Modeling Bilingual Children's Acquisition of Complex Sentences in German

Daniel Wiechmann (wiechmann@anglistik.rwth-aachen.de)
RWTH Aachen University
Department of English Linguistics

Judith Steinfeldt (judith.steinfeld@rwth-aachen.de)
RWTH Aachen University
Department of English Linguistics

Elma Kerz (kerz@anglistik.rwth-aachen.de)
RWTH Aachen University
Department of English Linguistics

Abstract

Although Bilingual First Language Acquisition research has increased considerably over the past few decades, there is still much controversy regarding the rate of development, i.e. the question whether bilinguals lag behind their monolingual peers in various aspects of language. Some studies have found similar rates of development, whereas others have found that bilingual children lag behind their monolingual peers. The current study contributes to this discussion of (dis)similar rates of development by investigating bilingual children's acquisition of German complex sentence constructions involving adverbial clauses (ACs). Our findings are consistent with usage-based approaches to language acquisition, which predict that bilingual acquisition should proceed slower due to learners having less exposure, on average, to each language.

Keywords: bilingual first language acquisition; language production; rate of development; complex constructions

Introduction

Bilingual First Language Acquisition (BFLA)

There has been an increasing interest in early bilingual language acquisition. Commonly this interest involves the question of whether the cognitive and developmental path (course of development) and time course of language learning by bilingual children is the same as that of their monolingual peers (rate of development). Although prevailing theoretical models of language acquisition have different views regarding the influence of endogenous and exogenous factors on the acquisition of abstract linguistic structures and patterns, they agree upon the idea that monolingual and bilingual language learning is qualitatively equivalent in that children go through the same series of developmental phases, starting off with single word productions, followed by two and multi-word utterances before they finally develop the capacity to produce complex sentences (Meisel 1986; de Houwer 1995, 2009). Prior research comparing the rate of development in monolingual and bilingual learners has produced somewhat mixed results. Some studies have found similar rates of development (cf. Pearson & Fernández, 1994; Paradis, Crago & Genesee, 2005/2006; Paradis 2010), whereas others have found that bilingual children lag behind their monolingual peers (Gathercole 2002a, 2002b, 2007; Nicoladis, Palmer & Marentette 2007; Pérez-Leroux, Pirvulescu & Roberge 2009). The current study contributes to this discussion of (dis)similar rates of development by investigating the bilingual acquisition of complex sentences involving adverbial clauses (ACs) in German, which mark the last stage in a series of milestones mentioned above.

Usage-based theory and BFLA

Usage-based (UB) theories belong to a family of emergentist models, which assume that (a) the development of language competence is contingent on the experience with language (Tomasello, 2003, Tomasello and Lieven, 2008, O'Grady, 2008). A conservative assumption about BFLA is that bilingual children, on average, receive less language input per language than their monolingual peers. UB-theories thus predict that reduction in overall exposure to a language should negatively affect children's rate of acquisition (Gathercole & Hoff, 2007, Paradis et al. 2011).

Usage-based theory and the acquisition of complex sentences

Complex sentences are grammatical assemblies consisting of multiple clauses. Two types of clauses are distinguished: (i) sentences including coordinate clauses and (ii) sentences including a matrix clause and a subordinate clause. Complex sentences containing subordinate clauses can be further subdivided into three basic sub-types: constructions with complement clauses, relative clauses and adverbial clauses. The most comprehensive study on the acquisition of complex sentences framed within UB-theory is Diessel (2004). Diessel proposes that complex sentences with ACs

develop through two different types of processes: Complex sentences involving complement and relative clause constructions develop through a process termed *clause expansion* (RC what is that). Complex sentences containing adverbial clauses develop through a process termed *clause integration*, in which two independent sentences are merged into a single bi-clausal unit. The earliest adverbial clauses produced by children are thus free-standing (isolated) clauses introduced by an adverbial subordinator, which are only pragmatically linked to a previous utterance. Over time, children learn to elaborate these structures and integrate them with a matrix clause. The last step in mastering complex sentences involves developing the capacity to produce sentence initial subordinate clauses, which impose greater demands on (verbal) working memory as initial clause require that the producer has planned the entire complex structure at the onset of the utterance (Hawkins 2004). Initial adverbial clauses thus develop later and their frequency, at first, is limited to specific subordinators. Another finding of Diessel's (2004) study is that children's early productions of complex sentences are tied to specific lexical expressions. The emergence of more schematic representations of such constructions takes place only after children have been exposed to a sufficient number of tokens to generalize over. This is reflected in the fact that children only gradually elaborate their repertoire of adverbial subordinators and overextend already learned types to situations where those types are semantically inadequate (e.g. use of a causal subordinator to express concessive or other adverbial relations). Two additional, more general indicators of language proficiency are the mastery of syntactic differences in German main and subordinate clause (verb second in main clauses vs. verb-final positioning in subordinate clauses) (cf. Roeper, 1973; Miller, 1976; Park, 1981; Clahsen, 1982) and mean length of utterance (MLU). MLU has been shown to be an important measurement of a child's gross language development and was found to correlate with the development of morphological and syntactic skills in young children (Brown 1973; Parker and Brorson 2005). Building on this research, the present study sets out to derive statements about differences in the rate of development of complex sentence constructions from measurements of five indicators of language proficiency:

1. Proportion of isolated/integrated adverbial clauses
2. Proportion of sentence-initial adverbial clauses
3. Proportion of misused subordinators
4. Proportion of correct verb position in sub clause
5. (Log) length of adverbial construction (MLU)

Method

All relevant data were elicited by having children watch a 6.5 minutes episode of a popular stop-motion animated children's television series. The children were then given a visual cue to a particular scene and asked to describe what happened in that scene. The children's' responses were audio-recorded and transcribed.

Participants

A total of 50 children from 4 to 6 years old participated in the study: 25 bilingual child participants (German in combination with another language) and a control group consisting of 25 monolingual children.[1] The children participating in the study were selected from several kindergartens with families from various social backgrounds and both, parents and kindergarten agreed to participate in the study.

Data

The elicitation procedure resulted in 27,301 word tokens produced by monolingual learners and 21,023 word tokens produced by monolinguals ones. From these corpora, all instances of the target constructions were extracted by way of manual inspection of the corpus data, yielding a total of 1,023 data points (601 from monolinguals, 422 from bilinguals). The extracted data were annotated with information pertaining to the indicators of language competency listed in the preceding section.

Table 1: Descriptive Statistics

	monolingual	bilingual
age (mean)	5;4	5;5
age (SD)	7.16	6.87
AC total	601	422
AC integrated	62.73%	50%
AC initial position	23.34%	13.74%
sub misused	1%	17.77%
correct verb position	72.38%	66.11%
length (mean)	12.36%	10.55
length (SD)	7.2	5.91

The language proficiency levels of the monolingual and bilingual children were compared with respect to five indicators of language proficiency. To test whether and to what extent the proficiency levels of mono- and bilingual speakers differ, we asked: Does competence indicator x [standing for competence type y] differ significantly between bilingual and monolingual children after controlling for age? The data were analyzed using linear and logit mixed effects models with SUBJECT as a random effect.[2]. We checked for normality and homogeneity by visual inspections of plots of residuals against fitted values. For all models, the significance of the predictor BILINGUAL

[1] There was are total of 12 different language pairs within the data. German was acquired in combination with one of the following languages: Albanian, Arabian, Bosnian, English, French, Hungarian, Kurdish, Persian, Russian, Spanish, Turkish or Vietnamese.

[2] All data were analyzed using R (R Development Core Team, 2009) and the R packages lme4 (Bates & Maechler, 2009) and languageR (Baayen, 2009; cf. Baayen, 2008)

was assessed through model comparison: For each model, we conducted likelihood ratio tests to see if a model including BILINGUAL is significantly better than the corresponding model containing only AGE and the random effect (SUBJECT).

Model 1: Length
Linear Mixed Model fit by REML approximation; p-values estimated via Markov Chain Monte Carlo (MCMC) sampling (n=10.000). Outcome variable (log) length of utterance.

Table 2: (Log) Length Model

Random effects:		**Variance**	**Std.Dev.**
child	(Intercept)	2.09	1.44
Fixed effects:	**Estimate**	**MCMCmean**	**pMCMC**
(Intercept)	1.21	1.23	0.0001
bilingual	**0.17**	**0.17**	**0.015**
age	0.01	0.01	0.0012

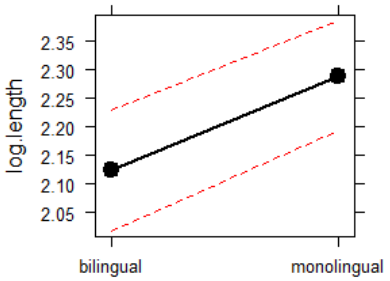

Figure 1: Effect of BILINGUAL on (log) LENGTH of construction

The analysis reveals that the there is a weak but statistically significant effect of BILINGUAL on the (logged) length of the construction (a log likehood ratio test comparing null model and model including bilingual yields Pr(Chi) < 0.05). The positive coefficient estimate in Table 2 indicates that the average construction length of monolingual learners is greater than that of bilingual learners, when age is controlled for.

Model 2: Integration
Mixed Logit Model fit by Laplace approximation. Outcome variable is proportion of integrated (= non-isolated) adverbial clauses.

Table 3: Integration Model

Random effects:		**Variance**	**Std.Dev.**
child	(Intercept)	2.09	1.44
Fixed effects:	**Estimate**	**SE**	**Pr(>|z|)**
(Intercept)	-5.64	1.87	0.00254
bilingual	**1.30**	**0.46**	**0.00461**
age	0.07	0.03	0.01469

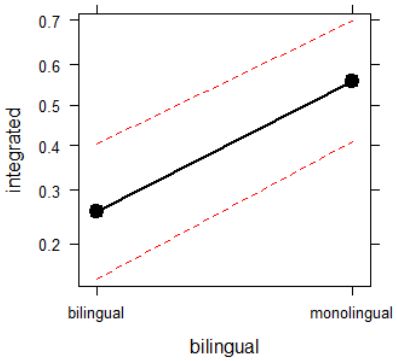

Figure 2: Effect of BILINGUAL on proportion of integrated adverbial clauses (AC)

The analysis reveals a weak but statistically significant effect of BILINGUAL on the proportion of integrated adverbial clauses (a log likehood ratio test comparing null model and model including bilingual yields Pr(Chi) < 0.01): Monolingual learners produce significantly more complex constructions (integrated ACs), when age is controlled for.

Model 3: Verb Position in Subordinate Clause
Mixed Logit Model fit by Laplace approximation). Outcome variable is proportion of AC with verb in correct (=clause final) position.

Table 4: Verb Position Model

Random effects:		**Variance**	**Std.Dev.**
child	(Intercept)	1.61	1.27
Fixed effects:	**Estimate**	**SE**	**Pr(>|z|)**
(Intercept)	0.94	1.61	0.56
bilingual	**0.46**	**0.40**	**0.25**
age	-0.01	0.03	0.75

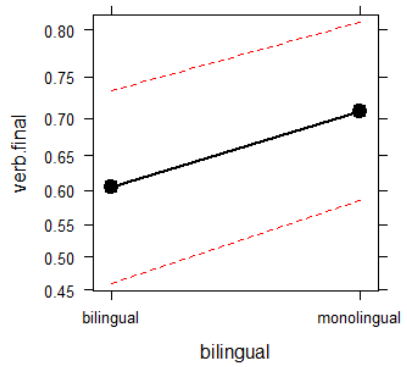

Figure 3: Effect of BILINGUAL on correct verb position in adverbial clauses (AC)

The analysis reveals a tendency for monolingual learners to produce a greater number of correct verb position but the effect is not statistically significant (a log likehood ratio test

comparing null model and model including bilingual yields Pr(Chi) > 0.25).

Model 4: Subordinator Misuse
Mixed Logit Model fit by Laplace approximation. Outcome variable is proportion of incorrectly used subordinators

Table 5: Subordinator Misuse Model

Random effects:		**Variance**	**Std.Dev.**
child	(Intercept)	5.90	2.43
Fixed effects:	**Estimate**	**SE**	**Pr(>|z|)**
(Intercept)	2.49	4.10	0.54
bilingual	**-3.74**	**1.05**	**0.00031**
age	-0.08	0.06	0.22

The analysis reveals medium sized and statistically significant effect of BILINGUAL on the proportion of correctly used subordinators (a log likehood ratio test comparing null model and model including bilingual yields Pr(Chi) < 0.001): Bilingual learners produce significantly more semantically inadequate subordinators.

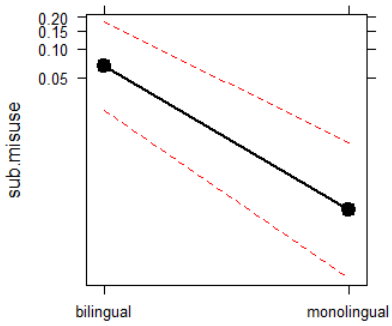

Figure 4: Effect of BILINGUAL on misuse of adverbial subordinator

Model 5: Adverbial Clause Position
Mixed Logit Model fit by Laplace approximation. Outcome variable is proportion of sentence initial adverbial clauses. This model was fit to the subset of the data that contains only those utterances that contain at least two clausal constituents, so that the adverbial clause can either precede or follow the main clause (N=588).

Table 6: AC Position Model

Random effects:		**Variance**	**Std.Dev.**
child	(Intercept)	5.39E-20	2.32E-10

Fixed effects:	**Estimate**	**SE**	**Pr(>|z|)**
(Intercept)	0.21	1.17	0.8546
bilingual	**0.56**	**0.24**	**0.0199**
age	-0.03	0.02	0.0755

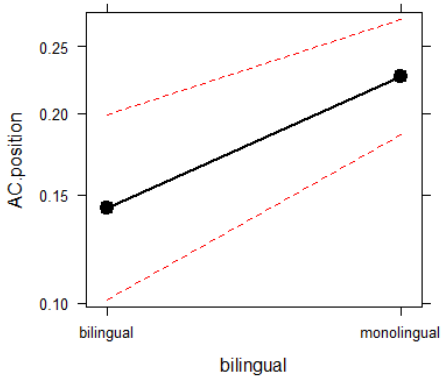

Figure 5: Effect of BILINGUAL on proportion of sentence-initial ACs

The analysis reveals a weak but statistically significant effect of BILINGUAL on the proportion of sentence initial adverbial clauses (a log likehood ratio test comparing null model and model including bilingual yields Pr(Chi) < 0.05): Bilingual learners produce significantly fewer sentence initial adverbial clauses.

Discussion

Prior research into the rate of bilingual and monolingual development has produced somewhat inconclusive results. While some studies have found similar rates of development, other studies found that bilingual children lag behind their monolingual peers in various aspects of language. Furthermore, the majority of research on the accuracy of bilingual production has been devoted to earlier phases of grammatical development such as the acquisition of the past tense (e.g. Paradis et al. 2011), the acquisition of mass/count nouns (Gathercole 2000a) or the acquisition of grammatical gender (Gathercole 2000b). Our findings contribute to this area of research by providing additional evidence from later stages of grammatical development, namely complex sentences, which constitute the last milestone in the acquisition of grammar (cf. Clahsen 1986). The research question guiding our analysis was as follows: Are bilingual children less accurate than their monolingual peers in the production of German complex sentences with adverbial clauses? *Experience-driven* or *usage-based* theories of language predict that bilingual children's acquisition of complex sentences should proceed slower due to them having less exposure, on average, to each language. We tested this general prediction across multiple dimensions. The five dimensions that served as responses in our models jointly define the space in which we measured language proficiency of monolingual and bilingual learners. We observed that bilinguals in fact lag behind in four out of five dimensions: their adverbial constructions are shorter, less often integrated into a complex sentential structure and when they are integrated, they are less often placed in

sentence initial position. Furthermore, bilingual productions exhibited a greater amount of violations of the semantic usage conditions of adverbial subordinators. This suggests that bilingual children have not yet developed a very nuanced set of words to link verbalizations of two events. Overall, our findings clearly indicate that bilinguals around age five have not caught up on their monolinguals peers in certain areas of grammar. The only dimension investigated here, where performance was equivalent across the two groups was the positioning of the finite verb in German subordinate clauses. However, as both groups are still quite removed from ceiling levels (< 80% correct usages), the equivalent performance cannot be attributed to the children's having mastered this grammatical domain. Our results also display a considerable amount of inter-individual differences as evidenced by rather pronounced intercept adjustments in the models: Some bilingual children even outperform some monolingual children across all dimensions. This gives rise to numerous questions for future research. Bilingual language learning may be dependent on factors not taken into account in the present study such as the typological distance between the two languages (interference effects) or the relative degree of balancing between the two languages being learned (language dominance).

Acknowledgments

We thank the children who participated in this study, their parents for their permission to conduct the study and all nursery school teachers for their cooperation.

References

Baayen, H. R. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, H. R. (2009). Corpus Linguistics in Morphology: Morphological productivity. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics. An International Handbook*. Berlin: de Gruyter.

Bates, D., & Maechler, M. (2009). *lme4: Linear mixed-effects models using S4 classes*. Available at: http://CRAN.R-project.org/package=lme4.

Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, Mass.: Harvard University Press.

Clahsen, H. (1982). *Spracherwerb in der Kindheit*. Tübingen: Narr.

Clahsen, H. (1986). *Die Profilanalyse. Ein linguistisches Verfahren für die Sprachdiagnose im Vorschulalter*. Berlin: Spiess.

Clahsen, H. & Muysken, H. (1986). The Availability of Universal Grammar to Adult and Child Learners: A Study of the Acquisition of German Word Order. *Second Language Research, 2*, 93-119.

De Houwer, A. (1995). Bilingual Language Acquisition. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of child language*. Cambridge, MA: Blackwell.

De Houwer, A. (2009). *Bilingual First Language Acquisition*. Bristol: Multilingual Matters.

Diessel, H. (2004). *The Acquisition of Complex Sentences*. Cambridge: Cambridge University Press.

Gathercole, V. C. M. (2002a). Command of the mass/count distinction in bilingual and monolingual children: An English morphosyntactic distinction. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children*. Clevendon, U.K.: Multilingual Matters.

Gathercole, V. C. M. (2002b). Grammatical gender in bilingual and monolingual children: A Spanish morphosyntactic distinction. In D. K. Oller & R. E. Eilers (Eds.), *Language and literacy in bilingual children*. Clevendon, U.K.: Multilingual Matters.

Gathercole, V. C. M. (2007). Miami and North Wales, so far and yet so near: A constructivist account of morpho-syntactic development in bilingual children. *International Journal of Bilingual Education and Bilingualism, 10*, 224-247.

Gathercole, V. M. & Hoff, E. (2007). Input and acquisition of language: Three questions. In E. Hoff & M. Shatz (Eds.), *The handbook of language development*. Oxford: Blackwell.

Hawkins, J. (2004). *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Lieven, E. & Tomasello, M. (2008). Children's first language acquisition from a usage-based perspective. In P. Robinson & N. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition*. New York: Routledge.

Meisel, J. M. (1986). Word order and case marking in early child language. Evidence from simultaneous acquisition of two first languages: French and German. *Linguistics, 24*, 123-183.

Miller, M. (1976). *Zur Logik der frühkindlichen Sprachentwicklung*. Stuttgart : Klett.

Nicoladis, E., Palmer, A. & Marentette, P. (2007). The role of type and token frequency in using past tense morphemes correctly. *Developmental Science, 10*, 237-254.

O'Grady, W. (2008). The emergentist program. *Lingua, 118*, 447–64.

Paradis, J., Crago, M. & Genesee, F. (2005/2006). Domain-specific versus domain-general theories of the deficit in SLI: Object pronoun acquisition by French–English bilingual children. *Language Acquisition, 13*, 33–62.

Paradis, J., Nicoladis, E. & Crago, M. (2007). French-English bilingual children's acquisition of the past tense. *Proceedings of the 31st Annual Boston University Conference on Language Development* (pp. 497-507) Somerville MA: Cascadilla Press.

Paradis, J. (2010). Bilingual children's acquisition of English verb morphology: Effects of language exposure, structure complexity, and task type. *Language Learning*

60 (3), 651–680.

Paradis, J., Nicoladis, E., Crago, M & Genesee, F. (2011). Bilingual children's acquisition of the past tense: a usage-based approach. *Journal of Child Language, 38*, 554-578.

Park, T. Z. (1981). *The development of syntax in the child with special reference to German*. Innsbruck.

Parker, M. & Brorson, K. (2005). A comparative study between mean length of utterance in morphemes (MLUm) and mean length of utterance in words (MLUw). *First Language 25*, 365-376.

Pearson, B. Z., & Fernández, S. C. (1994). Patterns of interaction in the lexical growth in two languages of bilingual infants and toddlers. *Language Learning, 44*, 617- 653.

Pérez-Leroux, A. T., Pirvulescu, M. & Roberge, Y. (2009). Bilingualism as a window into the language faculty: The acquisition of objects in French-speaking children in bilingual and monolingual contexts. *Bilingualism: Language and Cognition, 12*, 97-112.

Roeper, T. (1973). Connecting children's language and linguistic theory. In T. Moore (Eds.), *Cognitive development and the acquisition of language*. New York: Academic Press.

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