Copula omission: A deeper look into one explanation

Strohman, A.

Introduction

When a child acquires English as a first language, there is a long period of grammatical development that may stretch into the fourth year of life (de Villiers & de Villiers, 1973). During this period, there is a phenomenon in which children omit specific grammatical morphemes, one of which is copula *be*. This period of grammatical development has been characterized as the Optional Infinitive (OI) Stage (Wexler, 2011). This stage is a cross-linguistic phenomenon, but some languages tend to have more drop qualities than others. One of these languages is English. According to Rice, Wexler and Hershberger (1998), copula omission is part of the extended OI phase in children with Specific Language Impairment (SLI), and children omit due to a processing difficulty. Additionally, problems in the Optional Infinitive Stage have shown to be a clear clinical marker of language delay (Hadley & Short, 2005). According to this study, analyzing the onset of tense marking, specifically copula *be*, can be used to evaluate grammatical development in children.

Currently, there are numerous approaches as to why children omit copula. These approaches highlight the various theories of child language development. For example, *Function morphemes in young children's speech perception and production* by Gerken and Landau (1990) discusses the metrical template explanation as to why children omit. In this study, three experiments were performed to analyze how children distinguish between different function morphemes. The purpose was to observe the idea that children

may be able to detect these specific functors even though they omit them during speech. Mainly, the experimenters were curious as to what role stress played in the children's omission. In the first experiment, 9 girls and 7 boys between 23-30 months of age were asked to imitate simple four syllable sentences. The results showed that syllables with more stress were more likely to be preserved in children's speech. Additionally, these children utilized nonsense words in the sentences they repeated meaning they produced other syllables with strong stress. In the second experiment 8 boys and 7 girls between the ages of 24-30 months were asked to repeat simple sentences similar to those in experiment one. However, in experiment two the stimuli sentences were presented through an electronic device. The same results were found in this experiment: children tended to preserve the strong syllables and omit the weak ones. In the third and final experiment, it was tested to see how segmenting the sentences affected the patterns of omission. There were two groups of children. The first group made up of 6 boys and 2 girls, ages 24-27 months, were instructed to repeat natural voice. The second group made up of 5 boys and 3 girls, ages 26-30 months, were played a recording of synthetic voice. Yet again, this experiment yielded similar results in which strong stress was preserved and weak stress was not. Gerken and Landau use this idea of metrical template to explain this omission of weak syllables and how it is incorporated in language learning.

To further our discussion on the history of research, *Grammatical computation in the Optional Infinitive stage* by Wexler (2011) provides insight to this stage of development. While this paper does not offer a specific study, it details the major proponents of the Optional Infinitive (OI) stage. Wexler defines the properties of the OI stage as follows: "root infinitives are possible grammatical sentences for children in this stage, these infinitives

co-exist with finite forms, children know the relevant grammatical principles and have set the relevant parameters correctly" (Wexler, 2011). The article also claims that young children show poor performance in inflectional morphemes. According to Wexler, a major component of the OI stage is auxiliary and copula omission. One hypothesis of why auxiliaries are omitted is that they are only there to bind morphosyntactic features like tense and agreement. Furthermore, during the OI stage, children understand that *be* is required whenever tense exists and therefore do not omit in this context. However, more often than not the tense is omitted in the OI stage resulting in copula omission (Wexler, 2011). Wexler claims that these errors are due to some kind of genetically specific development, and there is little other evidence to explain the omission phenomenon.

This leads us to *Copula omission is a grammatical reflex* by Misha Becker (2004). This study evaluates both grammar-based and processing-based explanations for children's omission of the copula *be*. This study used transcripts from four children, ages 2;0 to 3;4, from the CHILDES database. The file used for each child was one in which copula was used in all construction types. Becker found that the grammar-based interpretation was supported by the data while the processing-based account was not. Particularly, the article states that "a correlation was found between omission of the copula and a semantic feature of the predicate" (Becker 2004). It is also observed that copula and tends to be omitted in utterances with locative predicate, or temporary properties, and has a tendency to be overt in utterances with nominal predicate, permanent properties.

While the findings in Becker's research offer insight to copula omission, there are aspects of her study that need further analysis. One of the participants was observed at a much later age range and for a longer span of time. While this could be to adjust for

developmental rates between the children, it does not provide consistent data across ages. Additionally, the ages of the children chosen for the study were inconsistent with one another. For example, the smallest age range between the first and last transcript was two months and the largest was nine months. Three of the four children's first transcript was chosen at 2;0 however one child was not observed until 2;7. Furthermore, the Mean Length Utterance (MLU) of the children ranges from 2.98 to 3.38. To allow for a conclusive study, it is important to study the transcripts of similar children based on age and MLU. However, since Becker only used four children in this study, it is possible that her results are accidental due to her small "n" size. The purpose of the following study is to retest Becker's hypothesis with a more explicit notion of probability; there is a chance that her results are not reliable due to her small sample size. This experiment studies a longitudinal sample of twelve children developing typically in the third year of life.

Method

Participants

Data for the current study were obtained from an existing longitudinal database (Rispoli & Hadley, 2008). Selection of participants began with a review of the average MLU of the four children studied by Becker (2004). The average MLU reported by Becker was 3.07. When this average was compared with the average MLU for ages reported in Chapman and Miller (1981), it was determined that our language samples should come from children 30-33 months of age. Twelve children were then randomly selected from the larger archive. Selection was random, based on numbers generated by the randomization feature of the Vassarstat website.

Procedure

The transcripts from the 60-minute language samples at age 30 and 33 months were coded for copula and copula omission by the author and other trained research assistants. They were coded for locative and nominal copula following the procedures specified in Becker (2004). In addition to coding specifications from Becker (2004), the author contacted Becker with additional questions concerning several sentence types that are not discussed in Becker (2004). From Becker (personal communication, April 25, 2012) the author learned that questions and existential constructions were never used in the original analysis. Nominal copula sentences included statements of intrinsic or inherent quality, color, size, category, form of statement, or a referential object. In contrast, locative copula sentences included food items (i.e. "this is for you"), furniture, temperature, body feeling, activity, or accomplishment. Then the number of productions and omissions were calculated for both contexts.

Two transcripts were chosen at random from the 24 coded by the author. Two trained research assistants were instructed in the coding procedures for locative and nominal copula. They were then given print outs of the copula sentences from the samples, and asked to independently code them as nominal or locative. The independent coders replicated the original coding with 100 percent agreement to the author's original coding.

Results

Recall that the purpose of this project was to replicate Becker's (2004) study with more participants. Our study tripled the number of participants, from 4 to 12. Transcripts were analyzed from the participants at 30 and 33 months of age for the number of nominal and locative statements produced and the rate of omission.

The average omission rate for nominal statements is 26.431%, and the average omission rate for locative statements is 28.691%. Table 1 shows the omission rates, averages, and the standard deviations for the twelve participants in both conditions. It shows that the standard deviation for nominal statements is 27.751%, and the standard deviation for locative statements is 23.680%. Six participants omit more in locative, one ties for omission rate in the two conditions, and three participants omit more in nominal.

To compute an accurate and stable omission rate, rates were calculated on a minimum of 10 sentences. Because of this, we eliminated two children from the study, making our final sample size 10. To test Becker's hypothesis that omission rates would be lower for nominal sentences than for locative predicates a correlated sample T-test was performed on predicate types across the two contexts. The t-test failed to find a difference t = -.53, p = > .05. Thus the null hypothesis, that there was no difference between the conditions was accepted.

Discussion

This study explored one of the current explanations for copula omission in children. According to Becker (2007), children are more likely to omit copula when the predicate a locative or temporary state than when the predicate is in a nominal or permanent state. However, Becker's support for her theory rested on only four children. This study increased the sample three-fold, to twelve children.

Figure 1 shows the distribution of nominal and locative omission with the standard deviation. The figure shows that the standard deviation of locative omission and the standard deviation of nominal omission are within 5.00% of one another. This similarity in

standard deviations between the two conditions shows that the type of omission, locative or copula, does not greatly affect the omission rate of the child.

The average omission rate for nominal statements in this study was 26.431% and the average omission rate for locative statements was 28.691%. Two children had too few locative predicates to calculate an omission rate. A correlated sample t-test with ten children revealed no significant difference in omission rate across these two conditions. Becker's hypothesis was not supported.

The results of this particular study encourage the researchers to look into the other explanations for copula omission. One idea, the Metrical Template explanation, provided by Gerken and Landau (1990), states that stress plays a significant role in the child's omission. According to the results of this study, the syllables with strong stress are preserved and the syllables with weak stress were omitted.

Another explanation for omission is Wexler's (2011) theory of the Optional Infinitive stage of development. This stage can be characterized by auxiliary and copula omission, both of which are caused by some kind of genetically specific development. During the OI stage, children understand that *be* is required whenever tense exists and therefore do not omit in this context. However, more often than not the tense is omitted in the OI stage resulting in copula omission (Wexler, 2011). Wexler claims that these errors are due to some kind of genetically specific development, and there is little other evidence to explain the omission phenomenon. The author recognizes the value of looking further into these other explanations for omission.

Clinical Relevance

The present study is clinically relevant to the grammatical profile of Specific Language Impairment (SLI). Copula omission is one hallmark characteristic of the grammar of children with SLI when they are entering school. Having a better understanding of the features of omission in children allows the clinician to better assess his or her clients. Copula omission should be seen as the symptom of a deeper problem in the learning and acquisition of grammar. While there are variations in the characteristics of children with SLI, forming a foundation of knowledge for omission rates can offer greater insight into the grammatical aspect of this disorder. Clinicians should familiarize themselves with the other explanations for omission to fully understand what implications this feature in development can have on the overall acquisition of grammar for children.

In summary, children acquiring English as a native language take time to acquire copula. During this period, omission of copula is typical. At present, we do not fully understand why children omit this important grammatical marker. This study adds a small piece of evidence that omission is not conditioned by the semantics of the predicate. Explanation for the errors points in the direction of either the Metrical Template explanation (Gerken and Landau, 199) or the Optional Infinitive Stage (Wexler, 2011). More research is clearly needed to understand this phenomenon.

Acknowledgements

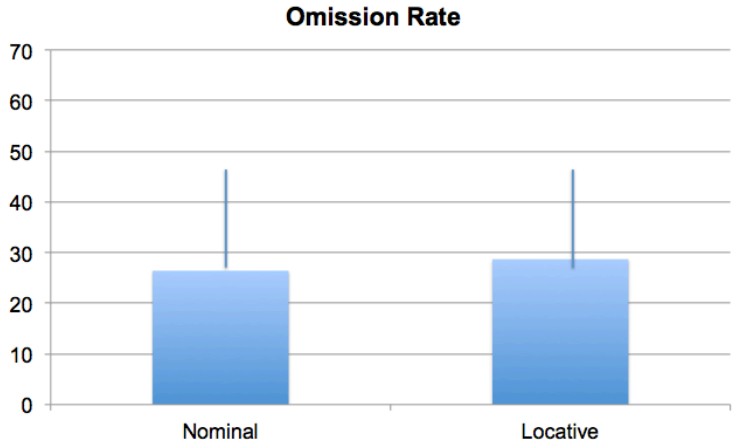
The author wishes to thank Dr. Matthew Rispoli and Dr. Pamela Hadley for their contributions. The author would also like to acknowledge the support of the members of the Applied Psycholinguistics Laboratory, and the support NSF, NSF BCS-082251.

Table 1*:* Rate of Omission in Nominal and Locative Copula of each participant

	N [C]	N [C:O]	Total	Rate of Omission	L [C]	L [C:O]	Total	Rate of Omission
18B	17	4	21	19.048	29	7	36	19.444
21B	2	12	14	85.714	4	6	10	60
26B	7	14	21	66.667	10	28	38	73.684
28G	10	2	12	16.667	16	6	22	27.272
33B	23	7	30	23.333	24	7	31	22.581
35G	19	0	19	0	21	1	22	4.545
39B	20	4	24	16.667	21	16	37	43.243
42B	15	0	15	0	12	0	12	0
47B	18	2	20	20	22	3	25	12
57B	31	6	37	16.216	66	21	87	24.138
Average				26.4312				28.6907
Stan. Dev				27.7511				23.6802

N = nominal, L = locative, [C] = copula produced, [C:O] = copula omitted

Figure 1: Distribution of Nominal and Locative Omission and Standard Deviation



References

Becker, M. (2004). Copula omission is a grammatical reflex. Language Acquisition, *12*(2), 157-167.

De Villiers, J., & de Villiers, P. (1972). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research*, *2*, 267-278.

Gerken, L., Landau, B., Remez, R. (1990). Function morphemes in young children's speech perception and production. *Developmental Psychology*, *26*(2), 204-216.

Hadley, P., & Short, H. (2005). The onset of tense marking in children at risk for specific language impairment. *Journal of Speech, Language, and Hearing Research*, *48*, 1344-1362.

Pine, J., Conti-Ramsden, G., Joseph, K., Lieven, E., & Serratrice, L. (2008). Tense over time: testing the Agreement/Tense Omission Model as an account of the pattern of tense-marking provision in early child English. *Journal of Child Language*, *35*, 55-75.

Wexler, K. (2011). Grammatical computation in the Optional Infinitive Stage. In J. de Villiers & T. Roeper (Eds.) *Handbook of Generative Approaches to Language Acquisition*. (pp. 53-118). New York, New York: Springer.

Wilson, S. (2003). Lexically specific constructions in the acquisition of inflection in English. *Journal of Child Language*, ***30***, 1–41.