To appear in: Bavin, E. & S. Stoll (eds.) *The acquisition of ergativity.* Amsterdam: Benjamins.

The acquisition of ergative case in Chintang*

Sabine Stoll & Balthasar Bickel

University of Zürich

1 Introduction

One of the earliest and most important challenges for language learners is to find out what role arguments play in their language, i.e. how does their language express 'who is doing what to whom?'. The key problem is how thematic roles (agents, patients, themes etc.) are linked to syntactic elements (noun phrases) and morphological markers (case and agreement markers). While there is disagreement as to the actual point in time when these linking patterns become established and when children have acquired the corresponding syntactic abstractions (Gertner et al. 2006, Dittmar et al. 2008), there is a general consensus that these are all relatively early achievements, at least in languages like English and German. These findings are surprising given the fact that the expression of arguments is an exceedingly complex phenomenon: there are often intricate role constellations in experiencer verbs (cf. e.g. *I am afraid of this* vs. *I fear this*), and there is substantial cross-linguistic diversity: unlike English and German, some languages mark an agent like *I* in *I worked* systematically differently from the agent in *I wrote a paper* and again differently from the agent in *I saw the paper.* In this paper we focus in particular on the challenge from cross-linguistic diversity.

The acquisition of argument linking patterns is traditionally accounted for in either of two types of theories, nativist theories that assume innate linking rules and usage-based theories that rely on general cognitive ablities such as pattern matching, imitation and generalization. These theories make different predictions on how cross-linguistic diversity plays out in acquisition.

* This research was supported by two grants from the Volkswagen Foundation (Grant No. II/79092, 2004-2009, PI Bickel; and Grant No. II/81 730, 2007-2012, PI Stoll) for research on Chintang, and two grants from the German Research Foundation for research on grammatical relations in Chintang and world-wide (Grant No. BI 799/5-1, 2009-2012; and Grant No. BI 799/3-1/2, 2006-2013; both PI Bickel). Additional support was provided by the Max Planck Society and the University of Zürich. Author contributions: Stoll designed the study and wrote the paper; Bickel contributed linguistic and statistical analyses. We thank the team of the Chintang Language Research Program (`http://www.spw.uzh.ch/clrp`) for their support in data preparation.

Within nativist theories, Pinker (1984) for example claims that children learn argument structure with the help of innate linking rules. These linking rules determine how thematic roles are mapped into syntactic expressions. According to this theory, the rules are embedded in the semantics of argument structure: children are hypothesized to always first focus on the most agentive argument in a sentence and link this argument to the subject of the sentence. Usually, the most agentive argument is the agent in a transitive construction (*I wrote a paper*), but the theory assumes that this role is generalized to agents in intransitive constructions as well (*I worked*) and further on to other intransitive arguments, which are less proto-typical agents (*I walked*, *I sat down*, *I slept* etc.). As a result, the theory predicts that all these argeuments are treated alike, and this matches with what is known as 'accusative' alignment: the sole argument *S* of a one-argument ("intransitive") verb is treated like the most agentive argument *A* of a two- or three-argument ("transitive" or "ditransitive") verb, in short *S*=*A*. English and German fit this pattern, but the question arises what an innate agent-based semantic generalization strategy predicts for a language that does not treat S and A alike, i.e. *S*≠*A* or what is traditionally called 'ergative alignment'. The theory predicts that when acquiring a language with such an ergative pattern, children will tend to overgeneralize the marker on A (the 'ergative') to at least the most agentive S arguments. Since an ergative pattern does not allow this, the result would be a case error. Interestingly, all available studies on the acquisition of ergative case markers so far contradict this hypothesis. Research on languages like Samoan (Ochs 1982), Kaluli (Schieffelin 1985), Basque (Ezeizabarrena & Larranaga 1996) or Hindi (Narasimhan 2005) revealed no evidence for an overgeneralization of ergative markers to agentive S arguments.

Usage-based theories take a radically different point of departure. Instead of assuming innate and universal learning strategies like Pinker's agent-based linking rules, these theories hypothesize that acquisition is driven by matching patterns in the speech they hear, gradually generalizing these patterns into abstract syntactic knowledge by probabilistic extrapolation (Tomasello 2003). For the acquisition of argument linking, these theories predict that children do not overgeneralize ergative markers to agentive *S* arguments simply because they hear no patterns of this kind in the input. In this regard, usage-based theories are more compatible with previous findings on the acquisition of ergative markers. However, usage-based theories also predict that children's use of noun phrases and their case marking closely mirrors what they hear in speech around them from early on. In order to test this, we need to systematically compare the behavior of the children to the use of case markers in their surrounding speech. This has not been done systematically on sufficiently large corpora of acquisition data and the present paper aims to fill this gap for one ergative language, Chintang (Sino-Tibetan, Nepal). In a longitudinal corpus of four children we first test whether the distributions found in the speech of the target children corresponds to those in their respective surrounding speech. Second we develop an account of how children generalize their patterns to the case marking rules of the adult language, based exclusively on general cognitive principles.

We proceed as follows: First we briefly sketch those aspects of Chintang grammar that are relevant for learning ergative case marking: verbs, agreement patterns and nouns (Section 2). We then introduce a series of hypotheses on how children can be assumed to learn ergative case assignment on the basis of current theories (Section 3). After explaining our data recording procedure and surveying the corpus (Section 4), we subject these hypotheses to em-

pirical testing in our corpus. We first measure the distributions of ergative case marking in the child-surrounding speech separately for each child and then correlate these distributions with the distributions found in the target children (Section 5). We then explore what these distributions tell us and how children actually learn the ergative. In Sections 6-8 we test three specific hypotheses on learning strategies: (i) functional learning based on semantic restrictions, (ii) lexically-restricted learning and (iii) interactive learning, i.e. learning based on the immediately preceeding discourse. Section 9 discusses general conclusions from these studies.

2 Chintang: grammatical background

Chintang is spoken on one of the lower foothills of Eastern Nepal by about 6000 speakers. All Chintang speakers are bilingual with Nepali, the Indo-European *lingua franca* of Nepal. Some speakers are trilingual with Bantawa as a third language, another Sino-Tibetan language of the region. Most children of the community still learn Chintang as their first language at home, and it is also the dominant language spoken in most homes, and definitely in the homes of the children we recorded in our study. However, all children are surrounded by Nepali from early on because parts of the Chintang community come from other ethnic subgroups that do not speak Chintang and resort to Nepali for daily interaction. Nepali also enjoys considerable prestige in the community and is considered a necessity for economic success. It is the sole medium of instruction in school.

2.1 Verbs and verbal agreement

Chintang verb morphology is polysynthetic, and there is obligatory inflectional marking of tense, aspect, mood and polarity. Verbs agree with one or two arguments and the morphology for this distinguishes between singular, plural and dual. In the first person plural and dual there is an additional division into 'us' including vs. excluding the addressee. Most of the complexity comes from the sheer number of possible category combinations, the long strings of morphemes and a highly intricate prosody/morphology interface that results in variable positioning of agreement markers (Bickel et al. 2007). In our corpus we found over 1,800 different verb forms. Thus, the task of the child in learning verb morphology is very demanding. Since a child encounters verb forms with different lexical roots, it will take some time before they realize the paradigmatic unity behind the forms and can start to productively produce their own forms. Given these formal complexities, we expect that children start becoming productive users of verb morphology relatively late, only at around age 3. This expectation was supported in an earlier study (Stoll et al. 2012a).

Verb agreement systematically distinguishes between *S* and *A* arguments as well as an object argument *O*. *A* is the most agentive argument of two-argument verbs; for three-argument verbs (like 'give', 'put', 'cover' etc.), the choice between themes and goals depends on the lexical root (Bickel et al. 2010). While all three argument roles *S*, *A* and *O* trigger distinct agreement, individual agreement markers align argument roles in specific ways depending on the person and number categories involved. For example, third person singular is expressed by different

agreement markers in all three roles (*S≠A≠P*) when we compare one-argument forms with two-argument forms involving a third person singular together with a second person argument:[1]

(1) a. ***Ø**-ep-no*
3s[S]-stand.up-NPST
'**S/he (S)** will stand up.'

b. ***u**-tup-ma-ʔã*
3[s]A-meet-1sO-NPST
'**S/he (A)** will meet me (O).'

c. ***na**-tup-no.*
3[s]A>2[s]O-meet-NPST
'**S/he (A)** will meet you (O).'

d. *tub-**u**-ku-ŋ*
meet-3[s]O-NPST-1sA
'I (A) will meet **him/her (O)**.'

e. *a-tub-**o**-ko*
2[sA]-meet-3[s]O-NPST
'You (A) will meet **him/her (O)**'

But when compared to transitive constructions with third person singular arguments in all roles, the expression of third person singular follows an accusative alignment, with a marker *-u* limited to P arguments; S and A are unmarked:

(2) a. ***Ø**-ep-no* (cf. 1a)
3s[S]-stand.up-NPST
'**S/he (S)** will stand up.'

b. ***Ø**-tub-**o**-ko.*
3s[A]-meet-3[s]O-NPST
'**S/he (A)** will meet **him/her (O)**.'

This constitues a trace of 'accusative' (*S=A≠O*) alignment. The opposite of this, ergative (*S=O≠A*) alignment is also found in verb agreement. For example, first person singular has the same markers for *S* roles as for *O* roles in all scenarios:

(3) a. *ep-**ma**-ʔã*
stand.up-1sS-NPST
'**I (S)** will stand up.'

b. *u-tup-**ma**-ʔã*
3[s]A-meet-1sO-NPST
'S/he will meet **me (O)**.'

c. *a-tup-**ma**-ʔã*
2[sA]-meet-1sO-NPST
'You will meet **me (O)**.'

d. *tub-u-ku-**ŋ***
meet-3[s]O-NPST-1sA
'**I (A)** will meet him/her (O).'

e. *tup-**na**-ʔã*
meet-1[s]A>2[s]O-NPST
'**I (A)** will meet you (O).'

[1] The empty set symbol represents morphological zeros, i.e. cases where a specific meaning is strictly entailed by the absence of any other available marker in a particular position. Interlinear glossing abbreviations follow the Leipzig Glossing Rules. Meaning components in square brackets are entailed by the overall combination of morphemes in a string but are not encoded in any given morpheme *per se*. For full paradigms, see Bickel et al. (2007) and the language documentation deposited in the DoBeS Archive accessible directly at `http://www.mpi.nl/dobes` or, with additional information, via the Chintang Language Research Program website.

The last example (3e) suggests that in some scenarios, *A* and *O* are coded by the same marker, distinct from the marking of *S*. This corresponds to what is sometimes called 'horizontal alignment' (*S≠A=O*). A final logical possibility is for all three roles to receive the same marking (*S=A=P*). This is also attested, but only for the marking of dual number in some scenarios:

(4) a. *a-ep-**ce**-ke*
2[S]-stand.up-d-NPST
'**You two (S)** will stand up.'

b. ***a**-tup-**ce**-o-ko* (> *atupcoko*)
2[A]-meet-3[s]O-NPST
'**You two (A)** will meet him/her (O).'

c. *na-tup-**ce**-ke*
3A>3[s]O-meet-d-NPST
'S/he/they will meet **you two (O)**.'

Computing the alignments for all (overt) markers in all person and number scenarios and across all tense, aspect and polarity paradigms suggests 60% S=A≠O, 17% S=O≠A, 15% S≠A=O, 6% S≠A≠P and 2% S=A=P alignments. If one also considers all zeros, there are many more cases of S=A=P alignment.

From an acquisitional point of view, this extreme diversity of alignment types suggests that verb agreement is of no help whatsoever in predicting or acquiring case marking. Unlike verb agreement, Chintang case marking is, as we will see in the next section, mostly based on an ergative *S=O≠A* pattern. This is radically different from case-marking languages like German or Russian where the pattern found in verb agreement shows virtually the same alignment pattern as case marking. Both are thoroughly accusative: verbs consistently register *S* and *A* and never *O* arguments and, likewise, the nominative case marking covers *S* and *A* and never *O*.[2]

2.2 Nouns and case marking of core arguments

Chintang case morphology is much less complex than verb morphology but compared to Indo-European languages there are many more forms a child has to learn. Nouns inflect for eleven cases and three numbers (singular, plural (inclusive, exclusive) and dual). In addition, nouns show agreement in person and number with any possessor that there may be, and some nouns obligatorily require possessors (e.g. kin terms like *-pa* 'father' or topological nouns like *-cik* 'side').

However, noun phrases in Chintang are always optional and the verb forms presented in the preceding section all constitute complete sentences. For example, a sentence like *tuboko* in (2b) would typically refer to previously mentioned referents and, unlike in English, these do not need to be overtly expressed by pronouns. The same sentence can even be used without any previous context, meaning something like 'someone met someone', i.e. there was a meeting between two

[2] barring possible exceptions in experiencer verbs like *mir gefällt es* 'to-me is.pleasent it', i.e. 'I like it', if the experiencer is analyzed as an A argument. But note that even here, case goes hand in hand with agreement: if the experiencer is not in the nominative, it no longer triggers agreement.

singular referents but the speaker is not interested in making any further specifications. This makes Chintang discourse even more 'elliptical' or more radically 'pro-drop' than, say, Italian, where pronouns are typically dropped only if more specific reference is already established. The scarcity of noun phrases manifests itself in a very low noun-to-verb ratio in adult speech (Stoll et al. 2012a) and poses the question of how children can acquire case morphology. We will return to this issue below, but before this, we discuss the nature of the morphology.

The basic case opposition for core arguments in Chintang is between an unmarked nominative and an overt ergative marker in *-ŋa*, but there are three complications that raise the acquisitional challenge. First, the ergative does not occur with all persons to the same extent. The ergative is obligatory with third person noun phrases of any kind. It is relatively frequent with first and second person plural forms but very rare with all other pronouns (Schikowski et al. in press). Possibly as a result of phonological merger (haplology), the ergative is ungrammatical with exlusive forms (first person plural and dual exclusive: *anaŋa* and *ancaŋa*, respectively).

A second complication stems from the syntax of the marker. The basic function is to mark noun phrases in the A role of two- or three-argument predicates, such as the following:

(5) a. *wa-**ŋa*** *∅-co-hatt-u-ce.* CLLDCh2R03S03.0828
hen-**ERG** [3sA]-eat.vt-COMPLETIVE-3O-nsO
'The hen (A) will eat them all up' (Adult speaker)

b. *kubi-**ŋa*** *u-khutt-a-ŋ=kha* CLLDCh1R03S02.0306
a.person-**ERG** 3[s]A-bring.sth.for.sb.vt-PST-1sO=FOC
'Kubi (A) brought it for me!' (Adult speaker)

However, the same ergative ending is also used for encoding instruments, causes and sources:

(6) a. instrument
*ba labar-**ŋa*** *kam-u-ku-ŋ=ta, them?* CLLDCh3R02S06.356
DEM.PROX rubber-**ERG** tie.vt-3[s]O-NPST-3[s]O-1sA=IPFV what
'I am fastening it with this rubber.' (Adult speaker)

b. cause
*wei?-**ŋa*** *basa-ko carko tuk-ma puŋs-e raicha.* CLLDCh1R07S07.050
rain-**ERG** DEM.PROX-GEN very ache.vi-INF begin-PST MIR
'Looks like his [throat] began to ache because of the rain.' (Adult speaker)

c. source
*huŋgo-i?-**ŋa*** *chukt-a-kha-ce!* CLLDCh1R07S07.050
DEM-LOC-**ERG** jump.vi-IMP-IMP-DUAL
'Jump from there!' (Adult speaker)

The agentive function is limited to core arguments (A). The instrumental function is sometimes associated with arguments that are licensed by a verb's argument structure (as is the case in 6a), but sometimes they underlie adjuncts (and sometimes the status is unclear). Cause and source functions are mostly associated with adjuncts. While it is likely that all functions are

historically related, the precise scope of the marker in Chintang is not predictable and children have to learn it. What the child encounters is a single form across various contexts.[3]

A third complication arises from the fact that two- and three-argument verbs can occur in an intransitive construction when the speaker wishes to leave unspecified the cardinality and/or the boundedness of the O referent (Bickel 2011, Schikowski in prep.). The formal consequences of this are that the A argument appears in the nominative case and that the verb shows intransitive agreement morphology, i.e. the A argument is treated as if it were an S argument. This is shown in the following minimal pair, where in the first sentence, the A argument is in the ergative and the verb shows transitive agreement; in the second example the A argument is in the unmarked nominative and the verb intransitively inflected:

(7) a. *huĩsa-ŋa maʔmi copt-o-k-o.*
DEM-ERG person[-NOM] [3sA-]look-3sO-NPST-3sO
'S/he (A) looks at the people.'

b. *hungo maʔmi cop-no.*
DEM[-NOM] person[-NOM] [3sS-]look-NPST
'He/she (A) looks at people.' (in general)

As a result of these three complications, ergative case marking appears to pose severe challenges for acquisition. Given the extreme extent of 'pro-drop' (to a point in fact where this term becomes a misnomer) and the restricted use with first and second person pronouns, ergative markers would seem to be very rare in the input that childrean here. And when they hear an ergative marker, the marker may not in fact encode an A argument but an instrument, a cause or a source. Conversely, when children hear A arguments, not all of them will have an ergative marker, and the presence of the marker depends on the referential properties of the object. The acquisitional task seems enormous, but in the following we show how children proceed and why the system is after all fully learnable on the basis of the input that children receive.

3 Theoretical expectations and hypotheses

One of the most pressing questions in the study of language acquisition is how children actually make use of patterns and signals provided in the input. This question is central for any actual study no matter whether one assumes that children map input signals into an innate and universal grammar format, or whether they map input signals into a language-specific grammar. In the following we discuss three major strategies that may play a key role in how children use input signals from the input when learning a language: (a) reliance on the frequency and saliency of forms; (b) reliance on semantic (functional) distinctions; and (c) reliance on repetitions from what interlocutors said.

[3] This is also the reason why we gloss the forms the same way even though one could of course use separate tags for each function.

3.1 Reliance on frequency and saliency

Especially in usage-based accounts, frequency in the input is often assigned a key role in learning constructions. However, an important methodological issue is how frequency is measured. Classical corpus-linguistic approaches focus on the count of a particular phenomenon (say, the ergative case) relative to its opportunities of occurrence. For the ergative in Chintang, this could mean for example the number of ergatives per noun and per pronoun, or per syntactically transitive construction. On such counts, ergatives are rare in the input, usually staying below 10% of the opportunities of occurrence (Stoll & Bickel in press). However, as Stoll & Bickel (in press) argue, opportunities of occurrence presuppose that children know the relevant contexts, e.g. that they are aware of the morphological affordances of nouns and pronouns and the syntax of transitive as opposed to intransitive constructions. Opportunities of occurrence are themselves only acquired over time, and knowledge of them cannot be presupposed. Therefore, from a psycholinguistic point of view, opportunities of occurrence are inappropriate for frequency measures.

What is more important from a psychological perspective is the question of how often children's memories are effectively trained by a phenomenon over a given time period: linguistic and nonlinguistic phenomena are easier to memorize and learn if they are repeated within sufficiently short periods (Schwartz & Terrell 1983, Childers & Tomasello 2002). In response, a more realistic measure of frequency is the number of ergatives that children hear *per time unit*, regardless of context. Using such a measure, Stoll & Bickel (in press) find that Chintang children hear an ergative on average every two minutes, sometimes even every minute. Such a repetition rate is sufficient for relatively fast and efficient memory training. This lowers considerably the acquisitional challenge.

An additional factor that alleviates the acquisitional challenge comes from the following observation: the massive 'pro-drop' that we find in Chintang has the side effect that overt noun phrase are particularly salient. They therefore provide easily accessible points of access to case marking. Based on these considerations, we expect children to learn the ergative relatively early and without major problems. To test this, we analyze in our first study below (Section 5) children's frequency distributions and those of their corresponding adults separately and then compare the distributions systematically over time.

Overall frequency distributions do not tell us, however, to what extent children make use of further contextual information, e.g. information derived from meaning differences or from the nominal and verbal constituents that co-occur with a given ergative case marker. These possibilities are taken up in the following.

3.2 Functional learning

Given the functional differentiation of the ergative into a marker for agents, for instruments, for causes and for sources (cf. 6 above), it is possible that children use the contextualized meaning of utterances when learning how to use the marker. If so, one would expect that not all functions are learned at the same time and to the same extent: any differences here are likely to depend on the frequency in which each function occurs. This leads us to hypothesize that children start out with the most frequent function and only later on they make use of less frequent functions.

To explore this hypothesis, we code each occurrence of the ergative marker according to its function.

3.3 Item-specific learning

Research on languages where argument structure is chiefly expressed through syntactic templates rather than case markers (i.e. English) suggests that children learn argument structure through what has come to be called item-specific learning: children first use a specific argument structure construction (say, a transitive 'noun phrase - verb - noun phrase' pattern) only with a heavy bias towards a small set of nominal and verbal items (Tomasello 1992, Lieven et al. 1997, Tomasello 2003). It is only gradually over time that children loosen the association of the construction with these lexical sets and develop an abstract representation of the construction. The reason for this process is that children heavily rely on imitation as learning strategy: they first aim to imitation adult speech as exactly as possibly and become sufficiently independent language users only over time.

A similar strategy may be involved in the acquisition of case markers such as the Chintang ergative. If so, we would expect children to bias the use of the ergative marker to selected nominal hosts and selected co-occurring verbs in the sentence. Given the complexity and constraints of the system, we would in fact expect relatively strong effects of this kind of item-specificity in the early phase of acquisition. To test this hypothesis, we employ below (Section 7) an information-theoretical method for estimating item-specificity in the corpus.

3.4 Interactional learning

Both functional learning and also item-specific learning is usually studied without focusing in detail on the content units in which the studied features occurs. However, communication consists of concrete instances of conversation, and there is good reason to assume that conversational units play a critical role for language acquisition: conversations require a coordinated and joint attentional frame (Tomasello 2003), and it has been shown that vocabulary development directly correlates with the extent to which children establish joint attention with their conversation partners (Tomasello & Todd, 1983; Carpenter, Nagell & Tomasello, 1998). In line with this, it has long been established that successful language acquisition depends on active conversational interaction; passive 'intake' is not sufficient (Ervin-Tripp 1973). Given these findings, the question is how precisely children exploit conversational units for learning purposes. Specifically, are there patterns that children rely on and does their usage of ergatives depend on — or is it even driven by — the ergative usage offered by children's conversation partners in dialogue units?

One key aspect of this is the extent of repetition within conversational units. There is evidence that repetitive patterns in the input are of considerable help in acquisition (Kuntay & Slobin 1996, Cameron-Faulkner et al. 2003, Stoll et al. 2009), but how exactly does repetition play out in conversational interaction. Among the strategies discussed by Brown (1998), dialogic

repetition is of particular relevance for the acquisition of a single marker such as the Chintang ergative.[4]

In dialogic repetition, parts of the preceeding utterances are repeated by the interlocutor. In some societies, e.g. among the Tzeltal, this type of repetition is a cultural practice that children learn very early Brown (1998). Brown hypothesizes that the cultural practice of repetition helps children to learn the complexities of morphology. So far this hypothesis has not been tested quantitatively, but in Section 8 we will test it for the acquisition of Chintang ergative.

Specifically, we hypothezise that in the early phase of ergative use children will repeat the forms provided by adults in dialogue more often than in later stages of development. At this age it is mostly others who direct communicative exchanges and the child will follow. We expect the child to be more active and better at initiating topics of conversation and using forms spontaneously only at later stages of development.

4 Participants, recordings and data overiew

We base our studies on a longitudinal corpus of four Chintang preschool children (Stoll et al. 2012a,b). Data collection took place within the framework of a large-scale interdisciplinary documentation project on Chintang and a related language (Gaenszle et al. 2005). Two children were aged 2;0 years at the first recording (one girl and one boy, here named 'Child 1' and 'Child 2') and two children aged 2;11 and 3;0 at the beginning of the recordings (one girl and one boy, named 'Child 3' and 'Child 4'). All children come from large families, and they have at least three siblings. Some of the target children are related and play together regularly. Children live in individual houses with their families. The houses are scattered around the hills with fields in between. Children were recorded by Nepali research assistants together with Chintang native speaker assistants, who were part of the natural environment of the children and worked for our project.

Recordings took place within one specific week per month and amounted to approximately 4 hours of recording per month. In total, the children were recorded for 18 months. The recordings took place in and around the home of the children, chiefly outside. The recordings were not restricted to specific people and activities but rather comprised the typical daily activities children are engaged in. These activities usually include a lot of other children. The number of interlocutors per recording varied but usually there were also a number of adults present who were busy with other activities and also talked a lot to each other. The number of people present ranged from 2 to extreme cases with up to 27 interlocutors, with a median of 8 interlocutors per individual recording session. However, not all participants were always present at the same time or during the entire session. This is also true of the closest caretakers who come and go during sessions and are sometimes completely absent from a session.

The data analyzed below consist of the monthly recording cycles that were ready for analysis at the time of the study. The data of all participants in the recordings was transcribed and

[4] Other repetition strategies, such as prompting routines ("Say 'X!'"), may play a role for more complex structures (cf. e.g. Demuth 1987, Ochs 1988, Schieffelin 1990), but in a study of the acquisition of Chintang noun phrase vs. verb usage we found no evidence of this (Stoll et al. 2012a).

translated by native speaker assistants and then annotated for morphological categories and parts of speech by student assistants. The amount of data available at the time of the study varied for the four children but this will be taken into account in the analysis.

As noted in the previous section, the natural setting in which Chintang children grow up includes substantial numbers of other people: e.g. from the extended family, friends and aquaintances from around the village. This means that children are exposed to a substantially more varied input and are surrounded by much more talk of adults among each other than is the case in the typical recordings of children in a Western urban context, with their traditional focus on isolated child/caretaker dialogues. Another consequence of the large number of adults is that Chintang children have less opportunity to say something than older peers or adults who were present at the recordings. As a result, in the Chintang setting more recordings are necessary to get the same amount of child data as gathered within a single recording session of a child/catetaker dyad in a Western setting. Figure 1 shows the amount of data (in number of words) per monthly recording cycles per child that we used for this study, separate for target children, other attended children and adults.[5]

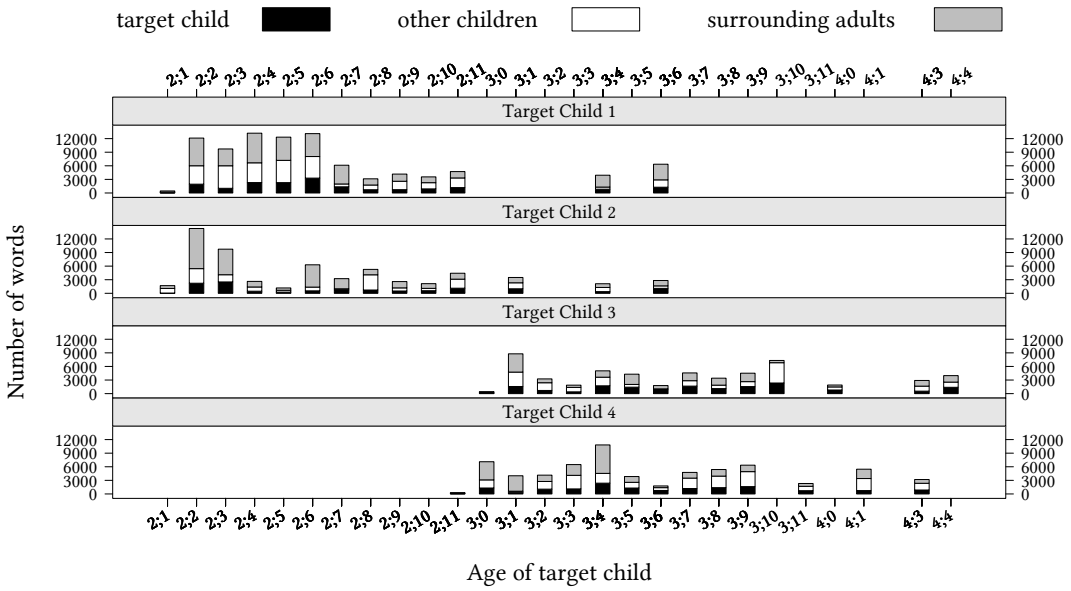

Figure 1: Data used in the present study in number of words (ages averaged within recording cycles)

[5] All quantitative analyses were performed using `R` (R Development Core Team 2011), with the additional packages `coin` (Hothorn et al. 2008), `entropy` (Hausser & Strimmer 2009), `gam` (Hastie 2010), and `lattice` (Sarkar 2008).

5 Overall usage of ergatives by children and adults

5.1 Methods

As argued in Section 3 above, the most appropriate frequency measurement for the input is frequency per time unit. However, we need a different method when assessing children's production and level of productivity. Especially in the Chintang context, children have considerably less opportunities to talk than adults. This will substantially lower their frequencies per time unit, irrespective of the extent to which children master the use of ergatives. In response, for assessing production and productivity, we compare children's and adult's usage on the basis of 'raw' usages per word. Functional, lexical and syntactic differentations are then measured in terms of more specific analyses (Sections 6 and 7).

Comparison between children and surrounding adults is done by dividing children's proportion of ergatives per word by adults's proportion of ergatives per word. This 'child-to-adult ratio' gives us an estimate on how many times less often children use the form. A value close to 1 suggests similar use, deviations below one suggest that children use ergatives less frequently than adults. (Other attendant children are excluded from the analysis here.)

5.2 Results

The distributions in our data do not show a difference between children and adults nor do we find a developmental trend in the use of ergatives in the four target children. As shown in Figure 2, children's use is very close to adults's use, the child-to-adult ratio never extending below .98 and indeed averaging very close to 1 (mean and median both .996)

However, even though the distributions are the same, it could of course be that children use the ergative in a completely different way than adults do. This question is addressed in the following studies.

6 Functional learning

6.1 Methods

For testing the hypothesis of Functional Learning, we coded all ergatives for all speakers for their function, as illustrated in (6) above. However, cause, source and a few other 'adverbial' functions are relatively rare, and we collapsed these into a general category. Our coding therefore distinguishes between agent, instrument and 'other' functions. Overall, the agentive function is much more common than the other functions, but what matters for our purposes is whether there is any difference in proportions between adults and children.

To explore this, we compute a similar 'child-to-adult ratio' as in the previous study: we divide the proportion of agentive functions used by adults by the proportion of agentive functions by children, and compute the same ratios for the other two functions. Values close to 1 of the ratio again mean that adults' and children's proportions are the same in a given function.

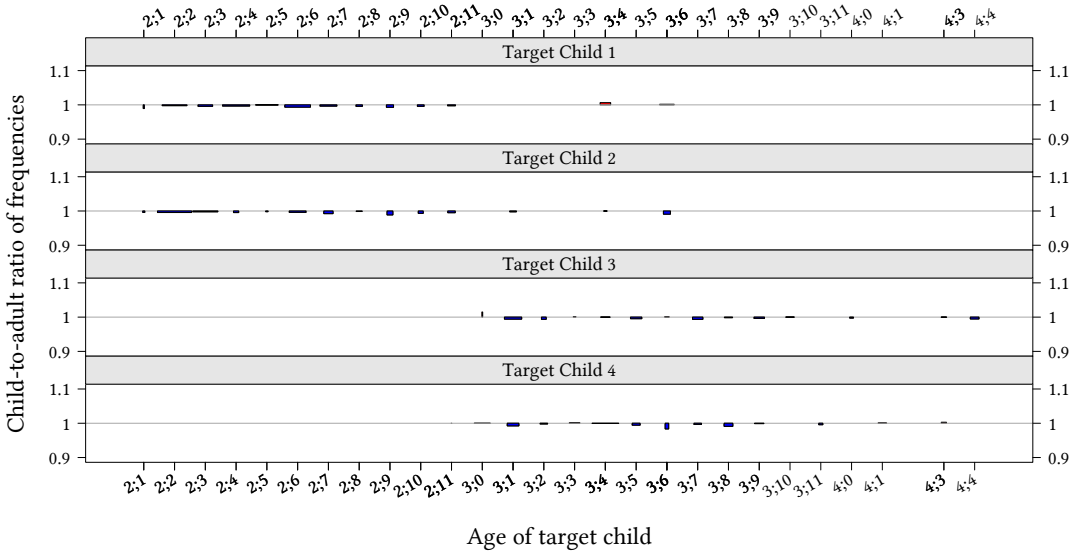

Figure 2: Usage distributions of ergatives per word by children and adults. (Bar width is proportional to the corpus size in a given recording cycle in number of words.)

6.2 Results

Figure 3 shows the child-to-adult ratios for three functions. All ratios are again very close to 1, with a median of 1.01 (Agents), 1.0 (instruments) and .95 (others). Fitting locally weighted smoothing (LOESS) regression lines reveals no evidence for a clear development over time. There is a slight tendency among the two 2-year olds for systematically lowered child-to-child ratios of non-agent functions, with a concomitantly increased ratio of agent functions. For Child 1 the difference is statistically significant, though only at .05 rejection level and with a relatively large confidence interval (Exact Wilcoxon Mann-Whitney Rank Sum Test comparing the proportion of agents vs. non-agents among adults and children up to age 3;0: median difference = .175, 95% CI = (.004, .271], p = .0215). For Child 2, the difference is not statistically significant (median difference of proportion agents = .163, p = .605). This tentatively suggests that at least one child starts by focusing on the agent function and assimilates her use of non-agent ergatives to that of adults only later.

If such a trend can be substantiated by more data, it will most likely be a relatively direct reflex of the fact that agents functions strongly outnumber all other functions in adults speech. This is demonstrated in Figure 4.

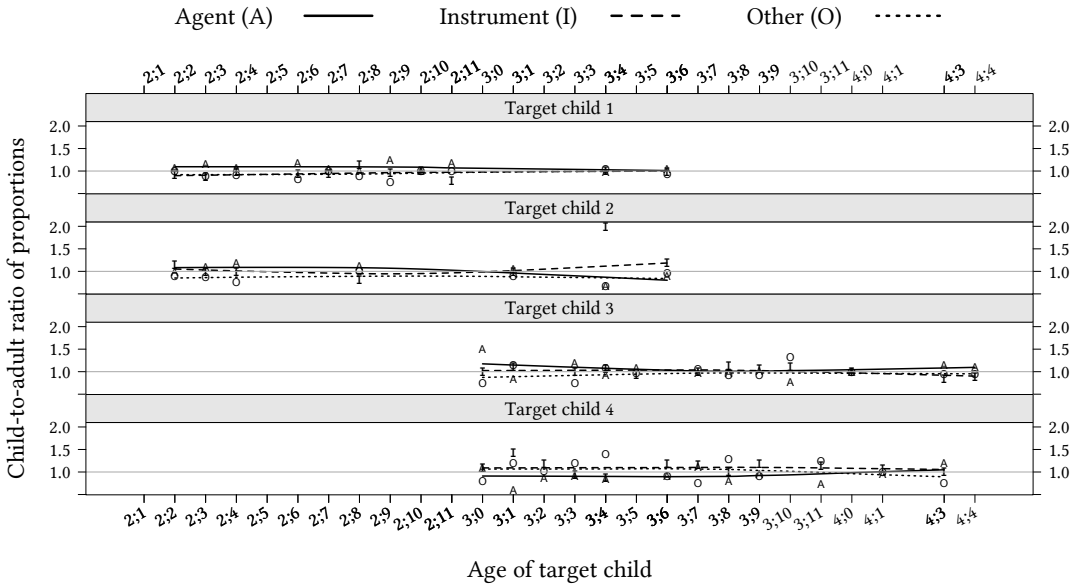

Figure 3: Usage distributions of ergative functions by children and adults. (The lines represent LOESS regression estimates.)

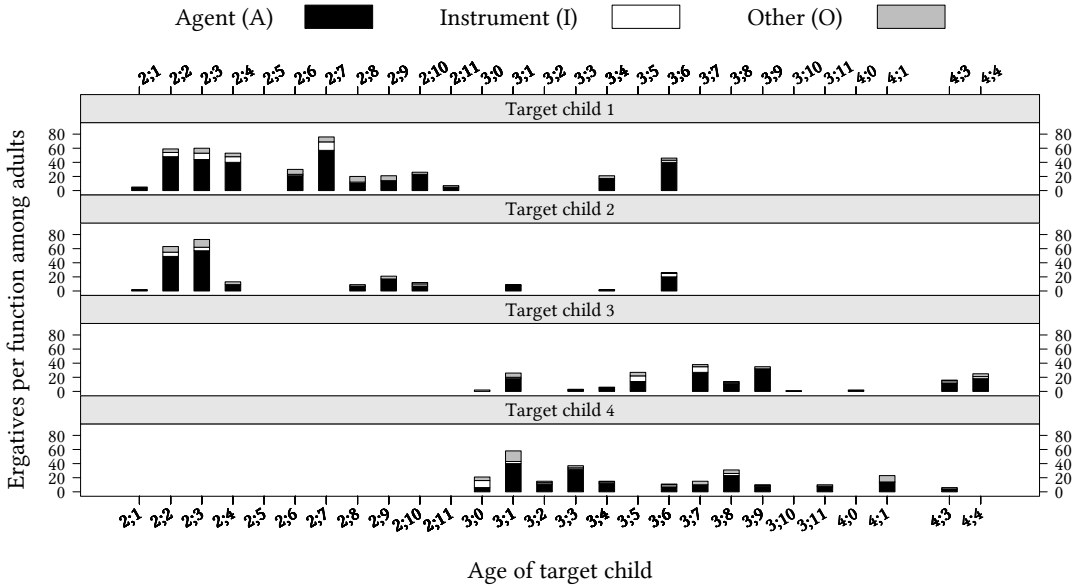

Figure 4: Frequencies of ergative functions by surrounding adults.

7 Item-specific learning

7.1 Methods

For a case marker, item-specificity has two aspects: (i) it refers to the host items, i.e. the range of and flexibility with the nominal and pronominal hosts that the case marker is attached to; (ii) item-specificity refers to the range of and flexibility with the verbal stems that the case marker is assigned by (as an an argument) or co-coccurs with (as an adjunct). If the acquisition of the ergative is guided by item-specificty to some appreciable extent, we expect children to limit both the nominal and the verbal range and flexibility of the marker in the beginning and to gradually assimilate to adult usage.

In order to measure nominal and verbal item-specificity, we rely on a general estimate of item-specificity based in information theory: a marker M is item-specific with regard to a context C to the extent that the Shannon entropy of C is minimal. Maximum entropy of C, by contrast, means that the marker is used across all contexts and in each context with equal probability. In general, the higher the entropy is the less we can predict the contexts in which a marker is used; conversely, low entropies means that a marker has limited distributions across contexts and/or that some contexts are preferred.

The entropy H of contexts C of a marker M can be estimated from the probabilities of all types of C_M (8a), and these can in turn be approximated via Maximum Likelihoof Estimation based on their frequencies in the corpus (8b).

(8) a. $$H(C_M) = -\sum_{C_i \in C_M} p(C_i) \cdot \log_2 p(C_i)$$

b. $$p(C_i) = \frac{N(C_i)}{\sum_{C_m \in C_M} C_m}$$

Thus, if there are three contexts $C_{1...3}$ so that, say, C_1 occurs 10 times, C_2 4 times and C_3 2 times, we can estimate each probability $p(C_i)$ from the proportion of the contexts of type C_i among all contexts. For C_1 this would be $\hat{p} = \frac{10}{16} = .625$, for C_2 $\hat{p} = \frac{4}{16} = .25$ and for C_3 $\hat{p} = \frac{2}{16} = .125$. These probabilities can then be used to estimate the entropy of all contexts.

For the ergative case, we estimate the entropy of the nominal context, i.e. of nominal hosts and the entropy of the verbal context, i.e. of verbs that can co-occur with an ergative. These entropy estimates capture the intuition that a child masters the ergative case better the more she uses the marker with different nominal hosts and verbs and the less she prefers some items over others.

It is important to emphasize that entropy estimates are based on relative frequencies, i.e. on the frequency of a context relative to all contexts (cf. the equation in 8b). Thus, if in the toy example above, corpus size had been bigger and instead of 10, 4 and 2 counts, one would have had 100, 40 and 20 counts, probabilities and therefore entropy estimates would have been exactly the same. Thus, entropy estimates are independent of sample size and the number of utterances a speaker produces (which varies widely).

For comparing adults and children we again use a child-to-adult ratio, here the ratio of the entropies of target children divided by the entropies of the surrounding adults during a given recording cycle.

7.2 Results

The graphs in Figure 5 displays the child-to-adult ratios of entropy estimates for nouns and verbs.[6] The LOESS regression lines suggest slight developmental trends between about two and half and fours years, except for Child 3. However, for Child 1 and 2 we do not have at present sufficient data to assess these trends statistically, and for Child 4 an F-test of the two regression fits is not significant, regardless of how small one chooses the local regression bandwidth (all $p > .1$).

The graphs also suggest that all children have lower entropies than adults, i.e. that their ergative use is generally more item-specific than that of adults. The differences seem stronger for verbal contexts than for nominal hosts. This is confirmed by the statistics reported in Table 1. The difference is particularly pronounced in the case of Child 3, who shows adult-like entropies for nominal hosts but still significantly lower entropies with regard to the choice of verb stems and their frequency distributions.

	Nominal hosts			*Co-occurring verbs*		
Child	Difference	95% CI	p	Difference	95% CI	p
1	−0.980	(-1.751, -0.396]	0.002	−1.845	(-2.571, -1.26]	0.000
2	−1.500	(-2.606, -0.616]	0.004	−2.181	(-2.807, -0.885]	0.003
3	−0.374	(-1, 0.22]	0.241	−1.000	(-1.682, 0]	0.032
4	−1.030	(-1.894, -0.4]	0.003	−1.068	(-2, -0.241]	0.007

Table 1: Median differences between children's and adults' entropies (Exact Wilcoxon Mann-Whitney Rank Sum Test)

8 Interactional Learning

8.1 Methods

For evaluating the predictions of the Interactional Learning hypothesis, we coded children's use of ergatives relative to the conversational environment. We define the environments by topic choice, i.e. all utterances that precede a child's ergative and bear on the same topic of conversation. Usually such an environment comprises several utterances. Children's ergative use was then coded relative to this utterance as to whether it represents (a) a repetition of what an adult or older child said, (b) a use analogically to a previous use by another speaker or (c) a spontaneous use. More specifically:

Exact repetition: an ergative form is coded as an exact repetition of an ergative if a child uses the exact same form that an interlocutor has used within the same conversational frame.

[6] It is not uncommon for ergatives to be used without verbs, e.g. when answering content questions. This is why the lower graphs contains less data than the upper graph and why entropy estimates are not always possible for verbs even when ergatives occur.

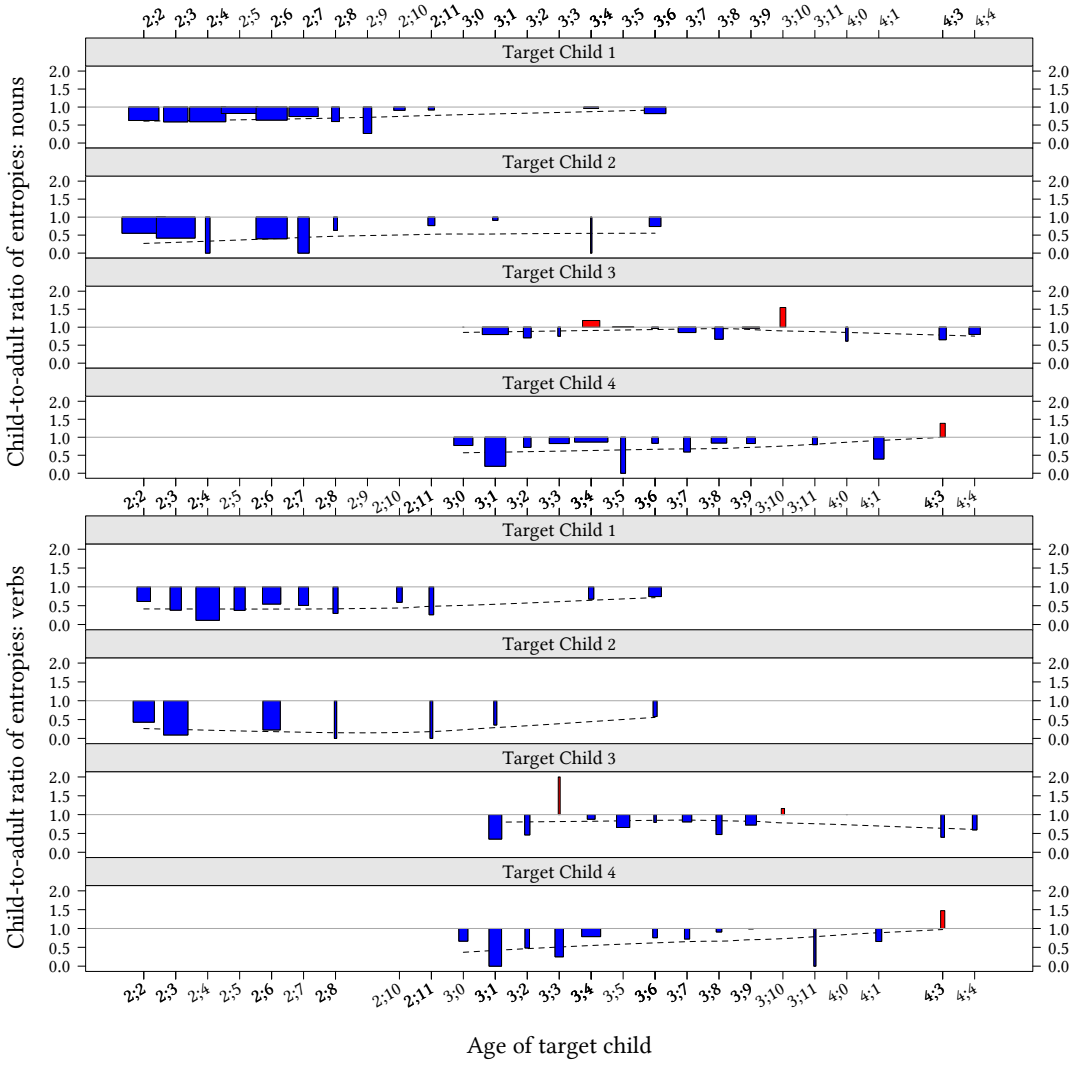

Figure 5: Entropies of nominal hosts (above) and verbs (below) co-occuring with ergative markers (in any function) by children and adults. (The lines represent LOESS regression estimates; bar width is proportional to the number of ergatives in a given recording cycle.)

This can occur when the child imitates an interlocutur's utterance as in (9a) or when the child is explicitly prompted to do so, as in (9b):

(9) Child, age 12: *la sipai-ce-ŋa na-kha=ne=na*
PTCL soldier-ns-ERG 3[SG]>2[SG]-take.vt=OPT=PTCL
'Well, so let the soldiers take you away!'

Child, age 2;3: *sipai-ce-ŋa na-kha?=ne=na* CLLDCh1R03S01.0089f
soldier-ns-ERG 3[SG]>2[SG]-take.vt=OPT=PTCL
'Let the soldiers take you away!'

(10) Child, age 12: *Rame-ŋa phil-o-ŋs-e=mo lud-a=na!*
Rame-ERG [3SG.A]pinch.vt-3[SG]P-PERF-PST=REP tell.vt-IMP=PTCL
'Say "Rame has pinched him"!'

Child, age 2;2: *Rame-ŋa phil-o-ŋs-e.* CLLDCh1R02S04a.0260f
Rame-ERG [3SG.A]pinch.vt-3[SG]P-PERF-PST
'Rame has pinched him.'

Analogical use: an ergative was coded as a case analogy if an interlocutor has used an ergative form in the same conversational frame and then the child uses subsequently an ergative but with a different host. This already shows some flexibility and ability to generalize in attaching the ending to different hosts than presented in the immediate context. Typical situations of this involve WH-questions as in the following examples, which illustrate different lengths of the conversational frame:

(11) Adult: ***sa-ŋa*** *na-khutt-e?* CLLDCh3R09S06.044f
who-ERG 3[SG]>2[SG]-bring.sth.for.sb.vt-PST
'Who brought it for you?'

Child, age 3;8: ***Dipe-ŋa*** *pid-a-ŋs-e-hẽ*
Dipe-ERG [3SG.A]give.vt-PST-PERF-PST-[1SG.P]EXCL.PST
'Dipe has given it to me.'

(12) Adult: ***sa-ŋa*** *ten-e?* CLLDCh1R02S05.0440ff
who-ERG [3SG.A]hit.vt-PST[SG.P]
'Who hit him?'

Child, age 2;2: *soita khi-no*
aimlessly [3SG.S]be.angry.vi-NPST
'He is being angry just like that!'

Adult: *hã?*
PTCL
'Huh?'

Child, age 2;2: ***Khele-ŋa***
Khele-ERG
'Khele [hit him]'

Spontaneous use: The use of an ergative by a child was coded as spontaneous use if (and only if) no interlocutor has used an ergative form in the same conversational frame. Once spontaneous use becomes established we can expect that the child is proficient with at least some subpart of the ergative system, and that the child has a higher proficiency than a child that uses only exact repetitions.

Self-repetitions were excluded from the analysis when they occured within the same conversational frame.

8.2 Results

Figure 6 demonstrates the development of exact repetitions, analogical and spontaneous uses of ergatives. While there is no apparent development in the case of analogical uses, the LOESS regressions suggest an almost linear decrease of exact repetitions and a concomitant increase of spontaneous uses at least for Child 1 and Child 3 up to three and a half years. In the case of Child 1, there is (barely) sufficient data to test the trend statistically: the LOESS regression of the decrease in exact repetitions accounts for $R^2 = .55$ of the data and fits decently well ($F = 5.4$, $df = 1.81$, $p = .034$); for the increase in spontaneous use the LOESS regression accounts $R^2 = .44$ of the data and fits marginally ($F = 3.52$, $df = 1.81$, $p = .081$). For Child 2, data are too sparse for analysis. Child 4, finally, seems to be an advanced level so that repetitions do not play a substantial role anymore.

These analyses also show that there is a substantial amount of individual variation. Not all children develop at equal pace or even in the same way.

9 Conclusions

We have shown that although the distributions of ergative case marking do not differ between the target children and their surrounding adults, there are differences in the way ergatives are used between adults and the children in our study. From a functional point of view the use of the different ergative functions seems to be similar to adult usage from early on as well. However, in an analysis of the contextual environment of the ergative, we do find differences between adults and children. Further between children there is strong individual variation in the way they learn the ergative case.

First, children and surrounding adults differ in the range and frequencies of hosts and also in the accompanying verbs used in an ergative construction, especially among the two-year-olds in our sample. Around age 3 the development is more similar to those of adults, and children are much less item-specific. The older the children, the less item-specific they become and thus the more proficient they are with the ergative case. Thus, there seems to be evidence that the youngest children in our sample approach the task of learning the ergative in an item-specific

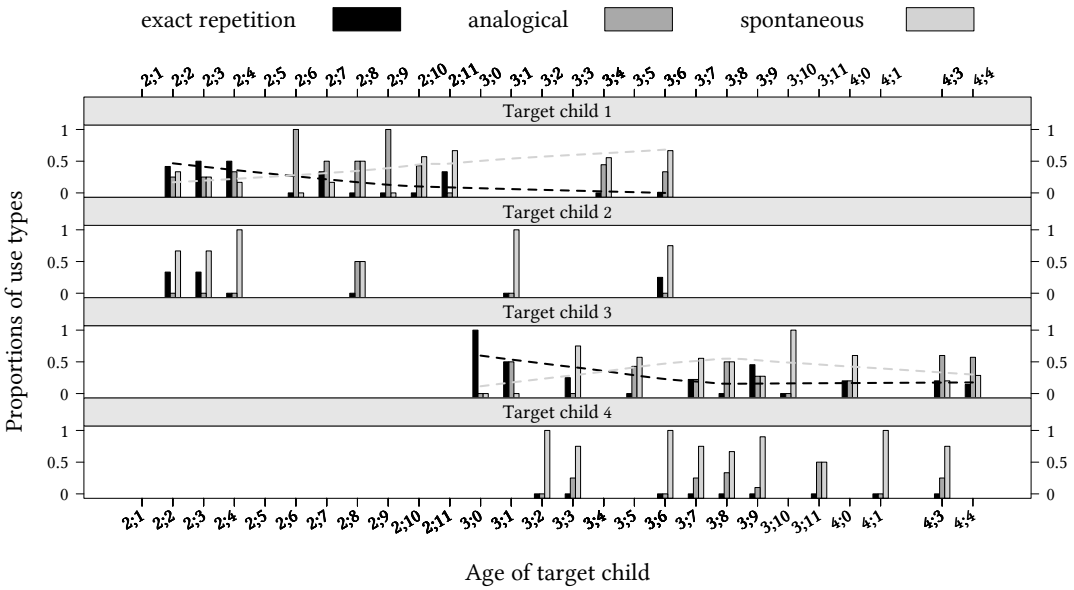

Figure 6: Types of ergative uses by target childern within conversational frames. (The lines represent LOESS regression estimates.)

way and they become more adult like as they grow older. Further, there is a clear development in the use of the ergative case even though this fact is not mirrored in the overall distributions.

Second, there also seems to be a development in the contextual use of ergatives in interactions, but more data is needed to assess these developments more carefully. In the one 2-year old child for which we have a lot of data, the data tentatively suggest a development from an initial preference for repetitions, and only later on in development the child used the ergative case more spontaneously. A similar development is visible in Child 3 who was 3 years old at the beginning of the study. At least in the first two recordings up to age 3;3 she uses many more exact repetitions and case analogies than spontaneous instances of the ergative case. Judging from the available data so far, Child 2 does not really seem to follow the learning strategy chosen by Child 1. This either can be due to individual variation and an altogether different approach that the child might take (which is very possible), or it might be the case that we do not have enough data available yet. This needs to be resolved in future research.

Another issue for future research is to assess how item-specific learning and interactional learning work together or are distributed across children. What is clear from a methodological perspective is that item-specificity can be fully assessed only once the effects of interactional learning are controled for: there is always possible that item-specificity is a side-effect of repetitions (imitations) in conversational frames. But further exploration of this possibility must be left for further research.

References

Bickel, B. (2011). Multivariate typology and field linguistics: a case study on detransitivization in Kiranti (Sino-Tibetan). In P. K. Austin, O. Bond, D. Nathan, & L. Marten (Eds.) *Proc. Conf. Language Documentation and Linguistic Theory 3* (pp. 3–13). London: SOAS.

Bickel, B., Banjade, G., Gaenszle, M., Lieven, E., Paudyal, N., Rai, I. P., Rai, M., Rai, N. K., & Stoll, S. (2007). Free prefix ordering in Chintang. *Language 83*, 43 – 73.

Bickel, B., Rai, M., Paudyal, N., Banjade, G., Bhatta, T. N., Gaenszle, M., Lieven, E., Rai, I. P., Rai, N. K., & Stoll, S. (2010). The syntax of three-argument verbs in Chintang and Belhare (Southeastern Kiranti). In A. Malchukov, M. Haspelmath, & B. Comrie (Eds.) *Studies in ditransitive constructions: a comparative handbook* (pp. 382–408). Berlin: Mouton de Gruyter.

Brown, P. (1998). Conversational structure and language acquisition: the role of repetition in Tzeltal. *Journal of Linguistic Anthropology 8*, 197–221.

Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science 27*, 843–873.

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology 38*, 967–978.

Demuth, K. A. (1987). Prompting routines in the language socialization of Basotho children. In B. B. Schieffelin, & E. Ochs (Eds.) *Language socialization across cultures* (pp. 51–79). New York: Cambridge University Press.

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German children's comprehension of word order and case marking in causative sentences. *Child Development 79*, 1152–1167.

Ervin-Tripp, S. (1973). Some strategies for the first two years. In T. E. Moore (Ed.) *Cognitive development and the acquisition of language* (pp. 261–286). New York: Academic Press.

Ezeizabarrena, M. J., & Larranaga, M. P. (1996). Ergativity in Basque: a problem for language acquisition? *Linguistics 34*, 955–991.

Gaenszle, M., Bickel, B., Banjade, G., Lieven, E., Paudyal, N., Rai, A., Rai, I. P., Rai, M., Rai, N. K., Rai, V. S., Gautam (Sharma), N. P., & Stoll, S. (2005). Research report: the Chintang and Puma Documentation Project (CPDP). *European Bulletin of Himalayan Research 28*, 95 – 103.

Gertner, Y., Fisher, C., & Eisengart, J. (2006). Learning words and rules. *Psychological Science 17*, 684.

Hastie, T. (2010). `gam`: Generalized Additive Models. R package version 1.03. URL `http://CRAN.R-project.org/package=gam`

Hausser, J., & Strimmer, K. (2009). Entropy inference and the James-Stein estimator, with application to nonlinear gene association networks. *Journal of Machine Learning Research 10*, 1469–1484.

Hothorn, T., Hornik, K., van de Wiel, M. A., & Zeileis, A. (2008). Implementing a class of permutation tests: the `coin` package. *Journal of Statistical Software 28*, 1 – 23.

Kuntay, A., & Slobin, D. I. (1996). Listening to a Turkish mother: Some puzzles for acquisition. In D. I. Slobin, J. Guo, & A. Kyratzis (Eds.) *Social interaction, social context, and language: Essays in honor of Susan Ervin-Tripp* (pp. 265–286). Hillsdale NJ: Erlbaum.

Lieven, E. V. M., Pine, J. M., & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language 24*, 187–219.

Narasimhan, B. (2005). Splitting the notion of 'agent': case-marking in early child Hindi. *Journal of Child Language 32*, 787–803.

Ochs, E. (1982). Ergativity and word order in Samoan child language. *Language 58*, 646–71.

Ochs, E. (1988). *Culture and language development: Language acquisition and language socialization in a Samoan village*. New York, NY: Cambridge University Press.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

R Development Core Team (2011). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, `http://www.r-project.org`.

Sarkar, D. (2008). *Lattice: multivariate data visualization with R*. Berlin: Springer.

Schieffelin, B. B. (1985). The acquisition of Kaluli. In D. I. Slobin (Ed.) *The Crosslinguistic Study of Language Acquisition* vol. 1: The Data (pp. 525–593). Hillsdale, NJ: Erlbaum.

Schieffelin, B. B. (1990). *The give and take of everyday life*. New York, NY: Cambridge University Press.

Schikowski, R. (in prep.). *Object-conditioned differential marking in Chintang and Nepali*. Ph.D. thesis University of Zürich.

Schikowski, R., Bickel, B., & Paudyal, N. P. (in press). Flexible valency in Chintang. In B. Comrie, & A. Malchukov (Eds.) *Valency classes: a comparative handbook*. Berlin: Mouton de Gruyter.

Schwartz, R. G., & Terrell, B. Y. (1983). The role of input frequency in lexical acquisition. *Journal of Child Language 10*, 57–64.

Stoll, S., Abbot-Smith, K., & Lieven, E. (2009). Lexically restricted utterances in Russian, German and English child-directed speech. *Cognitive Science 33*, 75–103.

Stoll, S., & Bickel, B. (in press). How to measure frequency? Different ways of counting ergatives in Chintang (Tibeto-Burman, Nepal) and their implications. In F. Seifart, G. Haig, N. Himmelmann, D. Jung, A. Margetts, P. Trilsbeek, & P. Wittenburg (Eds.) *Potentials of language documentation: methods, analyses, utilization*. Manoa: University of Hawai'i Press.

Stoll, S., Bickel, B., Lieven, E., Banjade, G., Bhatta, T. N., Gaenszle, M., Paudyal, N. P., Pettigrew, J., Rai, I. P., Rai, M., & Rai, N. K. (2012a). Nouns and verbs in Chintang: children's usage and surrounding adult speech. *Journal of Child Language 39*, 284 – 321.

Stoll, S., Bickel, B., Lieven, E., Banjade, G., Zakharko, T., Schikowski, R., Bhatta, T. N., Gaenszle, M., Paudyal, N. P., Rai, M., Rai, I. P., & Rai, N. K. (2012b). *Audiovisual corpus on the acquisition of Chintang. DOBES Archive*, `http://www.mpi.nl/DOBES`.

Tomasello, M. (1992). *First verbs: a case study of early grammatical development*. Cambridge [England] ; New York, NY, USA: Cambridge University Press.

Tomasello, M. (2003). *Constructing a language: a usage-based theory of language acquisition*. Cambridge, Mass.: Harvard University Press.