elpublishing

Language Documentation and Description

ISSN 1740-6234

This article appears in: *Language Documentation and Description, vol 3*. Editor: Peter K. Austin

Documenting child language

SONJA EISENBEISS

Cite this article: Sonja Eisenbeiss (2005). Documenting child language. In Peter K. Austin (ed.) *Language Documentation and Description, vol 3*. London: SOAS. pp. 106-140

Link to this article: http://www.elpublishing.org/PID/038

This electronic version first published: July 2014

Documenting Child Language

Sonja Eisenbeiss, University of Essex

1. Introduction

This paper argues for the integration of child language data into language documentation projects and shows the benefits that the documentation of child language can have for (i) acquisition researchers, (ii) descriptive, theoretical and historical linguists, and (iii) members of language maintenance or revitalisation projects. Moreover, it discusses which implications different user requirements have for data collection and provides pointers to resources, tools and stimuli.

The recent surge of interest in endangered languages has led to widespread efforts to document them; and a growing number of linguists have argued for the development of a new discipline of documentary linguistics (Himmelmann (1998), Lehmann (2001), Tsunoda (2001), Woodbury (2003)). For them, the aim of this new discipline is to create rich records of the linguistic practices of a speech community, which are not tied to particular linguistic approaches or even to the discipline of linguistics itself. Thus, they distinguish documentary linguistics from traditional linguistic description, which attempts to record a language as a system of abstract elements, constructions, and rules (Himmelmann (1998)). Despite the aim of a very broad orientation, the focus of current language documentation projects is still typically on the documentation of the language used by adult native speakers, and child language data only plays a minor role – or none at all (see e.g. the lack of relevant discussion in the references cited above).

In contrast, acquisition researchers have collected child language data for quite some time now, initially in the form of parental diaries (see Stern and Stern (1928) for an overview) and later through systematic audio or video recordings (see Slobin (1985, 1992) for an overview). Since the mid-1980s, language acquisition researchers have made their data available to others via the database of the CHILDES project (http://childes.psy.cmu.edu/, MacWhinney (2000), Sokolov and Snow (1994)). The earliest entries of this database are transcripts of spontaneous speech samples from Indo-European languages, but over the years the database has developed into a multi-media collection with transcripts and recordings from a growing sample of typologically diverse languages. At the same time, new databases are being built up, some of them using the same infrastructure, tools and archiving standards as language documentation projects which focus on the language of adult native speakers (see e.g. the language acquisition archive of the MPI for Psycholinguistics in Nijmegen, which is connected to the DOBES project, http://www.mpi.nl).

Acquisition researchers and field-working linguists with an interest in documentation already share an interest in the creation of (multi-media) corpora and the development of tools for the collection and analysis of such data. However, language documenters might still ask themselves whether it is wise to allocate some of their

Sonia Eisenbeiss (2005) Documenting Child language In Peter K. Austin (ed.) *Language Documentation and Description Vol 3*. 106-140 London: SOAS

limited resources to the documentation of child language. Conversely, acquisition researchers might be hesitant to integrate their data collection into larger documentation projects when this means that they have to conform to procedures, conventions and standards for data archiving which are not directly relevant for their own data management and analysis. Moreover, even if language acquisition researchers and language documenters agree that it is useful to integrate child language data into larger language documentation projects, they will still need more information about the aims, methods and needs of the other discipline to make efficient use of the resources available. In particular, the data collection and elicitation methods typically employed by fieldworkers cannot be simply transferred to the documentation of child language. Neither can they be simply replaced by the experimental methods used in child language laboratories because these are often not appropriate for use in the field. Thus, language documenters need more information about the methods used to collect child language data. Similarly, acquisition researchers require more information about field-compatible data collection methods and archiving standards for data from endangered languages. And researchers from both disciplines have to take into account what linguistic communities might expect from the documentation of their language and its acquisition. Therefore, I will discuss the potential benefits that the documentation of child language can have for (i) acquisition researchers, (ii) descriptive, theoretical and historical linguists and (iii) linguists and community members involved in language maintenance or revitalisation projects. For each of these groups, I will present typical uses of acquisition data and discuss which implications the needs and requirements of this user group have for data collection and participant selection.

2. Language acquisition researchers

Until the mid-1980s, nearly all child language studies involved a comparatively small set of Indo-European languages (see Slobin (1985) for an overview). Since then, acquisition researchers have collected and analysed child language data from a broader range of typologically diverse languages which are spoken in different cultural contexts (see e.g. Slobin (1992, 1997), Sokolov and Snow (1994), Fletcher and MacWhinney (1995), Ritchie and Bhatia (1999)). However, the set of languages for which acquisition data are available is still comparatively limited and contains only very few endangered languages. Nevertheless, crosslinguistic acquisition studies have already proved crucial for testing claims about potential universals of child language and child-directed speech, for investigating the temporal organisation of language development, and for studying the development of grammar, lexical items, conversational competence and narrative skills.

2.1 Universals of child language and child-directed speech

Potential universals of child language and child-directed speech play a central role in the debate about the role of nature and nurture in language acquisition, which has been driven by research on the so-called logical problem of language acquisition (see Pinker (1989), Bertolo (2001) for an overview): children only hear a finite number of target

language utterances, but they learn to produce, understand and judge an unlimited number of utterances. Thus, they have to generalise from individual utterances to the regularities of their target language. These generalisations could potentially deviate from the regularities of the child's target grammar in two ways: The child's generalisations might be too restrictive, or not restrictive enough. If the child's generalisation is too restrictive, the target might contain structures which are not covered by it. For instance, upon hearing input sentences with overt subject noun phrases (e.g. Giovanni canta 'John sings'), an Italian child might incorrectly assume that all Italian sentences must have an overt subject. This incorrect generalisation can easily be rejected on the basis of positive evidence (input sentences) as Italian children will hear a large number of sentences where the subject has been dropped (e.g. canta 'he sings').

In contrast, if the child's generalisation were not restrictive enough, it would allow for the production of sentences that are ungrammatical in the target language. For example, German children are confronted with input sentences where the subject noun phrase is omitted in topic position (e.g. Hast Du Hunger? Nein, ___ hab gerade gegessen. 'Are you hungry? No, ___ have just eaten'). Based on such sentences, a German child could arrive at the incorrect generalisation that subjects can be omitted in any sentential position (like in Italian). In this case, positive evidence is not sufficient to show that the child's generalisation is incorrect. Input sentences with omitted subjects in topic position as well as input sentences with overt subjects would be compatible with the child's incorrect generalisation. Rather, the child would require negative evidence, i.e. information about the ungrammaticality of the sentences that the child incorrectly expects to be grammatical (in this case: sentences with omitted subjects in non-topic position such as Gerade hab ___ schon gegessen. 'Just have ___ already eaten'). Crosslinguistic studies have shown that even though parents might sometimes correct their children, negative evidence is not systematically available to all children at all stages of development (see Marcus (1993) for an overview). Hence, acquisition researchers have to explain how children either avoid incorrect generalisations or recover from them if they cannot rely on negative evidence.

Generative linguists argue that children do not come up with random generalisations that then have to be rejected on the basis of negative evidence for the resulting errors. In their view, innate linguistic universals constrain all utterances of human languages so that generalisations which would violate these universals are never made and thus never have to be rejected. In order to support this assumption, generative acquisition researchers try to show that children may produce utterances that are not target-like, but that these utterances are possible structures of human languages and positive evidence is sufficient to overcome the non target-like generalisation. For instance, they have shown that English-speaking children sometimes fail to invert subjects and auxiliaries when they form questions (1), but when they are confronted with sentence pairs like (2) they do not seem to come up with the incorrect generalisation that questions in English are formed by fronting the first auxiliary of the sentence – even though this generalisation would be compatible with pairs of simple

sentences like (2) (Crain 1991). They never produce errors like (3), but form correct questions for sentences with embedding as soon as these structures can be tested (4).

(1) Where the chicken is running?

(2) The chicken is running. Is the chicken running?

(3) * Is the farmer who ___ running after the chicken is bald?

(4) Is the farmer who is running after the chicken ___ bald?

This can be captured by the so-called Structure-Dependency Principle (Chomsky (1966)). According to this principle, all grammatical operations refer to structural aspects of grammatical representations, and not to properties like linear order. Thus, children should entertain correct or incorrect generalisations about the auxiliary of the main clause, but no generalisations about the first auxiliary in a sentence. Recent minimalist versions of generative grammar try to derive principles like the Structure-Dependency Principle from more general principles of cognition (see Eisenbeiss (2005) for an overview). However, independently of the status of universal constraints, we need more data from typologically diverse languages to test the claim that such constraints prevent certain types of generalisations and errors in children's language.

Such data is also required to evaluate usage-based approaches to language acquisition (e.g. Tomasello (2003)). According to these approaches, very general cognitive principles are sufficient to constrain children's hypothesis space because children are quite conservative. They start out with limited generalisations, which are centred around individual lexical items and phrases - and then gradually extend these generalisations by analogy. Evidence for these claims comes from studies showing a high proportion of formulaic utterances in early child language as well as initial restrictions of morphemes and constructions to particular lexical items (see Tomasello (2003) for an overview). However, the studies available so far typically involve single children and their mothers and were conducted in Western societies. Thus, it is still open whether these results can be generalised to different cultural and family settings. In addition, recordings from situations which are less ritualised and more linguistically challenging than meal times, picture-book reading, and recurrent games are required to avoid any underestimations of children's linguistic productivity.

Cross-cultural and crosslinguistic studies have already provided a more nuanced picture of children's input. For instance, they refuted the claim that all children receive systematic implicit negative evidence, i.e. different types of responses of adults to children's incorrect and correct utterances (for instance more questions following incorrect children's utterances; see Marcus (1993) for discussion). However, there is initial evidence for other universal properties of children's input, in particular the predominant use of simple sentences, a high pitch and a high degree of variation in intonation (Ingram (1989), Locke (1995)). Moreover, child-directed speech is characterised by sequences of adult utterances with a constant communicative intention and different types of variation in form, e.g. lexical substitution and rephrasing, a shift

from full noun phrases to pronouns, the addition, deletion or reordering of constituents (Küntay and Slobin (1996), Eisenbeiss (2003), Bowerman et al. (2003)). Such 'variation sequences', together with syntactic simplification and strong prosodic patterns, can highlight constituent boundaries as well as morphological contrasts and provide evidence for word order flexibility, syntactic processes, and the optionality of particular constituent types. But, again, crosslinguistic and cross-cultural studies are needed to determine which of these properties are truly universal and how exactly they contribute to the process of language acquisition.

2.2 The temporal organisation of language development

Acquisition researchers not only have to find solutions for the logical problem of language acquisition and to test their claims about universals in crosslinguistic and cross-cultural studies. They also have to explain why a given grammatical property is typically acquired at a particular age, i.e., they have to account for the absolute timing of individual acquisition processes. Moreover, they must provide an account for the order in which grammatical phenomena are acquired, i.e. for the relative timing of language development.

With respect to the absolute timing of language development, some generative linguists argue that certain grammatical properties only become available after a certain point in development, activated by processes of neural maturation. For instance, Borer and Wexler (1987, 1992) have claimed that some of the syntactic processes required for the formation of passive sentences like The man was bitten by the cat "mature" around the age of four. This hypothesis, which was based on English, has been challenged by studies on other languages, where passives are more frequent in the input (see e.g. (Suzman (1985), Pye and Poz (1988), Demuth (1989, 1990), Allen (1994), Allen and Crago (1993, 1996)). In these studies even two-year old children have been found to produce passives productively, which suggests that the relevant syntactic processes are already available to them. Moreover, such findings show the potential impact of input properties like frequency.

With respect to the relative timing of acquisition processes, researchers have referred to both formal and conceptual factors. The most (in)famous form-oriented approach was the so-called derivational theory of complexity (Miller and Chomsky (1963), Brown/Hanlon (1970)). According to this approach, complex sentences are derived from basic sentences by transformations and the more transformations a construction involves the more complex it is - and the later it should be acquired. For instance, passives are assumed to be derived from active sentences by the passive transformation and should thus be acquired after active sentences. The derivational theory of complexity could not be confirmed empirically and proved to be too simple to capture the complexity of language development (Fodor et al. (1974), Ingram (1989: 435ff.)).

One of the best-known concept-based accounts of developmental orders was developed by Gentner (1982). According to her, nouns are universally acquired before

verbs because nouns can easily be mapped onto object concepts whereas the mapping of verbs onto action concepts is cognitively far more complex. However, recent studies have shown that children do not acquire verbs later than nouns when verb types are more frequent in child-directed speech than noun-types – as in Mandarin Chinese and Korean (Choi and Gopnik (1995); Tardif (1996)).

Taken together, the available results suggest that acquisition orders cannot simply be predicted on the basis of form or meaning alone, but result from a complex interplay of formal and semantic factors. For instance, studies on many Indo-European languages have documented a two-word stage in which grammatical morphemes and function words are frequently omitted, giving the impression of telegraphic speech (e.g. doggy run instead of the doggy is running). In contrast, children acquiring Greenlandic Eskimo, where the input consists mainly of polysynthetic verbs, do not exhibit a two-word stage, but go through a stage where most utterances do involve more than two morphemes (Fortesque and Lennert Olsen (1992)). This suggests that there might indeed be some maturational constraints on early child language, but that they do not simply affect all grammatical elements or particular syntactic operations. Rather, they seem to put an upper limit on the number of elements that children can produce in early stages of language development – and the morpho-syntactic properties of the target language then determine how this limitation affects child language. However, in order to test such accounts for the timing of language development, we need more studies which systematically vary the formal complexity of constructions and word forms within and across typologically diverse languages.

2.3 Grammatical development

In addition to the logical problem and the problem of development and order, acquisition researchers must explain how children acquire the form-meaning pairings encoded in the morphemes and constructions of their target language. In the 1980s and 1990s, several explicit models were developed. Proponents of the so-called semantic bootstrapping approach postulated an innate repertoire of grammatical categories, which are inherently linked to innate semantic categories (e.g. NOMINATIVE or ERGATIVE: Tr.AGENT, NOMINATIVE or ABSOLUTIVE: Intr.AGENT; ACCUSATIVE or ABSOLUTIVE: PATIENT; Pinker (1984). Moreover, children are assumed to use semantic information and their innate form-function links to find instances of their innate grammatical categories in the input. For instance, they could compare the markers for AGENT and PATIENT arguments of intransitive and transitive verbs to find out whether their target language had an accusative system (itr.AGENT= tr.AGENT) or an ergative system (itr.AGENT = tr.PATIENT). In such a model, children have to compare markers for arguments of transitive and intransitive verbs and thus must be able to unambiguously distinguish between them. However, in a language with argument ellipsis like Japanese, many transitive verbs frequently occur with just one argument or no arguments at all. Thus, Pinker would either have to explain how children manage to distinguish transitive and intransitive verbs or he would have to give up the assumption that children make a decision about the transitivity of every single verb they hear (see Lüpke, this volume on the relation

between lexical transitivity and argument occurrence in discourse in Jalonke). Moreover, he would have to extend his model to explain how children acquire languages with a split case system, where arguments with the same thematic role can carry different markers, depending on factors such as tense, aspect, person, animacy, pragmatic function, etc. (see Dixon (1987), VanValin, (1992), Blake (2001)). Such an extension would, for instance, be necessary to account for the acquisition of Hindi, where ergative markers are restricted to perfective contexts (see Kachru/Pandharipande (1978), Kachru (1987), Narasimhan accepted, a, b).

Similar problems arise for syntactic bootstrapping approaches, which claim that the syntactic frames of verbs can help children to home in on plausible verb meanings (see e.g. Gleitman (1990), Fisher et al. (1994), Naigles (1990, 1996)). For example, based on research with English-speaking children, Fisher et al. (1994:366) argue that children who hear a verb with three noun phrases in a sentence like She VERB-ed the chicken some food can infer that this verb expresses transfer. However, even the closely related language German exhibits sentences with three noun phrases which exhibit the same case markers and syntactic properties as the three arguments of the possession transfer verb geben 'to give', but do not encode transfer or possession (e.g. [Die Frau]NOM hackt [dem Huhn]DAT [den Kopf]ACC ab; [The woman]NOM hacks [the chicken]DAT [the head]ACC off; 'The woman is hacking the chicken's head off'). Moreover, not even the sentence as a whole needs to encode possession or transfer when a verb appears with three noun phrases. For instance, according to Wilkins (p.c.), the perception verb see/look in the Australian Aboriginal language Arrernte occurs with an ergative, an accusative and a dative noun phrase, but this verb does not express transfer and the dative noun phrase does not encode a GOAL or a POSSESSOR, but the endpoint of a path of vision (which is the location of the event participant that is encoded as an accusative noun phrase).

Findings such as those mentioned above have led to a number of revised models which take the crosslinguistic variability in argument realisation and form-meaning mapping into account and try to accommodate the linguistic properties and acquisition facts which have already been documented in crosslinguistic studies on adult and child language (see e.g. Eisenbeiss (2003), Fisher (1999, 2002) and the references cited there). These models cannot be discussed in detail here, but it is clear that any further development in this direction will require even more precise information about interlanguage variability and potential universals as well as information about acquisition paths and error patterns for typologically diverse languages. The same is true for models of the acquisition of other syntactic, phonological or semantic aspects of language.

2.4 Lexical development

Languages do not exhibit a simple 1:1 mapping of atomic concepts to lexical items. Rather, languages differ with respect to the way in which they divide their semantic space and package meaning components in lexical elements, and children must learn the language-specific semantic distinctions and packaging. For instance, languages

encode different amounts and types of event-participant information in their predicates (see e.g. Brown 2001), e.g., the roots of *holding* or *carrying*-verbs in Tzeltal Maya reflect where and how on the body the object is supported (on head, in arms, etc.), while speakers of Navajo must choose a verb root according to properties of the carried object itself (animate, a container with contents, etc.). Moreover, predicates impose constraints on participants and selection restrictions. For instance, the associated body part is not an argument of carrying verbs in Tzeltal.

Initial studies showed that children learning English and other European languages often fail to respect predicate constraints on event participants, e.g., they apply break to the tearing of cloth or paper and cut to smashing a walnut with a mallet or crushing ice with a rolling pin (Bowerman (1978), p.c., E. Clark (1993)). Such errors suggest that children start out with relatively undifferentiated event meanings and only later constrain them by adding in finer participant specifications. However, children acquiring Korean differentiate accurately among verbs for putting clothing on various body parts (Choi and Bowerman (1991)), and learners of Tzeltal discriminate among several verbs for carrying objects (Brown (2001)). Thus, more crosslinguistic studies are required to determine when and how children home in on the language-specific semantic distinctions of their target.

In addition, there are many studies on the acquisition of lexical packaging, most of them based on Talmy's typology of motion event encoding. According to Talmy (1985), motion events involve four basic components – MOTION, FIGURE (the moving entity), GROUND (the reference point object with respect to which the figure moves) and PATH (the course followed by the figure with respect to the ground). In addition, a motion event can have a MANNER or a CAUSE. Languages differ with respect to the way they package the different aspects of motion events for linguistic encoding. In so-called satellite-framed languages like English, the motion component is typically conflated, i.e. combined with the manner or cause of the motion and encoded in the verb, whereas the path is expressed by a 'satellite', e.g. a spatial particle or a prepositional phrase (*Jack walked/ran into the room*). In contrast, in verb-framed languages like Spanish, the path and motion tend to be conflated and realised in the verb whereas manner – if expressed at all, is encoded separately, e.g. in an adverbial phrase (*Jack entro al cuarto caminando/corriendo* 'Jack moved-into the room (walking/running)'). Since the 1990s, a growing number of studies have investigated how children learn the appropriate packaging of the motion event components in their target language (e.g. Berman and Slobin (1994), Bowerman et al. (2003), Slobin (2004)). These studies have shown that even the earliest word combinations are influenced by the lexical packaging strategy of the target language. However, more studies are required to determine how they arrive at these patterns and how they deal with lexical items in their language which do not exhibit the dominant conflation pattern of the language (e.g. the English verb *to enter*, which conflates MANNER and PATH like a verb from a satellite language). Moreover, we need more studies on the acquisition of languages that do not show a clear satellite-framed or verb-framed pattern, e.g. languages with serial verb constructions (e.g. Lao) where the different semantic components can be encoded in separate verbs.

2.5 The acquisition of narrative skills and conversational competence

Learning a language is not restricted to acquiring its grammar and lexicon. Speakers have to go beyond individual utterances and produce connected discourse and they must become a competent member of a particular linguistic community. Hence, a good documentation of child language must provide a basis for studies on narrative development and the linguistic socialisation process that leads to conversational competence in a particular community.

The starting point for studies on narrative development is the fact that the world does not present separate 'event units' which can then be encoded in discourse. Rather, events must be construed and verbalised through choice of perspective and through the grammatical options of the target language. Moreover, a narrative does not consist of a simple linear chain of events, but involves hierarchical structuring of the plot. Thus, children have to learn the language-particular means of event filtering, packaging and plot structuring; and a growing body of crosslinguistic studies have investigated how they achieve this (see e.g. Berman and Slobin (1994), Strömqvist and Verhoeven (2004)). Many of these studies involve similar or identical stimuli (e.g. wordless picture books) and thus allow for detailed crosslinguistic comparisons. These comparisons show that, independently of the target language, children initially tend to treat scenes as isolated events and only later start to relate events more by temporal links. However, by the age of nine, they are usually able to distinguish between foregrounded and backgrounded information, to encode temporal and causal connections, and to make evaluative statements about the states of mind of the characters in the narrative. Moreover, from the earliest stages, children are highly influenced by the linguistic means that their target language provides for perspectivisation and event packaging. Thus, there does not seem to be an initial universal stage in narrative development.

Studies on linguistic socialisation and the development of conversational competences investigate the factors that enable children to become a competent member of a particular linguistic community. Such studies have compared the interactions of children with their siblings, peers, older family members and other adults and they have shown the following factors to be crucial: explicit comments or differential responses for specific aspects of linguistic behaviour, the provision of adult and peer models, and the participation in particular gender- or class-specific subcultures which create and maintain particular styles of interaction (see e.g. Ochs and Schieffelin (1986, 1995), Coates (1998)).

2.6 Implications for data collection

The discussion in this section has shown that crosslinguistic studies on children's language and input are essential for a more nuanced picture of language acquisition. They have contributed to the debate about the logical problem of language acquisition by revealing potential universal properties of child language and child-directed speech.

For this it is necessary to obtain samples of children's interactions with primary caretakers, peers, siblings and other speakers in their community. Simply recording individual children's interactions with a researcher will not provide information about the relation between children's linguistic behaviour and the input they actually receive in their daily lives.

Crosslinguistic acquisition studies have also addressed the problem of development and order and investigated the relative and absolute timing of acquisition processes. Some of these studies involve longitudinal designs which focus on individual development and monitor the language development of individual children by recording them several times over a long interval. Others are cross-sectional studies with larger groups of children from different age ranges who are each recorded at one point in time only. These cross-sectional studies can supply acquisition researchers with representative information about the ages at which children typically master particular grammatical phenomena. However, at least three children per age group are required to capture inter-individual variation in the speed of linguistic development. Moreover, these children should be carefully chosen. For instance, one cannot study the crucial developmental processes between the age of two and three years if all three year olds in the sample are linguistically precocious and the two-year olds are developing comparatively slowly. Therefore, when planning a cross-sectional study, it is advisable not to rely on information about chronological age alone, but to supplement the information about age with a general measure of linguistic development for each child.

The most common general measure for linguistic development is the so-called MLU, i.e. the mean length of utterance (Brown (1973)). The MLU is calculated by determining the total number of morphemes in a sample with at least 100 utterances and dividing it by the number of utterances; see Brown (1973:53ff) for details of the original procedure and the CHILDES webpage for details on calculating MLU-values with the CLAN tools (http://childes.psy.cmu.edu/). The rationale for this measurement is based on the idea that an increase in average sentence length reflects an increase in morpho-syntactic complexity (see e.g. Brown (1973:53)). This assumption is not without problems. For instance, a 6-morpheme utterance like Sally want no more egg now, which lacks several grammatical morphemes, is morpho-syntactically less complex than a 5-morpheme utterance like Sally want-s egg-s, which exhibits target-like grammatical markers.

The exact procedures for MLU-calculation have to be adapted to the morphological properties of the respective language; and researchers have to decide (i) how speech should be segmented into utterances, (ii) which utterances should be included in the sample, and (iii) how the number of morphemes should be determined. All of these issues are a matter of debate (see e.g. Brown (1973), Eisenberg et al. (2001: 328ff) for discussion and further references.). Nevertheless, it is generally accepted that the MLU is a better predictor of linguistic behaviour than chronological age alone – at least as long as the MLU values are small. MLU-values above 3 or 4, which are typical for children of three years and older, tend to show a higher degree of variation and fluctuation (see Eisenberg (2001:325) for references and discussion). Hence, they are

assumed not to be reliable measures of linguistic development, but to reflect the nature of the particular communicative interaction (see e.g. Brown (1973)).

Thus, the mean length of utterance may not be a perfect measure of linguistic development and will have to be adapted to the language in question, but it can provide additional information for participant selection and help to achieve a more representative sample. Nevertheless, it is still difficult to draw direct inferences about developmental orders from cross-sectional samples as they do not provide data on individual development. A more reliable indicator of developmental orders is implicational relationships between the mastery of different grammatical phenomena. Take, for instance, a cross-sectional sample where (i) some children produce morpheme A as well as morpheme B, (ii) some children use only A, and (iii) no child shows mastery of B, but not A. Here, the acquisition of B implies the acquisition of A – which suggests that A is acquired before B – or at least not later.

Beyond providing information about universals and developmental orders, both longitudinal and cross-sectional acquisition studies with typologically different languages have helped to evaluate and modify models of grammatical development. For such studies, it is crucial to take into account that not all grammatical phenomena necessarily pose the same learnability problems. For instance, children might not learn the distribution of case markers, which are quite closely linked to thematic roles, in exactly the same way as the distribution of gender markers, which might have some semantic basis, but are ultimately based on purely distributional co-occurrence patterns in contexts of agreement (see Pinker (1984) and Eisenbeiss (2003, 2005)). Thus, these differences are reflected in acquisition models and in order to test such models, acquisition researchers need a data base that allows them to investigate a broad variety of grammatical phenomena. As not all of them are highly frequent, this might require either large samples or the use of elicitation techniques (see section 5). Moreover, children show considerable variation with respect to the absolute and relative timing of the acquisition process, the proportion of formulaic utterances etc (see e.g. Bates et al. (1995)). In addition, some children initially tend to replace function words or morphemes with phonologically reduced forms whereas others omit these elements completely or produce fully inflected - but often not target-like - forms from the beginning (Peters (2001a, b), Eisenbeiss (2003)). Thus, in order to distinguish between general acquisition mechanisms employed by all children and individual strategies or preferences, it is crucial to evaluate acquisition models on a representative sample of children which includes children from different genders and social groups as well as children with different personality types.

For studies on lexical development, data for different semantic domains are needed and these data must involve the linguistic encoding of similar, but slightly different events and objects to allow for detailed studies of children's semantic space and its development. This makes it advisable to use elicitation techniques to supplement spontaneous speech samples. Moreover, for fine-grained semantic analyses of words for events, topological relations and (parts of) objects, the intended referents of these words must be clearly identifiable from the non-linguistic context, which is easier with video recordings than with audio recordings or note-taking. This is also true

for studies on aspects of events that are not easily recoverable from audiotapes (e.g. the exact manner of motion of a motion event described by the child).

Finally, for crosslinguistic acquisition studies on conversational competence and narrative skills, one needs representative data samples which involve participants from different social groups and genders and cover interactions both within and across genders, social and age groups. Moreover, one should record narratives which come from different genres and are aimed at different audiences. In order to keep them comparable across cultures, one should consider the use of well-established elicitation techniques and materials (e.g. video-clips and picture-books that have been used in previous crosslinguistic studies; see section 5).

Thus, in sum, data for acquisition studies must come from a broad range of participants, genres, and discourse types and provide data for different grammatical phenomena and semantic domains. Moreover, studies on grammatical and lexical development tend to focus on children below the age of six; and studies on phonological development often even start with the very earliest sound productions of children. In contrast, studies on narrative development typically involve older children. Thus, data from a broad age range is required to allow for different types of analysis. Ideally, the documentation should cover the time span from birth to the pre-teenage years. This requires efficient corpus planning. In particular, it is advisable not to record a comparatively large group of individual children and to obtain input samples for each of these children from their primary caretakers. Rather, one should try to make recordings in a small number of families who come from different social backgrounds and each have several children. In this way, less time is required for travelling between recording sites and setting up equipment and one needs to record fewer adults for an analysis of children's input. In addition, the recording of all siblings provides a better basis for input analysis as it provides crucial information which is often neglected in studies where only the primary caretaker's conversations with the target child of the study are recorded and other sources of input are ignored. Moreover, factors of social group are easier to control if the sample involves several groups of children who share the same social and family background. However, this strategy might have to be modified for studies in cultures where small families are the norm or at least common. Here, it is advisable to record at least some single children in order to ensure representativeness. Moreover, some group recordings with several peers should be made to obtain a larger sample of children and to capture input from peers.

3. Descriptive, theoretical, and historical linguists

Careful corpus planning is also required if child language documentation is not only to make a contribution to acquisition research, but also to the description of adult language and its change over time. In the following, I will use a few examples to show how child language studies can supplement studies on morphological default operations, verb classes, implicational universals and language change. I will not argue that the acquisition results provide data that are "better" than data from studies on the adult linguistic system and, of course, the interpretation of individual findings is still a

matter of debate. However, I will argue that any claims about linguistic structure are stronger if they are based on converging evidence from typological and crosslinguistic acquisition studies. Moreover, I will show how the consideration of learnability issues which arises from acquisition studies might provide a new perspective that could lead to deeper insights into the linguistic systems of the world's languages and their development over time.

3.1 Morphological default operations

When several different morpho-syntactic operations compete, the status of these operations needs to be determined. Specifically, grammatical operations like the addition of an affix can be default operations; i.e. can apply to any member of the respective category, regardless of phonological or semantic properties, unless the application is blocked by the existence of a stored form. An example of this would be the English plural rule which adds –s to any noun without an irregular plural. Alternatively, an operation can be productive, but its application may be restricted to a particular subset of elements from this category which shares certain semantic, phonological or morphological properties, i.e. the operation can have an input condition. For instance, the affixation of the genitive singular marker –s in German is restricted to a particular declination class (so-called strong non-feminine nouns). And finally, a grammatical operation might be completely unproductive synchronically, like the affixation of *–en* in English irregular plurals (e.g. *oxen*).

Assigning a morphological operation to one of these three classes is not always straightforward. For instance, German noun plurals involve one of five plural endings (*-s, -(e)n, -e, -er, -ø*). Theoretical linguists agree that *–er* and *-ø* are not productive, but there is an ongoing debate about the status of the other plurals (see Clahsen (1999) and replies for an overview): *–(e)n* and *-e* are frequent and productive, but tend to appear with nouns that have particular morpho-phonological properties (e.g. feminine nouns ending in schwa for *-(e)n*). In contrast, *-s* has a low frequency, but applies to a broad range of nouns, including borrowings, acronyms and novel words. Therefore, there is a debate whether the more frequent or the more unrestricted plural – or none of them – is a default form. Similarly, Dutch has two plural affixes, *-s* and *–n*, which both apply to nouns that have particular phonological properties. Therefore, it is debated whether they both have the same status or whether one of them involves a morphological default operation.

These debates are far from settled and linguists have brought forward many arguments which are based on the adult language. Recently, acquisition studies have started to shed new light on these issues (see e.g. Clahsen (1999) and replies). In particular, researchers argue that children should overgeneralise default affixes to contexts in which a more specific form is required (e.g. ox-es instead of oxen). Moreover, default overgeneralisations should even occur when the respective operation has a low frequency (like the German –s plural) and they should not be restricted to items with particular phonological or semantic properties. In contrast, unproductive processes should only lead to few, similarity-based overgeneralisations. And processes

with input conditions should lead to generalisations which are constrained by these conditions. For instance, if the German -(e)n- plural is a process with an input condition (a restriction to feminine nouns ending in schwa), it should be generalised to all nouns fulfilling this condition – even if the child has not heard them yet. In addition, processes with input conditions might be influenced by input frequency because children need evidence for the acquisition of the process itself and for the acquisition of its input conditions. Thus, data about overgeneralisations cannot only inform us about the acquisition process, they can also help linguists to determine the status of grammatical representations. So far, this type of research has been restricted to a few Indo-European languages (most of them Germanic). Only recently, researchers have begun to study more complex morphological systems in non-Indo-European languages (e.g. the noun class systems of Bantu languages, see e.g. Demuth 1992, 1994).

3.2 Verb classes and form-meaning mappings

Linguists interested in form-meaning mappings use the syntactic behaviour of verbs to classify them and to derive information about form-meaning mapping from these classifications (see e.g. Levin 1993). For instance, linguistic studies on German and other languages (e.g. Basque) have distinguished two classes of two-place dative verbs (e.g. Wunderlich 1997, Joppen and Wunderlich 1995); and acquisition studies with 2-3 year old German children support these analyses: The dative argument of verbs of social interaction encodes a THEME and typically follows the nominative-marked AGENT noun phrase (see (5)). Moreover, the dative argument can become the nominative subject in so-called recipient passives (see (6)).

(5) Der Junge $_{NOM}$ hilft dem Huehnerzuechter$_{DAT}$.

The boy$_{NOM}$ helps the chicken farmer $_{DAT}$

(6) Der Huehnerzuechter $_{NOM}$ kriegt geholfen

The chicken farmer $_{NOM}$ gets helped

Therefore, it has been claimed that these dative THEME arguments carry a lexical case but otherwise behave like THEMEs which are realised as direct accusative objects. This is supported by the observation that children tend to initially replace the dative marker for the THEME with an accusative marker and start to produce the correct marking on a verb-by-verb basis as expected for an acquisition process that involves the learning of the item-specific properties of individual verbs (Eisenbeiss (2003), Eisenbeiss et al. to appear).

In contrast, dative EXPERIENCER/POSSESSORs tend to precede the nominative-marked STIMULUS/POSSESSUM (see (7)) and cannot become the subject in recipient passives (see (8)). Hence, it has been claimed that the dative EXPERIENCER/POSSESSOR is ranked higher than the STIMULUS/POSSESSUM of these verbs, but carries a lexical dative marker instead of the nominative marker for higher arguments. Evidence for this assumption comes from the observation that

children overgeneralise nominative markers to dative EXPERIENCER/POSSESSORs (Eisenbeiss (2003); Eisenbeiss et al. to appear).

(7) Dem Huehnerzuechter gefaellt die neue Huehnerrasse.

To-the chicken farmer $_{DAT}$ appeals the new breed of chicken $_{NOM}$

(8) * Der Huehnerzuechter kriegt gefallen.

The chicken farmer $_{NOM}$ gets pleased

Similar classes of dative verbs have also been documented for typologically different languages like Japanese and Basque (see e.g. Gamerschlag (1996), Joppen and Wunderlich (1995)); and child data could reveal whether similar error patterns can be observed for such languages. Similarly, theoretical linguists have provided a variety of analyses for the distinction between unaccusative and unergative verbs and these analyses have been the basis of studies on several Indo-European languages and should be tested on a larger variety of languages (e.g. vanHout (1996)).

3.3 Implicational universals

Typologists and acquisition researchers have a shared interest in implicational universals, which state that the presence of one linguistic property in a language implies the presence of another linguistic property in the same language. For instance, Jakobson (1942) claimed that the occurrence of fricatives should presuppose the occurrence of stops. Hence, one should not find a language that has fricatives, but no stops. Moreover, if children's grammars were subject to the same constraints that apply to adult speakers' grammars, there should be no stage in language development where children already produce fricatives, but no stops.

Thus, child language data could provide additional evidence for implicational universals which are established on the basis of typological studies. In addition, acquisition models can provide accounts for implicational universals. For instance, according to Greenberg's universal number 36, the occurrence of grammatical gender distinctions implies the occurrence of number distinctions (Greenberg (1963:74)). Similarly, crosslinguistic acquisition studies show that two- and three-year olds already exhibit gender distinctions, but gender acquisition does not seem to start before number distinctions appear (see Eisenbeiss (2003) for an overview).

This is unexpected if one adopts a semantic-bootstrapping analysis of gender acquisition like Pinker (1982). According to this analysis, children use the natural gender of noun referents and innate links between innate gender categories and concepts of natural gender to find instances of gender categories (i.e. forms with a particular gender) in the input. Thus, acquiring gender distinctions should be independent of the acquisition of other morphological distinctions and children should

be able to acquire linguistic systems that have gender, but no number distinctions.[1] Hence the implicational relations between gender and number cannot be captured if one assumes that gender distinctions are acquired by bootstrapping from natural gender.

In contrast, the observed implicational relations between gender and number can be captured if one considers the different status of number and gender features: number features do not encode an inherent property of a noun or verb. Rather, they can be treated as output features for grammatical operations which apply to these lexical elements. For instance, the addition of a noun plural marker can be interpreted as an operation which creates a plural form with output specification [+PLURAL] that contrasts with a singular form with the specification [-PLURAL]. Moreover, children should be able to determine the meaning distinctions that correspond to the contrasts between singular and plural forms independently of other grammatical features (see Eisenbeiss (2003) for details).

In contrast, gender distinctions may be semantically motivated, but they are ultimately based on purely distributional co-occurrence patterns in contexts of agreement (Corbett (1991)). Moreover, gender features (and noun class features) are inherent features of nouns and not assigned to them by a grammatical operation. Hence, they are not output features. However, they can impose restrictions on the application of grammatical operations, e.g., a [+FEMININE]-feature can restrict the affixation of the Latin nominative plural marker –ae to those adjectives which agree with a feminine noun. Adjectives agreeing with neuter or feminine nouns carry different affixes in the plural. Thus, gender features can be analysed as input features which restrict operations that assign output features. Eisenbeiss (2003) argues that children can only detect such a restriction on a grammatical operation once they have identified this grammatical operation in the first place. For instance, children can only acquire gender-based restrictions for plural affixes once they have discovered that plurals are marked differently from singulars. Only then can they establish a paradigm with the respective output forms – e.g. a simple adjective paradigm with a singular and a plural cell. If the language exhibits gender distinctions, they could find two or more gender forms competing for cells in this initial paradigm. This should then motivate the detection of the distributional patterns for the different gender forms that compete for one cell.

Note that this account for Greenberg's universal 36 is solely based on the claim that output features like NUMBER can be acquired independently of other features, whereas the acquisition of input features like GENDER requires the prior detection of output feature distinctions. This account does not make any predictions about GENDER and NUMBER per se. Other output features (e.g. case features) could also provide the basis for the discovery of gender distinctions. Hence, the fact that Greenberg only found an implicational relationship between gender and number - and not between gender and other output features like definiteness - might be an artefact of the pervasiveness of number distinctions in the world's languages. There might be

[1]Moreover, if Pinker's analysis were correct, children should show specific problems with the gender of inanimate noun referents as these do not have any natural gender, but this prediction could not be confirmed (MacWhinney 1978, Maratsos/Chalkey 1980, Levy 1983, Mills 1986, Müller 2000).

languages with gender distinctions which do not exhibit number distinctions, but given the learnability argument discussed above, such languages should have other nominal output features that result in paradigmatic contrasts. This hypothesis could be tested both in typological studies and in acquisition studies where implicational relations in the acquisition process could be investigated. Thus, in sum, data on developmental sequences in child language cannot only provide additional evidence for implicational universals which are established on the basis of typological studies. They can also provide a learnability account for such universals and a new perspective for further studies.

3.4 Language change

Language change occurs when particular linguistic properties are not transmitted across time, i.e. when older linguistic properties or elements are reanalysed and/or substituted by newer linguistic properties or elements. Adult monolingual native speakers exhibit such substitutions in the lexical domain, where they use new words to replace old words. In contrast, systematic grammatical innovations do not seem to be characteristic for adult monolingual native speakers (see Kroch (2001)). Rather, the lack of transmission for grammatical properties seems to be related to the process of language acquisition. Thus, the documentation of child language and child-directed speech is relevant to the study of language change. In particular, it can help historical linguists to determine whether linguistic change arises due to processes involved in first language acquisition or due to the way in which adults learn a second language in situations of contact.

Many historical linguists attribute language change to subtle changes in the evidence available to children learning their first language (see e.g. Lightfoot (1991, 1999)). For instance, the so-called principles and parameters model of generative grammar postulates innate universal wellformedness constraints for linguistic representations and innate "parameters", i.e. a finite set of values from which learners can choose on the basis of simple input data. For instance, the verb-second parameter determines whether finite verbs are positioned in the second position of the sentence and can be preceded by any other constituent (as in German) or not (as in English where the position immediately preceding the verb is restricted to subjects). The trigger, i.e. the relevant positive evidence for the [+VERB SECOND]-setting of this parameter would be sentences with a non-subject constituent in sentence-initial position and a finite verb in second position, e.g. object-verb-subject sentences. This trigger could become rare in children's input, for instance because their parents start to use fewer object-verb-subject sentences and instead produce more subject-verb-object sentences, which are compatible with both settings of the parameter. Such a lack of trigger data could then prevent children learning a verb-second language from setting the parameter to the target value – and thus lead to language change. Thus, adults might not change their grammar, but they might start to use certain linguistic structures less frequently, which might lead to their children building up grammatical representations that are different from the language of their parents.

Note that the idea that changes in input frequency can bring about syntactic change may be central for generative approaches, but it is also compatible with usage-based accounts of language acquisition. After all, such accounts assume that any type of language acquisition is driven by frequent patterns of form/meaning mappings in learners' input (e.g. Tomasello (2003)) – and thus it can be expected that changes in these patterns cause children to develop linguistic representations that differ from the ones of the older generation, who was exposed to data with different frequency distributions. Thus, independently of the linguistic framework, one would expect that at least some processes of linguistic change arise due to processes involved in first language acquisition. Then, syntactic innovations should be visible in children's data, but not in the data of their parents. However, the change should be caused by a low frequency of trigger data, which should be detectable by an analysis of children's input. Thus, an analysis of children's own language and their input could provide evidence for approaches that link changes in grammar to processes of first language acquisition.

However, many researchers assume that at least some language change processes are not caused by children, but can be attributed to the way in which adults learn a second language in situations of language contact, (e.g. Thomason and Kaufman 1988). In these situations, a group of adults might learn the second language imperfectly, due to interference from their first language. If their second-language usage then becomes the input for their children, they might pass on the interference-based properties of their own imperfectly learned second language. In such a situation, innovations should appear in the data of adult native speakers, and their children should pick up these changes in their own speech – a prediction that could be tested by analysing children's own language production and their input. Thus, in sum, child language data and children's input data can help historical linguists to distinguish between two different causes of acquisition-related language change: (i) frequency shifts in input data for first language acquisition and (ii) imperfect second language learning in contact situations.

3.5 Implications for data collection

As the previous discussion has shown, acquisition data can contribute to descriptive and theoretical linguistics by providing an additional source for testing claims about morphological default operations, verb classes and implicational universals and by providing the basis for learnability accounts for implicational universals. However, such a contribution to linguistic research is only possible if the documented child data sample involves a broad range of phenomena, including low frequency phenomena information about input frequencies for different forms and classes of linguistic elements. In order to achieve this, the range of activities and conversation topics for spontaneous speech samples should be quite broad. In addition, the individual samples should be large enough to allow for quantitative analyses of error data, e.g. for comparisons of error rates for different types of arguments or grammatical morphemes. Moreover, as I will argue in the following, spontaneous speech samples should be supplemented by elicited data that cover the core constructions of the language, ideally

with a focus on those phenomena where acquisition data might make a relevant contribution to language description and linguistic theory.

When it comes to implicational universals, longitudinal studies can provide fine-grained information about steps in development. However, just as claims about implicational universals in typology should be based on a sufficiently large sample of languages, claims about implicational universals in development should be based on (cross-sectional) studies with representative groups of children for different age ranges so that inter-individual variation is taken into account. Thus, a combination of longitudinal and cross-sectional designs would be optimal.

For potential contributions to historical linguistics, a good coverage of different phenomena is also crucial, but the path of development is slightly less important. Rather it is essential to provide child language data as well children's input data because both are required to determine whether ongoing processes of language change are driven by first or second language learners. Thus, in sum, a documentation of child language can make a contribution to descriptive, theoretical and historical linguistics if it provides a good coverage of the relevant grammatical phenomena in children's own speech and their input and shows the course of children's linguistic development.

4. Language maintenance projects and linguistic communities

Fieldworking linguists who are mainly interested in language maintenance and revitalisation might be very concerned by the idea of allocating their limited resources to the collection of child language data. And members of a linguistic community who are interested in resources for the maintenance of their language might find acquisition data less relevant than texts and recordings which represent more prestigious variants of the language. However, as I will argue in the following, documenting child language may be a good 'investment' for language maintenance projects as it can contribute to (i) the assessment of the endangerment status of languages (ii) a positive attitude towards the continued use of endangered languages and (iii) the development of resources for language maintenance and revitalisation projects.

4.1 Assessing endangerment status of a language

For an endangered language to survive, it must be passed on from parents to children. Hence, information about the degree to which this intergenerational transmission actually occurs is crucial for assessing how endangered a given language actually is (see Fishman (1991), Salminen (1999) for an overview of other endangerment criteria). Studies on intergenerational transmission typically use demographic data, questionnaires and interviews to investigate the size of linguistic communities, their attitudes to the language in question and its use, and these studies either focus on the official use of languages or they compare their use in different situations (see e.g. Fishman (1991) for an overview).

However, one could argue that the ultimate test and predictor of language maintenance is in parent–child interactions, because interactions between family members, in contrast to more formal situations, typically allow for real choice between languages. Thus, data on the actual use of different languages in child-adult interactions within the family could supplement the findings from demographic, questionnaire and interview studies. In my view this might be useful; the parents I have worked with in the creation of child language corpora were usually very keen to provide rich and varied linguistic input for their children and most of them were convinced that they fulfilled their own expectations to a large degree. But when they actually saw themselves on video, they were often surprised by how little they actually talked to their children – and how formulaic and full of linguistic routines their actual language use was. Hence, data from real child-adult interactions might provide a new perspective on statements offered in interviews and questionnaires and such data might help to develop better questionnaires with more detailed questions as they show areas where self-perception and actual language are most likely to differ. In addition, recordings of child-adult interactions where a particular language is chosen over others might also provide some information about the attitudes and other factors that influence this choice.

4.2 Fostering a positive attitude towards the continued use of endangered languages

Documenting acquisition data does not only provide a basis for assessing the state and development of intergenerational transmission, it can also foster a positive attitude towards the continued use of endangered languages. The very fact that a language is being documented already adds to its status and might thus make its preservation more desirable in the eyes of its native speakers. However, if only prestigious texts and the variety spoken by the elders and the elite of the language community were documented, this value might not be attributed to other variants and elitist ideologies might lead to a disengagement of many speakers who might otherwise become involved in language maintenance or revitalisation (see e.g. Flores Farfán (2001), Austin (2004) for a discussion of problems caused by linguistic purism and elitism). In contrast, it clearly signals the value of ALL variants of a language if researchers take an interest in documenting even those varieties which are clearly not "perfect" in the eyes of adult native speakers and spoken by low-ranking members of the community. It is probably no coincidence that for many European languages, the documentation of dialects and child language tended to go hand in hand when it started at the end of the 19th century (see e.g. Grimm (1864), Stern and Stern (1907, 1928)). Thus, the very fact that child language is documented might already challenge elitist and prescriptive views of language that stand in the way of language maintenance.

The real challenge, however, lies in convincing parents and educational professionals that the continued use of an endangered language does not limit the professional and social perspectives of children who grow up with this language as their native language or as one of the core languages in a multilingual setting. Here, the results of studies on multilingualism which clearly show its benefits could prove

helpful. Such studies show that providing children with input or instruction for the endangered language does not impede the acquisition of the national language and reading skill; rather it promotes the child's educational and cognitive development and at the same time it increases the active participation of parents in their child's schooling and the child's esteem for the parents (see e.g. Dolson and Mayer, 1992, Cummins, 1989, 1992, 2000). However, despite the fact that multilingualism is not the exception, but the norm for speakers across the world (see e.g. Cook 2003 for discussion), the vast majority of studies on bilingualism or multilingualism are based on pairs of languages that are closely related (e.g. English/German or German/Italian) and/or spoken in a Western society (e.g. Spanish/Basque, Welsh/English etc.; see e.g. De Houwer, A., 1995 for an overview and the journals *International Journal of Bilingualism, Bilingualism: Language and Cognition,* and *Bilingual Research Journal* as well as *Linguistics* for current studies). Thus, more data on multilingualism with different language combinations in different cultural settings are needed to eliminate negative myths about early bilingualism and to dispel the worries about bilingual language development that might prevent parents and education professionals from raising children with two languages (see e.g. the publications by *Multilingual Matters*). In addition, such studies could determine successful practices that foster the continued use of endangered languages rather than enhancing the existing shift towards a dominant language.

4.3 Resources

When it comes to resources for language maintenance or revitalisation projects, it is quite obvious that a good knowledge of children's lexical and grammatical abilities at particular stages is necessary for the development of age-appropriate teaching materials. Also, it is clear that curricula for language teaching might benefit from information about the order in which grammatical phenomena are typically acquired in first language acquisition. But the contribution of child language documentation and description to language teaching can go beyond these more obvious applications.

Firstly, studies suggest that it can help children to move from oral conversation to written text creation when they can transcribe their own speech (see e.g. Hidi and Klaiman (1984), Annany and Casell (2001)). Hence, audio and video recordings of children could become part of literacy programs. Secondly, at least some language teachers promote the use of children's songs or nursery rhymes in language teaching as these texts highlight the dominant stress patterns and the rhythmic properties of the language. Moreover, the embedding of linguistic activities in musical activities with children could help older learners in language revitalisation projects to overcome inhibitions in the use of a language they do not master fully. Finally, children's songs and nursery rhymes and indeed any (semi)ritualised linguistic practices involving children and adults are crucial for the intergenerational transmission of language and thus their documentation can help to support these practices.

The collection of such data requires only some basic knowledge of audio and video recording, which can be taught to members of the language community far more

easily than traditional linguistic elicitation or transcription techniques. Thus, it can become part of strategies for involving the language community in the documentation process instead of relying on a top-down approach to language planning and schooling (see e.g. Flores Farfán (2001), Austin (2004) for overviews and discussion of community involvement issues).

4.4 Implications for data collection

In the previous sections, I have tried to show how the documentation of child language can help linguists and linguistic communities to assess the endangerment status of a language, to foster a positive attitude towards the continued use of this language, and to develop resources for its maintenance. However, some of the data collection methods which acquisition researchers typically employ might need to be modified for such an enterprise. Firstly, the assessment of intergenerational transmission requires information about actual child-adult interactions in a broad variety of everyday situations – and not special 'free play' sessions between the child and a researcher or a parent, which are not typical for children's daily linguistic experience. Secondly, acquisition researchers typically focus on productive language use as they are interested in the child's developing linguistic abilities. Hence, they might not record children's songs, stories or nursery rhymes or other cultural and linguistic practices which are important for the creation of a language documentation that can inform and support language maintenance projects. Finally, the previous discussion has stressed the importance of community involvement in the documentation process as well as in the organisation and use of the resulting materials. As I have pointed out, one way of getting community members involved is to make them part of the recording process – even when this may mean investing more time for training than usual for a language acquisition research project where the researchers are doing most of the recordings themselves or with one or two assistants. Note, however, that a restriction of the sampling to a small number of extended families might mean that only comparatively few adults would have to be trained to do recordings.

Making a documentation useful for a community might also mean including more data from linguistic activities that are enjoyable to watch or listen to. This requirement clearly needs to be balanced with the needs of linguists and other language professionals who need information about a variety of linguistic elements and constructions. Elicitation games, which provide both information and a certain amount of entertainment, might provide at least a partial solution for this problem (see section 5).

5. Summary and discussion

In the preceding section I have shown how at least three types of potential users can benefit from the documentation of child language. Firstly, such documentation can help language acquisition researchers to investigate potential universals of child language and child-directed speech, the timing of acquisition processes, and the development of

grammar, lexicon, conversational competence and narrative skills. Secondly, descriptive, theoretical and historical linguists can use acquisition evidence for their studies on morphological default operations, verb classes, implicational universals and language change. The learnability issues which arise from such studies might provide a new perspective that could lead to deeper insights into linguistic representations and their development. Thirdly, the documentation of child language can help language maintenance projects to assess how endangered a language really is, to foster a positive attitude towards the continued use of this language, and to develop the resources for language maintenance and revitalisation projects.

The discussion in the previous sections has shown that this combination of benefits can only be achieved if the documentation involves participants from different social groups, age ranges and genders. As discussed above, this is easier if researchers focus on a few families with several children instead of recording individual children and their respective primary caretaker(s). Moreover, if the recording is focused on a few families it is easier to train the comparatively small number of adults in these families to do the recordings themselves. Thus, this organisation of the recording process might make it easier to involve members of the respective linguistic community in the data collection – which might make the documentation more valuable for them. The value of the documentation for the community can be further enhanced if community members are able to use the documentation for their own purposes, in particular for the creation of teaching materials and the support of cultural and linguistic practices. Such uses of a documentation require that it covers a broad range of grammatical phenomena, lexical elements, genres, and discourse types.

In this respect, the needs of community members and the needs of acquisition researchers and other linguists overlap. In addition, acquisition researchers and other linguists need data that are comparable enough to support crosslinguistic and cross-cultural studies. In order to achieve both broad coverage and crosslinguistic comparability, it is advisable to combine the four most common data types in acquisition studies and documentation projects: (i) observed communicative events, (ii) experimental data, (iii) staged communicative events, and (iv) data collected via linguistic elicitation.

5.1 Observed communicative events

In his article about documentary linguistics, Himmelmann (1998) defines communicative events as those in which external interference is limited to the fact that this event is being observed and/or recorded. Acquisition researchers typically call recordings of observed communicative events ‘spontaneous speech samples’ or ‘naturalistic data’. Despite the differences in terminology, recording spontaneous speech samples is common both in fieldwork and in acquisition studies because this data type is least likely to be affected by the observer's paradox (Labov (1972:113)), i.e. the fact that the objects of research are susceptible to change because of the ongoing research process, in particular the presence of researchers, stimuli or recording equipment. This is especially important for longitudinal studies, where stronger and

systematic observation effects can develop over time. Moreover, spontaneous speech samples can be obtained from speakers of all ages without any prior in-depth knowledge of the respective language and without having to create special stimuli or experimental items. Finally, spontaneous speech samples are required for MLU-calculations, can be analysed with respect to a broad range of phenomena, and can provide input data if children are recorded with their main communicative partners. Thus, spontaneous speech samples should be the basis of any documentation and they can also provide useful initial information for the creation of stimuli for elicitation and experiments.

However, recording procedures have to be adapted to the fact that the level of background noise tends to be louder and children are often less focused, cooperative, and well-articulated than adults (see Slobin (1967) and Demuth (1996) for many helpful suggestions). Moreover, when transcribing children's speech, researchers have to make decisions about conventions for transcribing and annotating utterances which are not target-like (e.g. utterances with baby words like *doggy* or words with omitted syllables like *(ba)nana*, and utterances with omitted grammatical morphemes like *chicken eat*.). The CHILDES webpage provides information about the CHAT format for transcription and annotation and the bibliography contains many useful references (http://childes.psy.cmu.edu/).

Even when these issues are taken into account, simply observing communicative events has its limitations. Firstly, one might not be able to obtain enough data for the study of linguistic phenomena which are comparatively rare in everyday conversations (for instance passive sentences with three-place verbs like The chicken was given some food by the farmer). Secondly, observing communicative events provides researchers with positive evidence, i.e. information about constructions that are grammatical for the speaker under study. But spontaneous speech samples do not supply negative evidence, i.e. they cannot tell researchers which constructions are ungrammatical for the speaker. After all, even if a certain type of utterance does not occur in a large spontaneous speech sample, this might simply be due to the discourse situation in the recording. Hence, the absence of particular constructions in spontaneous speech samples cannot be taken as evidence for their ungrammaticality. Thirdly, recording children (or other speakers) in routine situations might lead to an underestimation of their linguistic abilities as such situations do not require very sophisticated descriptions, negotiations or explanations that might involve more complex constructions. And fourthly, cultural practices and even the most basic everyday situations can differ widely from culture to culture – or even from family to family within the same culture. Thus, if one simply records those activities that are representative for the respective speaker or community, the communicative events that are observed might be very different from recording to recording, which makes crosslinguistic and cross-cultural comparisons difficult.

5.2 Experimental data

In (psycho)linguistic experiments, researchers carefully manipulate one or more variables (e.g. CONSTRUCTION TYPE, INFLECTION TYPE, or AGE) and measure whether any changes with respect to this variable have an effect on the speaker's or listener's behaviour. For instance, one could compare correctness rates for active and passives and for different age groups in an experiment where children have to point to pictures corresponding to active sentences (e.g. *The chicken is chasing the duck*) or passive sentences (e.g. *The chicken is being chased by the duck*). Experiments can be used to study a variety of linguistic behaviours (production, comprehension, grammaticality or truth-value judgment, imitation, etc.), and some types of experiments can provide negative evidence (see e.g. Crain and Thornton 1998, Lust et al. 1999, McDaniel et al. 1996, and Menn and Bernstein-Ratner 2000 for introductions to experimental methodology for child language studies).

Experiments involve more control, but also more external interference than the mere observation of communicative events. Thus, care must be taken to minimise unwanted task effects. More serious problems with the use of experiments result from the fact that many experiments can only be carried out with children that are at least three years old. Moreover, the design of successful experiments requires careful and systematic manipulation of linguistic variables and thus detailed prior knowledge about the respective grammatical phenomenon. Hence, experiments can only be used for the documentation of a language when detailed knowledge about the relevant construction is already available to the researcher – either from previous studies or at a later stage in a documentation project. However, it might be advisable to pilot the experimental technique and some initial stimuli at an earlier stage as this might show which modifications are necessary to adapt the technique and the materials to the research environment and the linguistic properties of the language.

5.3 Staged communicative events

Staged communicative events are communicative events which are enacted for the purpose of recording and thus do not serve any specific communicative purposes other than producing data (Himmelmann (1998)). Researchers staging such events might either give rather general instructions (e.g. 'tell me that story we talked about') or provide more specific stimuli (e.g. drawings, photographs, toys, or films). The use of elicitation stimuli is a comparatively recent development in descriptive and documentary linguistics (see Lüpke, this volume) as well as in acquisition research. Nevertheless, it is very promising for child language documentation projects.

Similarly to spontaneous speech samples, the application of elicitation games or stimuli is limited to the study of language production and can only provide positive evidence. However, the staging of communicative events can provide more crosslinguistic comparability as well as more control and richer data sets than observation of communicative events. Nevertheless, in contrast to experiments, the staging of communicative events tends to be exploratory in nature as no strict

experimental design is involved. For the fieldwork situation of many documentation projects, this can be an advantage as the staging of communicative events requires less detailed prior knowledge about the constructions of the language than the design of linguistic experiments.

Moreover, many elicitation stimuli are freely available and have been used for a large number of languages so that crosslinguistic comparisons are possible. Among the best known stimuli for children is the so-called "frog-story" stimulus, a word-less picture book that has been used in studies on narrative development (Berman and Slobin 1994, Strömqvist and Verhoeven 2004). Other stimuli that can be used with children are the "Fish Film" videos by Tomlin (1995, 1997), the "Pear Story" developed by Chafe (1980), the stimuli from the LEXLERN-project (Eisenbeiss et al. 1994), and many of the stimuli which can be found in the field manuals of the Language-and-Cognition group of the MPI for Psycholinguistics in Nijmegen (see http://www.mpi.nl).

5.4 Data collected via linguistic elicitation

Linguistic elicitation is a type of communicative event that has been invented for linguistic research and documentation, and three types of elicitation can be distinguished: (i) *contextualising elicitation*, where native speakers are asked to comment on or provide contexts for a given word or construction, (ii) *translation*, where native speakers are asked to translate a given form into their native language, and (iii) *judgment*, where native speakers are asked to evaluate the acceptability or grammaticality of a given form or construction (Himmelmann (1998)). The methodology of elicitation is still a matter of intense debate (see e.g. Matthewson (2004) and the references cited there for discussion). Moreover, linguistic elicitation requires the speaker under study to possess well-developed meta-linguistic abilities and is thus not easily applicable for child language studies. However, it might be possible to adapt some of these techniques for exploratory studies with children.

6. Conclusion

In summary, documentation of child language enables language acquisition researchers to investigate potential universals of child language and child-directed speech, the timing of acquisition processes, and the development of grammar, lexicon, conversational competence and narrative skills. It provides descriptive, theoretical and historical linguists with additional evidence for their studies on morphological default operations, verb classes, implicational universals and language change, and the learnability issues which arise from such studies might provide a new perspective that could lead to deeper insights into linguistic representations and their development. Moreover, the documentation of child language can help language maintenance projects to assess how endangered a language really is, to foster a positive attitude towards the continued use of this language and to develop the resources for language maintenance and revitalisation projects. In order for these benefits to be achieved, any

documentation of child language should involve children from different social groups, age ranges and genders as well as their parents, caretakers, siblings and peers. At the same time it should combine spontaneous speech data, experimental data, staged communicative events, and elicited data to capture a broad range of linguistic phenomena and the full range of children's linguistic abilities. Most importantly, however, the documentation of child language should be integrated into larger documentation projects and resources should be made easily accessible beyond the domain of language acquisition researchers. First steps in this direction have been made, see e.g. current developments in the CHILDES project (http://childes.psy.cmu.edu/), the collaboration of the Language Acquisition Group of the MPI for Psycholinguistics in Nijmegen and the DOBES project (http://www.mpi.nl) or Lust et al. (2005); but we still need to develop better common standards for elicitation, data collection, recording and archiving techniques and feasible solutions for the ethical, legal, and logistic problems that all language documentation projects have to deal with – whether they focus on adult or child speakers.

7. Acknowledgements

This paper could not have been written without the discussions I had with fieldworkers and acquisition researchers during my time at the MPI for Psycholinguistics in Nijmegen. For detailed comments, I would like to thank Peter Austin and the audience at the SOAS workshop on multidisciplinary approaches to language documentation, Juergen Bohnemeyer, Birgit Hellwig, Bettina Landgraf, Friederike Luepcke, and Ingrid Sonnenstuhl.

8. References

Allen, Shanley E. M. 1994. Acquisition of Some Mechanisms of Transitivity Alternation in Arctic Quebec Inuktitut. Doctoral Dissertation. Montreal: McGill University.

Allen, Shanley E. M. and Martha B. Crago. 1993. The Acquisition of Passives and Unaccusatives in Inuktitut. McGill *Working Papers in Linguistics/Cahiers Linguistiques de McGill 9,* 1-29. Montreal: McGill University.

Allen, Shanley E. M. and Martha B. Crago. 1996. Early Passive Acquisition in Inuktitut. *Journal of Child Language* 23: 129-155.

Ananny, Mike and Justine Cassell. 2001. TellTale: A Toy to Encourage Written Literacy Skills through Oral Storytelling. Paper presented at the Conference on Text, Discourse and Cognition, Jackson Hole, WY. Downloadable from: http://www.media.mit.edu/~ananny/papers/tdc2001paper.PDF

Austin, Peter K. (ed.) 2004. *Language Documentation and Description Volume 2*. London: SOAS.

Bates, Elizabeth., Philip S. Dale and Donna Thal. 1995. Individual Differences and their Implications for Theories of Language Development. In Paul Fletcher and Brian MacWhinney (eds.) *The Handbook of Child Language*, 96-151. London: Blackwell.

Berman, Ruth and Dan I. Slobin. 1994. *Relating Events in Narrative: A crosslinguistic Developmental Study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bertolo, Stefano (ed.) 2001. *Language Acquisition and Learnability*. Cambridge: Cambridge University Press.

Blake, Barry J. 2001. *Case*. Cambridge: Cambridge University Press.

Borer, Hagit and Kenneth Wexler. 1987. The Maturation of Syntax. In Thomas Roeper and Edwin Williams (eds.) *Parameter Setting*, 123-172. Dordrecht: Reidel.

Borer, Hagit and Kenneth Wexler. 1992. Bi-unique relations and the maturation of grammatical principles. *Natural Language and Linguistic Theory 10*, 147-189.

Bowerman, Melissa. 1978. The acquisition of word meaning: an investigation into some current conflicts. In Natalie Waterson and Catherine. Snow (eds.) *The development of communication*, 263-287. New York: Wiley.

Bowerman, Melissa, Penelope Brown, Sonja Eisenbeiss, Bhuvana Narasimhan and Dan I. Slobin. 2003. Putting Things in Places: Developmental Consequences of Linguistic Typology. Proceedings of the Child Language Research Forum 2002, Stanford University. Downloadable from: http://cslipublications.stanford.edu/CLRF/2002/CLRF-2002-title.html

Brown, Penelope. 2001. Learning to Talk about Motion UP and DOWN in Tzeltal: Is there a Language-Specific Bias for Verb Learning? In Melissa Bowerman and Steven C.Levinson (eds.) *Language Acquisition and Conceptual Development*, 512-543. Cambridge: Cambridge University Press.

Brown, Roger 1973. *A First Language: The Early Stages*. London: Allen and Unwin.

Brown, Roger and Camille Hanlon. 1970. Derivational complexity and order of acquisition in child speech. In John R. Hayes (ed.) *Cognition and the Development of Language*, 11-53. New York, NY: Wiley.

Chafe, Wallace (ed.) 1980. *The Pear Stories*. New York, NY: Able.

Choi, Soonja and Melissa Bowerman. 1991. Learning to Express Motion Events in English and Korean: The Influence of Language-Specific Lexicalisation Patterns. *Cognition 41*, 83-121.

Choi, Soonja and Alison Gopnik. 1995. Early Acquisition of Verbs in Korean: A Crosslinguistic Study. *Journal of Child Language 223*: 497-525.

Chomsky, Noam. 1966. *Cartesian Linguistics*. New York, NY: Harper and Row.

Clahsen, Harald 1999. *Lexical Entries and Rules of Language: A Multi-Disciplinary Study of German Inflection. Behavioral and Brain Sciences 22*: 991-1013.

Clark, Eve V. 1993. *The Lexicon in Acquisition*. Cambridge: Cambridge University Press.

Coates, Jennifer (ed.) 1998. *Language and Gender: A Reader*. Oxford: Blackwell.

Cook, Vivian. (ed.) 2003. *Effects of the Second Language on the First*. Clevedon: Multilingual Matters.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.

Crain, Stephen 1991. Language Acquisition in the Absence of Experience. *Behavioral and Brain Sciences 14*: 597-611.

Crain, Stephen, and Rosalind Thornton. 1998. *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax*. Cambridge, MA.: MIT Press.

Cummins, Jim. 1989 *Empowering Minority Students*. Ontario, CA: California Association for Bilingual Education.

Cummins, Jim 1992 Bilingual Education and English Immersion: The Ramirez Report in Theoretical Perspective. *Bilingual Research Journal 16* 1 and 2, 91–104.

Cummins, Jim 2000 *Language, Power and Pedagogy: Bilingual Children in the Crossfire*. Clevedon: MultilingualMatters.

DeHouwer, Annick. 1995. Bilingual Language Acquisition. In Paul Fletcher and Brian MacWhinney (eds.) *The Handbook of Child Language,* 219-250. London: Blackwell.

Demuth, Katherine 1989. Maturation and the Acquisition of the Sesotho Passive. *Language 65*: 56-80.

Demuth, Katherine 1990. Subject, Topic, and Sesotho Passive. Journal of Child Language 17: 67-84.

Demuth, Katherine 1992. Accessing Functional Categories in Sesotho: Interactions at the Morpho-Syntax Interface. In Jürgen M. Meisel (ed.) *The Acquisition of Verb Placement: Functional Categories and V2 Phenomena in Language Development, 83-107*. Dordrecht: Kluwer.

Demuth, Katherine 1994. On the 'Underspecification' of functional Categories in Early Grammars. In Barbara Lust, Margarita Suner and John Whitman (eds.) *Syntactic*

Theory and First Language Acquisition: Cross-Linguistic Perspectives Volume 1, 119-134. Hillsdale, NJ: Erlbaum.

Demuth, Katherine 1996. Collecting spontaneous production data. In Dana McDaniel, Cecile McKee and Helen Smith Cairns (eds.) *Methods for Assessing Children's Syntax,* 3–22. Cambridge, MA: MIT Press.

Dixon, Robert M. W. (ed.) 1987. *Studies in Ergativity*. Amsterdam: North Holland.

Dolson, David P. and Jan Mayer. 1992. Longitudinal Study of Three Program Models for Language-Minority Students: A Critical Examination of Reported Findings. *Bilingual Research Journal 16*, 105–157.

Sonja Eisenbeiss, Susanne Bartke, Helga Weyerts and Harald Clahsen (eds.) 1994. Elizitationsverfahren in der Spracherwerbsforschung: Nominalphrasen, Kasus, Plural, Partizipien. Arbeiten des Sonderforschungsbereichs 282 57. Düsseldorf: Heinrich-Heine-Universität.

Eisenbeiss, Sonja. 2003 Merkmalsgesteuerter Grammatikerwerb. Eine Untersuchung zum Erwerb der Struktur und Flexion der Nominalphrase 'Feature-Driven Structure Building. A Study on the Acquisition of Noun-Phrase Structure and Inflection', Doctoral Dissertation, University of Düsseldorf. Downloadable from:
http://www.ub.uni-duesseldorf.de/ebib/diss/liste_jahr?fak=4&year=2003

Eisenbeiss, Sonja. 2005. Generative Approaches to Language Learning. Ms. University of Essex.

Eisenbeiss, Sonja, Susanne Bartke and Harald Clahsen (accepted) Structural and Lexical Case in Child German: Evidence from Language-Impaired and Typically-Developing Children. *Language Acquisition.*

Eisenberg, Sarita. L., Tara McGovern Fersko and Cheryl Lundgren. 2001. The Use of MLU for Identifying Language Impairment in Preschool Children: A Review, *American Journal of Speech-Language Pathology,* 10: 323–342.

Fisher, Cynthia. 1999. Simple Structural Guides for Verb Learning: On Starting with Next to Nothing. In Eve V. Clark (ed.) *Proceedings of the 30th Stanford Child Language Research Forum*. Stanford, CA: CSLI Publications.

Fisher, Cynthia. 2002. Structural Limits on Verb Mapping: The Role of Abstract Structure in 2.5-Year-Olds' Interpretations of Novel Verbs. *Developmental Science* 5: 55-64.

Fisher, Cynthia., Geoffrey Hall, Susan Rakowitz and Lila R. Gleitman. 1994. When it is Better to Receive than to Give: Syntactic and Conceptual Constraints on Vocabulary Growth, *Lingua* 92:333-375.

Fishman, Joshua A. 1991. *Reversing language shift*. Clevedon: Multilingual Matters.

Fletcher, Paul and Brian MacWhinney (eds.) 1995. *The Handbook of Child Language*. Oxford: Blackwell.

Flores Farfán, José A. 2001. Culture and Language Revitalisation, Maintenance, and Development in Mexico. *The Nahua Alto Balsas Communities, International Journal of the Sociology of Language* 152: 185-197.

Fodor, Jerry A., Thomas G. Bever and Merrill F. Garrett. 1974. *The Psychology of Language: An Introduction to Psycholinguistics and Generative Grammar*. New York, NY: McGraw-Hill.

Fortescue, Michael D. and Lise Lennert Olsen. 1992. The Acquisition of West Greenlandic. In Dan I. Slobin (ed.) *The Crosslinguistic Study of Language Acquisition* Volume 3, 111-219. Hillsdale, NJ: Lawrence Erlbaum.

Gamerschlag, Thomas 1996. Kasus, Alternationen und Argumentlinking im Japanischen. *Arbeiten des Sonderforschungsbereichs* 282 80. Düsseldorf: Heinrich-Heine-Universität.

Gentner, Dedre. 1982. Why Nouns are Learned before Verbs: Linguistic Relativity versus Natural Partitioning. In Stan A. Kuczaj (ed.), *Language development: language, cognition and culture*, 301-334. Hillsdale, NJ: Erlbaum.

Gleitman, Lila R. 1990. The Structural Sources of Verb Meanings. *Language Acquisition* 1, 3-55.

Greenberg, Joseph H. Ed. 1963. *Universals of Language*. Cambridge, MA: MIT Press.

Grimm, Jacob. 1864. *Kleinere Schriften von Jacob Grimm Volume 1*. Berlin: Dümmler.

Hidi, Suzanne, and Roslyn Klaiman. 1984. Children's Written Dialogues: Intermediary Between Conversation and Written Text? In Anthony Pellegrini and Thomas Yawkey (ed.) *The Development of Oral and Written Language in Social Contexts,* 233-241. Norwood, NJ: Ablex.

Himmelmann, Nikolaus. 1998. Documentary and Descriptive Linguistics. *Linguistics*, 36:161-195.

Ingram, David. 1989. *First Language Acquisition: Method, Description and Explanation*. Cambridge: Cambridge University Press.

Jakobson, Roman 1942. *Kindersprache, Aphasie und allgemeine Lautgesetze*. Reprinted 1969. Frankfurt am Main: Suhrkamp.

Joppen, Sandra and Dieter Wunderlich. 1995. Argument linking in Basque, *Lingua* 97: 123-169.

Kachru, Yamuna 1987. Ergativity, Subjecthood and Topicality in Hindi-Úrdu, *Lingua* 71: 223-238.

Kachru, Yamuna and Rajeshwari Pandharipande. 1978. On Ergativity in Selected South Asian Languages, *Studies in the Linguistic Sciences* 8: 111-127.

Kroch, Anthony. 2001. Syntactic change. In Mark Baltin and Chris Collins (eds.) *The Handbook of Contemporary Syntactic Theory*, 699–729. Oxford: Blackwell.

Küntay, Aylin and Dan I. Slobin. 1996. *Listening to a Turkish Mother: Some Puzzles for Acquisition*. In Dan I. Slobin, Julie Gerhardt, Amy Kyratzis and Jiansheng Guo (eds.) Social Interaction, Social Context, and Language, 265-286. Mahwah, NJ: Erlbaum.

Labov, William. 1972. Some Principles of Linguistic Methodology, *Language in Society* 1: 97–120.

Lehmann, Christian. 2001. Language Documentation: A Program. In Walter Bisang (ed.) *Aspects of Typology and Universals*, 83-97. Berlin: Akademie Verlag.

Levin, Beth. 1993. *English Verb Classes and Alternations, a Preliminary Investigation. Chicago and London*: The University of Chicago Press.

Lightfoot, David. 1991. *How to set parameters: Arguments from language change*. Cambridge, MA: MIT Press.

Locke, John L. 1995. Development of the Capacity for Spoken Language. In Paul Fletcher and Brian MacWhinney (eds.) *The Handbook of Child Language*, 278-302. Oxford: Blackwell.

Lust, Barbara, Suzanne Flynn, Claire Foley and Yu-Chin Chien. 1999. How do we Know what Children Know: Problems and Advances in Establishing Scientific Methods for the Study of Language Acquisition and Linguistic Theory. In Ritchie, William C. and Tej K. Bhatia (eds.) *Handbook of Child Language Acquisition*, 427–56. San Diego, CA: Academic Press.

Lust, Barbara, Elaine Westbrooks, Suzanne Flynn and Theresa Tobin. (2005) Developing Adequate Documentation for Multifaceted Cross-Linguistic Language Acquisition Data. Paper presented at the Workshop on Language Documentation: Theory, Practice, and Values, Harvard University, Cambridge, MA.

MacWhinney, Brian. 2000. *The CHILDES Project: Tools for Analyzing Talk*. Third Edition. Mahwah, NJ: Lawrence Erlbaum.

Marcus, Gary F. 1993. Negative Evidence in Language Acquisition? *Cognition* 46: 53-85.

Matthewson, Lisa. 2004. On the Methodology of Semantic Fieldwork, *International Journal of American Linguistics* 70/4: 369-415

McDaniel, Dana, Cecile McKee and Helen Smith Cairns (eds.) 1996. *Methods for Assessing Children's Syntax*. Cambridge, MA: MIT Press.

Menn, Lise and Nan Bernstein Ratner (eds.) 2000. *Methods for Studying Language Production.* Mahwah, NJ: Erlbaum.

Miller, George A. and Noam Chomsky. 1963. Finitary Models of Language Users. In R. Duncan Luce, Robert R. Bush and Eugene Galanter (eds.) *Handbook of Mathematical Psychology Volume 2*, 419-491. New York, NY: Wiley.

Naigles, Letitia. 1990. Children Use Syntax to Learn Verb Meanings, *Journal of Child Language* 17: 357–374.

Naigles, Letitia. 1996. The Use of Multiple Frames in Verb Learning via Syntactic Bootstrapping, *Cognition* 58: 221-251.

Narasimhan, Bhuvana. (accepted, a) Case-Marking and Transitivity in Early Child Hindi. *Proceedings of the 2004 South Asian Languages Analysis Roundtable.*

Narasimhan, Bhuvana (accepted, b) Splitting the Notion of 'Agent': Case-Marking in Early Child Hindi, *Journal of Child Language.*

Ochs, Elinore and Bambi B. Schieffelin. 1986. *Language Socialisation across Cultures*. New York: Cambridge University Press.

Ochs, Elinor and Bambi B. Schieffelin. 1995. The Impact of Language Socialisation on Grammatical Development. In Paul Fletcher and Brian MacWhinney (eds.) *The Handbook of Child Language*, 73-94. Oxford: Blackwell.

Peters, Ann M. 2001a. Filler syllables: What is their Status In Emerging Grammar? *Journal of Child Language* 28: 229-242.

Peters, Ann M. 2001b. Response to Comments, *Journal of Child Language* 28: 283-289.

Pinker, Steven. 1982. A Theory of the Acquisition of Lexical Interpretive Grammars. In Joan Bresnan (ed.) *The Mental Representation of Grammatical Relations*, 655-726. Cambridge, MA: MIT Press.

Pinker, Steven 1984. *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

Pinker, Steven 1989. Learnability and Cognition: *The Acquisition of Argument Structure.* Cambridge, MA: MIT Press.

Pye, Clifton L. and P. Quixtan Poz. 1988. Precocious Passives and Antipassives in Quiche Mayan. *Papers and Reports on Child Language Development* 27, 71-80.

Ritchie, William C. and Tej K. Bhatia (eds.) 1999. *Handbook of Child Language Acquisition*. San Diego, CA: Academic Press.

Salminen, Tapani. 1999. UNESCO Red Book On Endangered Languages. Europe. Downloadable from: http://www.helsinki.fi/~tasalmin/europe_index.html.

Slobin, Dan I. (ed.) 1967. *A Field Manual for Cross-Cultural Study of the Acquisition of Communicative Competence.* Berkeley, CA: Language-Behavior Research Laboratory.

Slobin, Dan I. (ed.) 1985. *The Crosslinguistic Study of Language Acquisition* Volume 2. Hillsdale, NJ: Erlbaum.

Slobin, Dan I. (ed.) 1992. *The Crosslinguistic Study of Language Acquisition.* Volume 3. Hillsdale, NJ: Erlbaum.

Slobin, Dan I. (ed.) 1997. *The Crosslinguistic Study of Language Acquisition*. Volume 5. Mahwah, NJ: Erlbaum.

Slobin, Dan I. 2004. The Many Ways to Search for a Frog: Linguistic typology and the Expression of Motion Events. In Strömqvist, Sven and Ludo Verhoeven (eds.) *Relating events in narrative Volume 2: Typological and Contextual Perspectives*, 219-257. Mahwah, NJ: Lawrence Erlbaum.

Sokolov, Jeffrey L. and Catherine Snow (eds.) 1994. *Handbook of Research in Language Development Using CHILDES*. Hillsdale, NJ: Erlbaum

Stern, Clara and William Stern. ([1]1907, [4]1928) *Die Kindersprache*. Leipzig: Barth.

Strömqvist, Sven and Ludo Verhoeven (eds.) 2004. *Relating events in narrative: Typological and Contextual Perspectives*. Mahwah, NJ: Lawrence Erlbaum.

Suzman, Susan M. 1985. Learning the Passive in Zulu. *Papers and Reports on Child Language Development* 24: 131-137.

Talmy, Leonard. 1985. Lexicalisation Patterns: Semantic Structure in Lexical Forms. In Timothy Shopen (ed.) *Language Typology and Syntactic Description Volume 3*, 57-149. Cambridge: Cambridge University Press.

Tardif, Twila. 1996. Nouns are not always Learned before Verbs: Evidence from Mandarin Speakers' Early Vocabulary, *Developmental Psychology* 32: 492-504.

Tomasello, Michael. 2003. *Constructing a Language: a Usage-Based Theory of Language Acquisition*. Cambridge: Harvard University Press.

Thomason, Sarah Grey and Terrence Kaufman. 1988. *Language Contact, Creolisation and Genetic Linguistics*. Berkeley, CA: University of California Press.

Tomlin, Russell S. 1995. Focal Attention, Voice, and Word Order: an Experimental, Crosslinguistic Study. In Pamela Downing and Michael Noonan (eds.) *Word order in discourse*, 517-554. Amsterdam: John Benjamins.

Tomlin, Russell S. 1997. Mapping Conceptual Representations into Linguistic Representations: The Role of Attention in Grammar. In Jan Nuyts and Eric Pederson (eds.) *Language and conceptualisation*, 162-189. Cambridge: Cambridge University Press.

Tsunoda, Tasaku. 2001. Role and Ethics of Researchers and Method of Documentation. In Sakiyama, Osamu (ed.) *Lectures on Endangered Languages Volume 2*. From Kyoto Conference 2000, 261-268. Kyoto: Osaka Gakuin University Endangered Languages of the Pacific Rim Publication Series, C002.

van Hout, Angelique. 1996. Verb Frame Alternations and Event Semantics: A Case Study of Dutch and its Acquisition. Tilburg Dissertation in Language Studies, Katholieke Universiteit Brabant, Tilburg.

Van Valin, Robert D. 1992. An Overview of Ergative Phenomena and their Implications for Language Acquisition. In Dan I.Slobin (ed.) *The Cross-Linguistic Study of Language Acquisition* Volume 3, 15-37. Hillsdale, NJ: Erlbaum.

Woodbury, Anthony. 2003. Defining Documentary Linguistics. In Peter K. Austin (ed.), *Language Documentation and Description Volume 1*, 35–51. London: SOAS.

Wunderlich, Dieter 1997. Cause and the structure of verbs, *Linguistic Inquiry* 28: 27-68.