and cultural reproduction: Socialization, self, and syncretism in a Papua New Guinean village*. Cambridge: Cambridge University Press.
- Küntay, Aylın & Dan I. Slobin. 1996. Listening to a Turkish mother: Some puzzles for acquisition. In Slobin, Dan I., 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 & Communication* 6. 5–14.
- Lahaussais, Aimée & Marine Vuillermet (eds.). 2018. Methodological tools for linguistic description and typology. *Language Documentation & Conservation* [Special

- Publication 16]. Honolulu, HI: Hawai'i University Press. <http://nflrc.hawaii.edu/ld-c/sp16/>.
- Lehmann, Christian. 2004. Data in linguistics. *The Linguistic Review* 21(3-4).175–210. doi: 10.1515/tlir.2004.21.3-4.175.
- Levinson, Stephen C. 2006. On the human 'interaction engine'. In Enfield, Nick J. & Stephen C. Levinson (eds.), *Roots of human sociality. Culture, cognition, interaction*, 39–69. Oxford: Berg.
- Lieven, Elena V. 1994. Crosslinguistic and crosscultural aspects of language addressed to children. In Gallaway, Clare & Brian J. Richards (eds.), *Input and interaction in language acquisition*, 56–73. Cambridge: Cambridge University Press.
- Lieven, Elena & Sabine Stoll. 2009. Language. In Bornstein, Marc H. (ed.), *The handbook of cross-cultural developmental science*, 543–555. Mahwah, NJ: Erlbaum.
- Lüpke, Friederike. 2009. Data collection methods for field-based language documentation. In Austin, Peter K. (ed.), *Language documentation & description*, vol. 6, 53–100. London: SOAS. http://www.elpublishing.org/docs/1/06/ldd06_04.pdf.
- MacWhinney, Brian. 2000. *The CHILDES Project: Tools for analysing talk*. Mahwah, NJ: Erlbaum.
- Mastin, J. Douglas & Paul Vogt. 2016. Infant engagement and early vocabulary development: A naturalistic observation study on Mozambican infants from 1;1 to 2;1. *Journal of Child Language* 43(2). 235–264. doi: 10.1017/S0305000915000148.
- Majid, Asifa. 2011. A guide to stimulus-based elicitation for semantic categories. In Thieberger, Nick (ed.), *The Oxford handbook of linguistic fieldwork*, 54–71. New York: Oxford University Press. doi:10.1093/oxfordhob/9780199571888.013.0003.
- McConvell, Patrick & Felicity Meakins. 2005. Gurindji Kriol: A mixed language emerges from code-switching. *Australian Journal of Linguistics* 25(1). 9–30. doi: 10.1080/07268600500110456.
- McGregor, William. 1997. Functions of noun phrase discontinuity in Gooniyandi. *Functions of Language* 4(1). 83–114.
- Moran, Steven, Robert Schikowski, Danica Pajović, Cazim Hysi, & Sabine Stoll. 2016. The ACQDIV database: Min(d)ing the ambient language. In Calzolari, Nicoletta, Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asunción Moreno, Jan Odiijk, & Stelios Piperidis (eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, 4423–4429. Paris: European Language Resources Association (ELRA).
- Newport, Elissa L., Henry Gleitman, & Lila Gleitman. 1977. Mother, I'd rather do it myself: Some effects and noneffects of maternal speech style. In Snow, Catherine E. & Charles A. Ferguson (eds.), *Talking to children: Language input and acquisition*, 109–150. Cambridge: Cambridge University Press.
- Norcliffe, Elisabeth, Alice C. Harris, & T. Florian Jaeger. 2015. Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances. *Language, Cognition & Neuroscience* 30(9).1009–1032. doi: 10.1080/23273798.2015.1080373.

- Ochs, Elinor. 1988. *Culture and language development: Language acquisition and language socialization in a Samoan village*. Cambridge: Cambridge University Press.
- Ochs, Elinor & Bambi Schieffelin. 1984. Language acquisition and socialization: Three developmental stories and their implications. In Shweder, Richard A. & Robert A. Levine (eds.), *Culture theory: Essays on mind, self and emotion*, 276–319. Cambridge, UK: Cambridge University Press.
- Ochs, Elinor & Bambi Schieffelin. 1996. The impact of language socialization on grammatical development. In Fletcher, Paul & Brian MacWhinney (eds.), *The handbook of child language*, 73–94. Cambridge, MA: Basil Blackwell.
- O’Grady, William & Ryoko Hattori. 2016. Language acquisition and language revitalization. *Language Documentation & Conservation* 10. 45–57. <http://hdl.handle.net/10125/24686>.
- Onnis, Luca, Heidi R. Waterfall, & Shimon Edelman. 2008. Learn locally, act globally: Learning language from variation set cues. *Cognition* 109(3). 423–430. doi: 10.1016/j.cognition.2008.10.004.
- O’Shannessy, Carmel. 2015. Multilingual children increase language differentiation by indexing communities of practice. *First Language* 35(4–5). 305–326. doi: 10.1177/0260142723715609227.
- O’Shannessy, Carmel. 2012. The role of codeswitched input to children in the origin of a new mixed language. *Linguistics* 50(2). 305–340. doi:10.1515/ling-2012-0011.
- Pye, Clifton. 1986. Quiché Mayan speech to children. *Journal of Child Language* 13(1). 85–100. doi: 10.1017/S0305000900000313.
- Pye, Clifton, Barbara Pfeiler, & Pedro Mateo Pedro. 2017. Mayan language acquisition. In Aissen, Judith, Nora C. England, & Roberto Zavala Maldonado (eds.), *The Mayan languages*, 19–42. London: Routledge. doi: 10.4324/9781315192345.
- Rice, Keren. 1989. *A grammar of Slave*. Berlin/New York: Mouton de Gruyter.
- Rumsey, Alan. 2017. Dependency and relative determination in language acquisition: The case of Ku Waru. In Enfield, Nick J. (ed.), *Dependencies in language: On the causal ontology of linguistic systems*, 97–116. Berlin: Language Science Press. doi:10.5281/zenodo.573787.
- 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. doi:10.1371/journal.pone.0078103.
- San Roque, Lila, Lauren Gawne, Darja Hoenigman, Julia C. Miller, Alan Rumsey, Stef Spronck, Alice Carroll, & Nicholas Evans. 2012. Getting the story straight: Language fieldwork using a narrative problem-solving task. *Language Documentation & Conservation* 6. 135–174. <http://hdl.handle.net/10125/4504>.
- Saxton, Matthew. 2009. The inevitability of child directed speech. In Foster-Cohen, Susan (ed.), *Language acquisition*, 62–86. Basingstoke: Palgrave Macmillan.
- Schieffelin, Bambi. 1985. The acquisition of Kaluli. In Slobin, Dan I. (ed.), *The crosslinguistic study of language acquisition*, vol. 1, 525–593. Hillsdale: Erlbaum.
- Schieffelin, Bambi & Elinor Ochs (eds.). 1986. *Language socialization across cultures*. Cambridge: Cambridge University Press.

- Schultze-Berndt, Eva & Candide Simard. 2012. Constraints on noun phrase discontinuity in an Australian language: The role of prosody and information structure. *Linguistics* 50(5). 1015–1058. doi: 10.1515/ling-2012-0032.
- Seifart, Frank. 2008. On the representativeness of language documentations. In Austin, Peter K. (ed.), *Language documentation & description*, vol. 5, 60–76. London: SOAS. http://www.elpublishing.org/docs/1/05/ldd05_04.pdf.
- Shneidman, Laura & Amanda L. Woodward. 2016. Are child-directed interactions the cradle of social learning? *Psychological Bulletin* 142(1). 1–17. doi: 10.1037/bul0000023.
- Silva, Wilson & Scott AnderBois. 2016. Fieldwork game play: Masterminding evidentiality in Desano. *Language Documentation & Conservation* 10. 58–76. <http://hdl.handle.net/10125/24687>.
- Slobin, Dan I. 1985–1997. *The crosslinguistic study of language acquisition*, vol. 1–5. Mahwah, NJ: Erlbaum.
- Slobin, Dan I. & Melissa Bowerman. 2007. Interfaces between linguistic typology and child language research. *Linguistic Typology* 11(1). 213–226. doi: 10.1515/LINGTY.2007.015.
- Slobin, Dan I., Melissa Bowerman, Penny Brown, Sonja Eisenbeiß, & Bhuvana Narasimhan. 2010. Putting things in places: Developmental consequences of linguistic typology. In Bohnemeyer, Jürgen & Eric Pederson (eds.), *Event representation in language and cognition*, 134–165. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511782039.007.
- Snow, Catherine E. & Charles A. Ferguson (eds.). 1977. *Talking to children: Language input and acquisition*. Cambridge: Cambridge University Press.
- Stoll, Sabine. 2009. Crosslinguistic approaches to language acquisition. In Bavin, Edith L. (ed.), *The Cambridge handbook of child language*, 89–104. Cambridge: Cambridge University Press.
- Stoll, Sabine & Balthasar Bickel. 2013. The acquisition of ergative case in Chintang. In Bavin, Edith L. & Sabine Stoll (eds.), *The acquisition of ergativity*, 183–208. Amsterdam: Benjamins. doi: 10.1075/tilar.9.07sto.
- Stoll, Sabine, Balthasar Bickel, Elena Lieven, Netra P. Paudyal, Goma Banjade, Toya N. Bhatta, Martin Gaenzle, Judith Pettigrew, Ichchha Purna Rai, Manoj Rai, & Novel Kishore Rai. 2012. Nouns and verbs in Chintang: Children’s usage and surrounding adult speech. *Journal of Child Language* 39(2). 284–321.
- Strömquist, Sven & Ludo Verhoeven (eds.). 2004. *Relating events in narrative, vol. 2: Typological and contextual perspectives*. Mahwah, NJ: Erlbaum. doi:10.1017/S0305000911000080.
- Tomasello, Michael & Daniel Stahl. 2004. Sampling children’s spontaneous speech: How much is enough? *Journal of Child Language* 31(1). 101–121. doi:10.1017/S0305000903005944.
- Verstraete, Jean-Christophe & Dana Louagie. 2016. Noun phrase constituency in Australian languages: A typological study. *Linguistic Typology* 20(1). 25–80. doi:10.1515/lingty-2016-0002.

- Vogt, Paul, J. Douglas Mastin, & Diede M. A. Schots. 2015. Communicative intentions of child-directed speech in three learning environments: Observations from the Netherlands, and rural and urban Mozambique. *First Language* 35(4-5). 341–358. doi:10.1177%2F0142723715596647.
- Wirén, Mats, Kristina Nilsson Björkenstam, Gintarė Grigonytė, & Elisabet Eir Cortes. 2016. Longitudinal studies of variation sets in child-directed speech. *Proceedings of the 7th workshop on cognitive aspects of computational language learning*. 44–52. Berlin, Germany, 11 August, 2016. doi:10.18653/v1/W16-19.

Birgit Hellwig
bhellwig@uni-koeln.de

Dagmar Jung
dagmar.jung@uzh.ch