Studying Language Acquisition in Different Linguistic and Cultural Settings

Sabine Stoll

1 Introduction

It is undisputed that every child learns the language(s) of her social environment in the first years of life. However, what are the cognitive processes that enable children to learn language and which environmental processes are involved has been a matter of continuous dispute. This century-old debate of nature vs. nurture has also moved center-stage in the field of language acquisition research, resulting in two major strands of theories: nativist vs. empiricist/constructivist.

Nativist theories assume an innate linguistic module with an innate core grammar, which makes language acquisition possible. Assumptions about the nature of this linguistic module vary from a complex system of parameters (Chomsky 1981) to a single mechanism of recursion in the latest versions of generative theories of grammar (Hauser et al. 2002; for a summary see Ambridge & Lieven 2011). Even though the parameter-setting approach was abandoned in newer versions of generative theory, it is still assumed by some acquisition researchers (e.g. Hyams 2012).

The other major type, constructivist theories, also known as usage-based theories, assume a set of innate abilities, which in their combination enable humans to learn language, but these cognitive abilities are not assumed to be specific to language nor necessarily even specific to humans (Tomasello 2003). This approach emphasizes that these cognitive abilities are also found in other species, but the specific combination of them is responsible for the human-specific ability to understand, produce, and learn language in all its different varieties.

Currently there are around 7000 languages spoken and several hundred thousand have existed before they became extinct or developed into other languages. The ability of children to acquire any of these systems and contexts emphasizes the great flexibility of the human mind. This flexibility is the focus of this chapter.

Independent of the linguistic theory assumed, it is undisputed that language acquisition requires input. The amount of input necessary, however, is again a matter of theoretical dispute. Nativist accounts assume that minimal input is sufficient, whereas constructivist theories assume the necessity of an extensive and elaborate input to build up the vocabulary and the grammar of the language over several years. Both, languages in their grammatical build-up and the social environments in which they are learned, however, vary to an extreme degree and so far we
know relatively little about the impact of different cultural backgrounds and different linguistic systems on the acquisition process.

Thus, one of the challenges of language acquisition research is to find out exactly what are the common variables across cultures and languages that allow children to learn the linguistic system of their respective culture and socio-communicative environment. To find these variables has become the greatest challenge in acquisition research because according to recent work there are hardly any substantial linguistic universals (Bickel 2007, Evans & Levinson 2009) on which children can rely in learning language. If there are commonalities, they occur on the most abstract level, which would still not account for the abilities required to link these abstract concepts to the individual language-specific features expressing these concepts.

Further, so far, there is very little research comparing the socio-communicative environments in which children acquire their native languages. However, at least, there seems to be one crucial and universal feature of language acquisition common to all cultural contexts in which children grow up: the involvement of children in communicative interactions (e.g. Snow 1984). From a number of experimental studies we know that listening to communication alone without taking part in communicative interactions (or observing language in the case of sign language) is not enough to learn a language (e.g., Sachs & Johnson 1976, Kuhl et al. 2003). Thus, communicative interaction seems to be the driving force behind language learning and can even lead to the development of new linguistic systems. This has been shown in the development of Nicaraguan sign where, over several generations, children and older signers turned a rudimentary, inconsistent system of home signs into a fully fledged language (Senghas & Coppola 2001). Thus, there seems to be a human drive toward the development of systematic communication systems and transmitting such a system in personal interaction. The central research question is how children extract the necessary linguistic information from these interactions.

Acquisition research, therefore, aims to discover the cognitive, innate mechanisms in inter-relation with factors in the environment such as the linguistic system and the input that enables children to learn language. In other words, we still try to solve the old puzzle of the interaction of nature and nurture. Redirecting our focus on nurture, however, might help in better understanding the role of nature. Given the fact that interaction and input are foundational for acquisition, it seems unavoidable to take a more empirical bottom-up approach in comparing the acquisition of as many different languages as possible. This is a very demanding task if we take into account the large design space of languages without substantial linguistic universals (Bickel 2007, Evans & Levinson 2009). The first step in such an approach, then, is to analyze development in tracing naturalistic interactions in as many diverse languages as possible. The second step is to compare the different developments in the various languages. This chapter tries to trace these research tasks from a cross-linguistic perspective. The main focus is on the role of the ambient language, concentrating on features such as the amount and kind of input children receive. Two main topics of this chapter are:

- How can we systematically study the impact of grammar on the acquisition process?
- How can we study the role of the socio-communicative context for learning a language?

The chapter is organized as follows. To give an overview of how acquisition is studied, it starts out with a short presentation of the major methods commonly used in language acquisition studies. The focus is on sampling issues that are intertwined with these methods. Second, some potentially universal milestones of early language learning are presented before paying attention to environmental features relevant for learning language. Third, the chapter illustrates how
cross-linguistic research contributes to our understanding of the mechanisms of language acquisition and why the input plays a central role for understanding acquisition.

2 Main Research Methods and Sampling Issues

To study language acquisition both within and across languages there are several established methods to gather and analyze data. I discuss the major methods used in acquisition research and concurrent sampling issues: (1) naturalistic, observational data and (2) experiments. There is a third method, namely questionnaires, which are used to study specific questions such as early vocabulary. The most commonly used questionnaire is the MacArthur-Bates Communicative Development Inventories (Fenson et al. 1993), which has been adapted to a number of different languages. The focus in this chapter, however, is on naturalistic vs. experimental methods, whereby the following sampling issues take center-stage: choice of languages, choice and number of participants, and recording scheme.

2.1 Naturalistic/Observational Data

Naturalistic studies are conducted to learn about the child’s behavior in her natural environment. In contrast to experiments, naturalistic studies are correlational in nature and can merely contribute to hypothesizing about cause and effect. The main advantage of observational studies is that they provide insights into what children do spontaneously in their natural environment.

There are two main types of observational studies. In cross-sectional studies a large enough number of children of a specific age range is recorded, usually at one or a few points in time, to conduct a statistical test. In longitudinal studies, usually a small number of children is recorded over a longer time period at regular intervals. Both in cross-sectional and in longitudinal studies, recordings constitute a snapshot of a specific context, which can potentially bias our results.

Longitudinal studies are the main method for gathering comprehensive data about the developmental path of individual children and we will focus on this type of data collection in the following, since longitudinal studies are usually the starting point in cross-linguistic comparisons and the first type of data we collect in a hitherto unstudied language. Longitudinal studies are either diaries or audio-visual recordings.

2.1.1 Diaries

The earliest acquisition studies conducted around the end of the nineteenth century were diary studies of a small number of European languages such as French, Russian, English, and German (for a good summary, see Ingram 1989). These focused on the individual linguistic development of a child observed by a parent interested in language acquisition. The diary studies varied greatly in detail and focus and also in the age range of documentation.

Diary studies are especially useful in the early phases of acquisition when children speak very little, at unpredictable times and intervals during the day. The study needs to be conducted by the main caregiver who spends the day with the child. In diary studies of very early development nearly all utterances of the child can be recorded. This, however, becomes unfeasible as soon as the child starts to talk more. During this time, then, only newly emergent structures can be noted down. As a result diary studies usually cannot be used to answer questions about frequency distributions or quantitative questions in general. However, the exact qualitative development can be traced in noting down all occurrences of specific individual words or
constructions. This allows us to find errors, which are sometimes very rare and difficult to detect otherwise. Further, we gain a comprehensive picture of nearly every new construction or the precise development of a specific grammatical category.

The greatest problem with this method is the interpretation of how the child’s constructions relate to her environment. In diary studies there is generally a lack of information about the surrounding language, i.e., we usually do not know whether the utterance was a repetition of a preceding utterance by a conversation partner or indeed a spontaneous, unprompted occurrence. Further, we rely on the exact memory of the researcher who notes down the exact wording of the utterances. This can be challenging at times when the researcher is simultaneously the caregiver who has to look after the child. Modern recording devices, such as small dictaphones into which one can quickly repeat the exact utterance, can nowadays be of enormous help in the difficult logistics of such studies.

Diary studies have long been very useful in documenting the development of individual children, often resulting in hypotheses for developmental processes in general. Thus, even today such studies can have an immense impact, as shown by Tomasello’s diary study of his daughter’s language development (Tomasello 1992). This data led Tomasello to the insight that children start out with rigid, item-specific constructions, which only later in development become more flexible and are generalized to other constructions. These findings have strongly influenced subsequent experimental research and in fact have led to a paradigm change in acquisition research. Nowadays item-based learning is discussed on a par with rule-based learning, which was predominantly taken for granted before this study.

For comparative research diary studies are useful to acquire an overview of the first constructions learned. Further, they are very useful in learning about qualitative changes in the development of categories and constructions to compare them across languages. Thus, diary studies provide an excellent first step, especially in describing very early development. Further, they are useful for developing hypotheses about general learning processes, which then can be followed up in more detail and in a quantitative manner in longitudinal corpora or experiments.

2.1.2 Audio-Visual Corpora

In modern acquisition research, especially if we are interested in hitherto underresearched languages, we usually start out with longitudinal studies consisting of video/audio recordings of children in their natural contexts. Longitudinal studies usually follow the development of a handful of children at regular intervals over a specific time period. Ideally, the social and communicative environment of the child is covered in the recordings. This allows us to study the role of extralinguistic factors in language acquisition, such as the role of input and speech surrounding the child in general. Nowadays recordings are nearly exclusively made with video, but there are also some earlier studies that are audio only. Longitudinal studies are extremely time and labor intensive and consequently very costly. The recordings alone usually stretch from about a year to several years. Then, as a first step all the utterances by all the participants need to be transcribed, translated if necessary, and then annotated, e.g., morphological glossing, parts-of-speech glossing and/or other features of interest. These annotations involve both native speakers and linguists specialized in the language. In contrast to experiments, however, such studies can be used to investigate a wide range of research questions on various issues of grammar, vocabulary, socio-communicative behavior, pragmatics etc. There are several factors that play a crucial role in interpreting longitudinal studies. Sampling issues are relevant both for analyzing the acquisition process of an individual language and for cross-linguistic comparisons with other longitudinal
studies. There are three main types of variation, which have an important impact on the interpretation of longitudinal data in addition to the selection of target children:

- Individual variation: number of children recorded, age span of recordings etc;
- Contextual variation: number and kind of people present, types of situations recorded etc;
- Distributional variation: intervals and length of recording;
- Selection of target children.

**Individual Variation**

Longitudinal studies are always case studies of individual children. It is well known that there is large individual variation in development (e.g., Fenson et al. 1994, Bates et al. 1995) and without knowing the norm it is difficult to interpret what the behavior of an individual child tells us about the population. Just to give an illustration of the extreme variation shown in a large-scale cross-sectional study of English acquisition including 659 babies (8–16 months) and 1130 toddlers (16–30 months): “children reported scoring at the 10th percentile are reported to comprehend fewer phrases at 16 months than children scoring at the 90th percentile do at 8 months” (Fenson et al. 1994: 33). In this study similar variation was encountered for all linguistic variables tested. Thus, without knowing about the norm it is difficult to interpret the development of a small number of children, i.e., we do not know how their development compares to the population. These individual differences also make it problematic to pool the children of the same age and generalize the developmental path or time scale if we only have a small number of participants. Consequently it is difficult to generalize from the sample to the whole population for which the norm is unknown.

There are, however, ways to circumvent these difficulties. One possibility is to conduct incremental comparisons (Stoll & Gries 2009). In such an incremental approach, individual differences are placed center-stage. The idea here is to first analyze each child individually over the respective longitudinal period. Second, the child is compared to his/her surrounding adults. Then, in a third step, the individual child-adult dyads/polyads are compared. This allows us to determine whether there are similarities in the case studies both within and also across languages. This method also allows us to take into account the different age spans investigated in different studies. We know that age is a poor predictor of development and the incremental approach takes care of this variable by individually comparing the path of development independent of the exact age. Incremental comparisons thus allow a step-by-step comparison and subsequent potential generalizations of developmental paths.

**Contextual Variation in Recordings**

A further challenge with case studies is that recordings usually constitute only small snapshots of the activities in the daily life of a child. These specific contexts are not necessarily representative of the typical activities a child is engaged in during her day. This is highly relevant because we know that both conversational partners and activities taking place during recording have an impact on the speech of the child and the caregivers. They influence the type of constructions children and adults produce and they have a strong impact on how explicit the interaction partners are.

In an early study of communicative skills of 4-year-olds, Shatz and Gelman (1973) have found that the children in their study adjusted their speech to their interlocutors. If they talked to adults their constructions were longer and more complex than if they talked to their 2-year-old siblings. This shows that talk to peers or also younger siblings varies significantly from talk to
adults, not only in what is said but also in how it is said. It seems that this sensitivity starts very early on in development, even at a younger age than found in the Shatz & Gelman study. Ochs (1988) found that 2-year-old Samoan children switch register as indicated by different pronunciations depending on their interactional partners. This shows a very early meta-awareness of socio-communicative contexts. A recent experimental study by Hoff (2010) confirmed that both the context and the interactional partners had an influence on children’s speech, even though the individual differences remained constant across contexts. Compared to free play and mealtime, during book reading the language of the children (age range 1.5–3.0) was more complex with respect to vocabulary and overall more coherent. Further, there were significant effects on the speech of the children depending on their conversational partner. If children spoke to their mother their speech was much richer in vocabulary and they were more likely to reply to questions than if interacting with a sibling (Hoff 2010: 468). These observations show that the recorded contexts indirectly influence our results.

In many of the available longitudinal corpora, however, children interact exclusively with their mother or their primary caregivers. The caregiver plays with the child or engages him/her in conversation or play for the duration of the recording. Often no other adults or children are present. This makes the transcription and the preparation of the data easier but such a setting has the great disadvantage that the recordings are not necessarily representative of a child’s “normal” activities and conversation partners during her day. Further, all other contexts and conversational partners of the child and the parent, which are part of the natural context of growing up, are excluded. Thus, with such recordings we can only study child-directed speech in a very specific context, namely the child interacting with the caregiver. The role of other interactions or the influence of child-surrounding speech cannot be studied. Features of the context thus need to be counted as variables in comparing different corpora within and across languages.

**Distributional Variation**

We know that the frequency of occurrence of linguistic items and constructions correlates with acquisition. Tomasello & Stahl (2004) have illustrated impressively that the recording scheme and the amount of recording has an enormous impact on our results. Since linguistic features differ strongly in their frequency of occurrence in conversation, our sample size, in addition to the recorded variation of contexts, influences our chances of recording a phenomenon. Consequently, our results about the interrelation of acquisition and occurrences in the input largely depend on the reliability of our sampling regarding both content and size. If we are interested in a phenomenon that occurs only rarely, the probability of catching it when first produced correlates positively with the amount of recording per session in addition to the recording scheme.

Recording schemes vary from weekly to biweekly or monthly recordings. Usually, children are recorded for an hour per week or every fortnight at least. This recording scheme has severe sampling problems, as illustrated in detail by Tomasello & Stahl (2004). Rare phenomena are difficult to catch in such small corpora. As a result they are often caught only several months after their real first occurrence, which happens not to be during recording time. Rowland & Fletcher (2006) have confirmed the role of sampling size in the detection of errors and item-specificity in constructions. Following Tomasello & Stahl (2004), it turns out that the traditional recording scheme of 1 hour per week is much less suitable for characterizing overall development than, say, recording once for 5 hours per month. These five hours can be distributed over several recording sessions within a single week but the chance of catching rare phenomena rises significantly. The change in sampling strategy has the additional positive effect that we might
gather different situations in the daily life of the child. This recording scheme is also much easier to apply in fieldwork situations and will give a broader picture of the daily life of a child than weekly snapshots of 30 minutes to 1 hour, which tend to be repetitive in contexts if there is always a fixed schedule when the child is recorded.

Ideally of course, the amount of hours in weekly recordings is increased. This has been done in a few so-called high-density studies with up to 10 hours of recording per week (see Lieven & Behrens 2012). They even provide a better picture and less time lag in the detection of items learned by the child. Further, there is also one extreme study called the Human Speechome Project (Roy et al. 2009), in which all the language of one target child from birth to three years and the language he hears at home is recorded, resulting in several hundred thousand hours of home recording. This is an exceptional study with challenging data preparation, which is still ongoing. So far, it is not feasible to conduct such ultra-high-density studies in other languages or with a larger sample of children, which would of course constitute a great step forward in acquisition research. Our main challenge to date in working with naturalistic data is to take care of sampling variables and contexts of sampling to make sure the speech we collect is as representative as possible of the child and the environment of the child.

Selection of Target Children

We know from a number of studies that in addition to individual variation the socio-economic status (SES) of the families has an effect both on child-rearing behavior, which affects the development of the child (Hart & Risley 1995, Huttenlocher et al. 2002, Hoff 2003). Effects of SES on cultural attitudes and child-rearing behavior seem to occur across cultures. This, at least, is the case in societies with significant differences in income and education. The effects of SES might be less pronounced in societies with less educational stratification, such as rural societies living on subsistence farming. Nevertheless, this potentially biasing factor needs to be counted in when choosing children for a study.

2.1.3 Available Data: Cross-linguistic Sample of Longitudinal Corpora

A very important initiative by MacWhinney & Snow (1985) resulted in the collection of a large database of acquisition corpora that are openly available to the community (http://childes.psy.cmu.edu). This has been a great step forward for comparative language acquisition research and actually for the first time has made possible comparative studies on a wider range of languages. Still, the sample of languages for which we have corpora is relatively small, constituting not even 1 percent of the languages of the world (Lieven & Stoll 2010, Stoll 2009). For 39 languages corpora are openly available and for approximately a dozen other languages there exist corpora that are not open to the public. The goal of language acquisition research, however, is to make general statements about how children learn language. Thus, we need to be sure that our claims hold for all languages, and not only for a potentially biased subset.

In the currently available corpora there is indeed a strong family bias, of which we need to be aware in order to interpret and generalize our findings adequately. The family bias concerns the type of languages for which we have corpora. There are around 400 language families and a large number of isolates, all of which are learned by children growing up in their respective environments. Acquisition mechanisms or processes need to apply to any of these languages. Given the fact that only a tiny fraction of some features of the world’s languages have been studied with regard to their acquisition and 54 percent of the 39 open access corpora on Child Language Data Exchange System (CHILDES) are from the Indo-European language family,
mainly spoken in Western Europe, generalizations are tentative so far. These languages are very similar in many respects and this leads to an overrepresentation of structures typical of Europe but rare elsewhere (e.g., relative clauses introduced by relative pronouns, cf. Haspelmath 2001).

2.1.4 A New Approach to Cross-linguistic Sampling

Given these current data limitations, we are not yet in a position to make statements about language acquisition in general but, rather, primarily about some acquisition processes in a few languages, mainly of the same family. No acquisition theory has so far been tested, therefore, against a representative sample of languages. This, however, is definitely the main goal we need to aim for in future research. Since it is very unlikely that in the near future we will ever be in a position to have an adequately stratified sample of acquisition corpora, including all language families and important features, more creative sampling procedures are required. Only then will we be able to find out whether fundamental differences in grammar have repercussions on the nature of language acquisition processes. There are two constraints on sampling languages for comparative acquisition research:

- Maximize the diversity of linguistic structures in our sample. The reason is that we are interested in finding out what constitutes the flexibility of the human mind and whether there are some universal cognitive mechanisms that apply independent of the grammatical build-up of individual languages.
- Minimize the number of languages for which we have to sample new corpora. The reason behind this constraint is a practical one, because gathering a longitudinal corpus is a highly time- and cost-intensive endeavor, especially if fieldwork of hitherto undescribed languages is involved.

To escape this methodological impasse we recently proposed a new typological approach with a new sampling method (Stoll & Bickel 2012). The method is called Maximum Diversity Sampling with the purpose of maximizing the diversity of grammatical structures in our sample. For this we apply a fuzzy clustering algorithm (Kaufman & Rousseeuw 1990) to large typological databases (from the AUTOTYP system, http://www.autotyp.uzh.ch, and from the World Atlas of Language Structures, http://wals.info, following standard procedures in typology). This clustering algorithm focuses on ten variables that are important for acquisition, such as the presence and absence of agreement, case marking, word order, degrees of synthesis, polyexponence, inflectional compactness of categories, syncretism, and existence of inflectional classes. The clustering was set to find five maximally diverse language groups, which seems to be a reasonable number of language groups to deal with and, if necessary, to collect corpora (Stoll & Bickel 2012). This corpus of maximally diverse languages then allows us to test potentially universal acquisition mechanisms and draw more general conclusions than has been possible so far. Currently a large-scale cross-linguistic comparative project has started working with 10 languages from these five maximally diverse groups of languages (two languages per cluster so far). The project aims to find general features of the input and learning mechanisms that allow the child to learn languages by extracting information from the input (see www.acqdiv.uzh.ch/index.html).

2.2 Experiments

Experiments are the other main method used in acquisition studies and for a number of questions they are the only option. For instance, for comprehension studies testing prelinguistic
children, or studies about cognitive or socio-communicative abilities in very early development, experiments are unavoidable.

Experiments allow us to test clear hypotheses, clearly control for potentially relevant variables, and ultimately make claims about cause and effect. Further, we have large enough groups to gloss over individual differences and test for age differences, which is important in developmental studies. Controlling variables is an important and already difficult issue, but in cross-linguistic research, to hold this control constant across different cultural groups can be a challenge. For instance, in an experiment in which the stimuli are short video clips with people acting on some objects, we need to make sure that there is no difference in the familiarity with these objects, clothes people wear in the film, or even the fact of being familiar with films in general. Stimuli that are adequate for one cultural group are not necessarily equally adequate for other groups; this needs to be considered as well. Further, instructions need to be held constant across languages but this can be difficult with languages having or lacking specific grammatical features. If we study a language with classifiers vs. another language without classifiers, for example, the instructions themselves can influence the results. All these variables can have an impact on the answers of the child and hence on the results we obtain. Further, experiments about linguistic development are usually impossible to conduct in languages for which we have not yet studied the acquisition process at all. Experiments are, therefore, usually a second step after we have gained some overall knowledge about the acquisition process.

### 3 Critical Issues and Topics: Factors Influencing the Acquisition Process

We have seen that there are substantial individual differences in linguistic abilities across children of the same age groups within the same culture. These individual differences concern not only the size of the vocabulary but also when, how, and maybe even whether a grammatical construction is learned at all. It is well known that there are substantial differences in the vocabulary of children and even adult native speakers. Up to very recently, grammar was assumed to be uniform in all native speakers. Recent research, however, has challenged this claim by showing that we also find a large amount of heterogeneity in grammatical competence among adult native speakers (Dąbrowska & Street 2006). There are two potential reasons that have taken center-stage in the literature to account for such individual differences, both in language learning and in resulting adult language competence: innate differences in cognitive capabilities on the one hand and differences in the environmental variables of learning opportunities on the other. Innate abilities relevant to vocabulary and grammar learning could, for instance, consist of attention span, memory abilities including both short-term and long-term memory, pattern detection abilities, differences in processing speed etc., which are relevant for abstracting and generalizing lexical patterns into grammar. Other reasons might be differences in environmental factors such as the structure of the language to be learned and the amount and kind of input children receive. We start with some universal, probably innate milestones and then move on to environmental factors influencing acquisition.

#### 3.1 Milestones: Early Predispositions for Language

We know from acquisition studies of a number of typologically different languages that there are some very basic cognitive, linguistic, and socio-communicative milestones that seem to be reached around the same time window in development, no matter which language a child is
exposed to. These milestones mainly consist of prelinguistic abilities, on the one hand, and some very general observations about the time course of acquisition, such as the time span within which children learn the basic grammar of their language, on the other.

Good candidates for innate predispositions to acquire language are: categorical speech perception, pattern recognition, statistical learning, intention reading, joint attention, and imitation. Infants are born with the ability to perceive human speech sounds categorically (Eimas et al. 1971), but this ability is not specific to humans. However, the perceptual learning in the first year of life as a result of language input is presumably unique to humans (Fernald & Marchman 2006: 1038). This specific and constant input of their mother tongue/s presumably also leads to the loss of the earlier ability to discriminate all possible phonemes at around 10 months of age (Werker & Tees 1984). We also know that this massive input already has an impact on the phonetic perception of 6-month-old infants (Kuhl et al. 1992). Another extremely important capability has been discovered by Saffran et al. (1996), namely the ability, at 8 months, to learn statistically, i.e. to learn to segment syllables into words merely on the basis of transition probabilities. This is clearly a prerequisite for finding recurring patterns in the input, which is the basis of all word and grammar learning. Further, to ultimately learn language, patterns need to be remembered as well. There are some general developments of long-term memory in the first year of life. Rovee-Collier (1999) found that short-term memories of 2- and 18-month-old children tested in a nonlinguistic task were similar, but long-term retention develops significantly in the first year. Interestingly retention can be improved by adapting the training schemes (Rovee-Collier 1999: 81), which is very relevant for the role of linguistic input to the child.

At the same time a number of socio-cognitive abilities develop. This period has become known by the term ‘9-month revolution’ (Tomasello 1995). Around 9–12 months of age, children start to recognize symbolic gestures and that words are symbols (Bates et al. 1979), further understand joint attentional triadic frames, and develop the comprehension of intentions (Tomasello 1995). In addition, children around that age start imitating with role reversal (see Tomasello 2003 for a summary). This implies that children understand others as intentional agents, whose actions are then imitated. Role reversal is important for language learning: for instance, personal pronouns cannot be simply imitated if used by another speaker but need to be adapted to the perspective of the speaker. Another milestone relevant for communication is the child’s understanding and her own use of pointing, which is one of the earliest and probably universal steps in communicating with the environment. In an experimental study of children growing up in very different cultures, including two rural communities, one in India, the other in Peru, and an urban community in Canada, Callaghan et al. (2011) found very similar behavior regarding a number of socio-communicative skills, such as pointing and joint attention. Similar results were obtained in a longitudinal study comparing children in rural Germany with children growing up in Chintang, a village in Eastern Nepal (Lieven & Stoll 2013). In this longitudinal study of two very different cultures, no correlation was found between the onset and the amount of pointing, and the pointing directed to the child.

Usually around their first birthday, children also learn their first words, after going through some universally similar stages of babbling. Around 6 months of age they start with reduplicated babbling and around 9 months their babbling sounds somewhat similar to their native language (Oller 2000), i.e., they imitate its intonation and stress patterns.

Independent of language, vocabulary strongly increases in the second and third year of life. Around age 4 most children master the basic grammar of their native language. However, how children extract the necessary information from the input to build up grammar is far from being understood and research is just at the beginning.
3.2 Environmental Factors Influencing the Acquisition Process

The role of environmental influence on language learning has been one of the hot topics of language acquisition research. We know from studies of a wide range of languages and cultures that there is a large variation in the interactional patterns between adults and children, both in the type and amount of interactions and in cultural attitudes expressed in ethno-theories about child rearing (e.g. Ochs & Schieffelin 1984, Keller 2007). The relevant question is whether these differences have an impact on how and when language is learned. In the following I concentrate on two features that take center-stage in acquisition studies: (i) linguistic characteristics of the input and (ii) amount and type of input.

3.2.1 Linguistic Features

The world’s languages vary to an extreme degree on all linguistic levels. Even though there is a substantial number of language families with similar grammatical features, linguistic typology has recently forcefully questioned the idea of absolute universals that are valid for all languages (Bickel 2007, Evans & Levinson 2009). Even the most basic candidates for universals have been called into question lately.

Not even phonemes (duality of patterning, Sandler et al. 2011), nor words (Schiering et al. 2010), nor the distinction of nouns and verbs (Peterson 2010) are universal. It appears that each language has a unique combination of features, none of which seem to be universal. This means a huge design space for human language, which in principle every child can cope with. For instance, the number of sounds in languages ranges from a very minimal inventory of six consonants in Rotokas (West Bougainville, Papua New Guinea) to a very large inventory of 122 consonants in !Xóõ (Southern Khoisan; Botswana) (Maddieson 2005). Some of these sounds, such as clicks, are very complex to produce and are learned late. On the morphological level there is again huge variation ranging from languages with hardly any morphology, such as Chinese with zero verb forms, to extremely complex morphology such as Chintang (Nepal), with more than 1800 verb forms (Bickel et al. 2007). Both of these languages are from the Sino-Tibetan family, thus, even within language families, there can be extreme diversity. The challenges for children learning Chintang vs. Chinese vary to an extreme extent. Whereas a Chinese-learning child does not have to generalize a verbal paradigm at all, the task of a Chintang child is to extract the verbal paradigm in somehow encountering all 1844 forms and then generalizing the paradigm. Recent research on the acquisition of the verbal paradigm in Chintang suggests that children learn verb forms in an item-specific way before becoming productive verb users (Stoll & Bickel 2013).

Children learning languages with conjugation classes, as for instance in Polish (Dąbrowska 2001) or Finnish (Kirjavainen et al. 2012), still have different learning challenges to meet. These studies on languages with different types of morphologies suggest that complexity is an important factor for the learning process and needs to become the focus of future comparative work.

3.2.2 Input

One of the most burning questions in acquisition studies has been how much input is necessary for a child to build up the grammar and the vocabulary of her native language. There are several strands of research that provide insights into this question. First, there are cross- and within-cultural differences in how much children are addressed. Second, there are some potentially universal features in the structure of the input.
Amount of Input

We know from a number of influential longitudinal studies and also experiments focusing on the input children receive that there are enormous differences in the amount of language children hear per day. There are two main factors that have been shown to play a role in differences in input: cultural differences and socio-economic differences.

The way infants are addressed in different cultures and the ethno-theories behind these behaviors vary widely. As shown by Ochs & Schieffelin (1984) in an ethnographic study comparing how infants are addressed in three very different cultures during their first two years of life, “the biological predispositions constraining and shaping the social behavior of infants and caregivers must be broader than thus far conceived in that the use of eye gaze, vocalization, and body alignment are orchestrated differently in the social groups . . .” (Ochs & Schieffelin 1984: 299). In comparing the socialization of Anglo-American white middle-class babies with Samoan and Kaluli they find wide-ranging differences in the amount and kind of interactions with prelinguistic infants. Children growing up in US middle-class families are usually the center of attention and treated as full interactional partners right after birth. Adults in this culture accommodate their speech and behavior to the children and apply a self-lowering strategy in constantly interpreting possible intentions, even though the infants are far from being able to express any of these “intentions” themselves. This is very different from the environment of Kaluli and Samoan children, who are widely ignored in their prelinguistic phase and only start to be treated as interactional partners when they utter some culture-specific words (Ochs & Schieffelin 1984).

The studies of these three cultures were based on participant observation. These children, however, are not only surrounded by their primary caregiver and other adults but also by a large number of other children. So far, we know very little about how much interaction in direct input they indeed receive. We also know that there are at least some greeting rituals and there is also information about some explicit teaching strategies, which include prelinguistic infants in interactions. When other children address the infant in Kaluli, mothers face the baby toward the interlocutors and speak for the baby in a high-pitched voice (Ochs & Schieffelin 1984: 289). To fully understand the linguistic environment that infants in Kaluli and also Samoa encounter, quantitative studies, including the behavior of other children, would be instructive. Studies on a wider variety of cultures are clearly a desideratum in this area of research.

One other important factor found to correlate with the amount of input children receive is the socio-economic status (SES) of the families in which the children grow up. There has been a number of studies on the amount of input English children of families with different socio-economic status receive (Hart & Risley 1995, Hoff-Ginsberg 1991, 1998, Huttenlocher et al. 2002). Hart & Risley (1995), in a large-scale longitudinal study, have shown that the difference in terms of words that children of different SES groups encounter over a period of three years amounts to more than 3 million words for each child. Children of lower SES have a much lower input than children from higher SES groups. These studies also suggest that the overall amount of language children hear has an impact on the size of vocabulary and also on the grammatical development of the child. Hoff-Ginsberg (1991) has shown that there is not only a difference in speech directed to children in the different SES groups but she also found parallel social class differences in adult-to-adult speech. A further important finding of Hoff-Ginsberg’s study was that the context of recording (mealt ime, dressing, book reading and toy-play) had a significant effect on the language addressed to children. For instance, in contexts like book reading there were fewer differences among the different SES groups than in, for example, free play. Thus, the variable context plays an important role in the amount and also the type of language children receive. A recent study by Fernald et al. (2013) has shown that SES differences
even have a strong impact on early processing and also on vocabulary learning: differences in processing were already evident at the age of 18 months. At 24 months of age the differences in the size of vocabulary were striking. Children from lower SES families lagged around 6 months behind their peers, both in processing skills and in vocabulary. Also, at later stages, with respect to the development of complex syntax there, is a significant correlation between the number of complex sentences a child hears and the child’s mastery of multi-clause sentences (Huttenlocher et al. 2002).

All these studies strongly indicate that the amount of language a child hears has an impact on her linguistic development. This fact seems to be related to differences in adult competence, which turns out to be heterogeneous as well. Dąbrowska & Street (2006) have shown that we might have to revise the assumption that adults have a unified grammatical system. In comparing the ability to understand complex sentences, they found considerable individual differences correlating with education and socio-economic status of the speaker.

Type of Input

It is well known that not only the amount of input children receive varies considerably but also the kind of input for children of different ages and social groups.

One of the prerequisites for learning language seems to be participation in interactional communicative frames. As claimed by Snow (1984), frequent adjacent dyadic interactions with the child by the mother or a caregiver allow the child to predict the relevance of such interactions and ultimately understand the purpose of language. The question is whether these early interactions somehow facilitate language learning. It has been claimed that there are some prominent linguistic and potentially universal linguistic features of child-directed speech in these interactions, namely: frequent repetitions, prompting, slower speech, expansions of what the child said, shorter and fewer ungrammatical utterances (e.g. Snow & Ferguson 1977, Brown 1998). In many cultures very direct markers of such interactions in the form of prompts (‘say X’) and recasts (immediate use of the same word in a different construction or form) have been found; and features of early child-directed speech might help the child to parse and remember individual words and grammatical features (Farrar 1990, 1992). Interaction has thus been claimed to play a key role in helping children to understand that words have meaning. Then, later on, they learn words and grammatical constructions in repetitive interactions. Such interactions make use of all the of the milestones children acquire around 9–12 month of age, i.e. joint attention, intention reading and imitation.

As shown by Kuhl et al. (2003), even phonetic perception requires interaction. On the level of verb learning the importance of interaction for acquiring Tzeltal (Mayan, Mexico) verbs has been shown by Brown (1998), who states that Tzeltal mothers rarely talk to their prelinguistic infants and there is hardly any motherese observed. However, later on, verbal interactions between child and caregiver become crucial in learning Tzeltal grammar. Brown claims that to learn verbs children have to rely exclusively on “distributional facts about the linguistic context” (Brown 1998: 199). She hypothesizes that dialogic repetition is an important feature for learning the complex verbal system of Tzeltal. Prompts, expansions, and repetitions are a feature of the dialogic exchanges between the child and the caregiver. A similar observation was made by Küntay & Slobin (1996), focusing on what they call “variation sets,” i.e. repetitive variations of the same theme in interactions between a caregiver and a child (for further elaboration of this observation see Onnis et al. 2008). A recent quantitative study on the role of interactions looked at Chintang children learning the ergative case (Stoll & Bavin 2013). In this large-scale longitudinal study it was shown how repetition has a quantitative effect in learning the ergative case. First, children use the ergative case predominantly as an exact repetition of the form used by an interlocutor in the
same interactional frame. Then, later on, they use ergatives, which are not a direct repetition but prompted by the preceding utterance of an interlocutor that includes an ergative. Then, as a last step they use the ergative more frequently and spontaneously, without being prompted in any way. These direct interactions thus seem to have an impact on the learning process.

Further, the input itself seems to be structured in such a repetitive way that it potentially facilitates language learning. We know from the analysis of at least three languages (for English see Cameron-Faulkner et al. (2003), for a comparison of English, German, and Russian see Stoll et al. (2009)) that the input children receive is very repetitive, at least at the beginning of utterances. To see whether this is a specific feature of child-directed speech we would need some comparisons with the speech surrounding the child. These are tasks for the future, however, which will have an important impact in showing whether there are indeed strong linguistic differences between child-directed vs. adult-to-adult speech.

Not only do interactions themselves have structure but also the way in which these interactions are carried out. Since the early 1960s (Ferguson 1964) it has been well known that both the language addressed to infants (infant-directed speech), also known as motherese or baby talk, and the language addressed to small children, nowadays most often referred to as child-directed speech (CDS), differs significantly from the speech adults use among each other. These differences concern both prosodic and linguistic features and, of course, the content. The most prominent prosodic features are: higher pitch, larger pitch contours, longer pauses, slower tempo (Snow & Ferguson 1977, Fernald & Simon 1984, Fernald et al. 1989). Fernald & Simon (1984) found a significant difference between the language mothers addressed to their newborns and the language they used to address other adults; in particular, prosodic patterns rarely found in adult-to-adult speech were used frequently in infant-directed speech. Fernald et al. 1989 have argued for a universal set of prosodic modifications in infant-directed speech. In comparing speech modifications directed to British English, American English, German, French, and Japanese infants, they conclude that there is a universal set of features based on innate predispositions in the speech addressed to infants.

Bernstein-Ratner & Pye (1984), by contrast, argue that conventionalized, culture-specific features determine or at least influence the way adults address infants (for a summary of the controversy see Ingram 1995). In analyzing the speech addressed to Q’iche Mayan infants Bernstein-Ratner & Pye found no higher pitch in child-directed speech, but rather, on the contrary, higher pitch was used in adult-to-adult speech. Higher pitch in this culture is reserved for people with higher status, so there seems to be at least a tendency to use a different pitch for adults than for infants: whether it is higher or lower pitch depends on other cultural variables. Reports on Tzeltal (Mayan) by Brown (1998) suggest that mothers do not use motherese in this other Mayan language. Instead of talking in high pitch they whisper to their prelinguistic infants if they address them at all. Thus, these features of child-directed speech are not universal and hence also not necessary in the sense that not each and every one of them needs to occur in a language for children to be successful in learning this language. However, when occurring, these features are most likely to be helpful for the acquisition process (Nelson et al. 1989). At the very least they help catching and directing the attention of the child, which is a prerequisite to learning. This hypothesis is strengthened by the fact that infants seem to prefer infant-directed speech to adult-directed speech (Cooper & Aslin 1990, Fernald & Kuhl 1987) and this even holds if the adult speaks in a different language than the mother tongue of the baby (Werker et al. 1994).

Even if there is no clearly delimited number of features characterizing the input of any language, there is a strong suggestion that a combination of a number of these features in the input facilitates the acquisition process. To learn more about such a potential set of universal features that could be a construction kit for the input to the language-learning child, we need to know
more about the input in maximally diverse languages. This is indeed the most pressing task for future research as I see it.

4 Future Directions

Cross-linguistic language acquisition research is a relatively young field and it has made great progress in recent decades, starting out with initiatives like Slobin’s cross-linguistic studies of a wide range of typologically different languages (e.g., Slobin 1985, 1992, 1997), Berman & Slobin’s (1994) comparative study of narrative structures, and a large number of other cross-linguistic studies.

Further, the availability of a number of different corpora on CHILDES, including some languages from different language families, has had a major impact on the field. The available data is only a tiny sample, however, a fraction of what constitutes the design space for human languages. To truly understand acquisition we need a systematic picture of how the most diverse features of grammar are learned. Now, one of the most important future goals is to learn more about the flexibility of the language-learning mind. This entails concentrating systematically on languages with maximally diverse structures rather than studying languages of one family only or choosing languages by convenience sampling. This will then ultimately allow us to make claims about language acquisition processes in general rather than only about language-learning processes of individual languages. A first step in this direction was the development of a sampling method to identify maximally diverse languages (Stoll & Bickel 2012), which was a prerequisite to systematically studying the flexibility of the human mind. The next step, which has been undertaken, is a large comparative project actually comparing acquisition processes in these languages (see http://www.acqdiv.uzh.ch/index.html).

As detailed above, we know that children learn from their environment. To find out how they really do this, we need to know their socio-communicative environment in rich detail. To achieve this we should move away from classical studies focusing on semi-natural recordings of the child with her mother or some other close caregiver. Rather, we need to widen our research space by recording as much of the daily lives of the language-learning child as possible. This will expand our knowledge about the role and the structure of the input in general rather than focusing on a more or less artificial snapshot of dyadic interactions with her mother, a context that is often only a small part of the daily life of a child.

These goals, however, entail going out to the field and collecting acquisition data in as many diverse languages as possible, thereby taking into account the sampling issues mentioned in this chapter. For this we need close collaboration with linguists, psychologists, and anthropologists. Such an interdisciplinary approach was pioneered in a large-scale longitudinal study of Chintang, an endangered Sino-Tibetan language spoken in Nepal (cf., http://www.clrp.uzh.ch).

Related Topics

7 Language Ideologies (Kroskrity); 8 Social Subordination and Language (Huayhua); 9 Language Socialization (Paugh); 11 Language Socialization and Marginalization (García-Sánchez); 22 Language in the Age of Globalization (Jacquemet).

References

Studying Language Acquisition


Studying Language Acquisition


Further Reading