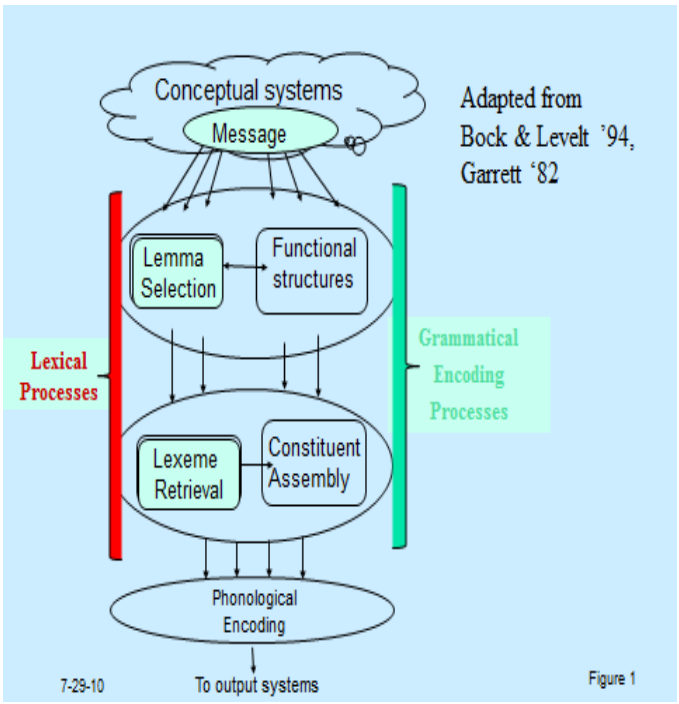
FLUENCY MARKERS FOR CHILDREN'S SENTENCE PLANNING: EARLY AND LATE STAGE PROCESSING

Dana McDaniel, University of Southern Maine
Cecile McKee and Merrill F. Garrett, University of Arizona

Language production

We compare here two studies of sentence planning in children and adults. Our index of planning was patterns of non-fluency in lexically and syntactically comparable utterances from both populations. We interpret these patterns in terms of current language production models.

Figure 1 illustrates the production architecture that we assume. The system represents mapping operations from a non-linguistic level of conceptual control (that gives rise to the intended Message) to a succession of *lexical* and *grammatical processes*. Two essential features of the figure relevant to our current work are:



- A 'horizontal' factor: Multi-level planning in concurrent processing stages that encompass different ranges of the elements occurring in a surface utterance. Early stage processing is multi-phrasal; late stage processing is local.

- A 'vertical' factor: Two processing streams, one lexical and one grammatical. These must be integrated in the course of utterance preparation.

We present analyses of several measures. These include overall production difficulty, the distribution of filled and unfilled pauses, cases where speakers stopped and restarted their utterances, and the use of complementizers in lower clauses. Our experiments used the same relative clause target sentences, but the first one elicited sentences that the speakers planned themselves (*No-Model*), and the second one used imitation (*With-Model*). A critical difference between these experiments is in the role of Message-level planning. With no model utterance to imitate, participants had to engage the full sentence-planning process. But with a model utterance to imitate, participants had to replan the target sentence but not to rebuild the message it conveyed.

Participants:

All participants were monolingual, native speakers of English with no known speech, language, or hearing disorders. The children attended schools and preschools in Portland, Maine, and

Tucson, Arizona; the adults were students in introductory-level linguistics courses at the University of Southern Maine and the University of Arizona. The children ranged in age from 3 to 8 years old. For the analysis, we divided the child participants into two age groups: 3- to 5-year-olds, the Young group; and 6- to 8-year-olds, the Older group. Table 1 shows the number of participants in each age group for each experiment.

Table 1: Participants in each experiment

	No-Model	With-Model
Young	23	35
Older	24	36
Adults	30	36

Production Tasks:

In both experiments, participants observed stories enacted with toys. In the No-Model experiment, they directed a blindfolded experimenter to pick up one of two identical toys. In the With-Model experiment, they repeated a puppet's request for one of the toys. The basic elements of the task are listed below.

- Scenario includes *two identical toys*, plus another different toy; one of the identical figures is identified by an event acted out by a storyteller (example below).
- **Storyteller**: This is a story about two fish, and they look exactly the same. This man is also in this story [puts man next to fish]. In this story, the man is going to lick one of the fish. [man says:] "I feel like licking a fish today ... hmm ... not this one ... I think I'll lick this fish!" lick, lick, lick
    - **No-Model**: [storyteller to partner experimenter] Now cover your eyes. [to participant] I'm going to point to one of the fish, and you tell [partner] to pick it up. [points to fish that the man is licking] **Participant** [targeted response]: "Pick up the fish that the man is licking."
    - **With-Model:** [puppet to child, referring to storyteller] Tell her, "Pick up the fish that the man is licking." [**Participant** repeats sentence for storyteller]

Materials:

The materials for both experiments were relative clause structures. We contrasted two syntactic variables: gap position (subject vs. object) and depth of embedding (one-clause vs. two-clause relatives), as shown in Table 2.

Table 2: Materials (three tokens of each type in each experiment)

Structural type	**Example**
One-clause subject relative (**S1c**):	Pick up the baby that is pulling the hen.
One-clause object relative (**O1c**):	Pick up the fish that the man is licking.
Two-clause subject relative (**S2c**):	Pick up the girl that Fred thinks was kicking the cow.
Two-clause object relative (**O2c**):	Pick up the cat that Belle guessed the boy was patting.

Coding and Analysis:

First, we coded the non-fluency types listed in (1) in each utterance.

(1) Non-fluency types:

- UNFILLED PAUSE: a moment of silence (for adult participants, criterion duration was 200 ms; for child participants, duration was 500 ms).
- FILLED PAUSE: 'filler' elements like *um or uh.*
- RESTART: (repetition or repair) speaker stops and resumes at an earlier part of utterance.
- OTHER: Stumbles and word lengthening.

Target utterances were divided into structurally-defined sections of interest (e.g., sentence, clause, phrasal onsets/offsets), and non-fluencies were tabulated within and between the sections. The sections are shown in Table 3.

Table 3: Sections used for analyses

Section #	**1**	**2**	**3**	**4**	**5**	**6***	**7***	**8**
Structure	onset	main clause verb	onset of NP	head	rel clause boundary	upper clause	compl clause boundary	lower clause
Example	[um]	pick up	[um]	the cat	[that]	Belle guessed	[that]	the boy was patting

* Only occur in the two-clause structures (S2c, O2c).

We computed a global difficulty score (DIFF) combining non-fluencies across the full sentence. This summed the number of silent pauses, filled pauses, stumbles, and lengthenings. A high DIFF score indicates a relative lack of fluency. We divided this score by the number of sections in the sentence in order to adjust for the length differences in the materials. We expected this measure to reflect overall processing load as structural type and experimental task varied.

Results and Discussion

DIFF ratios: Hierarchical general linear factors analysis (PROC-GLM in SAS) applied to DIFF ratios showed several significant contrasts. Across age groups and in both experiments, the two-clause utterances were, unsurprisingly, more difficult than the one-clause utterances. In many cases, there was also an effect of structure, with object relatives harder than subject relatives. This latter effect was significant in the No-Model experiment, but not in the With-Model experiment. A factor affecting this was the substantial reduction in DIFF scores from the No-Model to the With-Model experiment: Overall fluency was greater for the latter. Figures 2 and 3 illustrate DIFF patterns for the two experiments. The general correspondence of such effects across age groups is an important feature of the results. We will return to this point in the context of other measures below.

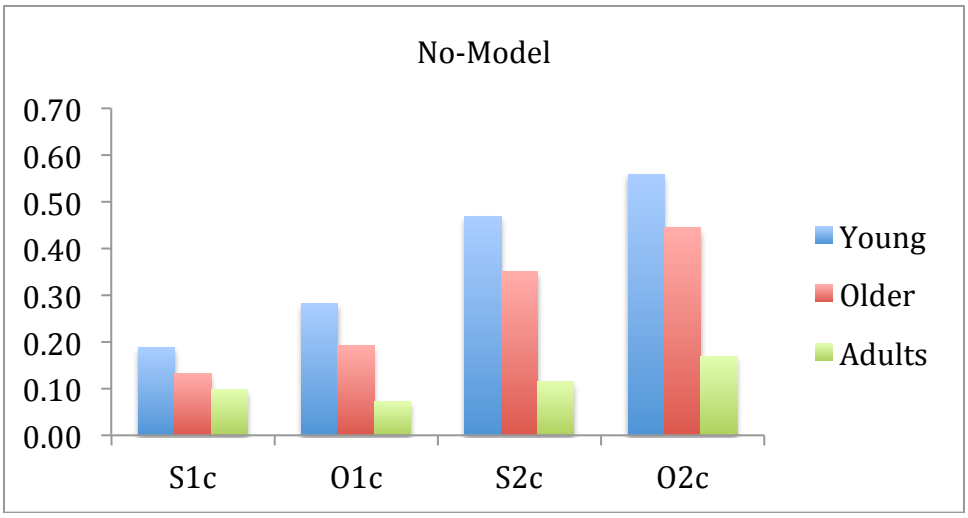


Figure 2: DIFF ratios for each sentence type (Subject/Object, 1clause/2clause) across age groups.

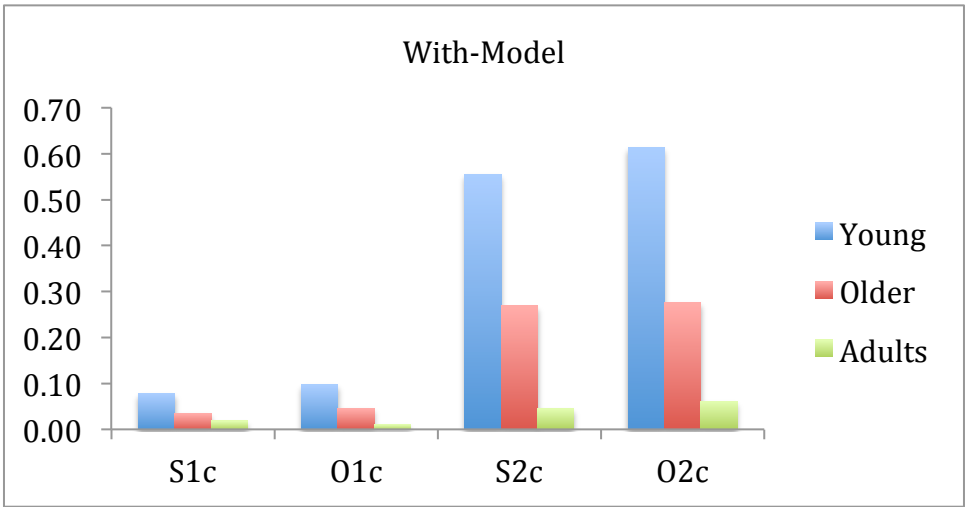


Figure 3: DIFF ratios for each sentence type (Subject/Object, 1clause/2clause) across age groups.

Unfilled pauses: We counted the number of unfilled pauses in each of the eight sections noted in Table 3 and divided by the number of utterances that included that section.

Just as the DIFF scores indicated, this measure showed greater fluency across all age groups in the With-Model experiment. The incidence of unfilled pauses was substantially reduced. Nevertheless, the loci for unfilled pauses were similar across experiments and across age groups. The latter is a key observation: Children and adults have similar planning domains on this evidence. To test the strength of that correspondence, we ranked the relative frequency of pause locations across the sentence and did so for each group. We then calculated Spearman rank correlations between age groups.

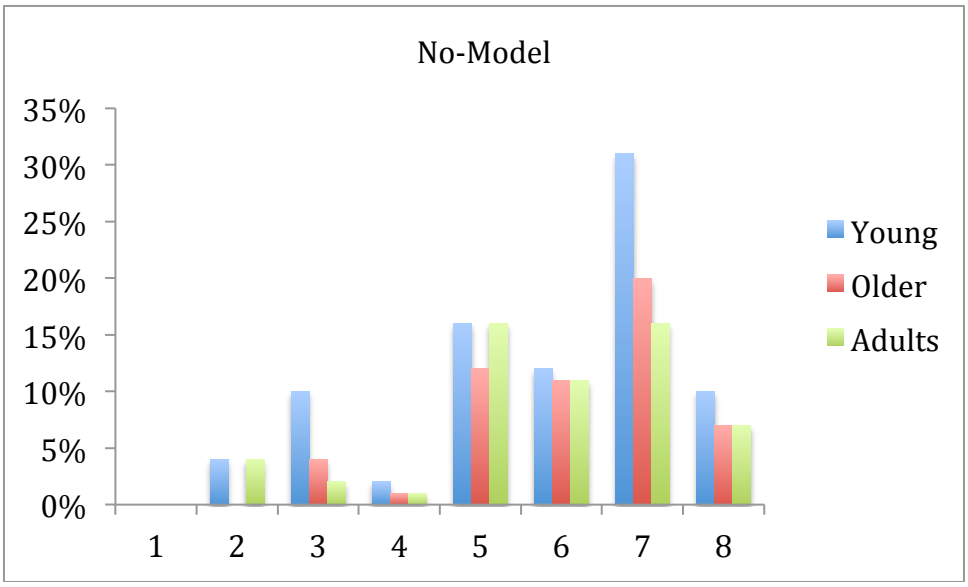


Figure 4: Percentage of utterances with unfilled pauses in each section of the sentence for each age group.

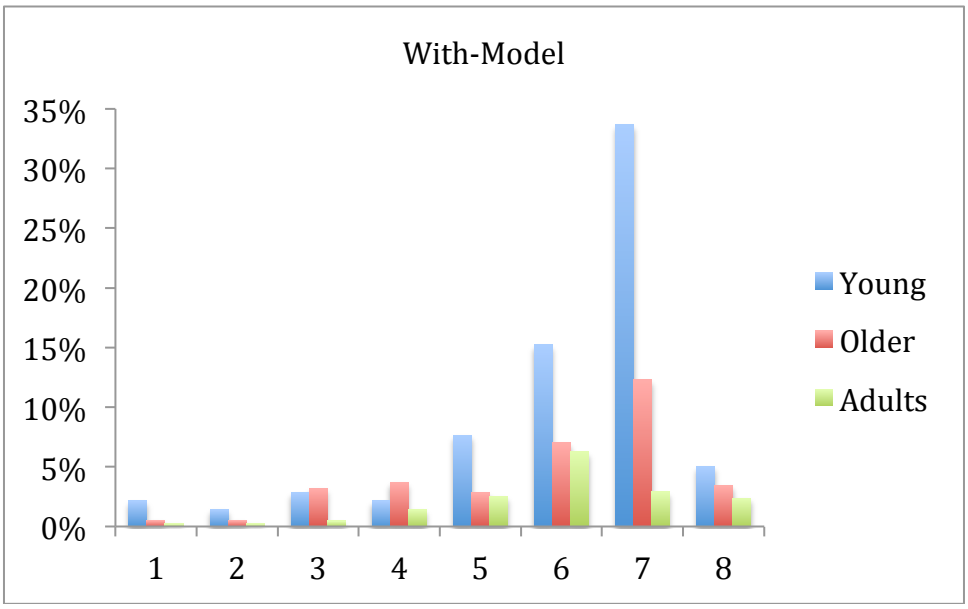


Figure 5: Percentage of utterances with unfilled pauses in each section of the sentence for each age group.

In the No-Model experiment, all Spearman rank correlations were significant: Adult/Older (r=0.88), Adult/Young (r=0.96), Older/Young (r=.93), ($p<0.05$, critical value 0.714 for n=7). In the With-Model experiment, two of the three were significant: Adult/Older (r=0.70), Adult/Young (r=0.88), and Older/Young (r=.75) ($p<0.05$, critical value 0.714 for n=7). The less robust contrasts in the With–Model experiment are probably due to the fact that the adults and older children paused at lower rates than in the No-Model experiment, and in the case of adults, substantially less, which reduced variation at the different sentence loci.

Generally, our experiments show that both children and adults prefer to pause at clause boundaries (Sections 5 and 7) and after the subject (first word) of the upper clause within the relative clause (Section 6). These structural loci for unfilled pauses have been identified in earlier work on adult language production (Beattie, 1980; Boomer, 1965; Butterworth, 1980; Ford, 1978).

Filled pauses: The rate of filled pauses was markedly lower in the With-Model experiment than in the No-Model experiment, and particularly so for adults, as shown in Table 4.

Table 4: Rate of filled pauses

	No-Model	**With-Model**
Young	64 of 242 target trials = **26%**	28 of 281 target trials = **10%**
Older	61 of 249 target trials = **24%**	18 of 382 target trials = **5%**
Adults	49 of 346 target trials = **14%**	5 of 425 target trials = **1%**

Within each experiment, the adults and children also behaved differently with respect to the location of filled pauses. In the No-Model experiment, adults produced filled pauses at a rate more comparable to children than they did in the With-Model experiment, but also with a striking bias toward sentence initial location, as Figure 6 shows. Adults do this primarily at utterance onset, while children distribute such planning across the sentence. We interpret these patterns as indicating that filled pauses are selectively responsive to Message level planning.

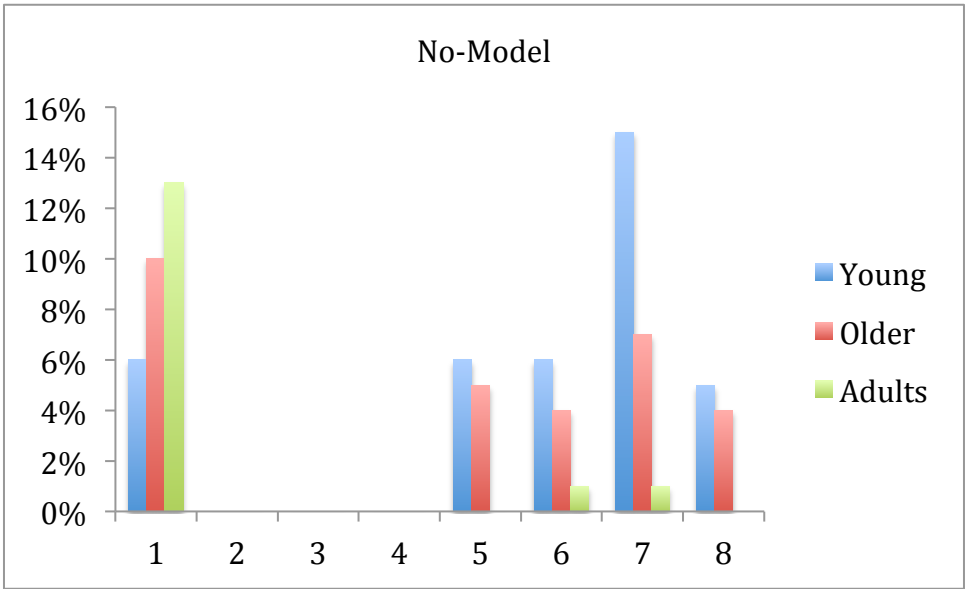


Figure 6: Percentage of utterances with filled pauses in each section of the sentence for each age group.

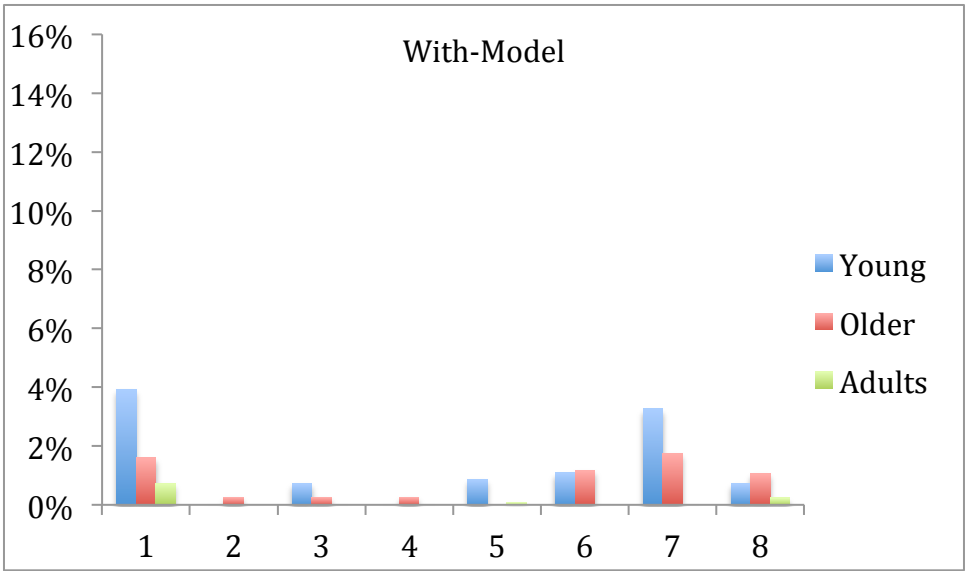


Figure 7: Percentage of utterances with filled pauses in each section of the sentence for each age group.

To further evaluate these patterns, we condensed the sections. This grouped together the filled pauses in loci prior to the relative clause (Sections 0-4), those for relativizer and upper clause (Sections 5 and 6), and those for complementizer and the lower clause (Sections 7 and 8). In the With-Model experiment, Adult and Older groups showed a significant contrast ($p<0.001$, Fisher's exact test, two-sided), as did Adult and Young groups ($p<0.001$, Fisher's exact test, two-sided). Older and Young groups did not differ ($p=0.06$, Fisher's exact test, two-sided), though there was the suggestion of greater dispersion in the Young group. In the With-Model experiment, similar comparisons did not produce similar outcomes: Groups did not differ significantly in their association with the three classes of sections ($p>.05$ for all values of Fisher's exact test, two-sided).

Tables 5 and 6 give the specific numbers of utterances with filled pauses to make the difference clear.

	Sp 0, Pickup, Sp1head	**Sp 2, 2.5, 3, Up Cl**	**Sp 4, 4.5, 5, Lo Cl**
Young	16	20	28
Older	26	18	17
Adults	45	2	2

Table 5: Number of utterances with filled pauses at three selected regions in the No-Model experiment

	Sp 0, Pickup, Sp1head	**Sp 2, 2.5, 3, Up Cl**	**Sp 4, 4.5, 5, Lo Cl**
Young	13	8	7
Older	9	2	7
Adults	3	1	1

Table 6: Number of utterances with filled pauses at three selected regions in the With-Model experiment

The general pattern that emerges is that adults use filled pauses almost exclusively at the onset of the utterance, whereas children use them throughout the utterance. Further, when there is less Message-level processing, as in imitation, adults use very few filled pauses, while the children's filled pauses tend more toward the onset of the utterance.

These differences between filled pause rates and the distribution of those pauses across the experiments supports analyses of filled pauses as indicators of Message-level planning. The difference between children and adults suggests that children generally do more mid-utterance message-level planning than adults do.

This view is reinforced by the distributional changes from No-Model to With-Model experiments. In the latter, adults' filled pauses disappear almost entirely. This suggests minimal Message-level planning. But the children's distribution for this experiment clearly shifts toward the adult pattern seen in the No-Model experiment, suggesting the With-Model task allows them to make a greater investment in Message level planning earlier in utterances.

Restarts: We also coded and counted repetitions and repairs (i.e., where a speaker stops and resumes at an earlier part of utterance). These non-fluencies were also substantially reduced in the With-Model experiment. And, as Table 7 shows, restart rates were highest for the Young and lowest for the Adult group, with the Older group intermediate, and this pattern held for both experiments.

Table 7: Restart rates across age groups

	No-Model	**With-Model**
Young	.31	.17
Older	.22	.13
Adults	.09	.03

Distributional differences among the participant groups are of most interest for the restart measure. For speakers in all age groups and in both experiments, most instances of this type of nonfluency began in the lowest clause. That is, when speakers stopped an utterance, they did so mainly in the lowest clause. Age groups differed in the domain of the restart. Adults tended to stop and restart in the same (lower) section of the utterance, whereas children often returned to earlier sections, as shown in Figure 8.

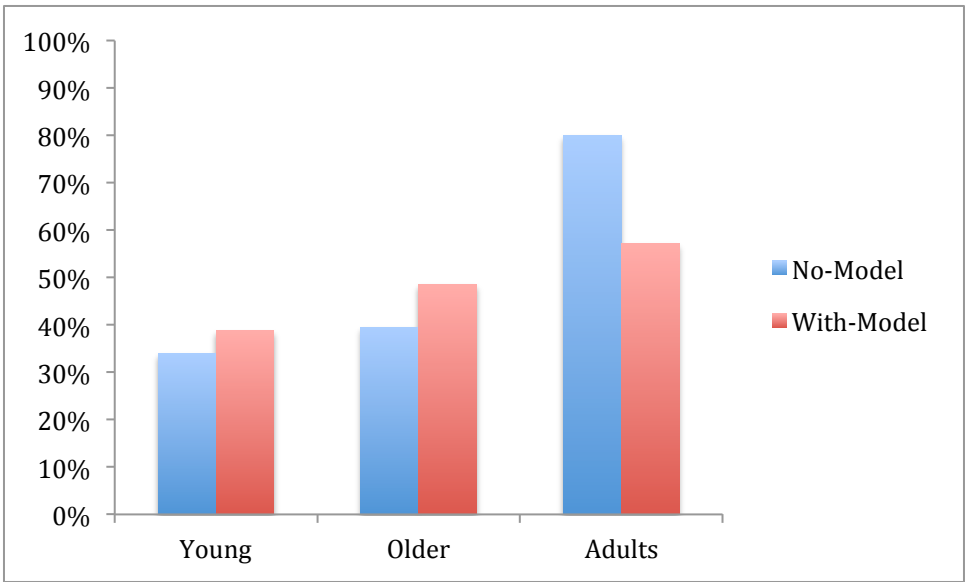


Figure 8: Percentage of restarts that remained in the lower clause out of the total number of restarts that began in the lower clause for each age group and in each experiment.

For the No-Model experiment, the lowest clause restarts stayed within the lowest clause most of the time in Adult utterances (81%). Young participants typically returned to an earlier section of the sentence: only 39% remained in the lowest clause. The Older group's restarts were intermediate: 55% remained within the lowest clause. These contrasts were significant only for the comparison of Adult and Young ($p=0.003$, Fisher's exact test, two-sided). The same pattern appears for the With-Model experiment, but somewhat less well defined. The difference is again significant only for the Adult/Young comparison ($p=0.04$, Fisher's exact test, two-sided). The very low incidence of adult restarts makes this comparison for the With-Model experiment less robust.

These findings again suggest a difference across the age groups in degree of mid-utterance message-level planning. For adults, the process of stopping to repair an earlier part of the utterance tends to be local, whereas children restart from further back (often restarting the whole sentence). A plausible account is that adults have the controlling representation, hold it and can refer to it to clean up the local unit; children don't have that record and its linkages to the sentence and have to regenerate from the Message level.

Use of complementizers in the lowest clause: The two-clause relative clause items included a complementizer position in the lowest clause. Grammatically, an overt complementizer is possible in the object-gap structure, but not in the subject-gap structure. In the With-Model experiment, the models never included an overt complementizer. The sentences in (2) illustrate the four logical possibilities:

(2a) S2c without overt complementizer: Pick up the girl that Fred thinks was kicking the cow.
(2b) S2c with overt complementizer: Pick up the girl that Fred thinks that was kicking the cow.
(2c) O2c without overt complementizer: Pick up the cat that Belle guessed the boy was petting.
(2d) O2c with overt complementizer: Pick up the cat that Belle guessed that the boy was petting.

Adults in both experiments avoided *that* in both subject and object gap structures. Children in both experiments used it approximately 1/3 of the time in both S2c and O2c. In the case of the With-Model experiment, this means that children inserted *that*, since it was not in the model. Figures 9 and 10 illustrate these patterns.

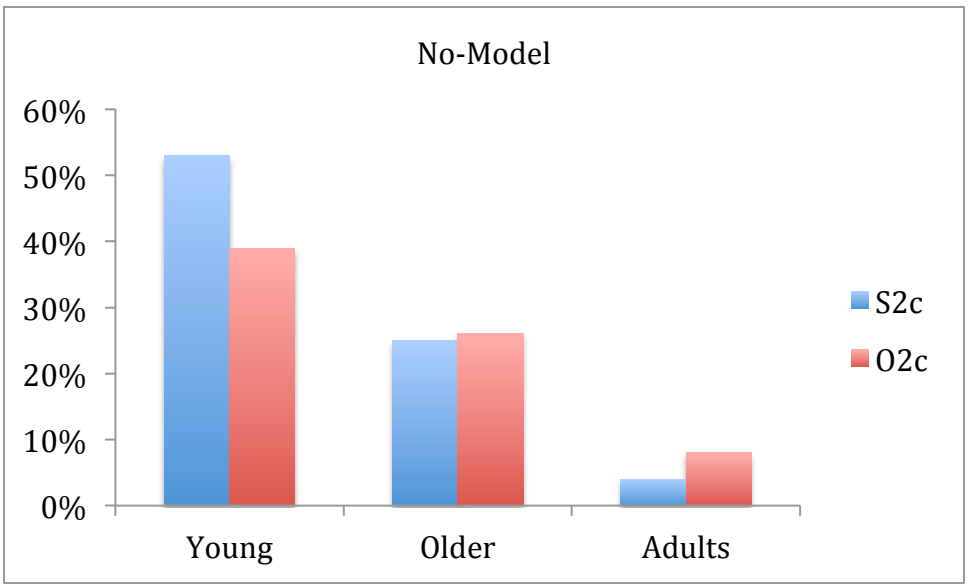


Figure 9: Percentage of Subject and Object two-clause relative clause utterances (S2c/O2c) that included the complementizer *that* in the lowest clause for each age group.

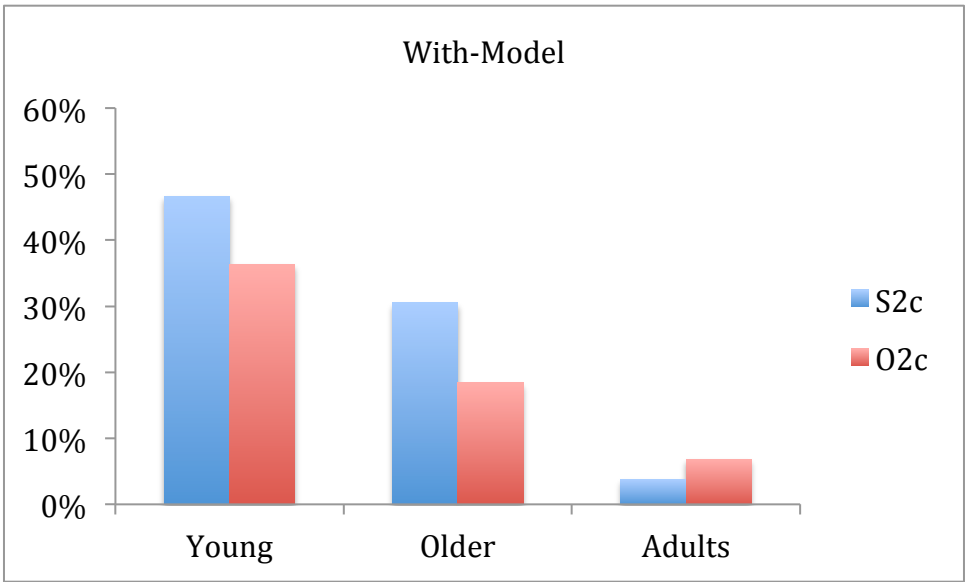


Figure 10: Percentage of Subject and Object two-clause relative clause utterances (S2c/O2c) that included the complementizer *that* in the lowest clause for each age group.

We suggest that children's frequent use of *that* may also be attributable to Message-level planning. Using a recall/imitation task with adults, Ferreira and Dell (2000) found a preference for the overt element in non-movement structures (e.g., *The boy thinks (that) the dog chased the cat*). Specifically, the complementizer was uttered more when the embedded material was not

easily accessible. The filler-gap structure in our experiments would encourage advance planning of the embedded clause, which would explain why the adults generally did not use an overt complementizer (including in the object case, where it is grammatical). The children, on the other hand, are not as able to plan the inside of an upcoming clause in advance and therefore often use the complementizer.

Conclusion

The comparison of fluency patterns in the No-Model and With-Model experiments across age groups sheds light on the development of the sentence planning process. Specifically, our findings indicate similar types of sentence planning for children and adults, but different degrees of advance planning. Adults tend to plan a whole message before beginning an utterance, and do little Message-level planning for imitation. Young children, on the other hand, often do Message-level planning mid-utterance and are not as able to take advantage of a modeled utterance.

References

Beattie, G. (1980). The role of language production processes in the organization of behaviour in face to face interaction. In B. Butterworth (ed.), *Language production: vol. 1, speech and talk*. London: Academic.

Bock K., Levelt W. J.M. (1994) Language production: Grammatical encoding. In: Gernsbacher M.A. (ed.) *Handbook of Psycholinguistics*. Academic Press, London.

Boomer, D. (1965). Hesitation and grammatical encoding. *Language and Speech **8***, 145-158.

Butterworth, B. (1980). Some constraints on models of language production. In B. Butterworth (ed.), *Language production: vol. 1, speech and talk*. London: Academic.

Ferreira, V. & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology **40***, 296-340.

Ford, M. (1978). Planning units and syntax in sentence production. Doctoral dissertation, University of Melbourne.

Garrett, M.F. (1982) Production of speech: Observations from normal and pathological language use. In: A. Ellis (ed) *Normality and Pathology in Cognitive Functions*. London: Academic Press.

McDaniel, D., McKee, C., and Garrett, M. F. (2010) Children's sentence planning: Syntactic correlates of fluency variations. *Journal of Child Language*, *37*, 59-94.

Rispoli, M. (2003). Changes in the nature of sentence production during the period of grammatical development. *J. of Speech, Language, and Hearing Research 46*, 818-830.

Wijnen, F. (1990). The development of sentence planning. *J. of Child Language 17*, 651-676.

Acknowledgments

This research was supported by the National Science Foundation (USM: BCS-0822558, U of Arizona: BCS-0822457). We are also grateful to:

- Wayne Cowart, for comments and suggestions.
- Student Research Assistants:
    - USM: Michele Bernier, Dawn Dodge, Andrew Fitzgibbons, Danielle Gonzalez, Erin Engelke, Cheri Lozoraitis, Abigail Mann, Joshua Plourde, Travis Roberts, Lena Taylor, K. Heather Thompson, Amy Williams
    - U of A: Julie Fratatoni, Kelly Gallardo, Alvaro Garza, Kara Hawthorne, Jawad Hussain, Kelsey Janet, Danielle Miller, Katrina Nicholas, Marie Roy, Michelle Sandoval
- Children, teachers, directors, and parents of schools and preschools in Portland, ME and Tucson, AZ.