"… this much is how much I'm taller than Joey…"
A Corpus Study in the Acquisition of Comparison Constructions

Vera HOHAUS and Sonja TIEMANN

Eberhard-Karls Universität Tübingen
Sonderforschungsbereich 833 „Bedeutungskonstitution"
Nauklerstraße 35 - 72074 Tübingen, GERMANY
Phone : 0049.7071.2977156 - Fax : 0049.7071.29775830

E-Mail : vera.hohaus@student.uni-tuebingen.de
E-Mail : sonja.tiemann@uni-tuebingen.de

ABSTRACT

Des recherches récemment effectuées dans le domaine de la sémantique formelle ont révélé trois paramètres dépendants qui renferment les clusters des qualités par lesquelles les constructions de comparaison diffèrent à travers les langues. Nous examinons les prédictions de ces paramètres en ce qui concerne l'acquisition du langage chez l'enfant et présentons les résultats d'une étude de corpus au sujet du cours de l'acquisition d'un nombre de constructions de comparaison en anglais et en allemand (dont figurent des comparatifs sous forme de phrase ou proposition, des superlatifs, des constructions qui contiennent des indications de mesures directes ainsi que des questions de degrés).

1. MOTIVATION

Crosslinguistic Variation in Comparison Constructions. Beck et al. [Becnt] - following work by Beck, Oda and Sugisaki [Bec04] - identify, for 17 languages, clusters of properties along which comparison constructions differ across languages (Figure 1), and suggest the three dependent semantic parameters in (I) to (III) to capture the variation.

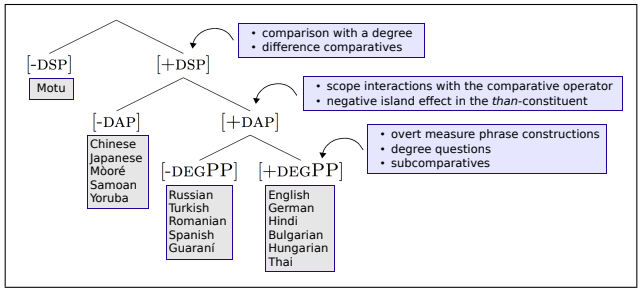

Figure 1: Parametric variation in comparison constructions across languages [Becnt]

(I) **Degree Semantics Parameter (DSP)**
A language {does/ does not} have gradable predicates, i.e. lexical items that introduce degree arguments.

(II) **Degree Abstraction Parameter (DAP)**
A language {does/ does not} have binding of degree variables in the syntax.

(III) **Degree Phrase Parameter (DEGPP)**
The degree argument position of an unmarked gradable predicate {may/ may not} be occupied by a syntactically visible element at a pre-LF level of syntax.

Examples of the relevant constructions and effects are provided in (1) to (7) in the next column.

(1) **comparison with a degree :**
Mary is taller than five foot.

(2) **difference comparative :**
Mary is two inches taller than John.

(3) **scope interactions with the comparative :**
This draft is 10 pages long. The paper is required to be exactly 5 pages longer than that. ("The minimum length required for the paper is 15 pages." / "The paper must have the length of 15 pages.")

(4) **negative island effect :**
**Mary is taller than none of her sisters is.*

(5) **measure phrase construction :**
Mary is six foot tall.

(6) **degree question :**
How tall is Mary ?

(7) **subcomparative :**
The knife is longer than the drawer is deep.

2. PREDICTIONS

Relating Variation to First Language Acquisition. A successful parametric theory of crosslinguistic variation should also be an adequate theory of the process of language acquisition in children. Assuming strong grammatical conservatism as in (IV) and a standard semantics of comparison constructions [vS84, Bec08] allows us to make predictions about the time course of first language acquisition for a number of frequent comparison constructions in English and German.

(IV) **Strong Grammatical Conservatism :**
Children do not begin making productive use of a new grammatical construction in their spontaneous speech until they have both determined that the construction is permitted in the adult language, and identified the adults's grammatical basis for it. [Sny08]

Predictions are that in spontaneous speech of a child learning English or German, degree morphology should occur first.

1. **degree morphology :**
   Sue is taller.
2. **phrasal and clausal comparatives :**
   Sue is taller than Joe.
   Sue is taller than Joe is.
   superlative :
   Sue is the tallest.
3. **overt and pronominal measure phrases :**
   Sue is six foot tall.
   Sue is this tall.
   degree questions :
   How tall is Sue ?

Phrasal and clausal comparatives should be acquired later (2.), and concurrently with the superlative. Overt and pronominal degree questions should be acquired last (3.), together with degree questions.

3. METHODOLOGY

In order to test the acquisitional predictions above, we selected transcripts from spontaneous speech of three American and three German children from CHILDES [Mac00]. The list of transcripts analyzed in our study is presented in the tables below.

Table 1: Corpora analyzed for American English

Child	Collected by	Ages	# child utterances
Adam	Roger Brown	2;3-5;2	90,852
Sarah	Roger Brown	2;3-5;1	31,369
Ross	Brian MacWhinney	2;6-7;5	30,912

Table 2: Corpora analyzed for German

Child	Collected by	Ages	# child utterances
Cosima	Rosemarie Rigol	0;0-7;2	76,888
Pauline	Rosemarie Rigol	0;0-7;7	83,572
Sebastian	Rosemarie Rigol	0;0-7;0	79,451

The CLAN program COMBO was used to identify potentially relevant child utterances, which were then searched by hand, checked against the original transcripts to exclude imitations, repetitions, and formulaic routines, and labelled according to the table in Figure 2 below.

level of analysis, labels and criteria				LABELS		
COMPARATIVE				**%comp**		
adjective	with morphology	without overt complement		**:adj**	**:morph**	**:-**
		with overt complement	phrasal			**:+** **:phr**
			clausal			**:cl**
	with *more*	without overt complement			**:more**	**:-**
		with overt complement	phrasal			**:+** **:phr**
			clausal			**:cl**
	irregular	without overt complement			**:irrg**	**:-**
		with overt complement	phrasal			**:+** **:phr**
			clausal			**:cl**
adverbial	with morphology	without overt complement		**:adv**	**:morph**	**:-**
		with overt complement	phrasal			**:+** **:phr**
			clausal			**:cl**
	with *more*	without overt complement			**:more**	**:-**
		with overt complement	phrasal			**:+** **:phr**
			clausal			**:cl**
	irregular	without overt complement			**:irrg**	**:-**
		with overt complement	phrasal			**:+** **:phr**
			clausal			**:cl**
SUPERLATIVE				**%sup**		
adjective	with morphology			**:adj**	**:morph**	
	with *most*				**:most**	
	irregular				**:irrg**	
adverbial	with morphology			**:adv**	**:morph**	
	with *most*				**:most**	
	irregular				**:irrg**	
MEASURE PHRASES				**%mp**		
overt measure phrase, including number and unit				**:pro**		
pronoun referring to measure phrase				**:overt**		
DEGREE QUESTIONS				**%dq**		
EXCLUSION				**%excl**		
imitation and repetition				**:im&rep**		
overlapping or unintelligible speech				**:overlap**		
routine				**:rout**		
ERRORS				**%err**		
omission of comparative morphology				**:om**		
use of comparative morphology for superlative				**:mix**		
overregularization				**:overreg**		
use of morpheme and any form of *much*				**:morph+much**		
use of complementizer other than *than*				**:comp**		
use of wrong unit of measurement				**:unit**		
omission of degree predicate				**:om-dp**		
other				**:other**		

Figure 2: Labels

The results were then analyzed for very first use and age of acquisition as well as for types of errors and their frequency. Frequency of grammatical constructions and errors was determined per 1,000 utterances. Irregular and periphrastic comparatives have not been taken into consideration when determining the age of acquisition for comparative morphology.

4. RESULTS

The results are summarized below. The table in Figure 3 presents the results for the age of very first use, the table in Figure 4 the results for the age of acquisition. Following Snyder [Sny07] and Stromswold [Str90], the age at which a child produced his or her first clear example of a construction followed soon after by regular use with a variety of lexical items was considered to be the age of acquisition for this construction (FIRST OF REPEATED USES). "Soon after" was here understood as within the next two months.

First Use	English			German		
	Adam (2;3-5;2)	Sarah (2;3-5;1)	Ross (2;6-7;5)	Cosima (0;0-7;2)	Pauline (0;0-7;7)	Sebastian (0;0-7;0)
comparative morphology	2;6	2;10	2;6	2;7	1;1	3;11
phrasal comparative	3;5	3;11	3;5	6;6	5;9	4;8
clausal comparative	3;8	>5;1[1]	3;5	>7;2	>7;7	>7;0
superlative	4;2	3;7	3;5	3;7	3;5	3;9
pronominal measure phr. construct.	4;3	3;0	2;10	3;3	2;4	2;7
overt measure phr. construct.	4;5	4;5	3;0	>7;2	6;5	>7;0
degree quest.	4;0	3;1	3;3	3;11	3;5	4;3

1 There were no occurences of the respective constructions in the transcripts.

Figure 3: Age of very first use

FRU	English			German		
	Adam (2;3-5;2)	Sarah (2;3-5;1)	Ross (2;6-7;5)	Cosima (0;0-7;2)	Pauline (0;0-7;7)	Sebastian (0;0-7;0)
comparative morphology	3;4	3;7	2;6	2;9	2;8	3;11
phrasal comparative	4;2	3;11[?]	3;5	Cannot be determined.[2]	6;6	6;3
clausal comparative	Cannot be determined.	>5;1	Cannot be determined.	>7;2	>7;7	>7;0
superlative	4;2	4;1[?]	4;8	3;7	4;5	4;3
pronominal measure phr. construct.	Cannot be determined.	4;0	Cannot be determined.	3;7	3;0	Cannot be determined.
overt measure phr. construct.	Cannot be determined.	4;5	Cannot be determined.	>7;2	Cannot be determined.	>7;0
degree quest.	Cannot be determined.	4;7[?]	Cannot be determined.	Cannot be determined.	6;10	Cannot be determined.

2 The low number of occurrences did not allow us to reliably determine the age of acquisition.

Figure 4: Age of acquisition

None of the children acquire all constructions before the end of the respective corpora, although Ross and Adam make use of all constructions before the end of the corpora. For both English and German, mean age of acquisition of regular comparative morphology was 3;1, with a range of 2;6 to 3;7 for English and of 2;8 to 3;11 for German.

For English, mean age of acquisition of the phrasal comparative was 3;8, with a range of 3;5 to 4;2. For the German children Pauline and Sebastian, however, mean age of acquisition of the phrasal comparative was 6;4, with a range of 6;3 to 6;6. For the superlative in English, mean age of acquisition was 4;3, with a range of 4;1 to 4;8. In German, mean age of acquisition for the superlative was 4;1, with a range of 3;7 to 4;5.

For English pronominal measure phrase constructions, only for Sarah could the age of acquisition reliably be de-

termined at 4 ;0. Mean age of acquisition for pronominal measure phrases for the German children Cosima and Pauline was 3 ;3, with a range of 3 ;0 to 3 ;7. In English, overt measure phrase constructions and degree questions were only acquired by Sarah, at 4 ;5 and 4 ;7 respectively. In German, degree questions were only acquired by Pauline by the end of the corpora, at 6 ;10.

In sum, acquisition of comparative morphology in English was earlier than acquisition of the superlative and the phrasal comparative, and earlier than of overt measure phrase constructions and degree questions. Acquisition of the superlative and the phrasal comparative was around the same time and earlier than of overt measure phrase constructions and degree questions. Gaps in the recordings at 3 ;8 as well as 4 ;0 and 4 ;7 probably account for the late age of acquisition of the superlative by Ross. Overt measure phrase constructions and degree questions were acquired last, and around the same time. This time course of acquisition corresponds to the predictions outlined above.

For German, acquisition of comparative morphology was earlier than acquisition of the superlative, and earlier than acquisition of overt measure phrase constructions and degree questions. Acquisition of overt measure phrase constructions and degree questions seems to be last. So far, the time course of acquisition in German, too, is in line with the predictions. However, German phrasal comparatives are not acquired concurrently with the superlative but considerably later than in English. Clausal comparatives in both languages were used too infrequent to gain much insight into their acquisition.

Unpredicted are the late acquisition of the phrasal comparative in German when compared to English, and the early acquisition of pronominal measure phrase constructions in both language when compared to the acquisition of overt measure phrase constructions.

Errors. The low proportion of errors among the search results seems to support our initial assumption of a conservative language learner. Reliable statements on the number of errors can only be made for English, as the German transcripts lacked a morphology tier, and the search thus did usually not return errors. Table 3 provides the number of correct comparison constructions with a child and the percentage of errors. Excluded material has not been counted towards the total.

Table 3: Error Rates

Child	# comparison constructions	# errors	% errors
Adam	225	15	5.24
Sarah	129	5	3.73
Ross	349	21	6.02

Errors in English included omission of comparative morphology, use of comparative morphology for the superlative, and frequently, the use of regular comparative morphology for irregular forms. Repeatedly, children also used both a comparative morpheme and *much*, as is illustrated with examples from Ross and Sarah below.

(8) *CHI : (be)cause it's gonna be more dirtier # huh Ma ?
(Sarah, age : 4 ;10.27, file : sarah129.cha)

(9) *CHI : it's more specialer [*] than that # because you love me and you gave me m_ and_ m's and I got the Attack_ Track # and # um # and that I +/ (Ross, age : 5 ;5.20, file : 61a2.cha)

Errors also included the use of a complementizer other than *than*. Below is an example from Adam.

(10) *CHI : it's bigger just like de [: the] truck .
(Adam, age : 3 ;5.01, file : adam30.cha)

5. DISCUSSION

The time course of the acquisition of the selected comparison constructions in English and German generally confirms the predictions. But the results also raise new questions for a comprehensive semantic theory of comparison constructions.

Rethinking phrasal comparatives in English and German. One of the questions relates to the analysis of phrasal comparatives in English and German. It is a universally made assumption that an analysis of the phrasal comparative should apply equally well to both languages. Our results however suggest that English and German phrasal comparatives are different.

We assumed initially that the phrasal comparative in both languages should receive the same semantics as the clausal comparative, thereby adopting a reduction analysis [Lec04]. On that analysis, the phrasal comparative is derived from a clausal via ellipsis. The acquisitional data suggest that in German, this might indeed be the case, and mastering the respective elliptical processes presents an additional challenge to the learner.

But since Heim [Hei85], semanticists have also considered a direct analysis of the phrasal comparative, i.e. an analysis without ellipsis, in which interpretation applies to the structure we see at the surface. Arguments in favor of a direct analysis for English go back to Hankamer [Han73]. Adopting a direct analysis for phrasal comparatives in English, and a reduction analysis for phrasal comparatives in German might explain the late acquisition of the latter, as suggested by Tiemann [Tie09]. It will be interesting to integrate further insights into the acquisition of the clausal comparative in both languages into such an account.

Rethinking overt and pronominal measure phrases. Both overt and pronominal measure phrase constructions have been analyzed as quantifiers over degrees. The early acquisition of pronominal measure phrases suggests that they should be assigned a simpler analysis instead. One possibility would be to explore an analysis of pronominal measure phrases as degrees, rather than as quantifiers.

6. CONCLUSIONS

We have presented acquisitional evidence that lends support to the proposed parameters of crosslinguistic variation in comparison constructions. At the same time, our results have shown that the time course of acquisition directly contradicts the view that English and German phrasal comparatives, as well as overt and pronominal measure phrases in both languages should receive a unified semantic analysis. Our study thus illustrates how formal semantics can be applied to – as well as benefit from – from language acquisition research, and *vice versa.*

RÉFÉRENCES

[Bec04] Beck, S., Oda, T. and Sugisaki, K. (2004), "Parametric variation in the semantics of comparison : Japanese versus english", *Journal of East Asian Linguistics*, vol. 13, pp. 289–344.

[Bec08] Beck, S. (2008), "Comparatives and superlatives", Manuscript, Universität Tübingen.

[Becnt] Beck, S., Krasikova, S., Fleischer, D., Gergel, R., Hofstetter, S., Savelsberg, C., Vanderelst, J. and Villalta, E. (in print), "Crosslinguistic variation in comparison constructions", *Linguistic Variation Yearbook*.

[Han73] Hankamer, J. (1973), "Why there are two *Than*'s in english", in *Proceedings of the 9th Annual Meeting of the Chicago Linguistics Society*, pp. 179–191.

[Hei85] Heim, I. (1985), "Notes on comparatives and related matters", Manuscript, University of Texas at Austin.

[Lec04] Lechner, W. (2004), *Ellipsis in Comparatives*, *Studies in Generative Grammar*, vol. 72, Berlin : De Gruyter.

[Mac00] MacWhinney, B. (2000), *The* CHILDES *Project : Tools for Analyzing Talk*, Hillsdale : Lawrence Erlbaum, 3rd edn.

[Sny07] Snyder, W. (2007), *Child Language : The Parametric Approach*, Oxford : Oxford University Press.

[Sny08] Snyder, W. (2008), "Relating language variation to language acquisition", Handout from CSAAL Workshop on Locating Variability, University of Massachusetts, Amherst.

[Str90] Stromswold, K.J. (1990), *Learnability and the Acquisition of Auxiliaries*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge.

[Tie09] Tiemann, S. (2009), *The Acquisition of Comparison : A Corpus Based Analysis*, Master's thesis, Eberhard-Karls Universität Tübingen.

[vS84] von Stechow, A. (1984), "Comparing semantic theories of comparison", *Journal of Semantics*, vol. 3, pp. 1–77.