

From Sound to Sentence

Studies on First Language
Acquisition

Editors

Mieke Beers
Beppie van den Bogaerde
Gerard Bol
Jan de Jong
Carola Rooijmans

Centre for Language and Cognition Groningen

2000

Table of Contents

Foreword

vii-ix

List of Contributors

xi-xiii

Mieke Beers

Phonological characteristics of the speech of deaf children

1-11

Julie Dockrell, David Messer, Rachel George & Gillie Wilson

Naming in children with word-finding difficulties:

Developmental delay or developmental difference?

13-29

Hilke Elsen

The acquisition of verb morphology

31-41

Steven Gillis & Dominiek Sandra

The influence of spelling rules on phonology

43-55

Annick De Houwer

Children's linguistic environments: A first impression

57-68

Melanie Jones <i>The focus of auxiliary verb learning</i> 69–79
David Malvern & Brian Richards <i>Validation of a new measure of lexical diversity: A pilot study</i> 81–96
Assimina Ralli & Julie Dockrell <i>Listening to stories and learning new words: What are the important factors?</i> 97–112
Dorit Ravid & Michal Nir <i>On the development of the category of adjective in Hebrew</i> 113–124
Annette Scheper, Anne Baker, Bart Siebelink & Philip Treffers <i>Missing arguments in conversation of Dutch-speaking children with psychiatric disturbance</i> 125–139
Petra Schulz <i>Children's interpretation of factive and non-factive verbs in English</i> 141–153
Joost van de Weijer <i>Language input and word discovery</i> 155–162
Frank Wijnen <i>Input, intake and sequence in syntactic development</i> 163–186

The acquisition of verb morphology

Hilke Elsen

In this paper, some predictions of connectionist models are investigated using data from a continuous diary study of a German girl. Several findings on the acquisition of German verbs are reported, e.g., the acquisition of verbs shows a sudden non-linear increase, the composition of the verbal lexicon undergoes changes, the first overregularizations are found when structural and quantitative changes in the verb vocabulary appear. For the production of overregularizations a wave-like development is found, that is, times of high and low production of overregularized forms alternate. The formation of participles shows various patterns, for example irregularizations.

The data presented here suggest a relationship between lexical and morphosyntactic development and the use of one associative learning mechanism for both regular and irregular inflection.

1. Introduction

The acquisition of regular and irregular morphology is the subject of several recent investigations. Researchers agree that children acquiring inflection show U-shaped behaviour curves. At first they use irregular as well as regular forms correctly. Then comes a period of overregularization and variation until the correct forms prevail. Researchers, however, do not agree on the explanation of this learning profile. Symbolists interpret overregularizations, forms like *goed*, *comed*, as evidence for the application of a rule or the insight into the idea of 'obligatory marking'. They are due to retrieval error. Connectionists hold that overregularizations appear because of quantitative and structural changes in children's verb vocabulary.

According to symbolists, a dual mechanism exists: rote, or more recently, associative and rule learning (e.g. Pinker & Prince 1988; Pinker 1991; Marcus, Pinker, Ullman, Hollander, Rosen & Xu 1992; Marcus 1995). They argue that irregular items are lexically represented and learned associatively or by rote. For regular inflection symbolic rules are needed. A U-shaped behaviour pattern results from the interaction between the two mechanisms. All early forms are produced correctly because each is stored as a whole in the lexicon. Then the child discovers the suffixation rule for regular inflection and applies it to irregulars as well. When irregular forms are used often, the application of the rule is blocked. Insufficient use of a form results in the failure of the blocking device. Hence, overregularizations occur. Finally, the exceptions to the rule are recognized. Irregulars are stored in the lexicon,

whereas regulars are formed using a rule. Properties of children's grammatical systems are responsible for the transition from rote to rule learning. Regulars and irregulars are differentiated qualitatively.

Recently, connectionists offered an alternative account for the acquisition of inflection (e.g. Rumelhart & McClelland 1986; Plunkett & Marchman 1991, 1993, Marchman & Bates 1994). They developed models that are capable of memorizing patterns as well as generalizing regularities with the help of one single mechanism. When simulating the acquisition of the English past tense, these models showed a U-shaped behaviour pattern comparable to the children's output (cf. Plunkett & Marchman 1993). Although only one associative mechanism was used, non-linear behaviour resulted. The onset of overregularizations was said to be triggered by a sufficient number of regulars in the lexicon which allowed for the abstraction of general patterns – the *'critical-mass' hypothesis* (Plunkett & Marchman 1993; Marchman & Bates 1994). Thus, a relationship between lexical and morphosyntactic acquisition was assumed. It was suggested that the development in both domains was governed by the same learning mechanism.

The first overregularizations may be related to growth in vocabulary size, as incremental increases in new regular verbs result in qualitative shifts in the way forms are treated. They may also be related to the proportion of regular and irregular verbs in the lexicon. Generalization of simulations was virtually absent as long as the percentage of regulars remained below 50%. In other words, when a sufficient number of verbs was reached and the proportion of regulars exceeded the 50% mark, the network model was able to generalize regular patterns so that overregularizations were produced. Quantitative increments in the size and structure of the training set triggered a shift in strategy: after a time of stem – past tense mapping by rote the lexicon was organized in terms of general patterns (system building). This reorganization was generated internally (Plunkett and Marchman 1993; Marchman and Bates 1994). Interestingly, relations between lexical and morphological development were not found in Marcus et al.'s (1992) investigation of the spontaneous speech of 83 children. Furthermore, evidence of a sudden non-linear increase of new verbs (a 'spurt'), or changes in the proportion of regulars has not yet been found in naturalistic studies of individual children (Pinker and Prince 1988: 140; Marcus et al. 1992) except for Stern and Stern (1928, 1965).

In this paper some predictions of connectionist models are investigated using data from one German girl. One purpose is to present data from this continuous observation of the acquisition of verb morphology. A further aim is to compare results obtained from different data samplings: continuous longitudinal data (Elsen 1991), longitudinal data from sporadically collected record-

ings (e.g. Marcus et al. 1992; Clahsen & Rothweiler 1993; Weyerts & Clahsen 1994) and parental report information on a large population (Marchman & Bates 1994). The results reported here are very similar to those predicted by the network models: There is a change in the proportion of regulars and irregulars. When this compositional change appears together with a sudden increase in new verbs, the first overregularizations are produced. The formation of past participles shows various error types. Furthermore, the acquisition of verbs shows a sudden non-linear increase. The production of overregularizations develops in a wave-like process, that is, high and low production of overregularized forms alternate.

Although the data presented here do not rule out a symbolic account which assumes two different mechanisms for the acquisition of verb morphology, they can easily be interpreted with the help of the more simple one-mechanism account. These findings shed new light on the relevance of network simulations. They illuminate the limits of fragmentary data samples.

The discussion of learning mechanisms has centered around the English past tense. In spoken German the past tense of verbs is rarely used, but rather the present perfect form. Therefore, we will focus on the acquisition of past participles. A rough sketch of German participle formation follows. Among other things, regular verbs form participles by adding a *-t* to the stem. This is called a weak inflection. By most irregular verbs a vowel change takes place and the suffix *-n* is added, this is a strong inflection. German has a few irregulars which form participles through a vowel change and the suffix *-t*, they are called mixed verbs. Some verbs exhibit still different patterns (for more detail see Elsen in press).

2. Method

The findings reported in this paper are based on the diary data of a German speaking girl A., (cf. Elsen 1991), collected continuously up to the age of 2 years and 5 months. All new words, word forms and novel pronunciations of established items were documented in IPA phonetic transcription together with essential linguistic and non-linguistic information. Afterwards, notes were taken, at first daily, then at wider intervals. The method is described in more detail in, e.g., Elsen (1996a, 1998, in press). Audiotaped data (from 0;4) were not used in this investigation.

3. Predictions

According to the symbolic dual mechanism approach, the onset of overregularizations should not be related to measurable increases in the number of regular verb types due to a clear dissociation between the mechanisms responsible for lexical and morphological acquisition. In the connectionist single-mechanism account, a non-linear critical mass relationship between lexical and morphological development should be found. Only the acquisition of a vocabulary of a certain size would allow for the abstraction of general patterns, resulting in the production of the first overregularizations, that is, productive usage. According to the model, overregularizations should not appear in small vocabularies where irregular verbs outnumber regulars (Plunkett & Marchman in press).

4. Results

My data show that total vocabulary does not increase at a constant rate according to types. Figure 1 presents the number of new words, new verbs resp., that were acquired during each third of a month: for the first ten days (**B**eginning), for the second ten days (**M**iddle) and for the rest of a month (**E**nd).

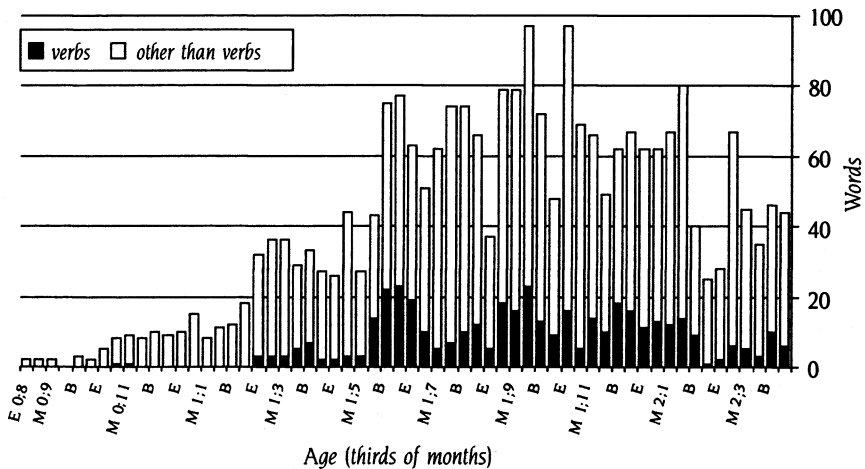


Figure 1. Number of A.'s new words

We can see several non-linear increases in this Figure (and cf. Elsen, 1996b). Nouns dominate the lexicon. The growth of verb vocabulary is not steady. That is, there is a 'verb spurt' at the end of 1;5. In general, early verb tokens are correctly inflected. In most cases, correct irregular forms precede overregularizations. Up to the end of the continuously collected data, about 25% of the irregular verbs show at least one token of overregularization (given in Elsen 1998, in press). Furthermore, the formation of incorrect past participles exhibits various patterns. We find overregularizations with different stems. They are irregular verbs with regular *-t*, e.g. *gehen* – *gegeht* (correct *gegangen*) 'to go', *nehmen* – *genehmt* (correct *genommen*) 'to take', *trinken* – *getrunkt* (correct *getrunken*) 'to drink'. There are irregularizations. These are regular verbs with irregular patterns, e.g. *schmecken*–*geschmoeckt* (correct *geschmeckt*) 'tasted', *aufkleben*–*aufgekleben* (correct *aufgeklebt*) 'glued on'. There are stems other than infinitive or participle or non-existing stems, both with regular *-t* or irregular *-n*, e.g. *gehen* – *gingt* (correct *gegangen*) 'to go', *wegnehmen* – *wegenimmt* (correct *wegenommen*) 'to take away', *lesen* – *gelosen* (correct *gelesen*) 'to read', *rausnehmen* – *rausnuhmen* (correct *rausgenommen*) 'to take out'. Furthermore, we find double-markings, *ziehen* – *geziehten* (correct *gezogen*) 'to pull', *abbrechen* – *abgebroschet* (correct *abgebrochen*) 'to break off', *tun* – *getant*, *getanen* (correct *getan*) 'to do' (for more examples see Elsen in press). What is more, different forms may coexist for several months. Some erroneous forms are more frequently applied than the correct participle.

In Figure 2, the accumulation of the spontaneously produced regular and irregular verbs is documented. We see that the rate of acquisition of regulars differs from that of irregulars. There is a point in time when regulars come to dominate: at the end of 1;5.

In Figure 3, you can see the proportion of regular and irregular verb types in the girl's verb vocabulary as a function of age (for the relative distribution of A.'s regular and irregular verbs as a function of verb vocabulary see Elsen in press).

After a period of predominance of irregular verbs and a time of relative equal distribution, the proportion of the regulars rises at the end of 1;5. These findings are quite similar to those of Marchman and Bates (1994) (cf. their Figure 1, p 353). They investigated 1130 children between 1;4 and 2;6 with a parental report technique. Parents were asked to check verb lists. Marchman and Bates found that 'the shape of the relationship between verb learning and the productive use of past tense morphology [was] consistent with the non-linear 'critical mass' assumptions of the single-mechanism theory' (Marchman & Bates 1994: 360). Note that the results from my continuous data collection match those from Marchman and Bates and thus support their findings with a different methodology.

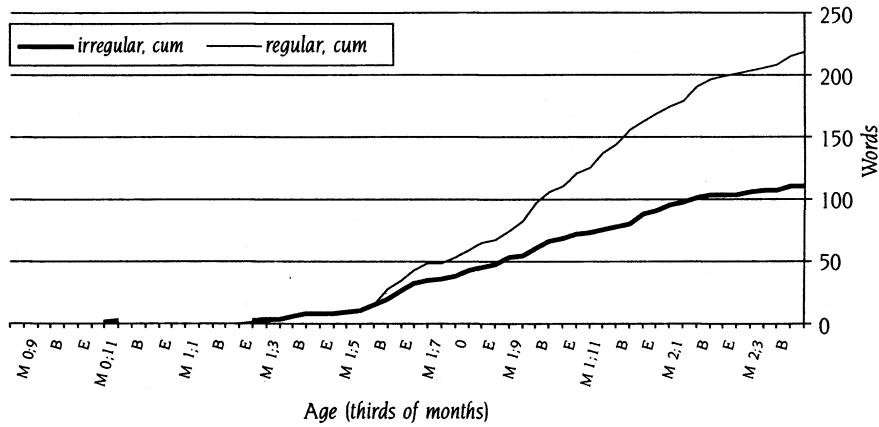


Figure 2. A.'s regular/irregular verbs, cum.

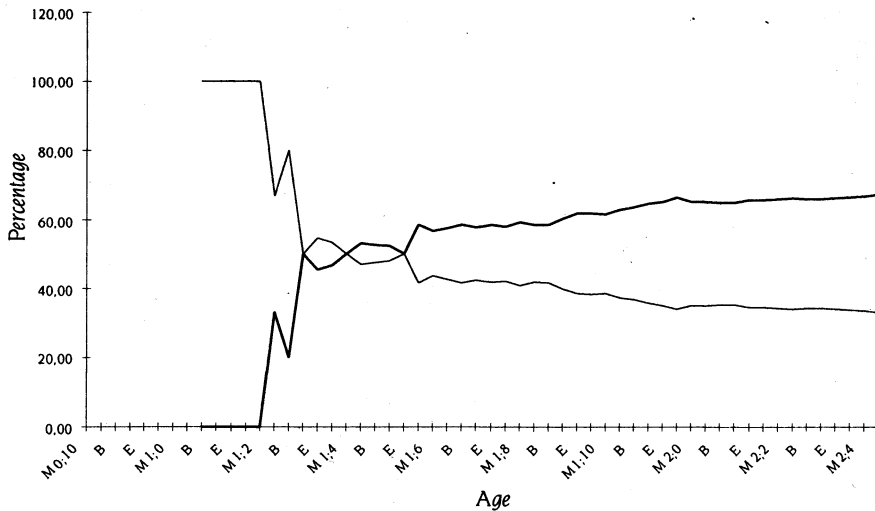


Figure 3. Proportion of A.'s regular and irregular by third of month

The graphs show that there is clear evidence that verb vocabulary composition is undergoing substantial reorganization. There is a distinct change in the proportion between regulars and irregulars. From the end of 1;5, regulars contribute more than 50% of the verbs overall.

When the child's lexicon contains 15 irregulars and 15 regulars at 1;5,27 (these verbs are given in Elsen 1998, in press), the first overregularized forms are found after some days of correct inflection at 1;5,30, [defa t] *gefällt* *failed.

The next ones appear at 1;6,2 [l̥snt] *lornt* *losted, at the end of 1;6 [mlten_mt] *mitgenehmt* *taken along with, and at 1;7,24 [vest] *west* *be-ed, participle stem. This happens at the time when the composition of the verb vocabulary changes and the proportion of regulars exceeds 50%, as predicted by the network. If the surge of morphological errors was simply due to an increase in opportunities for errors because of more new verbs, then the number of overregularizations should rise in later development as numbers of new verbs – and thus numbers of opportunities – rise, too. But this is not the case. If it was due to more talking in general, then rates of participle overregularizations should correlate to the rates of plural overregularizations, but they don't. So it is highly probable that the onset of overregularizations was triggered by quantitative and structural changes in the verb vocabulary.

Furthermore, after a period of few overregularizations they increase both in types and tokens by 1;9. Another peak is observed at 2;0/2;1. And looking beyond the time span of continuous note taking, periods of high and low production of overregularized forms alternate. Thus, overregularizations clearly come in waves. These waves are interesting in yet another respect – there seems to be systematicity in the way different patterns of erroneous participial formation are used over time periods (cf. Lindner in press).

Note – at the age of 1;9 as well as at 2;0 there is a definite increase in new regular verbs, see Figure 1. This indicates the possibility of several reorganizations of the system.

5. Discussion

Taken together, the present data support the 'critical mass' hypothesis. The results are congruent with a connectionist account of learning verb inflection. There is a definite relation between this child's change in vocabulary composition and the onset of overregularizations, as predicted by the single-mechanism approach. Furthermore, several patterns of over- and irregularization are found, as well as a discontinuity in expansion rate. Note that nonlinearities in vocabulary growth are a contributing factor to the onset of overregularizations, not a necessary or sufficient condition (Marchman & Bates, 1994; Plunkett & Marchman, in press).

These data are obviously not compatible with symbolists' results, whose data on individual children do not show sudden increases in new verbs and no relation between compositional changes in the verbal lexicon and first overregularizations (eg. Marcus et al. 1992: 86–88, 99; Weyerts & Clahsen 1994: 442). The explanation may be found in the method of data collection, as the data here were compiled continuously as opposed to the recordings of

several children during one hour a week or a month. I assume that they missed the first overregularizations and many low-frequency items and, more importantly, crucial non-linear development between recording sessions. It is highly probable that non-linearities in verb acquisition are missed when recordings for a child are too rare. Token frequency of many verbs, especially regulars, is initially extremely low (cf. Elsen, in press) so that numbers of types in a continuous sample will be very different from those in sporadically compiled corpora. There, the rise of new verbs depends upon an increase of token numbers. Strictly speaking, statements about type numbers are not possible for sporadically recorded samples. The same problem arises for early overregularizations. We cannot rule out the possibility that this child's second wave of overregularizations (cf. Elsen, in press), when type and token numbers rise, is equivalent to the first overregularizations found in other children. Thus, correlations between the onset of overregularizations and changes in vocabulary size and structure will hardly be found.

Non-continuous sampling might be the reason why other investigators of children's German past participles conclude that there is no irregularization of regular patterns: 'In all the data, there are no participles in which an irregular stem pattern has been extended to a weak verb. Rather, the only kind of stem error we found are regular stems replacing irregular ones.' (Clahsen & Rothweiler 1993: 1). 'Despite similar frequencies of regular and irregular participles in the input, it is only the regular -t suffix which is overregularized by children' (Weyerts & Clahsen 1994: 430). They interpret the networks' generalizing capacities as depending on the high frequency of regulars. As in German regular and irregular forms are said to be found in equal numbers and, at least according to Clahsen and colleagues, errors only affect irregulars, a qualitative distinction between regular and irregular inflection can be assumed. A symbolic inflectional rule must be responsible for the acquisition of regular patterns. But, as already pointed out, a change in the proportion of regulars and irregulars together with a growth in vocabulary size triggered the formation of regular patterns in the model. **And it is exactly this relation between changes of quantitative and structural properties of the verb vocabulary and first overregularizations that is found in the present data.** Therefore, the assumption of a symbolic rule is not necessary. Moreover, inflection patterns are less homogeneous and differences between regular and irregular patterns are less clear-cut than postulated by symbolists. The present data show that several irregular patterns were applied to regular and irregular verbs. Stems other than infinitive or participle as well as non-existing stems were combined with the regular -t or irregular -n suffix. Different patterns coexisted for months. Thus,

the claim that a child distinguishes qualitatively between regular and irregular inflection should be reconsidered. Though the concept of a rule is not explicitly excluded, it is not necessary to assume a symbolic rule for the formation of regular verbs. I suggest that both irregular and regular inflection is based on one associative mechanism – a lexical learning device.

6. Conclusion

The results presented here support a one-mechanism approach which suggests a relationship between lexical and morphosyntactic development and the use of one associative learning mechanism. Data show that there is a non-linear relationship between overregularizations to vocabulary size and structure. Various patterns of incorrect participle formation, which according to symbolists should not be found, are in line with the connectionist approach (Plunkett & Marchman, in press). Changes in the rate of verb vocabulary acquisition coincide with overregularization patterns, resulting in a wave-like generation of incorrect past participles. Overregularizations appear when the proportion of regulars exceeds the 50% mark, as predicted by the networks.

The divergent attitudes could converge, if we didn't have a fundamental problem with principles. As a prime goal, standard generative grammar creates a description of a universal grammar. The object of their investigations is theoretical Competence, as opposed to actual Performance. As a result, the question of how this competence gets into the head arises, and that subsequently assumes that Universal Grammar has psychological reality. Because of the strict differentiation between competence – as the subject of description – and performance – as the source of information -, rules and parameters which generate structure must be considered inborn so that they can be a part of competence. Since the data that the child hears belong to performance, input can only trigger innate mechanisms. This assumption of generative grammar – the strict differentiation of competence and performance -, together with the conception that generative principles and parameters are psychologically real, **forces** the belief that rules and categories are inborn. Thus, it is not the investigation of children's data, but the effort not to endanger the basic assumption that gives rise to the idea of innate rules and categories. One possibility is to redefine Universal Grammar as a system of description of language and to give up the strict separation of competence and performance as well as discrete, categorical rules. As was shown, no symbolic rules for regulars are necessary. Irregulars are not simply learned by rote. There is analyzing and overapplication to regular verbs. The difference in treatment of regular and irregular verbs is a gradual one.

Acknowledgments

A similar talk was given at the 21st Boston Conference on Language Development, November 1996. An elaborated version is published as Elsen (in press). The research presented here benefited from discussions with Katrin Lindner and Kim Plunkett and from grants from the Freistaat Bayern.

References

- Clahsen, H. and Rothweiler, M. (1993) Inflectional rules in children's grammars: evidence from German participles. *Yearbook of Morphology* 1992. 1-34.
- Elsen, H. (1991) *Erstspracherwerb*. Wiesbaden: DUV.
- Elsen, H. (1996a) Two routes to language: Stylistic variation in one child. *First Language* **16**, 141-158.
- Elsen, H. (1996b) Linguistic team-work – the interaction of linguistic modules in first language acquisition. E. Clark (ed.) *The Proceedings of the Twenty-seventh Annual Child Language Research Forum*. Stanford: CSLI.
- Elsen, H. (1998) *Ansätze zu einer funktionalistisch-kognitiven Grammatik. Konsequenzen aus Regularitäten des Erstspracherwerbs*. (Inaugural-) Habilitationsschrift, Universität München.
- Elsen, H. (in press) The acquisition of past participles: One or two mechanisms?. R. Fabri, A. Ortman and T. Parodi (eds.) *Models of Inflection*. Tübingen: Niemeyer.
- Lindner, K. (in press) Overgeneralization revisited: the case of German past participles. R. Fabri, A. Ortman and T. Parodi (eds.) *Models of Inflection*. Tübingen: Niemeyer.
- Marchman, V.A. & Bates, E. (1994) Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language* **21**, 339-366.
- Marcus, G. F. (1995) The acquisition of the English past tense in children and multilayered connectionist networks. *Cognition* **56**, 271-279.
- Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T.J. and Xu, F. (1992) Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*. 228.57.4.
- Pinker, S. (1991) Rules of language. *Science* **253**, 530-535.
- Pinker, S. and Prince, A. (1988) On language and connectionism: Analysis of a parallel distributed processing model of language acquisition. *Cognition* **28**, 73-193.

- Plunkett, K. and Marchman, V. (1991) U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition* **38**, 43-102.
- Plunkett, K. and Marchman, V. (1993) From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition* **48**, 21-69.
- Plunkett, K. and Marchman, V. (in press) Learning from a connectionist model of the acquisition of the English past tense. *Cognition*.
- Rumelhart, D. and McClelland, J.L. (1986) On learning the past tenses of English verbs: J.L. McClelland, D. Rumelhart, and the PDP Research Group (eds.) *Parallel Distributed Processing*. Vol. II. *Psychological and Biological Models*. Cambridge, MA: Bradford Books/MIT Press. 216-270.
- Stern, C. and Stern, W. (1928/1965) *Die Kindersprache*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Weyerts, H. and Clahsen, H. (1994) Netzwerke und symbolische Regeln im Spracherwerb: Experimentelle Ergebnisse zur Entwicklung der Flexionsmorphologie. *Linguistische Berichte* **154**, 430-460.