ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Decontextualized Utterances Contain More Typical and Stuttering-Like Disfluencies in Preschoolers Who Do and Do Not Stutter

Kathleen E. Oppenheimer,[a] Jessica Lee,[a] Yi Ting Huang,[a,b] and Nan Bernstein Ratner[a,b]

[a] Department of Hearing and Speech Sciences, University of Maryland at College Park [b] Program in Neuroscience and Cognitive Sciences, University of Maryland at College Park

ARTICLE INFO

Article History:
Received March 14, 2023
Revision received April 25, 2023
Accepted May 3, 2023

Editor-in-Chief: Cara E. Stepp
Editor: Raymond D. Kent

https://doi.org/10.1044/2023_JSLHR-23-00173

ABSTRACT

Purpose: Stuttering-like disfluencies (SLDs) and typical disfluencies (TDs) are both more likely to occur as utterance length increases. However, longer and shorter utterances differ by more than the number of morphemes: They may also serve different communicative functions or describe different ideas. Decontextualized language, or language that describes events and concepts outside of the "here and now," is associated with longer utterances. Prior work has shown that language samples taken in decontextualized contexts contain more disfluencies, but averaging across an entire language sample creates a confound between utterance length and decontextualization as contributors to stuttering. We coded individual utterances from naturalistic play samples to test the hypothesis that decontextualized language leads to increased disfluencies above and beyond the effects of utterance length. AQ1 AQ2

Method: We used archival transcripts of language samples from 15 preschool children who stutter (CWS) and 15 age- and sex-matched children who do not stutter (CWNS). Utterances were coded as either contextualized or decontextualized, and we used mixed-effects logistic regression to investigate the impact of utterance length and decontextualization on SLDs and TDs.

Results: CWS were more likely to stutter when producing decontextualized utterances, even when controlling for utterance length. An interaction between decontextualization and utterance length indicated that the effect of decontextualization was greatest for shorter utterances. TDs increased in decontextualized utterances when controlling for utterance length for both CWS and CWNS. The effect of decontextualization on TDs did not differ statistically between the two groups.

Conclusions: The increased working memory demands associated with decontextualized language contribute to increased language planning effort. This leads to increased TD in CWS and CWNS. Under a multifactorial dynamic model of stuttering, the increased language demands may also contribute to increased stuttering in CWS due to instabilities in their speech motor systems.

The Multifactorial Dynamic Pathways Theory suggests that the speech-motor system of individuals who stutter is susceptible to destabilization as linguistic demands increase, explaining the temporal overlap of the age of onset of stuttering and the rapid increase in utterance length in 2- and 3-year-olds (Smith & Weber, 2017). Consistent with this, children who do and do not stutter both display decreased motor stability when producing longer sentences compared to shorter ones, but the decrease in stability is greater for children who stutter (CWS; Usler & Walsh, 2018). Therapies for young CWS often involve reducing demands on children to produce complex language, which is thought to improve fluency by

Correspondence to Kathleen E. Oppenheimer: koppen@terpmail.umd.edu. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

placing less demand on their unstable speech motor systems. The resulting repeated practice producing fluent speech may train neural pathways to produce fluent speech (Smith & Weber, 2017). Additionally, since some CWS may have frank or subtle weaknesses in language proficiency that interact with a less-stable motor system to produce stuttering, thorough assessment of language would be necessary to identify and provide appropriate support for language development.

Children who do and do not stutter also produce typical disfluencies (TDs) that begin around the same time as stuttering-like disfluencies (SLDs; Ambrose & Yairi, 1999). Like SLDs, TDs occur more frequently in longer and more complex utterances than shorter or simple utterances (Buhr & Zebrowski, 2009; MacPherson & Smith, 2013; Melnick & Conture, 2000; Rispoli, 2003; Rispoli et al., 2008; Wagovich et al., 2009; Zackheim & Conture, 2003). Revisions are a category of TD defined as changes to at least one previously produced morpheme (Rispoli et al., 2008) and indicate that the speaker recognized and repaired a mismatch between the utterance that was produced and the intended message (Wagovich et al., 2009). Stalls are anticipatory interruptions that represent a glitch in utterance planning and include silent pauses, filled pauses, and repetitions (Rispoli, 2018; Rispoli et al., 2008; Wagovich et al., 2009).

Increased utterance length is also a characteristic of *decontextualized language*: extended, abstract discourse that is removed from the physical context of an interaction (Davidson et al., 1986; Rowe, 2012, 2013; Snow, 1983; Uccelli et al., 2019). Decontextualized utterances are longer, more syntactically complex, and contain more diverse vocabulary than contextualized utterances. Compared to contextualized utterances, they place higher demands on working memory, as the speaker recalls rarer words, produces more complex sentence structures, and plans utterances without the contextual support of the "here and now." Rowe (2013) describes three types of decontextualized language. Briefly, *causal explanations* make or request logical connections between events, objects, or concepts. *Pretend play* attributes actions, feelings, or identities to objects or involves acting out routines symbolically. Finally, *narratives* discuss past, future, or fictional events. Academic language is often decontextualized, and higher levels of decontextualized language use and exposure in early childhood are associated with higher vocabulary, narrative skill, and academic achievement later in development (Demir et al., 2015; Uccelli et al., 2019). Examples of these three types of decontextualized language are shown in Table 1. T1

The increased utterance length associated with decontextualized language would suggest that SLDs and TDs would increase in decontextualized compared to contextualized language. Two studies of English-speaking children provide insight into the association between decontextualized language and fluency, but different tasks used across studies make it difficult to compare findings. Masterson and Kamhi (1991) obtained language samples from children who do not stutter (CWNS) with typical development, reading disabilities, and language disorders and compared the presence of TDs across three sampling genres and two contextual support conditions in a fully crossed design. Genre did not impact fluency, contrary to our predictions that the explanation and narration conditions, which Rowe (2013) would classify as decontextualized, would contain more disfluencies than the description condition, which we considered contextualized. Contextual support, operationalized as the presence or absence of objects or pictured scenes to provide contextual support, impacted fluency in the opposite direction of our predictions: Children were less fluent when the referents were present. Trautman et al. (2001) obtained narrative and expository language samples from children in decontextualized and contextualized tasks, operationalized as the absence or presence of picture support. They found that children who do and do not stutter produced more TDs in the decontextualized condition, and CWS produced more SLDs in the decontextualized condition. Additionally, narrative language sampling conditions elicited more disfluencies than the expository task, even though both language sampling conditions might be considered decontextualized under the Rowe (2013) definition.

Defining decontextualized language as a characteristic of a language sampling task creates a confound with

AQ3 **Table 1.** Examples and definitions of decontextualized language.

Decontextualized language type	Definition	Example from Ratner–McWhinney corpus
Causal explanation	Making and requesting logical connections between objects, events, or concepts	"I pushed the button *so* now there's more."
Pretend play	Attributing actions, feelings, or identities to objects, and enacting scripts or routines	[about a baby doll] "He's crying so I need to pick him up. That will make him feel better."
Narrative	Discussing past, future, or fictional events	"And one time he stepped over my train track that xxx he stepped over it and said thank you."

Note. Definitions are from Rowe (2013).

utterance length. Masterson and Kamhi (1991) did not report on utterance length, but found that children used more compound or complex sentences in narratives compared to explanations and when speaking to a naïve listener compared to one who was already familiar with the material. Trautman et al. (2001) did not conduct syntactic analyses but acknowledged that their narrative condition may contain longer utterances than their expository condition based on prior research. The increased abstraction, length, and syntactic complexity associated with decontextualized utterances involve higher demands on mental representation that may tax working memory. Increased demands on working memory have been associated with increased disfluency in adults who do and do not stutter (Eichorn et al., 2015). Thus, decontextualized utterances may be more likely to contain disfluencies than contextualized utterances of the same length, given the increased demands on working memory. Coding and analyzing decontextualized language at the level of an utterance, rather than at the level of a language sample, can disentangle the relationship between utterance length, decontextualization, and fluency. One recent small study of seven Japanese-speaking preschoolers who stutter did just that and found that children stuttered more when discussing nonpresent objects compared to present objects, even when controlling for utterance length (Hara et al., 2022).

Our focus was to examine the role of decontextualized[1] language, following the definition provided by Rowe (2013), as a predictor of disfluency in children who do and do not stutter. This definition of decontextualized language is more expansive than distinctions based on the entirety of language sample contexts, such as the presence or absence of picture support (e.g., such as used by Trautman et al., 2001, and Masterson & Kamhi 1991). In contrast, Rowe's definition includes pretend play and causal explanations involving pictured or present objects.

FN1

We hypothesized that the increased representational demands associated with decontextualized language would result in increased stuttering above and beyond the increase associated with utterance length for CWS. These increased representational demands would also impact CWNS, but would be realized as increased TDs rather than stuttering. As an example, it is expected that a child who stutters who produces an eight-morpheme utterance and a five-morpheme utterance will be more likely to stutter on the eight-morpheme utterance. We hypothesize that if that child produces an eight-morpheme contextualized utterance (e.g., "Look at my new doll with pink hair") and an eight-morpheme decontextualized utterance (e.g., "Mom bought her because it was my birthday"), the child would be more likely to stutter on the latter due to the additional demands on working memory associated with decontextualized language.

Research Questions

AQ4

Our specific research questions were:

1. Are decontextualized utterances more likely to contain SLDs compared to contextualized utterances in preschool CWS?
2. If decontextualization leads to increased stuttering in preschool CWS, is the increase in stuttering greater than the increase expected from increased utterance length alone?
3. Are decontextualized utterances more likely to contain TDs compared to contextualized utterances in preschool children who do and do not stutter?
4. If decontextualization leads to increased TD in preschool children who do and do not stutter, is that increase in disfluency above and beyond that associated with increased utterance length?
5. Does the impact of decontextualization on TD differ between children who do and do not stutter?

Method

Participants

We utilized an archival repository of video-taped and transcribed interactions of children who did and did not stutter, interacting with adults in play with developmentally appropriate toys. The participants whose data we analyzed for this study were 15 preschool CWS (11 boys and four girls) between the ages of 28 and 47 months (M = 37 months, SD = 6.54 months). They were matched by age and sex to 15 CWNS between the ages of 28 and 50 months (CWNS; M = 37 months, SD = 6.89 months). The CWS were observed within 1 year of stuttering onset. Stuttering diagnoses were made by a speech-language pathologist. All children spoke English as their first and only language.

[1]We would like to note that the word *decontextualized* has multiple meanings in the intervention literature. It has also been used to refer to interventions that are provided outside of contexts that are meaningful or relevant to a client (e.g., Gillam et al., 2012), and to describe intervention techniques that are reported without adequate background information—such as service delivery systems, policies, and culture—that would allow the research to be generalized to different contexts (Mallick et al., 2021). In this article, we focus solely on decontextualized language as language that is not supported by the immediate context; specifically, we use the definition of Rowe (2013) to indicate language that describes events and concepts outside of the here and now.

Data Collection

Data for this study came from the Ratner–MacWhinney corpus (Bernstein Ratner & MacWhinney, 2018; Garbarino, 2021) available at FluencyBank (http://fluency.talkbank.org). In this section, we describe the procedures used for collecting and preparing the data prior to our analyses.

The language samples were collected at the University of Maryland's Language Fluency Lab as part of a larger project to examine fluency in preschool children who were typically developing, language delayed, bilingual, or stuttering. We discuss only data from the typically developing and stuttering children in this analysis. Participants were recruited through e-mails sent via community organizations, flyers posted in the community, a research participant database, and referrals from local speech-language pathologists. Two language samples were collected during play using a standardized set of toys: one language sample with an examiner, and one with the child's parents. Data had been transcribed into CHAT (Codes for the Human Analysis of Transcripts) format (MacWhinney, 2000), used by all TalkBank repositories. CHAT includes conventions for marking six types of SLDs (i.e., prolongation, blocking, broken word, repeated segment, lengthened repeated segment, and repeated word) and seven types of TDs (i.e., phrase repetition, word repetition, phrase revision, phonological fragment, pause, pause duration, and filled pause). Further details regarding the transcription and reliability procedures for this particular data set may be found in Garbarino and Bernstein Ratner (2022). Examples of these codes are provided in Table 2. T2

The transcribed CHAT files were first analyzed in CLAN (Computerized Language ANalysis) using the MOR command to separate the transcribed utterances into words and grammatical morphemes and tag them by part of speech. The output of the MOR command is referred to as the %mor line of the transcript, as it contains a breakdown by morphemes. This breakdown, or morphological parsing, allows further automatic analyses such as calculating utterance length in words or morphemes, and analysis of the types of structures in a sentence. CLAN is more than 99% accurate at syntactic tagging of adult utterances in CHAT transcripts and is a reliable method (95% accurate) for automated analyses of child language transcripts (MacWhinney et al., 2020). An example of a morphologically parsed section of a CHAT transcript used in this study is provided in Table 3. T3

Data Coding

The second author coded each intelligible child utterance as contextualized or decontextualized following the procedures outlined in Rowe (2013). The first author checked 10% of the transcripts for reliability. Observed agreement was 86%, yielding a Cohen's kappa of .72, which is considered substantial agreement (Cohen, 1960; McHugh, 2012). While Rowe further differentiates decontextualized utterances into pretend play, narratives, and

Table 2. Fluency codes for Computerized Language ANalysis (CLAN; reproduced with permission from Bernstein Ratner & Brundage, 2022).

Stuttering-like disfluencies (SLDs)	Code	Example	Notes
Prolongation	:	s:paghetti	Placed after prolonged segment
Broken word	^	spa^ghetti	Pause within word
Blocking	≠	≠butter	A block before word onset
Repeated segment	↫	↫r-r-r↫rabbit like↫ike-ike↫	The ↫ brackets the repetition; hyphens mark iterations
Lengthened repeated segment	↫ and doubling	↫rr-rr–r↫rabbit	The doubling of "r" indicates lengthening of the "r" segment
Word repetition	[/]	dog [/] dog	Further distinctions are done inside FluCalc (a CLAN function not used in these analyses)
Typical disfluencies	**Code**	**Examples**	**Notes**
Phrase repetition	<>[/]	<that is a > [/] that is a dog	< > is used to mark repeated material
Word revision	[//]	a dog [//] beast	Revision counts once
Phrase revision	<>[//]	<what did you> [//] how can you see it?	Revision counts once
Phonological fragment	&+	& + sn dog	Changes from "snake" to "dog"
Pause	(.) or (..) or (. . .)	(.)	Counts the number of short, medium, long pauses
Pause duration	(2.4)	(2.4)	Adds up the time values, if marked
Filled pause	&-	&-um &-you_know	Fillers with underscore count as one word

Table 3. Sample of morphologically parsed CHAT transcript from our corpus.

Utterance transcription	Morphological Parse (MOR line)	Notes
nobo^dy is gonna touch that?	pro:indef|nobody aux|be&3S part|go&PRESP~inf|to v|touch pro:dem|that?	"gonna," a single morpheme, is represented by "part|go&PRESP~inf|to."
it's very hot.	pro:per|it~cop|be&3S adv|very adj|hot.	The contraction "it's" is counted as two morphemes due to the two pipes in "pro:per|it~cop|be&3S"
he needs some corn!	pro:sub|he v|need-3S qn|some n|corn!	"needs," which contains two morphemes, is represented by "v|need-3S." The vertical pipe and the hyphen each indicate a morpheme to our algorithm.

Note. Morphemes are counted by adding the number of vertical pipes "|," which roughly correspond to words, and adding the number of hyphens, which indicate inflectional morphology. The "~inf" tag indicates part of a concatenative and is subtracted to avoid overcounting the number of morphemes in concatenated forms. CHAT = Codes for the Human Analysis of Transcripts.

causal explanations, our analyses do not make finer distinctions among types of decontextualized utterances. We chose the Rowe (2013) definition of decontextualized language for several reasons. Differing definitions of decontextualized language across studies makes it difficult to compare outcomes, and by using a definition not used in the fluency research we cite, we are adding to the problem. However, the Rowe (2013) definition of decontextualized language offers flexibility to apply to utterances from a variety of language sampling contexts. Decontextualized language use and exposure based on the Rowe (2013) definition has been linked to later language outcomes (e.g., Demir et al. 2015; Uccelli et al., 2019), so using that definition allows us to interpret our findings in light of implications for later language development. Additionally, since this research involves secondary data analysis, the Rowe (2013) definition allowed us to code utterances as decontextualized or contextualized within a single language sample that had not been collected specifically for that purpose.

To conduct the remaining analyses, CLAN output including the morphological parses (%mor lines) was imported into the R programming environment (R Core Team, 2022) using the packages dplyr (Wickham et al., 2022) to manipulate the data and tidyverse (Wickham et al., 2019) to visualize it. While calculations such as mean length of utterance and proportion of disfluent utterances can be done in CLAN, doing the calculations in R allowed us to calculate the relevant metrics for each utterance individually, rather than calculating an average over all the transcripts. All R code is available at https://osf.io/q95km/?view_only=e6c70b0f9dd04510b0f407345b9782d4.

Utterance length was measured in morphemes to approximate the linguistic complexity of the utterances in a way that is appropriate for the age of the participants in our sample, and because MLU-morphemes (MLU-m) is a commonly used measure of utterance length for preschool language sample analysis and the measure used by the KIDEVAL language sample analysis feature of CLAN (Brundage & Ratner, 2015). The %mor line, or the morphological parse, contains symbols to denote parts of speech, root words of inflected forms, and bound morphemes. These symbols can be used to calculate the number of morphemes in each utterance. To do this, we first counted the vertical pipes ("|"). The vertical pipe appears in each word in the morphological parse from CLAN, separating the word from a label for its part of speech (refer to Table 3 for an example). For example, the word "birds" in an utterance would appear as "n|bird-PL," with the vertical pipe separating the part of speech label, *n*, from the spoken word, *bird*. Hyphens ("-") indicate inflectional morphology, so we added the number of hyphens to the number of vertical pipes. To avoid overcounting cliticized productions of "to" as in *hafta, wanna, gonna*, we then subtracted the number of "~inf" tags, which indicated such clitics. AQ5

SLDs and TDs were counted for each utterance by counting the number of symbols denoting each type of disfluency. Accuracy of utterance length, verbs per utterance, and disfluency counts depended on the accuracy of the original CHAT transcripts, which is believed to be high based on the approaches the creators of the data set used to ensure data integrity described in Garbarino and Bernstein Ratner (2022).

Available Data Summary

Transcripts of CWS contained a total of 4,608 utterances before exclusion criteria were applied. Eliminating utterances with any unintelligible segment yielded 4,150 utterances. Finally, we eliminated utterances that did not contain at least one verb, because utterances that are too short could not be coded for contextualization. This left a total of 2,614 child utterances to be analyzed. For the final analysis, each child contributed between 33 and 343 utterances (M = 174 utterances per child, SD = 89).

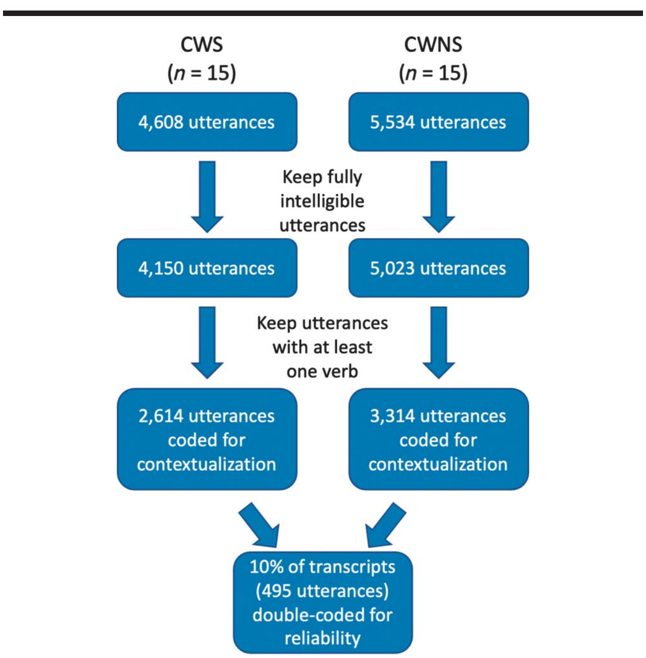
Transcripts of CWNS contained a total of 5,534 utterances before exclusion criteria were applied. Eliminating utterances with an unintelligible segment yielded 5,023 total utterances. After eliminating utterances that did not contain at least one verb, we analyzed a total of 3,315 utterances. Each child contributed between 110 and 375 utterances F1 (*M* = 221 utterances per child, *SD* = 72). Figure 1 depicts the coding process for CWS and CWNS.

Statistical Approach

All statistical analyses and data visualizations were conducted in the R programming environment (R Core Team, 2022) with the package lme4 (Bates et al., 2015) to run mixed-effects models. To confirm expected findings that longer utterances are more likely to contain SLDs in CWS, we predicted the occurrence of an SLD from the utterance length in mixed-effects logistic regression with random intercepts of participants. To confirm expected findings that longer utterances are more likely to contain TDs in CWNS, we predicted the occurrence of a TD from the utterance length in mixed-effects logistic regression with random intercepts of participants. In both models, we assessed significance by examining the coefficient of the fixed predictor with an alpha level of .05.

To answer our research questions, we followed similar statistical procedures. The outcome variable of interest (presence of TD or SLD in an utterance) was regressed on the relevant predictor variables in mixed-effects logistic regression with random intercepts of participants. Mixed-effects regression is appropriate when data are not independent, such as when we are analyzing multiple utterances individually that were produced by the same child in the same language sample. Logistic regression is used when the outcome is binary, such as containing a disfluency or not.

Figure 1. Flowcharts depicting the utterance coding process. CWS = children who stutter; CWNS = children who do not stutter.



When contextualization was a predictor, it was dummy-coded with values of 0 representing a contextualized utterance values of 1 representing a decontextualized utterance, so that positive slopes indicated more frequent disfluencies in decontextualized utterances compared to contextualized ones. Utterance length as a predictor was grand mean centered. We use the notation $\beta_{\text{morphemes}}$ to indicate the coefficient of the effect of utterance length, $\beta_{\text{decontextualized}}$ to indicate the coefficient of the effect of decontextualized utterances compared to contextualized ones, and $\beta_{\text{decon*morph}}$ to indicate the coefficient of their interaction.

We assessed statistical significance of predictors by examining the beta coefficients (β) with an alpha level of .05. We assessed model fit with a likelihood ratio test using the anova() function in R. Likelihood ratio tests compare a simpler model to a more complex model (i.e., one containing more predictors) by testing the null hypothesis that both models describe the data equally well. A significant result suggests that the more complex model is justified, as the added predictor(s) result in a model that better describes the data. A nonsignificant result of a likelihood ratio tests provides no support for adding the additional predictors in the complex model, so the simpler model is usually considered the better one for the sake of parsimony. The likelihood ratio test is only appropriate for comparing nested models, so we also used Akaike information criterion (AIC) values to compare nonnested models to each other. AIC measures how well a model describes the data with a penalty for additional predictors to discourage overfitting a model. Lower AIC values indicate better fit between the data and the model, and differences in AIC values larger than three are generally considered meaningful. AIC can only compare models that were fit to the same data set, so they should not be used to compare a model analyzing disfluencies in CWS to one analyzing disfluencies in CWNS, for example.

Results

Summary statistics for CWS and CWNS appear in Table 4. Although the number of utterances per child, T4 mean length of utterances in morphemes (MLU-m), and proportion of decontextualized language were numerically larger (with moderate effect sizes) for CWNS compared to those who do stutter, these differences were not statistically significant. As expected, children who stuttered

Table 4. Descriptive statistics for utterances spoken by 15 children who stutter (CWS) and 15 children who do not stutter (CWNS) that met inclusion criteria for analysis of decontextualized language (fully intelligible and contained at least one verb).

AQ6

Variable	*M* (*SD*), range		Difference	
	CWS	CWNS	Significance	*d*
Number of included utterances	134 (71) 27–282	176 (62) 79–326	$p = .09$	0.63
Mean length of utterance in morphemes (MLU-m) of included utterances	4.85 (0.77) 3.41–5.88	5.18 (0.63) 4.37–6.32	$p = .22$	0.47
Proportion decontextualized	0.33 (0.15) 0.00–0.58	0.39 (0.10) 0.25–0.62	$p = .38$	0.47
Proportion utterances containing stuttering-like disfluency	0.27 (0.16) 0.06–0.66	0.09 (0.06) 0.007–0.19	$p = .001$	1.49
Proportion utterances containing typical disfluency	0.16 (0.08) 0.04–0.32	0.15 (0.05) 0.08–0.27	$p = .70$	0.15

Note. Differences between groups were measured by Cohen's *d* and two-tailed *t* tests.

produced significantly more SLDs than CWNS, but, notably, the groups did not differ in the proportion of utterances containing TDs.

We replicated the expected finding that for CWS, longer utterances are more likely to contain stuttering. The results of a logistic regression predicting the likelihood of an utterance containing stuttering from its length in morphemes yielded a significant effect of utterance length ($\beta_{\text{morphemes}}$ = AQ7 0.29, *SE* = 0.02, odds ratio [*OR*] = 1.29, 95% confidence interval [CI] [1.23, 1.36], $p < .001$). Similarly, increases in utterance length also resulted in increased TD for both CWS ($\beta_{\text{morphemes}} = 0.13$, *SE* = 0.02, *OR* = 1.14, 95% CI [1.09, 1.20], $p < .001$) and CWNS ($\beta_{\text{morphemes}} = 0.14$, *SE* = 0.02, *OR* = 1.11, 95% CI [1.06, 1.16], $p < .001$).

Our first research question asked whether decontextualized utterances were more likely to contain stuttering than contextualized utterances in CWS. To determine the effect of contextualization on stuttering in CWS, we regressed SLDs on contextualization. The logistic regression model predicting the presence of stuttering in an utterance from decontextualization was significant based on comparison with the null model predicting stuttering from only the random intercept of participant, $\chi^2(1) = 60.1$, $p < .001$. The answer to the first research question is that decontextualized utterances were significantly more likely to contain stuttering than contextualized utterances ($\beta_{\text{decontextualized}} = 0.89$, *SE* = 0.12, $p < .001$, *OR* = 2.45, 95% CI [1.94, 3.09]). The results of this model are reported in Table 5. T5

Our second research question asked whether decontextualized utterances increased the likelihood of stuttering independent of utterance length. The logistic regression model predicting the presence of stuttering in an utterance from decontextualization and utterance length was significant based on comparison with the null model, $\chi^2(2) = 140$, $p < .001$, and was a better fit to the data than a model with decontextualization as a sole predictor, $\chi^2(1) = 79.9$, $p < .001$. It was also a better fit to the data than the model with utterance length as a sole predictor, $\chi^2(1) = 29.1$, $p < .001$. Each predictor, utterance length, and decontextualization explained unique variance in the likelihood of stuttering. When controlling for decontextualization, a one-morpheme increase in utterance length was associated with a significant increase in stuttering ($\beta_{\text{morphemes}} = 0.22$, *SE* = 0.03, $p < .001$, *OR* = 1.25, 95% CI [1.19, 1.31]). When controlling for utterance length, a decontextualized utterance is more likely to contain stuttering than a contextualized utterance ($\beta_{\text{decontextualized}} = 0.66$, *SE* = 0.12, $p < .001$, *OR* = 1.93, 95% CI [1.51, 2.46]).

Table 5. Model results for Research Questions 1 and 2.

AQ8

Model	$\beta_{\text{decontextualized}}$	$\beta_{\text{morphemes}}$	$\beta_{\text{decon}}{*}\beta_{\text{morph}}$	AIC
Null: SLD ~ (1|child)	—	—	—	2125
SLD ~ decon	0.89	—	—	2068
SLD ~ morph	—	0.29	—	2023
SLD ~ decon + morph	0.66	0.22	—	1989
SLD ~ decon * morph	0.72	0.33	–0.19	1977

Note. The random intercept of participant, written as (1|ChildID), is omitted for brevity in all but the null model. All coefficients were significant at $p < .001$. decon = decontextualized; morph = utterance length in morphemes; AIC = Akaike information criterion; SLD = stuttering-like disfluency.

Table 6. Model results for children who do not stutter for Research Questions 3 and 4.

Model	$\beta_{decontextualized}$	$\beta_{morphemes}$	$\beta_{decon}*\beta_{morph}$	AIC
Null: TD ~ (1|child)	—	—	—	2274
TD ~ decon	0.65	—	—	2242
TD ~ morph	—	0.14	—	2241
TD ~ decon + morph	0.50	0.11	—	2224
TD ~ decon * morph	0.53	0.14	−0.06, p = .22	2224

Note. The random intercept of participant, written as (1|ChildID), is omitted for brevity in all but the null model. All coefficients were significant at p < .001 unless otherwise noted. decon = decontextualized; morph = utterance length in morphemes; AIC = Akaike information criterion; TD = typical disfluency.

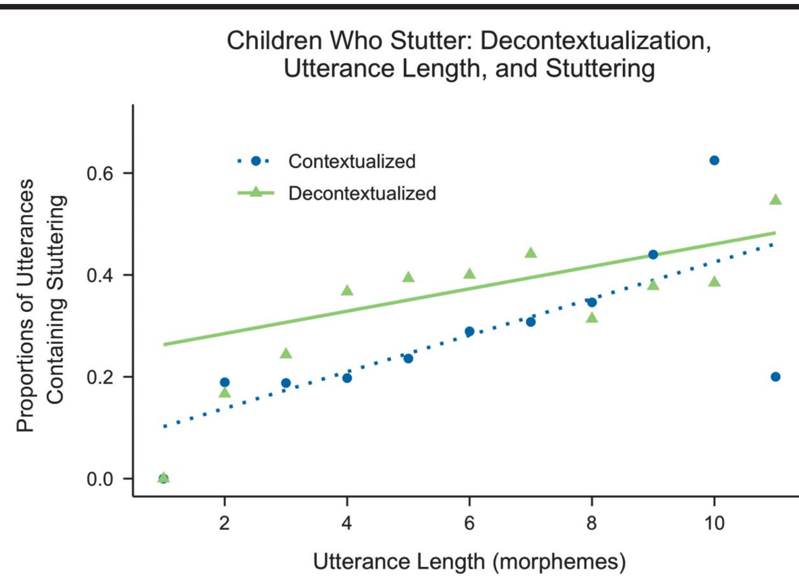
Adding an interaction between decontextualization and utterance length to the model increased model fit compared to the model without an interaction, $\chi^2(1)$ = 14.7, p < .001. In this model, a one-morpheme increase in utterance length significantly increased the likelihood of stuttering when decontextualization was held constant ($\beta_{morphemes}$ = 0.33, SE = 0.04, p < .001, OR = 1.39, 95% CI [1.29, 1.50]). Decontextualization remained a predictor of stuttering when utterance length was held constant ($\beta_{decontextualized}$ = 0.72, SE = 0.12, p < .001, OR = 2.06, 95% CI [1.61, 2.62]). The significant and negative interaction term ($\beta_{decon*morph}$ = −0.19, SE = 0.05, p < .001, OR = 0.82, 95% CI [0.75, 0.91]) meant that decontextualization had a stronger effect on stuttering for shorter utterances than for longer utterances. This model had a classification accuracy of 76% using a cutoff probability of 0.5. The results of the models tested for Research Questions 1
T6 and 2 are summarized in Table 6.

Based on the results of our model comparison, the answer to research question 2 is that utterance length and decontextualization both increase the likelihood of stuttering, but decontextualization has a larger effect on short utterances than long ones. This relationship is depicted in Figure 2. F2

Our third research question asked whether decontextualized utterances were more likely than contextualized ones to contain TDs. We analyzed the data for children who do and do not stutter separately. Analyzing data from CWNS, the logistic regression model predicting the presence of TDs from decontextualization was significant based on comparison with the null model predicting disfluency from only the random intercept of participant, $\chi^2(1)$ = 34.4, p < .001. Decontextualized utterances were significantly more likely to contain TDs than contextualized utterances ($\beta_{decontextualized}$ = 0.65, SE = 0.11, p < .001, OR = 1.91, 95% CI [1.51, 2.41]). The results of this model are reported in Table 6.

Children who stuttered followed a similar pattern, with a significant increase in TD for decontextualized compared to contextualized utterances. The logistic regression model

Figure 2. Proportion of decontextualized and contextualized utterances that contain stuttering-like disfluencies from children who stutter for each utterance length with more than 25 utterances.



predicting the presence of TDs from decontextualization was significant based on comparison with the null model predicting disfluency from only the random effect of participants, $\chi^2(1) = 40.8$, $p < .001$. Decontextualized utterances were significantly more likely to contain TDs than contextualized utterances ($\beta_{decontextualized} = 0.78$, $SE = 0.12$, $p < .001$, $OR = 2.19$, 95% CI [1.71, 2.80]). The results of this model are T7 reported in Table 7. The answer to our third research question is that for children who do and do not stutter, decontextualized utterances are more likely to contain TDs.

Our fourth research question asked whether decontextualized utterances contained more TDs than contextualized utterances when controlling for utterance length. As in our third research question, we analyzed the data for children who do and do not stutter separately. The logistic regression model predicting the presence of TD from decontextualization and utterance length was significant based on comparison with the null model for CWNS, $\chi^2(2) = 54.7$, $p < .001$, and for CWS, $\chi^2(2) = 54.7$, $p < .001$. Models that included utterance length were a better fit to the data than a model with decontextualization as a sole predictor for CWNS, $\chi^2(1) = 20.2$, $p < .001$, and for CWS, $\chi^2(1) = 12.9$, $p < .001$. Models including decontextualization were a better fit to the data than the model with utterance length as the sole predictor for CWNS, $\chi^2(1) = 19.1$, $p < .001$, and for CWS, $\chi^2(1) = 27.9$, $p < .001$. Each predictor, utterance length and decontextualization, explained unique variance in the likelihood of TD. When controlling for decontextualization, a one-morpheme increase in utterance length was associated with a significant increase in TD for CWNS ($\beta_{morphemes} = 0.11$, $SE = 0.02$, $p < .001$, $OR = 1.11$, 95% CI [1.06, 1.17]) and for CWS ($\beta_{morphemes} = 0.09$, $SE = 0.03$, $p < .001$, $OR = 1.09$, 95% CI [1.04, 1.15]). When controlling for utterance length, a decontextualized utterance is more likely to contain stuttering than a contextualized utterance for CWNS ($\beta_{decontextualized} = 0.50$, $SE = 0.11$, $p < .001$, $OR = 1.65$, 95% CI [1.31, 2.08]) and for CWS ($\beta_{decontextualized} = 0.67$, $SE = 0.13$, $p < .001$, $OR = 1.96$, 95% CI [1.52, 2.52]).

Adding an interaction between decontextualization and utterance length to the model did not significantly increase model fit compared to the model without an interaction for CWNS, $\chi^2(1) = 1.44$, $p = .23$, nor for CWS, $\chi^2(1) = 1.37$, $p = .4$. We cannot reject the null hypothesis that the effect of decontextualization on TDs is the same regardless of utterance length. The models with and without the interaction term had the similar AIC values (see Tables 6 and 7), providing further support that adding the interaction term did not improve model fit.

Based on the results of our model comparison, the answer to Research Question 4 is that, for both children who do and do not stutter, decontextualization does predict TD independent of utterance length, and there is no evidence for an interaction between those two predictors. This relationship is depicted in Figure 3. The classification F3 accuracy for the models predicting TD from utterance length and decontextualization was 82% for CWS and 84% for CWNS using a cutoff of 0.5.

Our fifth and final research question asked whether decontextualized language had the same effect on TDs for children who do and do not stutter. We first combined data from the two groups and regressed TDs on decontextualization and utterance length. Both predictors remained significant in the model of the combined data (utterance length: $\beta_{morpheme} = 0.10$, $SE = 0.02$, $p < .001$, $OR = 1.11$, 95% CI [1.07, 1.15]; decontextualization: $\beta_{decontextualized} = 0.58$, $SE = 0.09$, $p < .001$, $OR = 1.78$, 95% CI [1.50, 2.11]). We then added a diagnosis group to the model and each of the two-way interactions between diagnosis and the original predictors. In each of the three models with diagnosis group added, neither its coefficient nor the coefficient of the interaction between diagnosis and another predictor was significant (*ps* > .2) and none of the models with diagnosis as a predictor were a better fit to the data than the corresponding model without diagnosis as a predictor based on likelihood ratio tests (*ps* > .2). These findings are consistent with a claim that decontextualization has the same impact on TDs for children who do and do not stutter.

Exploratory Analyses

Based on recent findings that the two categories of TDs, stalls and revisions, are differentially impacted by

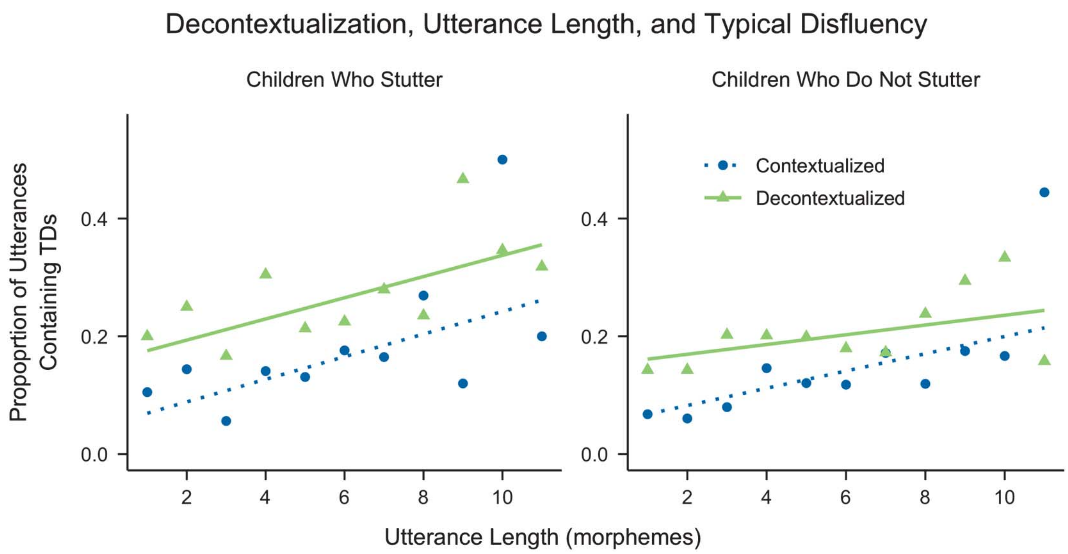
Table 7. Model results for children who stutter for Research Questions 3 and 4.

AQ10

Model	$\beta_{decontextualized}$	$\beta_{morphemes}$	$\beta_{decon}*\beta_{morph}$	AIC
Null: TD ~ (1|child)	—	—	—	1867
TD ~ decon	0.78	—	—	1828
TD ~ morph	—	0.13	—	1843
TD ~ decon + morph	0.67	0.09	—	1817
TD ~ decon * morph	0.69	0.13	−0.06, $p = .24$	1818

Note. The random intercept of participant, written as (1|ChildID), is omitted for brevity in all but the null model. All coefficients were significant at $p < .001$ unless otherwise noted. decon = decontextualized; morph = utterance length in morphemes; AIC = Akaike information criterion; TD = typical disfluency.

Figure 3. Proportion of decontextualized and contextualized utterances that contain typical disfluencies (TDs) from children who stutter and children who do not stutter for each utterance length with more than 25 utterances.



discourse factors (Garbarino & Bernstein Ratner, 2023), we re-examined Research Questions 3–5 about TDs looking at revisions and stalls separately. Since stalls reflect utterance planning effort, we might expect that decontextualized language would result in increased stalls and may be driving the effect of decontextualization on TD. Revisions reflect the speaker's awareness of a mismatch between their message and the utterance they produced, and we did not have a prediction about a relationship between decontextualization and revisions.

We repeated the statistical procedures for Research Questions 3–5 with stalls and revisions as outcome variables. In every case, we found the same pattern of significant effects in the same direction whether we used TD, revisions, or stalls as the outcome variable for both groups of children. Stalls and revisions were both more common in decontextualized utterances compared to contextualized utterances, and increased with utterance length.

A reviewer helpfully pointed out that utterance length in syllables may be a more appropriate measure of the motor complexity of an utterance than measuring utterance length in morphemes. To count syllables in each utterance of our transcript, we used the *nsyllable* function of the R package by the same name (Benoit, 2022), which uses the CMU Pronouncing Dictionary (Carnegie Mellon Speech Group, 2015) to look up syllable counts in words. We wrote a wrapper function around *nsyllable* to count syllables in utterances, while ignoring maizes, part-word repetitions, and whole-word repetitions. For 40% of utterances, length in morphemes was identical to length in syllables: for an additional 40% of utterances length in syllables within one unit of length in morphemes. AQ11 Length in morphemes and length in syllables was correlated strongly and significantly, r = .82, 95% CI [.81, .83], t(3422) = 84. Re-running our primary analyses with utterance length measured in syllables rather than in morphemes did not change the significance or direction of any main effects or interactions.

Discussion

In this study, we analyzed an existing data set of transcripts of children who do and do not stutter to assess a possible relationship between decontextualized language (discussion of topics outside of the "here and now") and disfluency. We found that decontextualized language was associated with increased SLDs in CWS and increased TDs in CWNS, and this increase could not be explained by increases in utterance length alone. We also found that the effect of decontextualization was strongest for shorter utterances.

Our findings add to our understanding of the relationship between decontextualized language and fluency. While different language sampling methods and different definitions of decontextualized language across studies make it hard to compare findings directly, converging evidence from different approaches all showing a link between decontextualization and fluency suggests that this is a robust finding for preschool-aged children (Nosek et al., 2022).

Our investigation differed from prior studies of decontextualized language and stuttering because we completed our analyses at the utterance level, by coding individual utterances for contextualization and controlling for utterance length. This allowed us to tease apart the influences of utterance length and decontextualization, which is important since decontextualized utterances are typically longer than contextualized ones, and longer utterances are associated with decreased fluency. Categorizing individual utterances as contextualized or decontextualized also allowed us to take advantage of data collected in naturalistic play settings, which may have more ecological validity than language samples taken during experimental tasks.

Our findings are in line with the claim that decontextualized language places increased demands on working memory as speakers formulate utterances outside of visual supports provided by the "here and now." Increased rates of TDs in decontextualized language were consistent with the greater planning effort required for decontextualized compared to contextualized language. The size of the effect of decontextualization on the presence of one or more TDs in an utterance was not statistically different between children who do and do not stutter, which is consistent with decontextualized language having the same effect on TD in both groups of children.

According to the Multifactorial Dynamic Pathways Model of stuttering, CWS have subtle weaknesses in motor planning that create susceptibility to stuttering. Subtle weaknesses in a variety of cognitive and linguistic skills in CWS may mean that the increased working memory demands of decontextualized language make it particularly demanding, contributing to additional SLDs (Ofoe et al., 2018). These findings are also consistent with other multifactorial models of stuttering, such as the CALMS model (Healey et al., 2004), and with the Demands and AQ12 Capacities Model (Starkweather, 1987), which suggests that limitations of a child's speech production capacity interacts with the demands of a speaking situation to produce stuttering. However, we note that the increased planning effort associated with decontextualized language leads to increased TDs in CWNS and leads to increases in both typical and SLDs in CWS.

Clinical Implications

The impact of decontextualized language on stuttering has potential implications for its clinical management. We begin this section with a reminder that the goal of therapy for a child who stutters is not to eliminate stuttering, but rather to promote communicative competence. Thus, therapy might target increased fluency, but might also target assisting the child to produce speech more comfortably, regardless of fluency. Regardless of therapeutic goal, our findings suggest that whether or not a task provides contextual support may impact the child's cognitive resources to either be fluent or manage other aspects of the interaction. We also note that communicating effectively includes using language for a variety of purposes, including narration, pretend play, and discussion of causal relationships—categories of decontextualized language. These language functions assume even greater importance as children move from preschool to academic language expectations.

First, we note that decontextualized language is important for child language growth, even if it may tax resources dedicated to fluent production. Additionally, child use of decontextualized language and its ensuing benefits are directly related to adult modeling and support of decontextualized topics by caregivers (e.g., Curenton et al., 2008; Leech et al., 2018; Rowe, 2013; Uccelli et al., 2019). The benefits of decontextualized language for language development and academic achievement suggest additional reasons that clinicians should be wary of recommendations to simplify language to reduce stuttering, as noted by Bernstein Ratner and Guitar (2006). CWS may have somewhat weaker language and metalinguistic skills than their typically developing peers (e.g., see summary in Bloodstein et al., 2021, but see Nippold, 2012). Exposure to parental decontextualized language has larger benefits to children with weaker language skills (Demir et al., 2015), so if it is the case that stuttering is associated with weaker language skills, simplifying input with the goal of reducing stuttering may reduce learning opportunities for these children by limiting their opportunities to hear and use the vocabulary and syntactic structures associated with decontextualized language.

Awareness that decontextualized language may increase stuttering may still be useful for parent counseling. Parents are often advised ways to reduce communicative pressure on a child, such as increasing wait time after a conversational turn or reducing their speaking rate. Knowing that decontextualized language may increase stuttering for their child can help parents target when to focus on implementing these indirect strategies in situations where their child might stutter more.

Clinicians can consider decontextualized language as one factor affecting stuttering in evaluation and treatment settings. When comparing the frequency of stuttering across two language samples, clinicians may take note of whether the samples are comparable in terms of decontextualized language. Language samples taken under similar conditions, such as during play or in a picture description task, may differ in how many decontextualized utterances they contain, and this could affect the speaker's fluency.

In treatment, some clinicians use a hierarchy of speaking situations in which a client might practice fluency-enhancing or stuttering management strategies. Considering the impact of decontextualized language may help with activity selection and treatment planning. For example, a conversation with a peer may be prone to more stuttering while talking about a recent event but may contain less stuttering while playing a board game where the discussion focuses on the here and now.

Our study included preschool children, and if decontextualized language continues to predict stuttering as they get older, that relationship may contribute to challenges with class participation in CWS. While not all decontextualized language is academic and not all academic language is decontextualized, they are meaningfully related. The amount of decontextualization increases as children progress in their academic careers (van Kleeck, 2014). Academic learning and assignments emphasize discussion of events and concepts that require decontextualized language, and students must often demonstrate their knowledge through oral explanations in small or large groups (Blood et al., 2001; van Kleeck, 2014). If increased stuttering resulting from the decontextualized nature of academic language limits a child's willingness to participate in class, SLPs may support their use of fluency strategies or desensitization during decontextualized language tasks.

Conclusions

Our study suggests that the increased language formulation and working memory demands associated with decontextualized language lead to increased typical and SLDs. This increase can be found even when controlling for utterance length, a confounding variable that is difficult to control for when analyzing whole language samples without utterance-level coding. TDs increased about the same amount in children who did and did not stutter when they produced decontextualized utterances, suggesting that both groups are impacted by the increased effort needed to plan decontextualized utterances.

Our findings align with a major multifactorial model of stuttering and identify decontextualized language as one of many linguistic factors that can lead to stuttering. Understanding the many ways that language and communication situations affect fluency and stuttering can help clinicians plan treatment and can help individuals who stutter and their families better understand their stuttering. Our sample included preschool-aged children, and extending these findings to older children may shine light on the relationship between decontextualized language, which is common in academic settings, and stuttering in the classroom.

Data Availability Statement

Videos and CHAT transcripts of the language samples used in this study are available on CHILDES at https://fluency.talkbank.org/access/UMD-CMU.html after password request to the site administrator. Data generated for this study (decontextualized language coding) are available from the author upon reasonable request. The code used to run the analyses is available on the Open Science Framework at https://osf.io/q95km/?view_only=e6c70b0f9dd04510b0f407345b9782d4>.

Acknowledgments

This work was funded by National Institutes of Health Grant R01DC018276 (K.O.), NSF NRT DESE Grant #1449815 (K.O.), National Institute on Deafness and Other Communication Disorders Grant R01DC015494 025949 (N.B.R.), NSF-BCS Grant No. BCS-1844194 (Huang), and a Maryland Summer Scholars grant to J.L. We are grateful to the participants and their families who generously gave their time to make this research possible. The authors would also like to thank Julianne Garbarino for helpful discussions in the early phases of this research. AQ13

References

Ambrose, N. G., & Yairi, E. (1999). Normative disfluency data for early childhood stuttering. *Journal of Speech, Language, and Hearing Research, 42*(4), 895–909. https://doi.org/10.1044/jslhr.4204.895

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Benoit, K. (2022). *nsyllable: Count syllables in character vectors*. R package version 1.0.1. https://CRAN.R-project.org/package=nsyllable

Bernstein Ratner, N., & Brundage, S. B. (2022). *A clinician's complete guide to CLAN and PRAAT*. https://www.talkbank.org/manuals/Clin-CLAN.pdf

Bernstein Ratner, N., & Guitar, B. (2006). Treatment of very early stuttering and parent-administered therapy: The state of the art. In N. Bernstein Ratner & J. Tetnowski (Eds.), *Current issues in stuttering research and practice* (pp. 99–124). Erlbaum.

Bernstein Ratner, N., & MacWhinney, B. (2018). Fluency Bank: A new resource for fluency research and practice. *Journal of Fluency Disorders, 56*, 69–80. https://doi.org/10.1016/j.jfludis.2018.03.002

Blood, G. W., Blood, I. M., Tellis, G., & Gabel, R. (2001). Communication apprehension and self-perceived communication competence in adolescents who stutter. *Journal of Fluency Disorders, 26*(3), 161–178. https://doi.org/10.1016/s0094-730x(01)00097-3

Bloodstein, O., Bernstein Ratner, N., & Brundage, S. B. (2021). *A handbook on stuttering*. Plural.

Buhr, A., & Zebrowski, P. (2009). Sentence position and syntactic complexity of stuttering in early childhood: A longitudinal study. *Journal of Fluency Disorders, 34*(3), 155–172. https://doi.org/10.1016/j.jfludis.2009.08.001

Carnegie Mellon Speech Group. (2015). *Carnegie Mellon University pronouncing dictionary* (Version 0.7b).

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Curenton, S. M., Craig, M. J., & Flanigan, N. (2008). Use of decontextualized talk across story contexts: How oral storytelling and emergent reading can scaffold children's development. *Early Education and Development, 19*(1), 161–187. https://doi.org/10.1080/10409280701839296

Davidson, R. G., Kline, S. B., & Snow, C. E. (1986). Definitions and definite noun phrases: Indicators of children's decontextualized language skills. *Journal of Research in Childhood Education, 1*(1), 37–48. https://doi.org/10.1080/02568548609594907

Demir, Ö. E., Rowe, M. L., Heller, G., Goldin-Meadow, S., & Levine, S. C. (2015). Vocabulary, syntax, and narrative development in typically developing children and children with early unilateral brain injury: Early parental talk about the "there-and-then" matters. *Developmental Psychology, 51*(2), 161–175. https://doi.org/10.1037/a0038476

Eichorn, N., Marton, K., Schwartz, R. G., Melara, R. D., & Pirutinsky, S. (2015). Does working memory enhance or interfere with speech fluency in adults who do and do not stutter? Evidence from a dual-task paradigm. *Journal of Speech, Language, and Hearing Research, 59*(3), 415–429. https://doi.org/10.1044/2015_JSLHR-S-15-0249

Garbarino, J., & Bernstein Ratner, N. (2022). What is the role of questioning in young children's fluency? *American Journal of Speech-Language Pathology, 31*(5), 2061–2077. https://doi.org/10.1044/2022_AJSLP-21-00209

Garbarino, J., & Bernstein Ratner, N. (2023). Stalling for time: Stall, revision and stuttering-like disfluencies reflect language factors in the speech of young children. *Journal of Speech, Language, and Hearing Research, 66*(6), 2018–2034. https://doi.org/10.1044/2023_JSLHR-22-00595

Garbarino, J. T. (2021). *Utterance-level predictors of stuttering-like, stall, and revision disfluencies in the speech of young children who do and do not stutter* [Doctoral dissertation, University of Maryland]. https://drum.lib.umd.edu/bitstream/handle/1903/27378/Garbarino_umd_0117E_21511.pdf

Gillam, S. L., Gillam, R. B., & Reece, K. (2012). Language outcomes of contextualized and decontextualized language intervention: Results of an early efficacy study. *Language, Speech, and Hearing Services in Schools, 43*(3), 276–291. https://doi.org/10.1044/0161-1461(2011/11-0022)

Hara, Y., Hata, W., Mizuto, Y., Murakami, T., Ichikawa, M., Kamioka, S., & Ishizaka, I. (2022). Relationship between stuttering symptoms and speech content in early childhood: Explanation of "here-and-now" events and "non-here-and-now" events. *The Kitasato Medical Journal, 52*(2), 112–118.

Healey, E. C., & Trautman, L. S., & Susca, M. (2004). Clinical applications of a multidimensional approach for the assessment and treatment of stuttering. *Contemporary Issues in Communication Science and Disorders, 31*(Spring), 40–48. https://doi.org/10.1044/cicsd_31_S_40

Leech, K. A., Wei, R., Harring, J. R., & Rowe, M. L. (2018). A brief parent-focused intervention to improve preschoolers' conversational skills and school readiness. *Developmental Psychology, 54*(1), 15–28. https://doi.org/10.1037/dev0000411

MacPherson, M. K., & Smith, A. (2013). Influences of sentence length and syntactic complexity on the speech motor control of children who stutter. *Journal of Speech, Language, and Hearing Research, 56*(1), 89–102. https://doi.org/10.1044/1092-4388(2012/11-0184)

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk* (3rd ed.). Erlbaum.

MacWhinney, B., Roberts, J. A., Altenberg, E. P., & Hunter, M. (2020). Improving automatic IPSyn coding. *Language, Speech, and Hearing Services in Schools, 51*(4), 1187–1189. https://doi.org/10.1044/2020_LSHSS-20-00090

Mallick, R., Kathard, H., Thabane, L., & Pillay, M. (2021) A scoping review of the school-aged stuttering intervention literature. *Evidence-Based Communication Assessment and Intervention, 15*(4), 194–206. https://doi.org/10.1080/17489539.2021.1976964

Masterson, J. J., & Kamhi, A. G. (1991). The effects of sampling conditions on sentence production in normal, reading-disabled, and language-learning-disabled children. *Journal of Speech and Hearing Research, 34*(3), 549–558. https://doi.org/10.1044/jshr.3403.549

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica, 22*(3), 276–282. https://doi.org/10.11613/BM.2012.031

Melnick, K. S., & Conture, E. G. (2000). Relationship of length and grammatical complexity to the systematic and nonsystematic speech errors and stuttering of children who stutter. *Journal of Fluency Disorders, 25*(1), 21–45. https://doi.org/10.1016/s0094-730x(99)00028-5

Nippold, M. A. (2012). Stuttering and language ability in children: Questioning the connection. *American Journal of Speech-Language Pathology, 21*(3), 183–196. https://doi.org/10.1044/1058-0360(2012/11-0078)

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Kline Struhl, M., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology, 73,* 719–748. https://doi.org/10.1146/annurev-psych-020821-114157

Ofoe, L. C., Anderson, J. D., & Ntourou, K. (2018). Short-term memory, inhibition, and attention in developmental stuttering: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 61*(7), 1626–1648. https://doi.org/10.1044/2018_JSLHR-S-17-0372

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rispoli, M. (2003). Changes in the nature of sentence production during the period of grammatical development. *Journal of Speech, Language, and Hearing Research, 46*(4), 818–830. https://doi.org/10.1044/1092-4388(2003/064)

Rispoli, M. (2018). Changing the subject: The place of revisions in grammatical development. *Journal of Speech, Language, and Hearing Research, 61*(2), 360–372. https://doi.org/10.1044/2017_JSLHR-L-17-0216

Rispoli, M., Hadley, P., & Holt, J. (2008). Stalls and revisions: A developmental perspective on sentence production. *Journal of Speech, Language, and Hearing Research, 51*(4), 953–966. https://doi.org/10.1044/1092-4388(2008/070)

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development, 83*(5), 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Rowe, M. L. (2013). Decontextualized language input and preschoolers' vocabulary development. *Seminars in Speech and Language, 34*(4), 260–266. https://doi.org/10.1055/s-0033-1353444

Smith, A., & Weber, C. (2017). How stuttering develops: The multifactorial dynamic pathways theory. *Journal of Speech, Language, and Hearing Research, 60*(9), 2483–2505. https://doi.org/10.1044/2017_JSLHR-S-16-0343

Snow, C. (1983). Literacy and language: Relationships during the preschool years. *Harvard Educational Review, 53*(2), 165–189. https://doi.org/10.17763/haer.53.2.t6177w39817w2861

Trautman, L. S., Healey, E. C., & Norris, J. A. (2001). The effects of contextualization on fluency in three groups of children. *Journal of Speech, Language, and Hearing Research, 44*(3), 564–576. https://doi.org/10.1044/1092-4388(2001/044)

Uccelli, P., Demir-Lira, Ö. E., Rowe, M. L., Levin, S., & Goldin-Meadow, S. (2019). Children's early decontextualized talk predicts academic language proficiency in midadolescence. *Child Development, 90*(5), 1650–1663. https://doi.org/10.1111/cdev.13034

Usler, E. R., & Walsh, B. (2018). The effects of syntactic complexity and sentence length on the speech motor control of school-age children who stutter. *Journal of Speech, Language, and Hearing Research, 61*(9), 2157–2167. https://doi.org/10.1044/2018_JSLHR-S-17-0435

van Kleeck, A. (2014). Distinguishing between casual talk and academic talk beginning in the preschool years: An important consideration for speech-language pathologists. *American Journal of Speech-Language Pathology, 23*(4), 724–741. https://doi.org/10.1044/2014_AJSLP-14-0032

Wagovich, S. A., Hall, N. E., & Clifford, B. A. (2009). Speech disruptions in relation to language growth in children who stutter: An exploratory study. *Journal of Fluency Disorders, 34*(4), 242–256. https://doi.org/10.1016/j.jfludis.2009.09.004

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemnund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., . . . Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), Article 1686. https://doi.org/10.21105/joss.01686

Wickham, H., François, R., Henry, L., & Müller, K. (2022). *dplyr: A grammar of data manipulation*. R package version 1.0.10. https://CRAN.R-project.org/package=dplyr

Zackheim, C. T., & Conture, E. G. (2003). Childhood stuttering and speech disfluencies in relation to children's mean length of utterance: A preliminary study. *Journal of Fluency Disorders, 28*(2), 115–142. https://doi.org/10.1016/s0094-730x(03)00007-x