

BRIEF RESEARCH REPORT

**Error patterns in young German children's  
*wh*-questions\***

DANIEL SCHMERSE, ELENA LIEVEN  
AND MICHAEL TOMASELLO

*Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany*

(Received 29 April 2010 – Revised 6 September 2011 – Accepted 16 February 2012 –  
First published online 28 May 2012)

ABSTRACT

In this article we report two studies: a detailed longitudinal analysis of errors in *wh*-questions from six German-learning children (age 2;0–3;0) and an analysis of the prosodic characteristics of *wh*-questions in German child-directed speech. The results of the first study demonstrate that German-learning children frequently omit the initial *wh*-word. A lexical analysis of *wh*-less questions revealed that children are more likely to omit the *wh*-word *was* ('what') than other *wh*-words (e.g. *wo* 'where'). In the second study, we performed an acoustic analysis of sixty *wh*-questions that one mother produced during her child's third year of life. The results show that the *wh*-word *was* is much less likely to be accented than the *wh*-word *wo*, indicating a relationship between children's omission of *wh*-words and the stress patterns associated with *wh*-questions. The findings are discussed in the light of discourse–pragmatic and metrical accounts of omission errors.

Children's production of errors in *wh*-questions provides an interesting test case for theoretical approaches to language acquisition and syntactic development. Construction-based accounts argue that children formulate questions using lexically specific frames (e.g. *What is [THING] [PROCESS]?* or *What does [THING] [PROCESS]?*) and that error patterns reflect the item-specific nature of children's *wh*-questions (Ambridge, Theakston, Lieven & Tomasello, 2006). Based on data from

[\*] Address for correspondence: Daniel Schmerse, Max Planck Institute for Evolutionary Anthropology – Developmental and Comparative Psychology, Deutscher Platz 6 Leipzig 04103, Germany. e-mail: daniel\_schmerse@eva.mpg.de

Norwegian, a Germanic V2 language, Westergaard (2009) argued that some non-target consistent forms such as omission of verbs or *wh*-words are incompatible with constructivist accounts. According to Westergaard's account, omission errors disconfirm the constructivists' assumption that children primarily rely on frequent input patterns (specific *wh*-word + verb frames) when formulating interrogatives.

In the current study, we present data from German children's production of *wh*-questions to investigate whether *wh*-omission errors occur in lexically specific frames and whether input properties such as prosodic patterns or discourse givenness may influence their production.

In reports of German-learning children's *wh*-questions it has been found that children omit the utterance initial *wh*-word especially during early stages of development (Clahsen, Kursawe & Penke, 1995; Tracy, 1994). Interestingly, *wh*-word omission has also been reported to occur frequently in other Germanic V2 languages such as Dutch (Van Kampen, 1997), Swedish (Santelmann, 2004) and Norwegian (Westergaard, 2009), but, to our knowledge, has not been reported systematically for English. Other examples of errors in German children's early *wh*-questions include verb doubling (Penner, 1994), subject omission (Hamann, Penner & Lindner, 1998), verb omission (Steinkrauss, 2009) and non-inversion errors (Wode, 1975).

However, overall little is known about the frequency with which German children produce these different types of errors in formulating questions. Therefore, we performed a detailed longitudinal analysis of the different types and rates of errors in German children's *wh*-questions. Our hypothesis was that German children would show high error rates of verb omission and omission of *wh*-words, as this seems to be the most common error in *wh*-questions across typologically similar languages.

The second objective of the current study concerned the factors that might explain *wh*-omission errors. We wanted to know whether *wh*-omission occurs in lexically specific frames and whether input properties such as prosodic characteristics of caregiver speech may have an influence on error patterns. We hypothesized that if lexical specificity constrains the production of errors, we would not expect lexical overlap between those questions with omission and those without. However, if lexical specificity alone does not constrain the production of errors, additional factors have to be considered in order to explain under what conditions utterance-initial *wh*-elements are omitted. Two possibilities are that omission errors might be influenced by pragmatic factors (given information in utterance-initial position being omitted more often than new information), and prosodic factors (unstressed elements being omitted more often than stressed).

## STUDY 1

## METHOD

*Speech corpus*

All *wh*-questions were extracted from the longitudinal data of six children from the Szagun corpus (Szagun, 2004), available from the CHILDES database (MacWhinney, 2000). The analyzed utterances were only taken from typically developing children. The recordings were made every six weeks during the children's third year of life. The results reported here are based on data collected when the children were between 2;0 and 3;0. The data for Emely were taken from recordings between 2;0 and 3;4 in order to obtain a higher number of utterances. The mean length of utterance (MLU) was calculated in words per utterance by the first author. MLUs for the children ranged between 1.03 and 2.10 at age 2;0 and between 2.16 and 4.31 at age 3;0 (Table 1). We classified all *wh*-questions according to MLU stages: stage I (1–1.99), stage II (2.0–2.49), stage III (2.5–3.0) and stage IV (> 3.0).

*Error coding*

The first author coded the following types of errors: *wh*-omission errors, verb-omission errors, subject-omission errors and non-inversion errors. Verb doubling errors, which have also been reported, were not found in the data.

- (a) WH-OMISSION. *wh*-omission errors were defined as interrogative question structures containing a verb in initial position (their status as questions was determined from the context, including intonation and the interlocutor's response).
- (1) *macht das pferd?*  
doing the horse?  
'(What) is the horse doing?' (Ann, 2;5)
- (b) VERB-OMISSION. *wh*-questions containing a *wh*-word but no verb were coded as verb-omission errors.
- (2) *wo die pfanne?*  
where the pan?  
'Where (is) the pan?' (Soe, 2;6)
- (c) SUBJECT-OMISSION. *wh*-questions containing a *wh*-word and a verb, but which did not have a subject were coded as subject-omission errors.
- (3) *wo is?*  
where is?  
'Where is (X)?' (Lis, 2;4)
- (d) NON-INVERSION ERRORS. Non-inversion errors were determined by position of the finite verb. Since in German *wh*-question finite verbs

occur in second position (i.e. after the *wh*-word), questions that deviated from the verb-second word order were coded as word order errors.

(4) *warum eine frau das ist?*

why a woman that is?

'Why is that a woman?' (Eme, 2;9)

- (e) INTERROGATIVE CONTEXTS. In order to give percentages of error types, we summarized erroneous structures and correct questions as INTERROGATIVE CONTEXTS. For the purposes of *wh*-word specific analysis, we furthermore identified two subcategories of interrogative contexts: *wo*-contexts ('where'-contexts) and *was*-contexts ('what'-contexts). Single-word *wh*-questions, embedded *wh*-questions and fragments were excluded from the analysis (e.g. *Welche X?* ('Which X?'), *Wie bitte* ('Pardon?'), *Was für ein X?* ('What kind of X?'), *Warum nicht* ('Why not?'), *[Wh] denn?* ('[Wh] denn<sub>[particle]</sub>?'). We also excluded seventy-seven questions that contained neither a verb nor a *wh*-word. These structures were only marked with the modal particle *denn*, which is commonly used in German *wh*-questions.

### Coding reliability

A second rater was trained in error coding by the first author and coded a total of 180 interrogative contexts from all six children (12% of the data). The level of agreement between coders was 97.2% (Cohen's Kappa = 0.93,  $p < 0.001$ ,  $N = 180$ ).

## RESULTS

### Overall error rates

Children produced errors in approximately 30% of their questions. Table 1 shows a clear pattern of these errors. First, the omission of utterance-initial elements (*wh*-words and verbs) is the most common error for all children. Verb-omission occurred in 15.2% and *wh*-omission occurred in 13.4% of all *wh*-questions, which is comparable to the findings of Clahsen *et al.* (1995), who reported a *wh*-omission rate of 19%. Second, non-inversion errors and subject-omission errors are extremely rare and do not occur in all children. Most of the children's non-inversion errors contained a finite verb in final position. However, there were also four instances in which children produced a non-finite verb in final position, but no finite verb in second position (e.g. *Wo der passen?* 'Where this fit?'). Furthermore, there are large individual differences in both the production rate of *wh*-questions as well as occurrences of particular types of errors. For example, Emely produced the lowest number of *wh*-contexts, but showed the highest rate of

TABLE I. *Total numbers and percentages of errors produced in children's wh-questions*

Child	Age	MLU range	Int. Contexts	<i>wh</i> -omission		Subject-omission		Verb-omission		Non-inversion	
				%	(N)	%	(N)	%	(N)	%	(N)
Ann	2;0-3;0	1.23-3.00	640	8.1	(52)	0.3	(2)	20.2	(129)	0.5	(3)
Eme	2;0-3;4	1.03-2.16	72	62.5	(45)	0.0	(0)	2.8	(2)	2.8	(2)
Fal	2;0-3;0	1.66-4.31	137	5.1	(7)	0.7	(1)	19.0	(26)	2.2	(3)
Lis	2;0-3;0	2.05-3.05	192	16.1	(31)	0.5	(3)	8.3	(16)	0.0	(0)
Rah	2;0-3;0	1.36-3.73	103	16.5	(17)	4.8	(5)	15.5	(16)	1.9	(2)
Soe	2;0-3;0	2.10-3.38	360	13.6	(49)	0.0	(0)	11.1	(40)	0.8	(3)
		TOTAL	1504	13.4	(201)	0.7	(11)	15.2	(229)	0.9	(13)

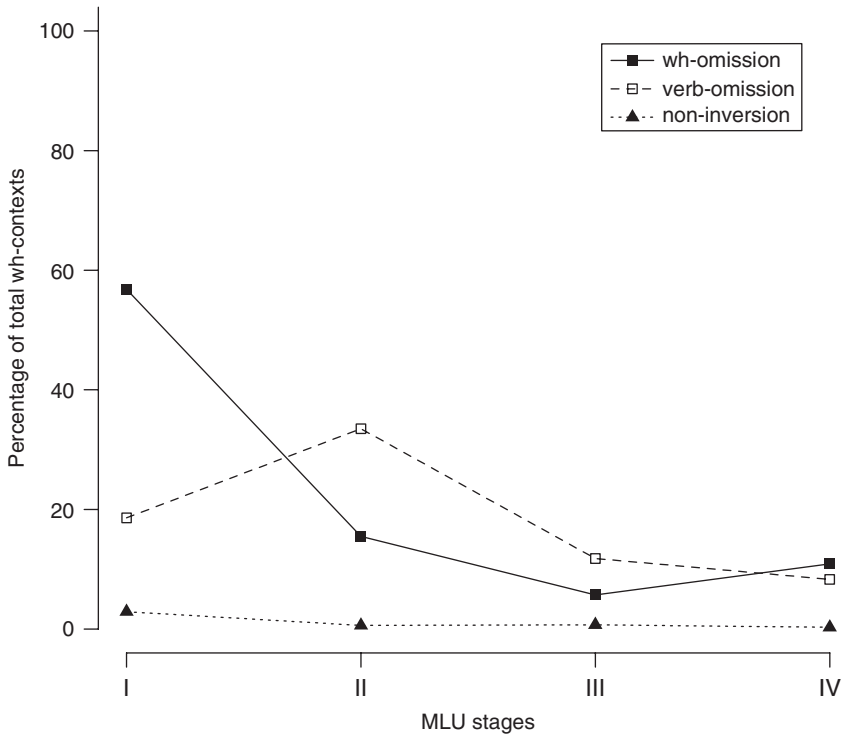


Fig. 1. Error rates for *wh*-omission, verb-omission and non-inversion errors across MLU stages.

*wh*-omission errors (62.5%). However, a closer look revealed that 92.5% of Emely's *wh*-less questions consisted of only two lexical formulas: (*Was*) *ist das?* 'What is that?' and (*Wie*) *heisst der?* 'How is he called?', each of which occurred in a single but different recording session.

Figure 1 shows the distribution of error rates across MLU stages. The omission of *wh*-elements and verbs is a phenomenon that is particularly characteristic of the MLU stages I and II. One reason for the decline in error rates might be that omission of elements is bound to specific lexical frames that constitute the majority of question constructions in early phases, but the proportion of these decreases as children acquire more types of *wh*-questions.

We compared the use of verb types in *wh*-overt and *wh*-less questions across MLU stages. Figure 2 shows the mean number of different verb types produced at each MLU stage. The results indicate that an increase in the use of verb types was seen only for *wh*-overt questions, but not for constructions in which the *wh*-word was omitted. A lexical analysis showed

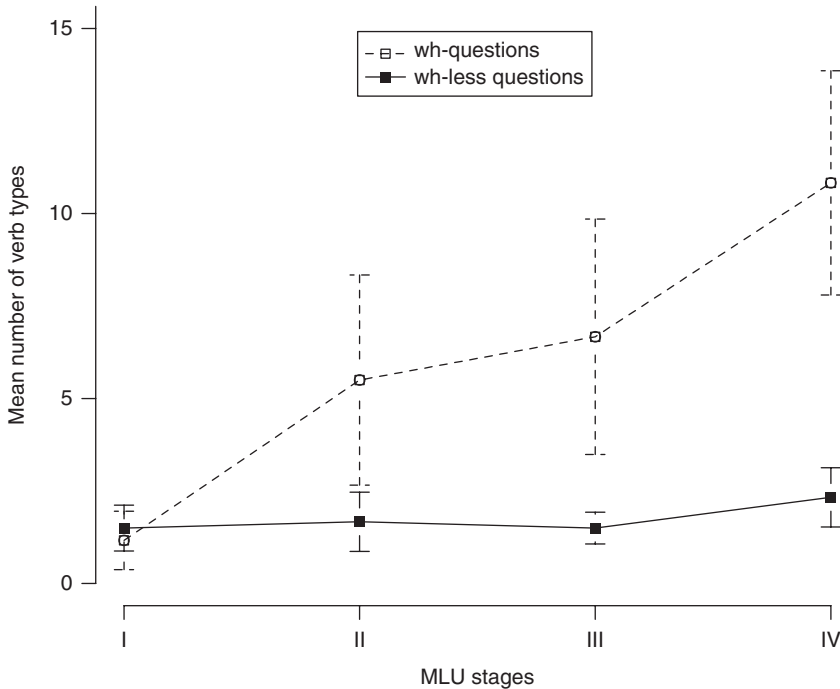


Fig. 2. Mean number of verb types used in *wh*-less and *wh*-overt questions across MLU stages. Error bars show standard errors.

that children omit *wh*-words only with a limited set of verbs. All *wh*-omission errors occurred with one of the following eight verbs (including their inflectional forms): *machen* ('do'), *passen* ('fit'), *kommen* ('come'), *gehören* ('belong'), *haben* ('have'), *heissen* ('be called'), *gehen* ('go'), and the copula. Furthermore, 78% of all *wh*-omission errors occurred in a set of five lexical frames: (*Was*) *ist das* '(What) is that' (N = 97), (*Wie*) *heisst der/die/das* '(How) is he/she/it called' (N = 28), (*Was*) *machst du* '(What) are you doing' (N = 17), (*Was*) *macht der/die* '(What) is he/she doing' (N = 8), (*Wo*) *kommt das hin* '(Where) does this go' (N = 7). What these frames have in common is that they contain a pronoun in subject position. Next, we checked whether *wh*-less questions and *wh*-overt questions differ with respect to lexical specificity in the verb position and the subject-NP for each child individually.

#### *Verb use and type of subject-NP*

From the sample of *wh*-overt questions, we extracted all non-subject questions and all instances of the construction [WH COP NP] and coded

TABLE 2. Total numbers of verb types and type of subject-NPs in *wh*-questions and *wh*-less questions

Child	<i>wh</i> -questions ( <i>N</i> = 957)			<i>wh</i> -less questions ( <i>N</i> = 195)		
	Verb types	NP-subjects	Pn-subjects	Verb types	NP-subjects	Pn-subjects
Ann	31	94	314	6	3	46
Eme	6	9	14	4	1	40
Fal	14	47	39	2	2	5
Lis	15	41	95	3	10	20
Rah	8	23	33	5	1	16
Soe	13	60	188	5	0	52
TOTAL	43	274	683	8	17	179

how many different verb types occurred in the position following the *wh*-word and whether the subject was realized pronominally or as a full NP. We applied the same coding procedure for *wh*-less questions.

We found that all children use more verb types in overt *wh*-questions than in *wh*-less questions (see Table 2). For Ann, Fal, Lis and Soe this difference was more pronounced than for the other two children (Eme and Rah), who also showed the lowest production rate of interrogative contexts overall. However, although children tended to produce fewer verb types in *wh*-less questions, these verbs were not restricted to *wh*-less questions. Four of six children (Ann, Fal, Lis and Soe) used all verbs from their *wh*-less questions in *wh*-overt questions as well. Rahel produced two verb types in both structures (copula, *kommen* 'come') and Emely three verb types (copula, *heissen* 'be called', *gehören* 'belong'). With respect to subject-NP type, children predominantly used pronoun subjects in *wh*-questions as well as in *wh*-less questions (71.4% and 91.3%, respectively). However, the type of subject realization did not distribute equally over the two structures with pronouns being more frequent in *wh*-less questions ( $\chi^2(1, N = 1152) = 28.59$ ,  $p < 0.001$ ).

#### *Omission as complexity reduction?*

Bloom (1990) proposed that the omission of sentential elements could be explained as a general cognitive strategy to reduce the complexity of an utterance. He found that sentences with longer verb phrases (VP) tend to be produced less often with a subject than sentences with shorter verb phrases. Therefore, we tested whether such a VP length effect could also be found in the case of omission errors in *wh*-questions. The MLU for *wh*-questions was 4.28, and 2.93 for *wh*-less questions. Subtracting the *wh*-word from every overt *wh*-question yielded a VP length of 3.28, a number that was still



TABLE 3. Children's omission errors for different *wh*-contexts (*Szagan corpus*)

	<i>Wo</i> -contexts	<i>Was</i> -contexts
Verb-omission	124	93
<i>wh</i> -omission	23	147

significantly larger than the MLU of *wh*-less questions ( $t(293)=4.54$ ,  $p<0.001$ ). Therefore we conclude that children do not omit the *wh*-word in order to reduce the length of the utterance. This argument is further supported by the fact that 96% of all *wh*-less questions contain a semantic 'light' verb (so they are not more semantically difficult either).

#### *wh*-specific errors

When considering errors in children's language we must always note whether the incidence of errors patterns uniformly across different types of lexically specific structures. To answer this question, we analyzed the rates of omission errors in *wo*-contexts and *was*-contexts. Overall, there were 750 *wo*-contexts and 614 *was*-contexts, which together constituted 90.7% of all interrogative contexts.

Table 3 shows the number of omission errors for *was*-contexts and *wo*-contexts. The rates of verb-omission are very similar in both contexts, with 16.6% of all *wo*-contexts ( $n=124$ ) and 15.1% of all *was*-contexts ( $n=93$ ) missing a verb. Although it is not possible to determine which verb the child intended to produce, it seems that in the vast majority of cases it is the copula that has been dropped. It should be pointed out that in *was*-contexts the rate of copula omission might be overestimated. This is because the singular form of the copula *ist* and the *wh*-word *was* are often reduced into one contracted form for which it becomes hard to distinguish whether the copula is present or not. But the key finding is that a large difference in error rates was found for the omission of different *wh*-words. The data shows that the *wh*-word *was* is significantly more likely to be dropped than the *wh*-word *wo* ( $\chi^2(1, N=387)=75.13$ ,  $p<0.001$ ). The *wh*-word is missing in only 3.1% of all *wo*-contexts ( $n=23$ ), compared to 23.9% of all *was*-contexts ( $n=147$ ).

#### DISCUSSION

Summarizing the results, we find a clear pattern of errors that young German-learning children produce when forming *wh*-questions. First, omission of utterance-initial elements such as the *wh*-word or the verb in second position can be identified as the main source of error. Taken

together, in 28.6% of all interrogative contexts, either the *wh*-word or the verb is omitted. Second, non-inversion errors constitute a rare phenomenon. Interestingly, similar low rates of word order errors have also been reported for other V2 languages (for Swedish, see Hansson & Nettelbladt, 2006).

An analysis at the lexical level revealed that *wh*-omission errors are restricted to particular lexical items. First, children mostly drop the *wh*-word *was* but preserve the *wh*-word *wo*. Second, *wh*-omission errors occur only with a small set of verbs throughout all MLU stages. Third, almost all *wh*-less questions contain a pronoun subject following the verb, indicating that the referent of the subject-NP is given in the discourse or even present in the interaction. Finally, the majority of children used verbs occurring in *wh*-less questions in *wh*-overt questions as well, suggesting that additional factors other than lexical specificity must be involved in the omission of *wh*-words.

Notice that the lexical pattern of *wh*-less constructions has important implications for the prosodic structure of these frames. In our data, the vast majority of *wh*-less questions have the lexical form [(WH) VERB Pn]. According to Lambrecht (1994), in constructions of the form [WH VERB NP], NPs whose referents have already been established in discourse (as indicated by their pronominal form) are unlikely to receive an accent. The sentence accent in these constructions therefore falls onto the verb, giving *wh*-less questions a strong–weak stress pattern. It is important to point out that this is true for semantically light as well as semantically heavy verbs, since utterances must have at least one accent to be informative (Lambrecht & Michaelis, 1998).

These information structure considerations as well as the observation that lexical specificity alone cannot explain *wh*-omission point to the possibility that omission errors might result from a prosodic constraint, as proposed by Gerken (1991; 1994). According to Gerken's account children tend to omit weakly stressed elements and favour the production of strong–weak sequences over weak–strong sequences. We investigated this hypothesis in a second study by analyzing the prosodic characteristics of different types of *wh*-questions in German child-directed speech (CDS).

## STUDY 2

Gerken (1991) proposed a metrical account to explain English children's omission of sentential subjects. According to her hypothesis, the speech production system of English-learning children around the age of two is influenced by a prosodic constraint favouring strong–weak sequences over weak–strong sequences. More precisely, children tend to omit weakly stressed syllables that cannot be parsed into a trochaic foot. Furthermore,

Gerken (1994) has shown in an imitation experiment that weak syllables in utterance-initial position are more likely to be dropped than weak syllables in utterance-internal or -final position. Since German, like English, shows a predominant trochaic stress pattern in multisyllabic words, it is likely that young German-learning children operate with similar production constraints. We reasoned that if the prosodic characteristics of the ambient language influence children's production, the different omission rates of *wh*-words in *was*- and *wo*-questions might be traced to specific stress patterns associated with *wh*-questions in German child-directed speech.

## METHOD

### *Speech corpus*

Since the Szagun corpus analyzed in Study 1 is not linked to sound files, *wh*-questions for acoustic analysis were taken from the dense German child language corpus of Leo (Behrens, 2006), for which the sound files were available to us. To check whether Leo himself produced *wh*-less questions, a short analysis of Leo's *wh*-questions from age 2;1 to 2;5 (MLU range 1.67–2.12) was performed according to the same coding scheme described above. During this period Leo produced forty-three *wh*-questions. In five of these questions Leo omitted the *wh*-word (11.3% of all *wh*-contexts). No instances of *wh*-omission errors were found in the months 2;4 and 2;5. Thus, Leo showed a compatible developmental trajectory to that of the children in the Szagun data. We therefore analyzed the input Leo received to check how characteristics of maternal utterances might have influenced the observed patterns.

### *Sentence material*

We extracted sixty maternal *wh*-questions that the mother produced during the child's third year of life. In order to match the use of verb forms and to obtain a sample that was large enough for statistical analysis, we chose the three most frequent *wo*+verb combinations and the matching *was*+verb combinations (10 instances per combination):

- *Wo ist* ('Where is')                      *Was ist* ('What is')
- *Wo sind* ('Where are')                    *Was sind* ('What are')
- *Wo kommt* ('Where comes')           *Was kommt* ('What comes')

### *Coding*

A professional phonetician (native speaker of German) coded the sound files for presence of stress on the *wh*-word. The files were presented in a randomized order. The first author (also a native speaker of German) coded

TABLE 4. *Distribution of accents on wh-words in child-directed speech (Leo corpus)*

	<i>wh</i> -word accented	<i>wh</i> -word not accented
<i>Was</i> -questions ( $n=30$ )	9	21
<i>Wo</i> -questions ( $n=30$ )	19	11

the complete set of sound files for reliability. The agreement rate was 85% (Cohen's  $\kappa=0.81$ ,  $p < 0.001$ ,  $N=60$ ). This agreement rate lies within a well-established range for judging the presence of accents on words in sentential contexts for German (Grice, Reyelt, Benz Müller, Mayer & Batliner, 1996).

#### RESULTS

Table 4 shows the distribution of accents across *wo*-questions and *was*-questions. The results indicate that the *wh*-word in *wo*-questions is accented significantly more often than the *wh*-word in *was*-questions (Fisher's exact test;  $p=0.02$ ).

#### DISCUSSION

An intonation unit in spoken language contains more than just one accent, i.e. pre-nuclear and nuclear pitch accents. Thus, an accented *wh*-word is not automatically the most prominent part of the utterance. However, speakers strongly tend to avoid accenting two adjacent lexical elements (Selkirk, 1984; Speyer, 2008). With respect to *wh*-questions this means that if the *wh*-word is accented, this automatically prevents the verb following the *wh*-word from being accented as well. Furthermore, in *wo*-questions the NP following the verb prototypically is a full NP (e.g. *Wo ist X?* 'Where is X?'), which receives a topic-establishing accent (Lambrecht, 1994). In *was*-questions, on the other hand, the NP following the verb prototypically is a pronoun whose referent has been activated in previous discourse and it is therefore not stressed, i.e. the accent falls onto the verb (e.g. *Was MACHT der?* 'What is he DOING?').

These information structure considerations help to explain the main result of Study 2, that is, that *was*-questions in German CDS have a tendency to show an initial weak-strong pattern, whereas *wo*-questions do not.

#### GENERAL DISCUSSION

The current study addressed whether German children's errors in *wh*-questions reflect prosodic patterns of the input and whether errors

distribute uniformly across different types of *wh*-constructions. With respect to the latter question, we found clear evidence that errors of *wh*-omission tend to occur more frequently in *was*-questions than in *wo*-questions. The explanation for this pattern of errors may be found in both pragmatic and prosodic factors, which, of course, are intimately intertwined in the information structure of sentences in all languages. In the following, we discuss the discourse–pragmatic and the prosodic account of omission errors and evaluate which explanation accounts best for our data.

#### *Where do omission errors come from?*

Two types of factors are often assumed to influence omission errors in child language: the metrical structure of the target utterance and the givenness status of information (Gerken, 1994; Hughes & Allen, 2006). The present analysis suggests that prosodic factors do have an influence on omission errors in *wh*-questions, although experimental data is needed to distinguish more clearly between the two accounts.

A discourse–pragmatic explanation of omission errors would predict that children drop the *wh*-word in contexts where the designatum of the *wh*-word is inferable from the preceding discourse or present situation. Santelmann (2004), for example, proposed for *wh*-less questions in child Swedish that children produce these structures in analogy to so-called topic-drop constructions. Similarly, Jordens and Dimroth (2006) suggest that children acquiring Dutch order constituents sequentially on the basis of functional categories such as topic, linking element and predicate. Thus, in the child's grammar the utterance-initial position is identified as the topic-position, which has an anchoring function towards discourse. According to Santelmann's account, as well as Jordens and Dimroth's proposal, children learn that the utterance-initial position (topic position) can be left empty if the topic expression is inferable from the context and they overgeneralize this knowledge to *wh*-questions.

Although German, as Swedish and Dutch, allows topic-drop constructions and they seem to occur frequently in CDS (Hamann & Plunkett, 1998; Hamann, 2002), it is not clear whether children derive a functional interpretation of the utterance-initial position for *wh*-questions based on declarative topic-drop constructions. In fact, *wh*-questions are functionally quite distinct from declarative topicalized structures and are generally analyzed as focus–argument constructions rather than topic–argument constructions (see Lambrecht, 1994). What is pragmatically presupposed in *wh*-questions is not the *wh*-expression (the initial position), but the open proposition (everything following the *wh*-word). Thus, by asking *Who is coming tonight?* the speaker presupposes that someone will be coming tonight (and that the hearer is in a position to identify that person).

Crucially, an account of *wh*-less questions in analogy to topic-drop constructions would have to explain the asymmetry in omission errors between the *wh*-words *was* and *wo*. Such an explanation would have to show that in child-directed speech the designatum of the *wh*-word *wo* (i.e. the location or moving direction of an entity) is not inferable from the context, whereas the designatum of the *wh*-word *was* is inferable.

Alternatively, the metrical account predicts that during the early stages of development children omit unstressed syllables, which cannot be parsed into the prosodic constituent foot (Demuth, 1994). This prosodic constraint influences children's production of multisyllabic words as well as grammatical morphemes, such as determiners. Gerken (1991) showed that children are more likely to omit the second weak syllable (definite article) than the first weak syllable (*-es*) in the phrase *KISSes the PIG*, because it is not part of a foot. Furthermore, Freudenthal, Pine and Gobet (2007) used a computational modelling approach to demonstrate that the rate of English children's optional infinitive errors (e.g. *he go home*) can be simulated as the omission of function words, i.e. the omission of the modal *will* from *he will go home*, based on the metrical characteristics of the children's input. These findings demonstrate how a model that relies on learning from frequent input patterns in child-directed speech including the prosodic characteristics of that input can successfully account for the production of errors. This is generally in line with the results of the current study.

A metrical account might also explain why we do not see similar rates of *wh*-less questions in English children's early speech. The metrical account would predict that the sequence Wh+AUX does not exhibit a weak-strong pattern, because the auxiliary or copula in English non-subject *wh*-questions is almost never stressed. In line with this hypothesis is the observation that English children frequently omit the copula or the auxiliary verb in *wh*-questions (e.g. *where he going?*). For example, Rowland, Pine, Lieven and Theakston (2005) report for the Manchester corpus auxiliary/copula omission rates between 24% and 50% across MLU stages. These authors have also found some lexically specific omission patterns and speculate that phonological factors might explain their findings (e.g. copula *are* being more likely to be omitted than copula *is*). However, it remains an empirical question whether the omission of auxiliary verbs and copula forms in English children's *wh*-questions and the omission of *wh*-words in German children's *wh*-questions can be accounted for by the same prosodic constraint operating on children's early production.

In sum, the two studies presented here indicate a relationship between children's omission rates and the prosodic characteristics of interrogative questions and suggest that construction-based approaches have to take into account prosodic structure in addition to discourse-pragmatic factors and lexical specificity to explain the full range of the data.

## REFERENCES

- Ambridge, B., Theakston, A., Lieven, E. & Tomasello, M. (2006). Comparing different accounts of non-inversion errors in children's non-subject wh-questions: 'What experimental data can tell us?' *Journal of Child Language* **30**, 519–57.
- Behrens, H. (2006). The input–output relationship in first language acquisition. *Language and Cognitive Processes* **21**, 2–24.
- Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry* **21**, 491–504.
- Clahsen, H., Kursawe, C. & Penke, M. (1995). Introducing CP: wh-questions and subordinate clause in child language. Paper presented at the Groningen Assembly on Language Acquisition, Groningen, Center for Language and Cognition, September, 1995.
- Demuth, K. (1994). On the underspecification of functional categories in early grammars. In B. Lust, M. Suñer & J. Whitman (eds), *Syntactic theory and first language acquisition: Cross-linguistic perspectives*, 119–34. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Freudenthal, D., Pine, J. & Gobet, F. (2007). Meter based omission of function words in MOSAIC. In R. L. Lewis, T. A. Polk & J. E. Laird (eds), *Proceedings of the 8th International Conference on Cognitive Modelling*, 219–24. New York: Psychology Press.
- Gerken, L. A. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language* **30**, 431–51.
- Gerken, L. A. (1994). Young children's representation of prosodic structure: Evidence from English-speakers' weak syllable omissions. *Journal of Memory and Language* **33**, 19–38.
- Grice, M., Reyelt, M., Benz Müller, R., Mayer, J. & Batliner, A. (1996). Consistency in transcription and labeling of German intonation with GToBI. In T. Bunnell & W. Isardi (eds), *Proceedings of the 4th International Conference on Spoken Language Processing*, Philadelphia, 1716–19. Philadelphia, PA: I. E. E. E. Press.
- Hamann, C. (2002). *From syntax to discourse: Pronominal clitics, null subjects, and infinitives in child language*. Dordrecht: Kluwer.
- Hamann, C., Penner, Z. & Lindner, K. (1998). German impaired grammar: The clause structure revisited. *Language Acquisition* **7**, 193–245.
- Hamann, C. & Plunkett, K. (1998). Subjectless sentences in child Danish. *Cognition* **69**, 35–72.
- Hansson, K. & Nettelbladt, U. (2006). Wh-questions in Swedish children with SLI. *International Journal of Speech-Language Pathology* **8**, 376–83.
- Hughes, M. & Allen, S. (2006). A discourse-pragmatic analysis of subject omission in child English. In D. Bamman, T. Magnitskaia & C. Zaller (eds), *Proceedings of the 30th Annual Boston University Child Language Conference*, 293–304. Somerville, MA: Cascadilla Press.
- Jordens, P. & Dimroth, C. (2006). Finiteness in children and adults learning Dutch. In N. Gagarina & I. Gülzow (eds), *The acquisition of verbs and their grammar: The effect of particular languages*, 173–200. Dordrecht: Springer.
- Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents*. Cambridge: Cambridge University Press.
- Lambrecht, K. & Michaelis, L. A. (1998). Sentence accent in information questions: Default and projection. *Linguistics and Philosophy* **21**, 477–544.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*, 3rd edn. Mahwah, NJ: Lawrence Erlbaum Associates.
- Penner, Z. (1994). Asking questions without CPs? On the acquisition of root wh-questions in Bernese Swiss German and standard German. In T. Hoekstra & B. Schwartz (eds), *Language acquisition studies in generative grammar*, 177–214. Amsterdam: John Benjamins.
- Rowland, C., Pine, J., Lieven, E. & Theakston, A. (2005). The incidence of error in young children's wh-questions. *Journal of Speech, Language, and Hearing Research* **48**, 384–404.
- Santelmann, L. (2004). Wh-questions in early Swedish. In G. Josefsson, C. Platzack & G. Håkansson (eds), *The acquisition of Swedish grammar*, 261–308. Amsterdam: John Benjamins.

- Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Speyer, A. (2008). Topicalization and clash avoidance: On the interaction of prosody and syntax in the history of English with a few spotlights on German. Unpublished PhD dissertation, University of Pennsylvania.
- Steinkrauss, R. (2009). *Frequency and function in wh-question acquisition: A usage-based case study of German L1 acquisition*. Groningen dissertations in linguistics, University of Groningen.
- Szagan, G. (2004). Learning by ear: On the acquisition of case and gender marking by German-speaking children with cochlear implants and with normal hearing. *Journal of Child Language* 31, 1–30.
- Tracy, R. (1994). Raising questions: Formal and functional aspects of the acquisition of wh-questions in German. In R. Tracy & E. Lattey (eds), *How tolerant is Universal Grammar: Essays on language learnability and language variation*, 1–34. Tübingen: Niemeyer.
- Van Kampen, J. (1997). *First steps in wh-movement*. Delft: Eburon.
- Westergaard, M. (2009). Usage-based vs. rule-based learning: The acquisition of word-order in wh-questions in English and Norwegian. *Journal of Child Language* 36, 1023–51.
- Wode, H. (1975). Some stages in the acquisition of questions by monolingual children. *Word* 27, 261–310.