15 Dense Sampling

Elena Lieven and Heike Behrens

Summary

This chapter describes the methods used to develop much denser samples of children's naturalistic speech than have previously been available. Most longitudinal corpora capture an estimated 1–2% of children's speech. Depending on the exact sampling regime, the new dense corpora capture an estimated 7–15%. Dense sampling is important in assessing the productivity of children's grammars and in the collection of rarer structures. It also allows much more reliable quantitative comparisons between the input and the child's developing system as well as the use of computational and modeling methods that cannot be used with corpora of smaller sizes. The collection and transcription of these denser corpora is complete or under way for English, German, Polish, Japanese, Estonian, and Finnish. The limitations of the method are the extensive resources that are required and, given this, the fact that corpora can only be collected from a very small number of children.

Research Aim

Naturalistic recordings of children's speech are an essential part of the study of child language development. Experiments, of their very nature, control what it is the child has to either understand or produce, and diary studies can only focus rather narrowly either in time and/or on a particular construction or phenomenon. Therefore periodic longitudinal recordings of children talking have always had a central place in the study of language acquisition, especially since they provide

Research Methods in Child Language: A Practical Guide, First Edition.
Edited by Erika Hoff.

information not only about the children's own language production (arguably an insufficient cue for the assessment of their competence) but also about the input the child receives as the basis for the generalization over language, and the interactional processes that may help the child to identify the relevant information in adult speech. Thus a corpus, once transcribed, can serve as the resource for addressing many different research questions, including the interaction between different strands of development, whereas an experiment needs to reduce the number of variables involved. The main constraint in setting up a corpus is the time involved in transcription, especially when the children are young and their speech is not very clear. This has meant that almost all longitudinal corpora can only capture a very small proportion of a child's waking life, typically 1–2 hours every 2–3 weeks at best. Dense databases (DDBs) aim to achieve a much higher sampling level (of between 5% and 20%).

Although DDBs also collect only a proportion of the child's active and passive linguistic experience, they can fulfill an important role in dealing with the problem of thin sampling. Thin sampling presents two major problems. First, there is the possibility that the absence of a structure in the child's transcribed speech is due to its rarity rather than to the fact that the child has not yet acquired that structure. Clearly researchers could radically underestimate the child's developmental level in these circumstances. A second problem acts in the opposite direction: the rare appearance of a complex utterance in the child's corpus could be taken as evidence of acquisition when in fact it is partially or wholly rote learned, but there are not enough data from either child or interlocutors to check this possibility (Rowland and Fletcher, 2006; Rowland, Fletcher, and Freudenthal, 2008; Tomasello and Stahl, 2004). Of course the ideal would be to collect everything that a child says, but this is an even more resource intensive undertaking which only the *Human Speechome Project* has realized (Roy, 2009; see Naigles, Chapter 16 this volume).

For What Population Is It Suited?

Dense databases can be collected for all populations and types of language acquisition. For practical reasons (see below), dense sampling is easier and timewise better if children spend more time at home rather than full-time nursery or school. It could be particularly suited to the longitudinal study of children who talk and communicate little: they will familiarize to the recording situation, and dense sampling will yield a better picture of their communicative and linguistic profile than thin sampling. Another application is the recording of children growing up multilingually, if the sampling was well distributed over the various linguistic environments. However, collecting DDBs, especially continuous ones, requires a deep commitment by the families involved. Thus when recruiting families, attention has to be paid to the stability of life circumstances. The requirements of such a study self-select families who take an intense interest in the development of their children and are willing to arrange their lives according to the requirements of the study.

What Can Be Learned from the Application of This Technique?

DDBs make possible a number of analyses that cannot be obtained with experiments or thinner samples. Obviously, DDBs are extremely useful for tracing the acquisition of infrequent structures such as relative clauses or complement clauses (e.g., Brandt, Diessel, and Tomasello, 2008; Brandt, Lieven, and Tomasello, 2010). They provide the data necessary for good descriptive accounts and fine-grained analyses of developmental processes.

DDBs also allow one to check the conclusions of previous studies based on much thinner sampling. Apart from just providing more extensive data, DDBs are needed to assess hypotheses derived from linguistic and developmental theory. A long-standing debate in acquisition theory concerns the interpretation of first and single occurrences. While researchers from a nativist perspective argued that a limited number of instances were sufficient evidence of the child's competence to produce the underlying structure, developmental psychologists and usage-based linguists claimed that more evidence is needed to support such far-reaching conclusions: in terms of development, it was observed that children generalize to new structures only very carefully such that early instances do not automatically provide evidence for a full-fledged adult-like representation. Here, the distributional analyses possible with DDBs allow for a fine-grained assessment of productivity, because we can not only identify first usages of a particular linguistic structure, but also assess whether a structure develops in a lexically specific fashion or in a more general fashion. Of particular interest are the changing representations within individual children (not just averages over different groups): only dense data allow us to analyze whether a particular development takes place across the board or where it starts. Regular corpora soon become surprisingly thin once the level of analysis reaches a certain fine-grainedness (e.g., correlation of morphology and lexicon) and this makes it difficult to produce reliable quantitative assessments of the relative frequency of different forms.

Results showing the lexical specificity of child language are in line with findings from usage-based and corpus linguistics in which actual language use is studied based on "real-world" texts (e.g., newspapers, internet, conversations). First, usage-based theories of language claim that language structure is shaped by language use (e.g., Bybee, 2010). This implies that the input data available to the child should provide the necessary positive and indirect negative evidence for the structures of the target language. Second, findings from corpus linguistics show that, for many linguistic structures, we find a skewed distribution of lexical items, and it has been argued that this is important for the process by which children develop a particular structure. For instance, the verb *get* allows multiple syntactic frames, so-called alternations of argument structure, and Goldberg, Casenheiser, and Sethuraman (2005) found that *get* predominantly occurs as verb–object–locative construction when the meaning is "caused motion" (*Bob got the ball over the fence*), but as verb–object–object construction when the meaning is "transfer" (*Bob got him a cake*).

Such statistical skewings may facilitate learning and, indeed, studies show that children's early production relies on such frequency information. Training studies support the conclusion that children learn constructions best when trained with skewed input (Goldberg, 2006). The interrelationship between the lexicon and the form–meaning correspondences of syntactic constructions and their effect on children's acquisition of such constructions cannot be revealed with thin samples.

Dense databases will give a much more accurate assessment of the relationship between types and tokens in the child's speech. An example is the issue of "blocking" in the debate on morphological overgeneralization in English. Do children stop over-generalizing irregular verbs in the past tense (e.g., *goed* for *went*) as soon as they have learned the correct form – an essential part of the "dual route" model (Marcus, 1995) – or do instances of the over-regularization slowly reduce, which is more compatible with the "single route" model (Marchman, Plunkett, and Goodman, 1997)? On the basis of estimates from "thin" longitudinal corpora on CHILDES, Maratsos (2000) calculated that the latter was more likely to be the case. Using the Thomas-Brian DDB, Maslen *et al.* (2004) were able to show that Maratsos's hypothetical calculations were correct. Also relevant to this debate is the issue of whether German has a single "default" plural (governed by a rule in the "dual route" model). Using the Leo DDB, Behrens (2002) showed that the child's error patterns as well as their frequency changed over time, reflecting gradual abstraction over the phonotactic factors that determine the choice of the plural allomorphs rather than the acquisition of a "once and for all" rule or rules. This provided detailed support for previous studies with less dense data such as that of Köpcke (1998) and Szagun (2001).

Denser data also allow us to investigate the relationship between different strands of development and between the child's speech and the input. This makes an interesting comparison with the "getting it all" topical diary approach (Naigles, Chapter 16 this volume). Topical diaries are exhaustive for the items that parents are asked to record and this may allow for the assessment of productivity across that particular group of items. However they are probably of limited use for identifying the mechanisms of development in terms of the mutual influence of lexical items and constructions on one another, let alone any influences of the input which, of course, is not recorded. A good example is the use of growth curves in determining the relationship between the learning of constructions. Ruhland and van Geert (1998) point out that densely sampled data are essential for this. Abbot-Smith and Behrens (2006) used the Leo corpus to investigate the question of how learning one construction may support or hinder the learning of another – the construction conspiracy hypothesis (Morris, Cottrell, and Elman, 2000). Abbot-Smith and Behrens investigated Leo's development of German passive and future constructions containing a lexical verb and the auxiliaries with *sein* (to be) or *werden* (to become). Because Leo had acquired *sein* in the copula construction months earlier than he started using the passive, the authors predicted that the *sein*-passive would develop earlier and faster than the *werden*-passive, and this was found to be the case. Leo became (morphologically) productive with the *sein*-passive earlier than *werden*-passives, according to the criteria of higher cumulative type frequency, type/token ratios, and active–passive alternation, and his rate of assimilation of new verb types into the former was twice that of the latter. When investigating patterns of development, the

authors predicted that the form of a "nonsupported" construction should be exponential (with a slow start and the child showing an increasing ability to use the construction the more it has been previously used). By contrast, a supported construction should demonstrate sudden acquisition and a subsequently linear function, similar to that of adult baseline in which cumulative frequency grows linearly to a point where it tails off when the mother has said most of the verb types most frequent in the particular construction. Again this is what was found.

A major issue in child language is the extent to which children are innovative and productive with language. How much of what they say is simple repetition of what they have heard others say, how much of it is novel, and what is the basis of this novelty? With dense sampling the researcher has a much better chance of discovering whether there are prior instances of an utterance and, therefore, assessing the child's level of productivity. Lieven and colleagues examined this issue in a number of studies using the English DDBs with a method they called "traceback" (Dąbrowska and Lieven, 2005; Lieven, Salomo, and Tomasello, 2009). Each 6 week DDB was divided into a "test corpus" of the last 2 hours of recording in the 6 weeks and a "main corpus" consisting of the previous 28 hours. All multiword utterance types were collected from the test corpus and their relationship to utterances in the main corpus was then examined. This involved identifying possible "component units" from the main corpus that could have been used to construct the utterances in the test corpus. Two types of component units were defined: "fixed strings" and "schemas with slots." Fixed strings were fully lexically specific, while schemas contained slots as well as fixed lexical material (e.g., *There's an X, I'm X-ing it*). Slots could be identified in schemas only if they were semantically coherent, and the single or multiword strings that filled the slots also had to match the semantics of the slot. All possible matches were identified using a computer program. The program output was then used to identify semantically coherent strings and slots and to derive the potential tracebacks. Lieven, Salomo, and Tomasello (2009) were able to trace back a very large proportion (58–92%) of the four 2-year-olds' utterances in the test corpus either to exact repetitions in the main corpus or to schemas that required only one substitution into a slot. The vast majority of these substitutions were of single nouns into "referent" slots. With developing mean length of utterance (MLU), the range of other semantic slots developed, as did the variety of strings. Some of these schemas seemed highly productive, but the schemas themselves were fairly simple and the scope of that productivity was, at first, limited both semantically and structurally.

In a related study Bannard, Lieven, and Tomasello (2009) used a computational method to generate grammars from the main corpora for two of the same children at 2:0 and 3:0 (Thomas-Brian and Annie). The grammars consisted of lexically specific constructions, with or without slots, and contained no fully abstract rules. They were extracted by aligning all strings that overlapped in their lexical content. This resulted in potential slots in the alignments and the process was repeated until no further alignments were possible. The result was a large set of alternative candidate analyses which were probabilistically sampled and used to parse all the unique multiword utterances in the test corpora. At 2:0, the mean performance of these sampled grammars showed wide coverage and good predictive fit as compared with fully abstract grammars, with

the children showing radically limited productivity. The researchers also added more abstract categories such as "noun" and "verb" to the grammars. At 2:0 only the addition of the category of noun improved coverage. At 3:0, the children's productivity had sharply increased and the addition of abstract linguistic information over verbs as well as nouns markedly improved performance. Another important finding came from parsing the different corpora with the different grammars. Not surprisingly the 2:0 grammars show poor coverage and fit when parsing the 3:0 grammars. In addition, when one child's grammar was used to parse the other child's corpora, much better results were found for the 3:0 data than for the 2:0 data, suggesting that initially the children's grammars were very idiosyncratic but that by 3:0 they were starting to converge on the adult system. The important point to note is that the larger the corpora over which this procedure operates, the more reliable the results.

Procedure

The DDBs were collected and transcribed by staff of the Max Planck Institute for Evolutionary Anthropology in Leipzig and associated scientists. Families were recruited through advertisement and/or informal contacts. The sampling regimes vary in terms of both the periods covered and the number of hours of recording, though in all cases the aim was to record for at least 5 hours per week over the period of the study. We have done this for a number of English-speaking children as well as for children learning German, Estonian, Polish, Japanese, and Finnish. In the case of the last four of these languages, we collaborated with researchers working on language development in these languages (for details, see Table 15.1). The most extensive in terms of time and age range cover 5 hours per week of one German and one English boy from ages 2:0 to 3:0. Table 15.1 shows all the DDBs that have been collected and their current status in terms of hours and ages covered and transcription. The aim, eventually, is to deposit them all with the open-source CHILDES database (see Corrigan, Chapter 18 this volume).

DDBs can be combined with other means of assessment. For example, tests of the children's language development can be made (this was done at the onset of the English studies, but not for the other languages as standardized tests were not available). For the Thomas-Brian, Annie, and Leo corpora, the parents also kept a daily diary with the most precocious utterances of the day. Thus the emergence of new structures on the tape-recorded sequences can be checked against the parental notes.

Implementation

In order to avoid intervention in regular family life, dense recording regimes work best when parents can carry out the recordings themselves. The families participating in the Max Planck DDBs were equipped with a recording device and two high-quality

Table 15.1 The Max Planck dense databases

Child	*Language*	*Age*	*No of hours*	*Status of data*	*Principal investigators*
Annie	English	2:0–2:1:14	5 per week	Transcribed	Lieven/Goh
		3:0–3:1:14	5 per week		
Thomas-Brian[a]	English	2:00:12–3:00:12	5 per week	CHILDES	
		3:03:02–4:11:20	5 in 1 week/month		
Eleanor	English	2:0–2:01:08	10 per week	Transcribed	
		2:2:01–2:11:09	10 in 1 week/month		
		3:0–3:01:15	10 per week		
Fraser	English	2:0–2:01:10	10 per week	Transcribed	
		2:2:01–2:11:06	10 in 1 week/month		
		3:0–3:01:11	10 per week		
Cathy	English	3:0–3:1:14	5 per week	In progress	
		3:01:15–3:11:29	5 in 1 week/month	In progress	
		4:0–4:1:14	5 per week	In progress	
		4:01:15–4:11:29	5 in 1 week/month	In progress	
		5:0–5:1:14	5 per week	In progress	
Hannah	English	3:0–3:1:14	5 per week	In progress	
		3:01:15–3:11:29	5 in 1 week/month	In progress	
		4:0–4:1:14	5 per week	In progress	
		4:01:15–4:05	5 in 1 week/month	In progress	
Leo	German	2:0–2:11:29	5 per week	CHILDES	Behrens
		3:1:14–5:0	5 in 1 week/month		

Marysia	Polish	2:0–2:1:14	5 per week	Transcribed and glossed	Dąbrowska
		3:0–3:1:14	5 per week	Transcribed, glossing in progress	
Julia	Polish	2:0–2:1:14	5 per week	Transcribed, glossing in progress	
		3:0–3:1:14	5 per week	Transcribed, glossing in progress	
Andreas	Estonian	2:0–2:1:12	5 per week	CHILDES	Vihman/Vija
		2:3:26–2:8:13	≈ 1 per month		
		3:0–3:1:13	5 per week		
Child J	Japanese	2:0–2:01:14	5 per week	In progress	Matsui
		2:01:15–2:11:29	1 per week	In progress	
		3:0–3:01:14	5 per week	In progress	
Piia	Finnish	1:07:21–4:00:13	3 per week	In progress	Kirjavainen

[a]"Brian" is this child's pseudonym, but just before the corpus was deposited with CHILDES, the family gave permission for the use of the child's real name, "Thomas," and for the deposition of sound files as well as transcriptions.

wireless microphones that could be situated wherever the activity took place. This mobility makes a flexible schedule possible, as the recordings can be interrupted when the child is no longer willing to engage in the situation or wishes to change location. For the English and German DDBs, one recording in each week was videotaped. Experience showed that 1 hour recordings are very exhausting for 2-year-olds, and parents often sampled half an hour in the morning and half an hour in the afternoon. To guarantee privacy, parents were given the right to withhold recordings of family situations they did not feel comfortable with making publicly available. Families were paid the equivalent of 50% of a clerical salary.

In addition to sampling density, a choice about the time frame has to be made. Continuous dense recordings constitute the most labor intensive sampling regime, but provide a clear developmental picture as the development can be traced over time. Continuous recording is especially valuable if it covers a time frame of several years in order to be able to analyze later stages of development as well. The pitfall of case studies – lack of inter-individual reliability checking – can be overcome if the language development of the child in question is compared to the development of other children's data.

Dense databases can also be implemented with temporal gaps. For a number of the DDBs, 6 week periods of dense (5 h/week) or double dense (10 h/week) recordings were followed by intervals without recordings. Short periods of dense data can be analyzed for purposes of testing the productivity and distribution of particular structures.

The quality of the settings is also very important. For practical reasons, the families participating in the DDBs were asked to refrain from noisy activities, to avoid background noise, and to keep the television, radio, and other "noisy" appliances off during recording. This means that we mainly covered dyadic, sometimes triadic play situations within the families or with a few, usually well-known visitors.

Are There Any Likely Pitfalls?

The biggest pitfall is that the original scheme for transcribing and coding the data is too ambitious given the amount of time and personnel available. Thus careful choices have to be made regarding the degree of detail of the transcription (phonetic or orthographic) and the amount and detail of coding. Recent technological advances provide (partial) solutions for all these issues: if the transcripts are linked to the digitized sound file or video, a standardized orthographic transcription can always be amended with a more detailed transcription when needed. Regarding the coding of data, for English the morphosyntactic annotation program has been completed. It provides part-of-speech tagging and lemmatization (providing the "base form" of the word to facilitate lexical searches, e.g., "walk" for "walked," "walking," and "walks": MacWhinney, 2008). For other languages or other areas of linguistic interest, the situation is less good, though we are well advanced with glossing the Polish data and are starting on the Finnish and Japanese data. Standardized

orthographic transcription seems to be the only choice for dense databases, as the size of the database makes searches by reading through the transcripts impossible (Behrens, 2008).

Developments in speech and speaker recognition technology will probably help in the future to identify those segments in the recordings where speech occurs, and it may even be possible to identify the speaker. This would speed up and facilitate the transcription process (see Roy, 2009, for the implementation of such ideas). Currently, such technology is most advanced when only a limited number of speakers is present. The future will probably lie in a smart mix of automated speech recognition and perhaps partial automatic transcription and coding, combined with additional manual detailed transcription and annotation of the relevant sequences for the research question at stake.

Because of the high investment in recording and transcribing the data, the collection, transcription, and glossing of DDBs impose unique requirements in terms of sources of support, and often only more permanent institutions like the Max Planck Institutes can back such a long term investment. However, support might be more readily available from funding programs for the collection of DDBs from children who are in one way or another atypical, or for languages other than English.

It could be argued that the demands on the family of DDB recording limit the types of social and family backgrounds represented in DDBs. Despite the fact families were paid the equivalent of a 50% clerical position, there are bound to be sociological factors involved in the self-selection of parents for such a project. While this is true, it is also not obvious how it can be remedied. Again, it is always possible to check the data obtained from DDBs against corpora from other children to identify possible biases and test these experimentally.

Finally, recordings usually took place inside the house during play, book reading, or snack time. As a consequence we know little about how the child performs in unknown situations or with unknown interlocutors. Eisenbeiss (2009) argues that recording the same situations may emphasize the impression that children's linguistic performance is largely stereotypical and formulaic, and suggests that researchers bring in new stimulus material (interesting new and complex objects, new games) to stimulate talk about unfamiliar objects or situations.

Interrater Reliability

Transcribing dense databases requires a team of transcribers who must be trained and supervised throughout, as the child's language development may lead to new transcription problems. For the DDBs we opted for orthographic, not phonetic transcription and linked the utterances to the sound wave. Linking is an option in the Child Language Data Exchange System (CHILDES: see Corrigan, Chapter 18 this volume). The main problems remaining were establishing standards for domains without standardized orthography (interjections, discourse markers, but also multiword units like compounds and proper or brand names: see MacWhinney, 2008).

Before the databases were published in the CHILDES archive (again see Corrigan, Chapter 18 this volume), word lists were created to check for spelling variants. A domain that proved to be notoriously difficult to gain agreement on is the demarcation of utterance boundaries because spoken language tends to be continuous, and the speech stream does not always signal clause or sentence boundaries. Contrary to the CHILDES recommendation to transcribe in clauses (defined by the presence of a finite verb), we decided to represent the flow of the discourse and transcribed in turns rather than in clauses.

If one wants to indicate imitations and self-repetitions in order to exclude them from further analyses, criteria for what counts as repeated – just whole utterances or partial utterances (chunks), within a range of how many utterances – have to be developed, implemented, and checked thoroughly.

In case of the Leipzig–Manchester databases, the principal investigator or an experienced lab coordinator supervised the project over the whole period and made sure that standards did not change implicitly. As they were also familiar with the families and the typical topics of conversation, they helped resolved ambiguities the transcribers might have.

Future Directions for the Use of DDBs

As can be seen from Table 15.1, there is still a long way to go in the transcription and glossing of many of the DDBs. Thus the research reported here gives only an initial idea of the ways in which they can be used. In comparing adult and child it is often important to control the adult data down to the level of the child, for instance in terms of the types used and the number of tokens, otherwise one is inevitably likely to end up with the child looking less productive than the adult. DDBs allow us to do this without reducing the amount of data to almost nonexistence (Aguado-Orea, 2004; Krajewski, Lieven, and Theakston, submitted). Similarly, large corpora can be used to derive frequency and distributional information when designing stimuli for experiments, with the frequency counts much more closely matched to the ages of the children in the proposed experiment. Computational linguists are also starting to make use of the DDBs to build and test probabilistic and algorithmic models of various aspects of language learning (Bannard, Lieven, and Tomasello, 2009; Chang, Lieven, and Tomasello, 2008).

Since dense databases cover not only the child's productions but also large portions of the input in a naturalistic setting, we can study the effect of input and interaction on the emerging child language directly. A particular child's development can be studied against the backdrop of the input s/he received up to that point in time, and we can also study the effect of interaction on acquisition. For example, the role of variation sets (small adaptations of the utterances that highlight and contrast the formal properties of a particular structure: Küntay and Slobin 2002) could be explored and the size of the database would allow for more control and quantitative testing.

Another potentially important direction is that of children's semantic and pragmatic development, where thinner data may only record the most frequent usage of a word or structure. Thus Cameron-Faulkner, Lieven, and Theakston (2007) were able to analyze Thomas-Brian's uses of multiword negation in great detail and compare it to his mother's use of negation, even though some form–function mappings were quite rare. Another example is Steinkrauss's (2009) study of Leo's development of *wh*-questions. While his analysis by and large confirmed previous findings that the emergence of *wh*-word–auxiliary pairings is item based and reflects the distribution in the input, he also found that some high-frequency question types were not picked up by the child because, from the child's point of view, they were pragmatically irrelevant (e.g., rhetorical questions using mock-surprise tone).

In sum, DDBs take corpus-based studies on language acquisition to a new level because they provide the richness of data necessary to assess current hypotheses on learning processes in light of new insights from developmental and linguistic theories.

Key Terms

Corpus linguistics In corpus linguistics, samples of spoken and/or written naturalistic language use are analyzed to characterize properties of the language system and aspects of language change. This is by contrast with the method of introspection, where competent speakers judge whether linguistic structures are acceptable or not.

Crosslinguistic studies In crosslinguistic studies, a linguistic phenomenon is compared across languages to assess aspects like cognitive complexity and processability and their effect on language development.

Linguistic productivity Productive structures can generate more types of the same underlying structure. For example, in English morphology, the affix *-ly* can be used to make adjectives into adverbs. In child language, productivity measures are used to assess whether the child reproduces forms from memory or is able to create new forms based on knowledge about the underlying structure. Productivity is typically assessed by contrast within the morphological paradigm, by percentage of correct utterances in obligatory context, or by experimental studies that test children's productivity with nonce words, i.e., words they have not heard before and cannot have memorized.

Longitudinal corpora Data samples that track the development of individuals or groups over time. The sampling rate can vary from daily recordings to recordings every couple of weeks or months.

Usage-based theory Usage-based theories assume that language structure is shaped by communicative processes in so far as all linguistic units are considered to be symbolic, i.e., form–function pairings of varying degrees of abstraction. In acquisition, usage-based theories assume that language is acquired from the language use the child is exposed to.

References

Abbot-Smith, K., and Behrens, H. (2006) How known constructions influence the acquisition of other constructions: the German passive and future constructions. *Cognitive Science*, 30, 995–1026.

Aguado-Orea, J. (2004) The acquisition of morpho-syntax in Spanish: implications for current theories of development. Unpublished PhD dissertation, University of Nottingham.

Bannard, C., Lieven, E., and Tomasello, M. (2009) Modeling children's early grammatical knowledge. *PNAS*, 106, 17, 284–17, 289.

Behrens, H. (2002) Learning multiple regularities: evidence from overgeneralization errors in the German plural. In A.H.J. Do, L. Domínguez, and A. Johansen (eds), *Proceedings of the 26th Annual Boston University Conference on Language Development* (pp. 72–83). Somerville, MA: Cascadilla.

Behrens, H. (2008) Corpora in language acquisition research: history, methods, perspectives. In H. Behrens (ed.), *Corpora in language acquisition research: finding structure in data* (pp. xi–xxx). Amsterdam: Benjamins.

Brandt, S., Diessel, H., and Tomasello M. (2008) The acquisition of German relative clauses: a case study. *Journal of Child Language*, 35 (2), 325–348.

Brandt, S., Lieven, E., and Tomasello, M. (2010) Development of word order in German complement-clause constructions: effects of input frequencies, lexical items, and discourse function. *Language*, 86, (3), 583–610.

Bybee, J. (2010) *Language, usage, and cognition*. Cambridge: Cambridge University Press.

Cameron-Faulkner, T., Lieven, E., and Theakston, A. (2007) What part of *no* do children not understand? A usage-based account of multiword negation. *Journal of Child Language*, 34, 251–282.

Chang, F., Lieven, E., and Tomasello, M. (2008) Automatic evaluation of syntactic learners in typologically-different languages. *Cognitive Systems Research*, 9, 198–213.

Dąbrowska, E., and Lieven, E. (2005) Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics*, 16, 437–474.

Eisenbeiss, S. (2009) Contrast is the name of the game: contrast-based semi-structured elicitation techniques for studies on children's language acquisition. *Essex Research Reports in Linguistics* 57. http://www.essex.ac.uk/linguistics/publications/errl/errl57-7.pdf.

Goldberg, A.E. (2006) *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.

Goldberg, A.E., Casenhiser, D., and Sethuraman, N. (2005) The role of prediction in construction learning. *Journal of Child Language*, 32, 407–426.

Köpcke, K.-M. (1998) The acquisition of plural marking in English and German revisited: schemata vs. rules. *Journal of Child Language*, 25, 293–319.

Krajewski, G., Lieven, E., and Theakston, A. (submitted) Productivity of a Polish child's inflectional noun morphology: a naturalistic study. *Morphology*.

Küntay, A.C., and Slobin, D.I. (2002) Putting interaction back into child language: examples from Turkish. *Psychology of Language and Communication*, 6, 5–14.

Lieven, E., Salomo, D., and Tomasello, M. (2009) Two-year-old children's production of multiword utterances: a usage-based analysis. *Cognitive Linguistics*, 20, 481–508.

MacWhinney, B. (2008) Enriching CHILDES for morphosyntactic analysis. In H. Behrens (ed.), *Corpora in language acquisition research: finding structure in data* (pp. 165–197). Amsterdam: Benjamins.

Maratsos, M. (2000) More overgeneralizations after all: new data and discussion on Marcus, Pinker, Ullman, Hollander, Rosen and Xu. *Journal of Child Language*, 27, 183–212.

Marchman, V., Plunkett, K., and Goodman, J. (1997) Overregularization in English plural and past tense inflectional morphology: a response to Marcus (1995). *Journal of Child Language*, 24 (3), 767–779.

Marcus, G.F. (1995) Children's overregularization of English plurals: a quantitative analysis. *Journal of Child Language*, 22, 447–459.

Maslen, R., Theakston, A., Lieven, E., and Tomasello, M. (2004) A dense corpus study of past tense and plural overregularization in English. *Journal of Speech, Language and Hearing Research*, 47, 1319–1333.

Morris, W.C., Cottrell, G.W., and Elman, J.L. (2000) A connectionist simulation of the empirical acquisition of grammatical relations. In S. Wermter and R. Sun (eds), *Hybrid neural symbolic integration* (pp. 175–193). Berlin: Springer.

Rowland, C., and Fletcher, S. (2006) The effect of sampling on estimates of lexical specificity and error rates. *Journal of Child Language*, 33, 859–877.

Rowland, C., Fletcher, S., and Freudenthal, D. (2008) How big is enough? Assessing the reliability of data from naturalistic samples. In H. Behrens (ed.) *Corpora in language acquisition research: finding structure in data* (pp. 1–24). Amsterdam: Benjamins.

Roy, D. (2009) New horizons in the study of child language acquisition. *Proceedings of Interspeech 2009*. Brighton, UK.

Ruhland, R., and van Geert, P. (1998) Jumping into syntax: transitions in the development of closed class words. *British Journal of Developmental Psychology*, 16 (part 1), 65–95.

Steinkrauss, R. (2009) Frequency and function in wh-question acquisition. Doctoral dissertation, University of Groningen (GRODIL 75).

Szagun, G. (2001) Learning different regularities: the acquisition of noun plurals by German-speaking children. *First Language*, 21, 109–141.

Tomasello, M., and Stahl, D. (2004) Sampling children's spontaneous speech: how much is enough? *Journal of Child Language*, 31, 101–121.

Further Reading and Resources

The read.me files for the DDBs already deposited in the CHILDES database (see Table 15.1) are a useful source of information.