ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

Stalling for Time: Stall, Revision, and Stuttering-Like Disfluencies Reflect Language Factors in the Speech of Young Children

Julianne Garbarino[a] and **Nan Bernstein Ratner**[a]

[a] Department of Hearing and Speech Sciences, University of Maryland, College Park

ARTICLE INFO

Article History:
Received October 16, 2022
Revision received January 4, 2023
Accepted February 16, 2023

Editor-in-Chief: Cara E. Stepp
Editor: Ayoub Daliri

https://doi.org/10.1044/2023_JSLHR-22-00595

ABSTRACT

Purpose: Disfluencies can be classified as stuttering-like disfluencies (SLDs) or typical disfluencies (TDs). Dividing TDs further, stalls (fillers and repetitions) are thought to be prospective, occurring due to planning glitches, and revisions (word and phrase revisions, word fragments) are thought to be retrospective, occurring when a speaker corrects language produced in error. In the first study assessing stalls, revisions, and SLDs in matched groups of children who stutter (CWS) and children who do not stutter (CWNS), we hypothesized that SLDs and stalls would increase with utterance length and grammaticality but not with a child's expressive language level. We expected revisions to be associated with a child having more advanced language but not with utterance length or grammaticality. We hypothesized that SLDs and stalls (thought to be planning-related) would tend to precede grammatical errors.
Method: We analyzed 15,782 utterances from 32 preschool-age CWS and 32 matched CWNS to assess these predictions.
Results: Stalls and revisions increased in ungrammatical and longer utterances and with the child's language level. SLDs increased in ungrammatical and longer utterances, but not with overall language level. SLDs and stalls tended to occur before grammatical errors.
Conclusions: Results suggest that both stalls and revisions are more likely to occur in utterances that are harder to plan (those that are ungrammatical and/or longer) and that, as children's language develops, so do the skills they need to produce both stalls and revisions. We discuss clinical implications of the finding that ungrammatical utterances are more likely to be stuttered.

Disfluencies are typically broken down into two major subtypes: stuttering-like disfluencies (SLDs) and typical disfluencies (TDs). SLDs include part-word or monosyllabic whole-word repetitions, blocks, prolongations, and broken words (e.g., Ambrose & Yairi, 1999; Logan & Conture, 1995). These disfluencies, particularly part-word and monosyllabic whole-word repetitions, can and do occur in the speech of people who do not stutter (PWNS), but they are more frequent and often involve more iterations (e.g., "a-a-a-a-a-and" vs. "a-and") in the speech of people who stutter (PWS; Ambrose & Yairi, 1999). SLDs are often accompanied by muscle tension and awareness when they are produced by PWS (Ambrose & Yairi, 1994; Tichenor et al., 2017). TDs include multisyllabic whole-word repetitions, phrase repetitions, word and phrase revisions, word fragments, and fillers such as "um" and "uh" (Ambrose & Yairi, 1999; Logan & Conture, 1995; Yaruss et al., 1999), and they occur in the speech of PWS and PWNS (e.g., Ambrose & Yairi, 1999; Buhr & Zebrowski, 2009).

Disfluency–Language Relationships

SLDs and Utterance-Level Language Factors

Decades of research have demonstrated that SLDs are more likely to occur in longer and/or more syntactically complex utterances. The relationship between SLD

Correspondence to Julianne Garbarino: jgarbari@umd.edu. ***Disclosure:*** *The authors have declared that no competing financial or non-financial interests existed at the time of publication.*

 Copyright © 2023 American Speech-Language-Hearing Association

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

production and longer utterances has been shown with length measured in syllables (Logan & Conture, 1995; Yaruss, 1999), morphemes (Richels et al., 2010; Yaruss, 1999; Zackheim & Conture, 2003), words (Buhr & Zebrowski, 2009; Gaines et al., 1991; Hollister et al., 2017; Wagovich et al., 2009; Yaruss, 1999) and syntactic constituents[1] (Melnick & Conture, 2000; Yaruss, 1999).

If disfluencies result in part during problems in the utterance planning process, then they should occur at locations where psycholinguists believe that planning loads are high. SLDs do, in fact, tend to occur at the beginnings of syntactic constituents (Bernstein, 1981), clauses (Wall et al., 1981), and utterances (Buhr & Zebrowski, 2009; Choi et al., 2020; Gaines et al., 1991). Considered altogether, both length and complexity studies and loci studies have found SLDs to occur where planning loads are highest—in longer and more complex utterances and early within the utterance or other planning unit (e.g., Bernstein Ratner & Sih, 1987; Buhr & Zebrowski, 2009; Choi et al., 2020; Wagovich et al., 2009; Yaruss, 1999).

TD–Language Relationships

Similar to SLDs, when all TD types are measured as a group, they tend to occur in longer (Buhr & Zebrowski, 2009; Yaruss et al., 1999; Zackheim & Conture, 2003) and more complex (Buhr & Zebrowski, 2009; Gordon & Luper, 1989; Haynes & Hood, 1978; Yaruss et al., 1999; Zackheim & Conture, 2003) utterances. However, despite the tendency for TDs to be analyzed as one type of disfluency, it is not clear that all types of TDs actually relate to language factors similarly. Rispoli and colleagues have differentiated between *stalls* and *revisions* (Rispoli, 2003, 2018; Rispoli et al., 2008). Stalls simply delay and do not change any words or morphemes, and they include fillers and word and phrase repetitions. They are "prospective" and occur due to glitches in planning language not yet produced (Rispoli, 2003). Revisions, in contrast, are "retrospective," involving changes to previously produced words or morphemes, and include word and phrase revisions and fragments (Rispoli, 2003). We will refer to the group of TDs including word and phrase revisions and fragments as "revisions." The more specific subtypes will be described as "word revisions," or "phrase revisions." We will refer to fillers, word repetitions, and phrase repetitions as a group as "stalls."

An informative discrepancy has emerged in studies that have separated stalls and revisions. Stalls do not have an apparent relationship with a child's overall language abilities; a longitudinal study tracking children from ages 1;9 to 2;9 [years;months] found that for different children, stalls increased, decreased, or remained constant as mean length of utterance in morphemes (MLU-m) and expressive syntax scores increased (Rispoli et al., 2008). These results converged with evidence from an earlier cross-sectional study that enrolled children ages 1;10–4;0 (Rispoli, 2003) and found that revision rate correlated with MLU-m and expressive syntax but that stall rate did not correlate with either of these. While they do not appear to relate to language development, stalls do have a relationship with utterance planning load; stalls are more likely to occur in longer utterances (Rispoli, 2003; Rispoli et al., 2008). In contrast, revisions have not been found to have a significant relationship with utterance length but do increase with expressive language level (as measured by MLU-m) and verb phrase (VP) complexity (Rispoli, 2003; Rispoli et al., 2008).

Rispoli (2003, 2018) has hypothesized about why stalls and revisions have different relationships with utterance length and complexity and language development, pointing to differing prerequisite skills for these TD types. For a stall to occur, all that is required is a "glitch" in sentence planning. In contrast, revisions also require that that the child has a repertoire of other words or phrase structures to substitute. Support for this explanation was provided by Rispoli's (2018) finding that the number of different subject noun phrases that children produced in their language samples (i.e., subject diversity) was a significant predictor of subject revision rate. Children who had more available substitutes for that specific sentence role made more revisions in that position. Rispoli (2003, 2018) also suggests that sufficient monitoring skills are prerequisites of revising, which further explains increasing revisions with language development.

In a longitudinal study following nine children who stutter (CWS) for approximately 10 months, Wagovich et al. (2009) used the stall–revision framework. SLDs and stalls occurred more often in longer utterances (as measured by MLU-m) and were variable over time, whereas revision rate did not relate to utterance length but increased over time. The results of this study, along with those of Rispoli (2003) and Rispoli et al. (2008), suggest that the best way to understand TDs and their relationship to utterance-level and developmental language factors may include looking at TD subtypes.

Children's Sentence Planning

Disfluency–Planning Relationships

The study of disfluency in children's speech and children's sentence planning has been closely linked with and has built on some of the early work in adult sentence planning (e.g., Boomer, 1965; Goldman-Eisler, 1968; Maclay &

[1]Melnick and Conture (2000) and Yaruss (1999) defined the number of syntactic constituents by counting the number of subject, verb, object, complement, and adverbial clauses.

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Osgood, 1959). Clause junctures appear to be important planning locations for children, as they are for adults (McKee et al., 2017). While children's sentence planning shares many features with adult planning, children may have to stall to plan more often (McDaniel et al., 2010).

One theoretical approach to understanding how TD production interacts with language development is that TDs, particularly stalls, occur in utterances that are at the leading edge of what a child is able to produce; under the leading edge theory, children have the linguistic knowledge to produce these utterances but do not have the processing capacity to do this entirely fluently (Rispoli, 2003; Rispoli & Hadley, 2001; Rispoli et al., 2008). Under this theory, children have a comfort zone of utterances that they can produce fluently and a leading edge of utterances that are still within their reach but that are more likely to be disfluent. In other words, disfluent utterances are likely to be the most syntactically advanced utterances within a child's output.

Grammatical Errors and Sentence Planning

We will use the term *grammatical error* in the sense of the works of Lee (1974) and Eisenberg et al. (2012), which is deviation from adult grammar, specifically, from the grammar of adult speakers of the same dialect. This use of "error" only indicates that a production is non-adultlike, and not that a child's utterance diverges from what is expected at their age. Errors, under this definition, are common in the speech of young children with typical language development. A study of children ages 2;6–6;6 reported that young typically developing children produced errors in 11% of utterances on average, with a range across children of 0%–35% (Dunn et al., 1996). Throughout the article, we use "ungrammatical" to refer to utterances containing at least one grammatical error and "grammatical" to refer to utterances without grammatical errors. We use "grammaticality" to refer to the property of an utterance defined by whether it is grammatical or ungrammatical.

McKee et al. (2017) describe the following apparent irony in the study of children's expressive language development: Advances in the complexity of spoken language are assumed to reflect growth in competence, whereas there is in fact evidence that performance factors impact children's language production in multiple ways. What a child produces does not always reflect their highest level of language knowledge. McKee et al. (2017) note that children understand grammatical morphemes before they consistently produce them. In addition, utterance length and complexity, which impact fluency (e.g., Buhr & Zebrowski, 2009), determine whether an inconsistently produced morpheme is realized in a specific utterance (see Charest & Johnston, 2011, for a review). Similarly, prosodic factors impact whether an emerging structure is produced within a particular utterance. For example, stressed, utterance-initial lexical nouns are more likely to be produced than unstressed, utterance-initial subject pronouns (Gerken, 1991; McGregor & Leonard, 1994).

We know of three reports exploring the relationship between grammaticality and disfluency, all focused on SLDs in CWS, without including TDs or children who do not stutter (CWNS). Using language samples from 15 CWS, Bernstein Ratner (1997) found that a higher proportion of ungrammatical utterances than error-free utterances contained SLDs. This suggested a relationship between grammaticality and SLD production but left several unanswered questions, as length and complexity were not controlled. More recently, Watson et al. (2011) assessed language samples from 11 monolingual Spanish-speaking CWS and found that the odds of an utterance containing an SLD was higher for ungrammatical utterances when controlling for age, utterance length in syllables, and utterance complexity. They concluded that syntactic complexity is an important contributor to stuttering and that their findings support the hypotheses that both stuttering and grammatical errors result from attempting to produce difficult structures. In contrast, Yaruss (1999) did not find a significant relationship between grammaticality and stuttering in the speech of 12 CWS.

Whether grammatical errors temporally follow disfluencies that are related to planning (stalls and SLDs) remains unaddressed. Because stalls and SLDs are associated with planning (e.g., MacWhinney & Osser, 1977; McDaniel et al., 2010; Wagovich et al., 2009), they would have to precede an omission or other error or occur at the beginning of a word containing an error to potentially be caused by difficulty planning that portion of the utterance. Of the studies we discussed (Bernstein Ratner, 1997; Watson et al., 2011; Yaruss, 1999), none considered the location of disfluency. We expect that stalls and SLDs will tend to occur before, or on the same word as errors, rather than after them. This prediction is based on the expectation that stalls and SLDs will occur where a child is planning the part of the utterance that is produced inaccurately.

Hypothesis 1: Differential Predictors of SLDs, Stalls, and Revisions

Findings that SLDs tend to occur early in utterances and in longer and more complex utterances have been taken to indicate that SLDs are related to planning, at least in children (both those who do and do not stutter) between 3 and 5 years of age (e.g., Buhr & Zebrowski, 2009; Yaruss, 1999). They have also been incorporated into the multifactorial dynamic pathways (MDP) model of stuttering, which proposes that SLDs occur when high loads in other domains destabilize the already vulnerable speech-motor systems of PWS (Smith & Weber, 2017). Some prior studies on TD production by young children

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

have considered all TDs together and found that TDs have similar relationships to language variables as SLDs (Buhr & Zebrowski, 2009; Gordon & Luper, 1989; Haynes & Hood, 1978; Yaruss et al., 1999; Zackheim & Conture, 2003). Other studies have separated stalls from revisions and found that stalls occur more often in longer and more complex utterances but that revisions do not (Rispoli, 2003; Rispoli et al., 2008; Wagovich et al., 2009). This background motivates us to ask whether SLDs and stalls differ from revisions in their relationships with length, complexity, grammaticality, and language level in the speech of young children who do and do not stutter.

We expect that SLDs and stalls will occur more often in longer and more complex utterances, and in utterances containing grammatical errors, and that they will be variable across different children but not show a clear relationship with a child's overall language level. In contrast, we predict that revisions will increase with a child's overall language level but will not show a relationship with length, complexity, or grammaticality. This would be the first effort to compare revisions to SLDs and stalls in matched groups of CWS and CWNS and assess all in the same (large) study. If SLDs and stalls, both of which are thought to serve a prospective purpose, relate to length, complexity, grammaticality, and language level in similar ways, but revisions, thought to serve a retrospective function, do not (Rispoli & Hadley, 2001), this would lend further support to SLDs reflecting some aspect of sentence planning.

Hypothesis 2: Grammaticality as a Predictor of SLDs and Stalls

Prior research with relatively small samples of 11 CWS (Watson et al., 2011) and 15 CWS (Bernstein Ratner, 1997) has suggested that CWS do produce more SLDs in ungrammatical utterances, but another study with 12 CWS did not find this relationship to be significant (Yaruss, 1999). Only one of these studies controlled for utterance length and complexity (Watson et al., 2011), and no study has looked at matched CWS and CWNS and both SLDs and TDs or looked at stalls and revisions separately. Doing so would provide a broader picture of how SLDs, stalls, and revisions fit into the leading edge theory (Rispoli & Hadley, 2001). The leading edge theory is certainly consistent with the findings that length and complexity are related to disfluency, particularly stall, production. However, length and complexity are clearly not the only indicators and, arguably, not the best structural indicators of what a child is able to produce easily. Whether the child can produce an utterance accurately or whether the child is attempting structures they have not yet mastered and thus produce ungrammatically (at least in that particular utterance) may be a better indicator of which utterances are in the child's comfort zone or leading edge. If SLDs and stalls occur more frequently in utterances containing grammatical errors, this would allow for an expansion of the leading edge theory. It would suggest that grammatical utterances generally occur within the comfort zone, and grammatical errors generally occur at the leading edge.

The MDP model (Smith & Weber, 2017) does not currently include grammaticality as a linguistic factor that may destabilize the motor system, leading to stuttering. Finding that grammaticality relates to SLD production would contribute to our theoretical understanding of stuttering by telling us that structures known to be hard for a given child (indicated by an error) are where SLDs occur. Further, if SLDs occur before the error, this would be additional support that planning of difficult structures may contribute to SLD production (Watson et al., 2011).

This background motivates the question behind our second hypothesis, which asks how the grammaticality of an utterance produced by a young child who stutters or child who does not stutter relates to the likelihood that it contains a stall or SLD. We expect that utterances containing grammatical errors will be more likely than grammatically correct utterances to contain stalls and SLDs. We make this prediction in part based on reports from Bernstein Ratner (1997) and Watson et al. (2011; though they contrast with Yaruss' [1999] results). We extend this expectation to stall-containing utterances as well.

Hypothesis 3: SLDs and Stalls Preceding Grammatical Errors

The question underlying our third hypothesis asks whether SLDs and stalls tend to precede grammatical errors. We hypothesize that stalls and SLDs in utterances containing grammatical errors will tend to occur before and not after the error. Because children can plan more than one word at a time (MacWhinney & Osser, 1977), "before" does not necessarily mean "immediately preceding." If disfluencies tend to occur before errors, this would strengthen any conclusions drawn about the role of planning difficult (i.e., error-containing) upcoming portions of the utterance in SLD and stall production.

Method

Participants

Corpora and Matching

We analyzed language samples from 32 CWS and 32 matched CWNS who enrolled in one of three previous or ongoing studies on childhood stuttering and/or language development. Seventeen CWS are from the Ratner corpus (Wagovich & Bernstein Ratner, 2007). Some of the Ratner corpus CWS enrolled in this study were part of

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

the Bernstein Ratner (1997) study. Fifteen CWS are from the first year of the University of Maryland (UMD)–Carnegie Mellon University (CMU) corpus (Bernstein Ratner & MacWhinney, 2018). Original data were collected using institutional review board (IRB)–approved procedures. The secondary data analysis involved in this study was determined exempt by UMD's IRB.

All CWS were matched to CWNS on age (within 4 months) and gender and as closely as possible on maternal education. UMD–CMU CWS were matched to CWNS from the same corpus during the recruitment process. Ratner corpus CWS were matched by Luckman et al. (2020), either to CWNS from the Ratner corpus (10 pairs) or to CWNS from the Weismer corpus (seven pairs; Ellis Weismer et al., 2013).

Twelve participants were female, and 52 participants were male. They were monolingual English speakers ranging from 28 to 50 months old. Maternal education levels were available for UMD–CMU participants. For these CWS, 14 mothers had college degrees or higher and one mother had some college courses. For these CWNS, all 15 mothers had college degrees or higher.

Diagnosis of Stuttering

All CWS were diagnosed by a speech-language pathologist from the original cadre of American Speech-Language-Hearing Association Board Recognized Specialists in Fluency Disorders. Ratner corpus CWS were tested within 4 months of stuttering onset, and UMD–CMU CWS were tested within 1 year of stuttering onset.

Testing for Typical Language Development

We required that each participant had language skills falling within normal developmental ranges, which was operationalized as having a majority of their available standardized test scores and MLU-m falling 1.5 *SD*s below the mean or higher.[2] This guideline was generally based on Tomblin et al.'s (1996) recommendation that kindergartners score at least 1.25 *SD*s below the mean on two of five composite scores for diagnosis of language disorders.

[2]Different sets of measures were available for each of the three corpora. Available tests were the Peabody Picture Vocabulary Test–Fourth Edition (L. M. Dunn & Dunn, 2007), the Clinical Evaluation of Language Fundamentals: Preschool–2 (CELF-P2) Sentence Structure subtest, and the CELF-P2 Concepts and Following Directions subtest (Semel et al., 2004) for the UMD–CMU corpus; some combination of the Peabody Picture Vocabulary Test–Revised (L. M. Dunn & Dunn, 1981), the Expressive One-Word Picture Vocabulary Test–Revised (Gardner, 1990), the Clinical Evaluation of Language Fundamentals: Preschool (CELF-P) Linguistic Concepts subtest, and CELF-P Word Structure subtest (Wiig et al., 1992) for the Ratner corpus; and the MacArthur Communicative Development Inventory–Words and Sentences (Fenson et al., 1993), Preschool Language Scale–Third Edition (PLS-3) Auditory Comprehension subtest, and the PLS-3 Expressive Language subtest (Zimmerman et al., 1992) for the Weismer corpus. MLU-m, based on all fully intelligible utterances, was compared to norms from the CHILDES database (Bernstein Ratner et al., 2020).

Language Sample Collection and Preparation

Language samples were collected in similar play-based contexts in the three original studies, with participants playing with examiners and/or caregivers. While seven of the CWS–CWNS pairs contained children from two different original studies, there is no reason to believe that the toys provided would have impacted the relationships between language and fluency. Transcripts used CHAT format, for analysis with CLAN language sample analysis software (MacWhinney, 2000).

Disfluency Coding

Disfluencies were coded according to CHAT's fluency coding conventions (Bernstein Ratner et al., 2020). Part-word repetitions, prolongations, blocks, broken words, and monosyllabic word repetitions were considered SLDs. Fillers "um" and "uh," word fragments, multisyllabic word repetitions, multisyllabic word revisions, phrase repetitions, and phrase revisions were considered TDs.

In order to avoid overmarking disfluencies when a child was intentionally saying a word more than once, the following was done. When there was a repetition of (a) a name; (b) "look," "hey," or "oh"; and (c) or "yes," "no," or another agreement/disagreement marker, or when an adjective was said multiple times for emphasis, and coders perceived the repetition to be markedly intentional, they omitted the repetition marker. This resulted in these utterances being classified as fluent unless they contained disfluencies elsewhere. This affected 99 of the 16,328 fully intelligible utterances.

Segmentation Procedures

The *two out of three criteria* rule (Bernstein Ratner et al., 2020) was used in segmenting utterances. Utterance boundaries were placed when at least two of the following occurred: (a) a perceptible pause, (b) terminal intonation contour, and (c) complete grammatical structure. Preposed elements such as "yes" were attached to the rest of the utterance unless at least two out of three criteria were met. Following Rispoli (2003), no more than two independent clauses could be connected by a coordinating conjunction within one utterance. After this, an utterance boundary would be placed and the next independent clause would begin a new utterance.

Consensus Reliability Procedures

This study used consensus reliability procedures similar to those used in previous studies looking at disfluency

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

using language sample analysis (Rispoli, 2003; Watson et al., 2011). These focused on morpheme-by-morpheme transcription, utterance segmentation, and fluency coding.

Because over 25 people were involved in creating the initial transcripts from the three original studies, the first author reviewed and corrected all transcripts to complete the first pass. Then, research assistants (RAs) who were not involved in the original transcription, and who had completed training and passed coding tests, completed the second passes. In the second pass, RAs marked any disagreements with the first pass transcription. So that RAs would not defer to the first author's coding, the identity of the first pass coder was not disclosed. In consensus meetings, the second pass RA directed a consensus rater (another RA) to these marked utterances. With this process, 0.3% of utterances were marked and discussed for segmentation, as were 1.8% for fluency coding and 2.4% for morpheme-by-morpheme transcription. Coders reached consensus for 100% of utterances.

Available Data Summary

Fully and partially unintelligible utterances were excluded from all analyses. There were 16,328 fully intelligible utterances, with a mean of 255.1 per participant (range: 41–514). The utterances included in the analyses addressing each hypothesis will be later described.

Variable Coding

Length

Length was computed using MOR, the automatic parser in CLAN. MOR is a program available within CLAN, and MOR for English has been evaluated as 95% accurate for part-of-speech labeling and more highly accurate for identifying segments within utterances (MacWhinney et al., 2020).

Developmental Sentence Score Complexity Level

Developmental Sentence Score (DSS) was chosen as a measure of syntactic complexity because it has been used as the complexity in previous studies finding relationships between more syntactically complex utterances and stuttering. DSS examines 50 utterances that contain a noun and verb in a subject–predicate relationship and determines which of a set of score-weighted language structures are included in each utterance (Lee, 1974). Under traditional DSS guidelines, error-free utterances receive an additional "sentence point." The sample's DSS score is the mean score for the 50 utterances. Because grammaticality was its own predictor in this study, our *DSS complexity* score was defined as the DSS level for the utterance, without the sentence point.

Language Level

The child's overall language level was represented by our MLU-m-alternative (MLU-m-a). The utterances used in computation of MLU-m-a were those included in the full analyses addressing Hypotheses 1 and 2. We eliminated utterances not considered to reflect the child's ability to spontaneously generate language. Similar to prior works by Rispoli (2003), Hollister et al. (2017), and Watson et al. (2011), we excluded immediate self-repetitions, exact repetitions of the immediately preceding adult utterance, isolated filler words or word fragments, utterances that trailed off or were interrupted, and rote utterances (e.g., singing or counting).

We excluded additional utterances that might have inflated or deflated the impression of a child's language ability. These were (a) single-word responses to questions; (b) utterances consisting of only "yes," "no," or another word meaning "yes" or "no" (e.g., "sure"); and (c) following Rispoli et al. (2008) and Wagovich et al. (2009), utterances consisting of only "I don't know." After excluding these utterances, 11,142 utterances remained (for individual participants, M = 174.1, range: 41–343).

MLU-m-a was different from the child's overall MLU in morphemes (MLU-m) that was used to determine eligibility. The overall MLU-m used for eligibility determination was based on all fully intelligible utterances so that it could be compared to available norms. Derived MLU-m-a was, on average, 0.8 morphemes higher than MLU-m.

Grammaticality

The "intended utterance" (utterance after removing disfluencies) was coded for grammaticality. For children who spoke mainstream American English (MAE), grammaticality coding followed a guide based on the study of Eisenberg et al. (2012). For the three CWS and two CWNS who produced both MAE and African American English (AAE), grammaticality was based on the works of Oetting and McDonald (2001) and Oetting and Pruitt (2005).

Coding Reliability for Length, DSS Complexity, and Grammaticality

For MLU-m and MLU-m-a to be reliable measures, morpheme-by-morpheme transcription and utterance boundary placement had to be reliable. These were all checked through the consensus process. MLU-m-a also required accurate and reliable coding of utterances for exclusion. Of the fully intelligible, excluded utterances, 81.5% were excluded automatically through the use of functions available in CLAN. The other 18.5% were excluded through hand-coding by RAs. The first author checked the first five files coded by each research assistant, and all subsequent files were checked by a second RA. DSS complexity was computed automatically by CLAN.

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 1. Examples of disfluency-error order coding for an MAE-speaking child.

Utterance	Disfluency-error order coding
That that that _ for me.	Disfluency first or on same word as error
Those babies *i-i-is* mine.	Disfluency first or on same word as error
I _ gonna put that *in-inside.*	Error first
Her *d-d-doesn't* want two doll_.	Excluded: One error is before the disfluency and one is after.

Note. Disfluencies are italicized. Error locations are indicated by underlining or an underscore. MAE = Mainstream American English.

Trained RAs hand-coded grammaticality and then completed consensus reliability procedures for this variable. These RAs passed tests of grammaticality coding for samples from participants who spoke MAE only. RAs who worked on grammaticality coding for participants who spoke in AAE and MAE received additional training and also passed tests of grammaticality coding for bidialectal speakers. The first pass involved RAs coding the transcript for grammaticality. In the second pass, the first author reviewed the transcript and marked any utterances for which she disagreed with the first pass. Lastly, a third coder (not given information about coder identity) reviewed these marked utterances. Through this process, a consensus decision was made for 100% of utterances.

Disfluency-Error Order Coding

For any disfluent utterances containing at least one SLD or stall (the disfluencies expected a priori to relate to planning), the order of the SLD(s)/stall(s) and grammatical error(s) was coded. SLDs or stalls must occur on or before an error for it to be possible that planning the error portion of the utterance is a contributing factor to production of the SLD or stall. Therefore, disfluencies occurring before or on the same word as the error were grouped together in the analysis. Errors of omission were considered to occur on the word following the omission. Utterances were excluded from this analysis if (a) the location of the error could not be definitively determined to be before or after the SLD or stall, such as an utterance with incorrect word order or one that was clearly ungrammatical but where the intended utterance was unclear (e.g., "same i-i-ice cream I got"); (b) there was more than one SLD/stall and at least one occurred before the error and at least one occurred after; or (c) there was more than one error and at least one occurred before the SLD(s)/stalls(s) and at least one occurred after. See Table 1 for examples of order coding.

Disfluency-Error Order Reliability

Two separate coders were used. There were 17 disagreements out of 718 qualifying utterances (2.3%). A third coder, not given information about coder identity, reviewed utterances with these disagreements. Consensus was reached for all utterances through this process.

Analysis and Results

Hypotheses 1 and 2: Analysis Plan

All statistical models were run in R (Version 3.6.1; R Core Team, 2019), using the lme4 package (Version 1.1–21; Bates et al., 2015). Hypotheses 1 and 2 were addressed using the same models. We were interested in the following factors as they related to the likelihood of disfluency: grammaticality, complexity, length, and the child's expressive language level. Planned logistic mixed-effects models included fixed effects (a binary grammaticality value, cluster-mean centered DSS complexity,[3] cluster-mean centered length in morphemes, grand-mean centered MLU-m-a) and a random effect of participant to account for the nesting of utterances within participants. Predicted binary outcomes, across the three models, were (a) whether the utterance contained at least one SLD or not, (b) whether the utterance contained at least one stall or not, and (c) whether the utterance contained at least one revision or not. Positive regression coefficients, therefore, would reflect increasing likelihood of disfluency with an increase in the value of the predictor. There were 11,142 fully intelligible utterances that met the inclusion criteria for models addressing Hypotheses 1 and 2 as outlined previously. α was set to .017 because of the three models for the three disfluency types. See Table 2 for descriptive information about the utterances and predictors.

We used DSS complexity as our measure of complexity, as this is a measure that can be computed at the utterance level and because DSS has been used in previous studies finding associations between utterance complexity and disfluency. One limitation of this approach is that only 5,641 of the 11,142 utterances otherwise available for these analyses were DSS-eligible. Two sets of models were run on the 5,641 DSS-eligible utterances. In the first set, DSS complexity was included, and in the second set, it was omitted ("set of models" here refers to three models with the three disfluency types as outcomes). Results of likelihood ratio tests indicated that DSS complexity did not significantly improve model fit for the SLD model,

[3]Clustering occurred at the level of the participant.

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 2. Length and grammaticality of fluent utterances and utterances containing three disfluency types.

Utterance type	*n*	% of Utts.	Length *M (SD)*	% Ungram.
CWS				
Fluent	3,554	66.8	3.4 (2.0)	16.7
SLD-containing	1,332	25.0	5.0 (2.5)	34.7
Stall-containing	538	10.1	4.8 (2.6)	36.1
Revision-containing	312	5.9	5.5 (2.7)	38.8
Total	5,322	100	3.9 (2.3)	22.2
CWNS				
Fluent	4,852	83.3	4.0 (2.3)	17.8
SLD-containing	410	7.0	5.6 (2.9)	25.4
Stall-containing	365	6.3	5.2 (2.8)	21.4
Revision-containing	346	5.9	5.5 (2.8)	28.0
Total	5,820	100	4.3 (2.4)	18.8
All participants				
Fluent	8,406	75.4	3.8 (2.2)	17.3
SLD-containing	1,742	15.6	5.1 (2.6)	32.5
Stall-containing	903	8.1	5.0 (2.7)	30.1
Revision-containing	658	5.9	5.5 (2.7)	33.1
Total	11,142	100	4.1 (2.4)	20.4

Note. Some utterances are represented more than once in this table because they contained more than one type of disfluency. Therefore, percentages do not add up to 100%. % of Utts. = percentage of utterances produced by CWS, CWNS, or all participants that were fluent or contained each disfluency type; % Ungram. = % ungrammatical; CWS = children who stutter; CWNS = children who do not stutter; SLD = stuttering-like disfluency.

$\chi^2(1) = 0.01$, nonsignificant (ns); stall model, $\chi^2(1) = 1.16$, ns; or revision model, $\chi^2(1) = 0.11$, ns. Because inclusion of DSS complexity did not improve model fit and its inclusion greatly reduced the data set, DSS complexity was omitted and all 11,142 utterances were included in the final models.

All stalls were considered as one group in the analyses that follow, as were all revisions. For descriptive purposes, the types of stalls and revisions represented in the data set are shown in Tables 3 and 4, respectively.

Hypotheses 1 and 2: Results

Ungrammatical utterances had higher odds of containing all three disfluency types, when holding utterance length and MLU-m-a constant ($z = 8.34$, $p < .001$ for SLDs; $z = 5.06$, $p < .001$ for stalls; and $z = 5.97$, $p < .001$ for revisions). The presence of a grammatical error was associated with a 1.76-times increase in the odds it would contain at least one SLD, a 1.53-times increase in the odds it would contain at least one stall, and a 1.73-times increase in the odds it would contain at least one revision, all when holding length and MLU-m-a constant. Similarly, increasing length increased the odds that utterances would contain each disfluency type, when holding grammaticality and MLU-m-a constant ($z = 19.88$, $p < .001$ for SLDs; $z = 8.67$, $p < .001$ for stalls; and $z = 12.98$, $p < .001$ for revisions). An increase of one morpheme in the length of an utterance was associated with a 1.29-times increase in the odds that it would contain at least one SLD, a 1.14-times increase in the odds it would contain at least one stall, and a 1.23-times increase in the odds it would contain at least one revision, all when controlling

Table 3. Types of stalls in stall-containing utterances.

Stall type	*n*	% of stalls
Filler(s) only	597	66.1
Phrase repetition(s) only	246	27.2
Multisyllabic word repetition(s) only	21	2.3
Phrase repetition(s) & filler(s)	33	3.7
Multisyllabic word repetition(s) & filler(s)	3	0.3
Multisyllabic word repetition(s) & phrase repetition(s)	3	0.3
Total	903	100

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Table 4. Types of revisions in revision-containing utterances.

Revision type	*n*	% of revisions
Phrase revision(s) only	214	32.5
Word revision(s) only	205	31.2
Fragment(s) only	128	19.5
Phrase revision(s) & fragment(s)	78	11.9
Word revision(s) & phrase revision(s)	24	3.6
Word revision(s) & fragment(s)	5	0.8
Word revision(s), phrase revision(s) & fragment(s)	4	0.6
Total	658	100

Note. Utterances in which the phrase being replaced (the reparandum) contained a fragment, such as, "the chair is too s- the chair is too big" were classified under "phrase revision(s) & fragment(s)" or under "word revision(s), phrase revision(s) & fragment(s)" if they also contained at least one word revision.

for grammaticality and MLU-m-a. MLU-m-a was unrelated to the odds that an utterance would contain at least one SLD, when holding utterance length and grammaticality constant ($z = 0.46$, ns). Utterances produced by participants with higher MLU-m-as had higher odds of containing both stalls and revisions, when holding utterance length and grammaticality constant ($z = 3.41$, $p < .001$ for stalls and $z = 3.09$, $p = .002$ for revisions). An increase in one morpheme in a child's MLU-m-a was associated with a 1.58-times increase in the odds an utterance that child produced would contain at least one stall and a 1.29-times increase in the odds an utterance they produced would contain at least one revision, when controlling for utterance grammaticality and length. See Table 5 for model output and odds ratios (*ORs*) for the models used to address Hypotheses 1 and 2.

Table 5. Results from mixed-effects models of predictors of disfluency for three disfluency types.

Model and effect	*B*	*OR*	*z*	*p*
SLD				
Fixed effects				
Intercept	–2.35	0.10	–16.06	< .001
Grammaticality	0.57	1.76	8.34	< .001
Length	0.26	1.29	19.88	< .001
MLU-m-a	0.08	1.09	0.46	.646
Random effect	*Var.*			
Participant	1.21			
Stall				
Fixed effects				
Intercept	–2.90	0.06	–26.61	< .001
Grammaticality	0.43	1.53	5.06	< .001
Length	0.13	1.14	8.67	< .001
MLU-m-a	0.46	1.58	3.41	< .001
Random effect	*Var.*			
Participant	0.55			
Revision				
Fixed effects				
Intercept	–3.10	0.04	–42.67	< .001
Grammaticality	0.55	1.73	5.97	< .001
Length	0.21	1.23	12.98	< .001
MLU-m-a	0.26	1.29	3.09	.002
Random effect	*Var.*			
Participant	0.13			

Note. Outcomes coded as 0 = no SLD in utterance, 1 = SLD in utterance, 0 = no stall in utterance, 1 = stall in utterance, and 0 = no revision in utterance, 1 = revision in utterance. Grammaticality coded as 0 = grammatical, 1 = ungrammatical. Number of participants = 64. Number of utterances = 11,142. *OR* = odds ratio; SLD = stuttering-like disfluency; MLU-m = mean length of utterance in morphemes; MLU-m-a = MLU-m-alternative; Var. = variance.

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

A series of post hoc tests were conducted to explore the role of group in the results that we found, particularly because some prior studies that reported results contrasting with some of the findings here have only enrolled CWNS (Rispoli, 2003; Rispoli et al., 2008) or CWS (Wagovich et al., 2009). Thus, assessing these relationships within each group would enable closer comparisons to this previous work.

In order to determine whether the factors that predicted SLD, stall, and revision production for all participants together were also significant predictors of the disfluency types for each group, models were run on utterances from CWS and CWNS separately. It was possible that production of disfluencies, particularly SLDs, may have been influenced by different factors across groups. For these post hoc tests, α was set to .017 due to the three models run for each group.

Results of the three models run on utterances produced by CWS looked very similar to those reported for the full set of participants. As with the models run using utterances from all participants, ungrammatical and longer utterances had higher odds of containing SLDs (z = 7.70, $p < .001$ for grammaticality and z = 17.83, $p < .001$ for length), stalls (z = 4.82, $p < .001$ for grammaticality and z = 6.59, $p < .001$ for length), and revisions (z = 5.03, $p < .001$ for grammaticality and z = 9.19, $p < .001$ for length), holding all other factors constant. As with the models run using utterances from all participants, utterances produced by CWS with higher MLU-m-as had higher odds of containing stalls and revisions, holding utterance length and grammaticality constant (z = 3.00, p = .003 for stalls, and z = 3.35, $p < .001$ for revisions).

Results from the models run only on utterances produced by CWNS differed in a few ways from those run on utterances from both groups. First, utterances produced by CWNS with higher MLU-m-as had higher odds of containing SLDs, holding utterance length and grammaticality constant (z = 2.78, p = .005). A one-morpheme increase in the MLU-m-a of a child who does not stutter was associated with a 1.47-times increase in the odds that an utterance they produced would contain an SLD, holding grammaticality and length constant. Second, grammaticality was not a significant predictor of whether an utterance produced by a child who does not stutter would contain a stall, holding utterance length and MLU-m-a constant (z = 1.80, p = .072). If an unadjusted α of .05 had been used, there would have been a trend toward an effect in which ungrammatical utterances produced by CWNS had higher odds of containing stalls, controlling for grammaticality and MLU-m-a. Ungrammatical utterances produced by CWNS had 1.29-times higher odds of containing stalls compared to grammatical utterances. This contrasts with the 1.68-times increase for CWS. Third, for utterances produced by CWNS, an increase in MLU-m-a was not associated with a significant increase in the odds that it would contain a revision, controlling for grammaticality and length (z = 1.03, ns).

Finally, group was added to the three original models to determine whether the odds of each disfluency type occurring differed by group when controlling for all other predictors. When controlling for grammaticality, length, and MLU-m-a, utterances produced by CWS (unsurprisingly) had greater odds of containing SLDs (z = 8.59, $p < .001$) and stalls[4] (z = 2.71, p = .007), but there was no significant difference between groups in the odds of revisions occurring (z = 0.06, ns). If an utterance was produced by a child who stutters rather than a child who does not stutter, there was a 5.61-times increase in the odds that it would contain an SLD and a 1.73-times increase in the odds that it would contain a stall, holding grammaticality, length, and MLU-m-a constant.

Hypothesis 3: Analysis Plan

To determine whether SLDs and stalls tend to precede grammatical errors, the 718 utterances that were eligible for analyses addressing Hypotheses 1 and 2 and contained at least one SLD or stall and a grammatical error were identified. Then, 106 of these were excluded because there was no clear error–disfluency order. This left 612 utterances remaining, across 59 participants, 551 having the disfluency first on the same word as the error and 61 having the error first. A logistic mixed-effects model was run, predicting order (disfluency first or on the same word as the error vs. error first) with a random intercept for participant and no fixed effects.

Hypothesis 3: Results

There was a significant intercept in the model, reflecting greater odds of the stall(s) and/or SLD(s) occurring before or on the same word as the error than of the error occurring before the stall(s) and/or SLD(s) (z = 9.92, $p < .001$, OR = 11.57).

Hypothesis 3: Additional Post Hoc Analysis

Disfluencies tend to occur in the utterance-initial position (e.g., Buhr & Zebrowski, 2009), so it was possible that the significant effect was driven by this tendency.

[4]While a complete discussion of this topic is beyond the scope of this article, we encourage readers and researchers to consider the role of stalls as postponement behaviors. Stalls serving to postpone SLDs may have impacted the rate of stall production in our sample and may play a role in samples collected as part of future research projects.

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Therefore, an additional analysis was run only on the 263 ungrammatical utterances across 51 participants that had at least one non-initial stall or SLD. The model run on this limited set of utterances also had a significant intercept (z = 5.05, p < .001, OR = 3.81), reflecting greater odds of the stall(s) and/or SLD(s) occurring before or on the same word as the error(s) than of the error(s) occurring before the stall(s) and/or SLD(s).

Discussion

Predictors of SLD Production

First, we expected that ungrammatical and longer utterances would be more likely to contain SLDs, but we did not expect an association between the language level of the child who produced the utterance and whether their utterances contained an SLD. These predictions were supported; such results are consistent with a large body of previous research showing that longer utterances are more likely to contain SLDs (e.g., Richels et al., 2010; Yaruss, 1999; Zackheim & Conture, 2003) as well as with two of three previous studies assessing the role of grammaticality in SLD production (Bernstein Ratner, 1997; Watson et al., 2011). These findings are also consistent with the long-held proposal that SLDs occur in part due to language planning demands (e.g., Hall et al., 2007). The lack of an association between language level and SLDs, controlling for length and grammaticality, is not definitive evidence of no association. The apparent absence of an association between the number of SLDs and language level is consistent with the conception of disfluency as highly variable (e.g., Yaruss, 2004).

Predictors of Stall Production

Next, we expected that stalls, like SLDs, would be more likely to occur in ungrammatical and longer utterances because of previous suggestions that stalls are related to planning difficulty (Rispoli, 2003; Rispoli et al., 2008; Wagovich et al., 2009). These predictions were supported. Because of prior indications that a "glitch" in sentence planning is all that is needed for a child to produce a stall—and because of previous studies failing to find an association (Rispoli, 2003; Rispoli et al., 2008; Wagovich et al., 2009), we did not expect a relationship between language level and stall production. Contrary to predictions, more advanced expressive language was associated with greater odds of stall production, holding grammaticality and utterance length constant. It was not simply that children with more advanced language development attempted longer utterances and that this drove increased stall production. Because length and grammaticality were controlled in these models, the result means that if two children produced utterances of the same length, and the utterances were both grammatical (or ungrammatical), the utterance from the child with higher mean utterance length was more likely to contain a stall. For additional context, as shown in Table 3, 66.1% of the stall-containing utterances only had a filler or fillers, and 27.2% of the stall-containing utterances only contained phrase repetition and no other stalls. Therefore, looking at filler development may be useful in explaining this finding.

There are several possible explanations for the increase in stall (particularly filler) behavior for children with more advanced language development. One possibility concerns the child's ability to monitor for upcoming pauses. In an influential study, Clark and Fox Tree (2002) found that adults tended to produce "um" before longer pauses (where syntactic planning is thought to occur more often) and "uh" before shorter pauses (where lexical access is thought to occur more often). Hudson Kam and Edwards (2008) describe the production of "um" and "uh" as "deceptively complicated," (p. 315) as this requires a child to anticipate a delay. They found that 3- and 4-year-olds produced "um" and "uh" in ways that were close to adultlike patterns (before pauses) but that they did not yet differentiate between the two fillers by placing "um" before longer pauses and "uh" before shorter pauses. Children with more advanced language might be better at monitoring their speech for upcoming delays and therefore be better at knowing when to insert a filler. This could be tested by assessing whether children with less advanced expressive language use more silent pauses without preceding fillers.

A second possibility concerns the insertion of the filler into the sentence plan. Fraundorf and Watson (2014) asked adults to retell three stories while viewing bulleted lists of major plot points. They hypothesized that fillers would tend to be used at the beginnings of new messages, where the speaker would not be required to replan in order to put a filler into the utterance. They found evidence supporting this hypothesis; participants tended to use fillers when encountering message-level, rather than grammatical or phonological-level, difficulties. If children with more advanced language are better able to insert fillers into the sentence plan, this could partially explain the increase in stalls with language level.

SLDs and Stalls and Disfluency-Error Order

SLDs and stalls had much greater odds (11.57 times greater) of preceding errors or occurring on the same word as the error than of following errors. Disfluencies tend to be utterance-initial (e.g., Buhr & Zebrowski, 2009); this could have fully or partially accounted for the

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

result. Thus, utterances with only utterance-initial SLDs or stalls were removed for a second model. The significant result remained, with the *OR* reduced to 3.81. The overall effect, therefore, appears to be partially, but not fully, accounted for by the tendency of disfluencies to be utterance-initial. These findings help to understand the effect of errors on the types of disfluencies that were expected, a priori, to be associated with planning demands (SLDs and stalls). As far as we know, this was the first analysis of the relative ordering of disfluencies and grammatical errors within utterances. It strengthens the interpretation that planning of the error-containing portion of the utterance could be influencing SLD and stall production.

Predictors of Revision Production

We expected that revisions would increase with language level but be unrelated to factors measuring utterance-level demand (grammaticality and length). Revisions did increase with MLU-m-a, and this may be attributable to one or more of the skills Rispoli (2003) discussed: language comprehension or monitoring, the ability to stop oneself from speaking and make a quick change, and/or having sufficient alternative words or structures to substitute. Future experimental work measuring attention to a child's own output, their comprehension skills, and their error-monitoring skills is needed in order to better understand these relationships.

This study also produced the unexpected results that both ungrammatical utterances and longer utterances had higher likelihoods of containing revisions. Revisions were varied across fragments, word revisions, phrase revisions, and combinations of these, with no single category making up more than a third of the revision-containing utterances. Possible explanations are that (a) more-difficult-to-plan utterances are more likely to contain production slips (not necessarily grammatical errors) that the child might want to revise or (b) production slips that are produced equally across utterances are more likely to be noticed and corrected if the utterance is longer and/or ungrammatical. The first of these possibilities seems more probable, but it is not possible to differentiate between these based on the current data.

No Difference Between Language Factors Predicting Stalls and Revisions

Previous research has suggested a dichotomy between the linguistic factors associated with stall production and those associated with revision production (Rispoli, 2003; Rispoli et al., 2008; Wagovich et al., 2009). However, the current data indicated that both stalls and revisions were associated with planning-related factors (length and complexity) and the child's language level (MLU-m-a). Language level was particularly surprising with regard to how it impacted stalls compared to revisions. While MLU-m-a was a significant predictor of both stalls and revisions, the *OR* was higher for stalls; a one-morpheme increase in MLU-m-a was associated with a 1.58-times increase in the odds of at least one stall occurring and a 1.29-times increase in the odds of at least one revision occurring. The higher *OR* for stalls than for revisions was surprising, since no relationship was expected for stalls and a relationship was expected for revisions. The current data, therefore, suggest that there should be a reconsideration about whether the dichotomy in predictors of stalls and revisions generalizes to a broad range of utterance types (compared to the analysis only of active declarative sentences by Rispoli, 2003; Rispoli et al., 2008; and Wagovich et al., 2009).

Theoretical Implications

According to the leading edge theory, children develop a comfort zone of utterances that they can typically produce fluently. When they attempt more recently acquired or advanced utterances, children are producing utterances at their leading edge and are more likely to be disfluent (Rispoli & Hadley, 2001). Rispoli and Hadley (2001) measured utterance length and VP complexity and considered longer utterances or those with more complex VPs to be at the child's leading edge. Current results suggest that grammaticality should be added to the list of considerations. Grammatical utterances are more likely to be within the comfort zone of utterances and produced without disfluencies, whereas ungrammatical utterances that a child attempts are more likely to be in the disfluent leading edge.

Current findings also suggest that children have some level of knowledge that their ungrammatical utterances are ungrammatical. If children were completely unaware at all levels of language production that their utterances did not match the adultlike target, then there would not be a mechanism by which ungrammatical utterances would have increased disfluency. The current findings provide another example of why children's expressive language should not be considered to reflect their highest level of competence.

The final set of theoretical implications pertains to the MDP (Smith & Weber, 2017). The MDP makes a key connection between linguistic and motoric factors by incorporating research findings showing decreased motor stability in the speech of PWS as length and complexity increase (Kleinow & Smith, 2000; MacPherson & Smith, 2013; Maner et al., 2000; Usler & Walsh, 2018). If current findings are replicated, grammaticality should be considered for incorporation into the MDP as a factor representing language demand alongside length and complexity.

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Current findings that length continues to be a predictor of disfluency, even when controlling for grammaticality and the child's language level, support the MDP's current incorporation of linguistic factors as impacting SLD production.

A question arises about the mechanism by which grammatical errors become associated with SLDs. One possibility is that ungrammatical utterances pose production difficulty for CWS, resulting in lower motor stability in utterances with grammatical errors. Because of the magnitude of the association between grammaticality and SLD production (for CWS, a 1.89-times increase in the odds of SLD production for ungrammatical utterances), better understanding the mechanism underlying this relationship seems particularly important to understanding how language factors relate to motor stability.

The behaviors that we detail in this report can be seen as consistent with at least one theoretical account of the ontogenesis of stuttering. There are few theories about either stuttering or language production that are developmental in nature, as opposed to accounts of full-fledged language use and monitoring. A rare exception is the vicious circle/vicious cycle hypothesis (Bernstein Ratner & Wijnen, 2007; Vasiç & Wijnen, 2005). In this model, precocious development of self-monitoring skills leads some children to be atypically attuned to errors (at any level of execution) in their speech output and attempt repairs. This hypothesis posits a specific asynchrony impacting fluency and repair in which most children are typically insensitive to errors in their own production (as has been repeatedly demonstrated in early development); disfluency in early production, whether stuttered or more typical in nature, is a marker of the development of self-monitoring skills. Further asynchronous development in profiles will then distinguish between those children with better or worse linguistic encoding skills in the face of superior detection of such errors, leading to subgroups of children with what used to be called developmental disfluency and children whose repeated attempts to repair errored output lead to increasing levels of tension and frustration.

Clinical Implications

In discussing clinical implications, the most relevant statistical model is the one looking at which factors are related to SLD production for CWS; this model indicated that longer utterances and ungrammatical utterances produced by CWS were more likely to contain SLDs, with *OR*s of 1.89 for grammaticality and 1.36 for length. Because of the different ways that length and grammaticality are measured, it is not possible to say that grammaticality is necessarily a stronger predictor of whether a child who stutters will stutter on an utterance than length is. It is, however, possible to say that when a child attempts an utterance they produce with an error, this has a greater impact on the likelihood that it will be stuttered than increasing the length of the utterance by one morpheme does. The increase in odds of stuttering with a two-morpheme increase in length is similar to the increase associated with an utterance being ungrammatical rather than grammatical.

The two previous studies identifying associations between grammaticality and stuttering and the current one have controlled for different factors and used different statistical approaches. Two have enrolled monolingual English-speaking CWS (Bernstein Ratner, 1997, and this study), whereas one enrolled monolingual Spanish-speaking CWS (Watson et al., 2011). The diversity in methods and approaches suggests that there may be a robust relationship between grammaticality and stuttering, though additional research on interactions between language and fluency should continue to measure grammaticality (and the contrasting finding of Yaruss, 1999, should also be considered). The clinical applications outlined below may be considered based on the current state of the evidence and would be strengthened by replication of these findings.

The finding that grammatical errors are related to disfluency production, and, most importantly here, SLD production, underscores the need for thorough language assessments in evaluations for young children seen for concerns about stuttering. Comprehensive evaluations have long been recommended, given that we have known about associations between linguistic factors and SLDs for decades and that the prevalence of language disorders appears to be higher among CWS than in the general population (Hall et al., 2007). With recent findings indicating that language growth is associated with recovery from stuttering (Hollister et al., 2017; Leech et al., 2017, 2019), this recommendation appears even more critical. If language skills are weak for a child who stutters, then language development must be supported. Further, if a speech-language pathologist wants to know what aspects of a child's expressive language are weak, they have to conduct an in-depth language assessment. Because each child's leading edge, and which utterances are likely to be ungrammatical, change as their language develops, frequent language sampling or other monitoring of expressive language skills is required for clinicians to have up-to-date knowledge about the child's leading edge. Finally, dynamic assessment may be a valuable tool as it can provide information about whether structures unattested in language sampling data can be produced grammatically by a child who stutters.

Another clinical implication of the association between ungrammatical utterances and SLDs is that speech-language pathologists should consider incorporating

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

grammaticality into treatment hierarchies for preschool-age CWS when modeling and teaching new speech skills. After their thorough language evaluations, clinicians would have information about which structures a child is more or less likely to produce grammatically. If following previous findings, clinicians may currently be increasing length and complexity as CWS show improved performance of a new skill such as easy, relaxed speech (e.g., Richels & Conture, 2007). If the grammaticality finding were incorporated, then a hierarchy may involve first working on structures that CWS can produce grammatically before moving to utterances with a higher likelihood of grammatical errors. This can be done by modeling or through a structured elicitation task; with both of these methods, a clinician who has a thorough understanding of the child's expressive language skills can better design activities when teaching any fluency skills, whether they involve fluency shaping or modification of stuttered events: Both impose yet additional demand on a fragile language production system. Certainly, this and prior work suggest that failing to control what the child is expected to say during any fluency intervention activities increases the likelihood that elevated language challenge will increase the likelihood of stuttering during an activity. With this said, comprehensive evaluations with CWS should also assess the impact of stuttering across domains, and instruction in new speech skills or a focus on reduction in the rate of stuttering may or may not be a target of therapy for any particular CWS.

Further, clinicians should keep in mind that language growth is associated with recovery (Hollister et al., 2017; Leech et al., 2017, 2019), and children produce ungrammatical utterances as they gradually learn new structures (Eisenberg et al., 2012). Analysis of disfluent events may be useful in designing practice with emerging language structures.

The finding that grammatical errors increase the likelihood of stuttering indicates that special care should be taken with CWS who have concomitant language disorders (for whom the amount of clinical guidance is already limited). Severity of overt stuttering behaviors (i.e., SLD production and accompanying behavioral concomitants) will vary across CWS who also have language disorders, just as it varies across all children (e.g., Yaruss, 2004). It is therefore not possible to say that the presence of a concomitant language disorder will necessarily cause overt stuttering behaviors to be more severe for a particular child, even though children with language disorders will produce more grammatical errors (Eisenberg & Guo, 2016, 2018). However, because children with language disorders not only produce more grammatical errors but also persist in producing ungrammatical utterances through later ages (Eisenberg & Guo, 2016), clinicians should consider that CWS who also have language disorders may need additional attention to the development of skills to aid them when they stutter (including learning to tolerate the stuttering). In fact, as percent grammatical utterances, a language-sampling based measure, has been demonstrated to show acceptable to good diagnostic accuracy for children up to age 8;11 (Eisenberg & Guo, 2016), associations between grammaticality and stuttering may need to be considered in particular detail for school-age CWS with concomitant language disorders. These children may be working toward more advanced fluency skills while still producing grammatical errors; it may be difficult for a child to manage both fluency and grammar at the same time.

Limitations and Future Directions

Although we focused on grammatical/syntactic complexity as a determinant of disfluency (and particularly, its location), there is research that identifies lexical properties as determinants of disfluency as well, for example, frequency and phonological neighborhood (e.g., Anderson, 2007; Tsai, 2018). These attributes are computed at the individual word level, whereas we examine the utterance level. Few measures exist to examine lexical properties across a child language sample, and most involve lexical diversity indices that compute the number of unique words used in the sample (such as the type–token ratio [TTR], number of different words/100 [NDW], or vocabulary diversity [VocD]). Each of these has numerous psychometric weaknesses as well as dubious relationship to developmental psycholinguistic complexity (see Yang et al., 2022). Furthermore, diversity of word choices across a sample is not a parallel to our focus in this project, which examined fluency profiles that accompany well-formed and ill-formed utterances; errors in lexical selection might be (Charest & Skoczylas, 2019). This could be a valuable inquiry for future study but would require targeted analysis of whether or not a child's use of words was appropriate to context as well as lexical selection constraints. The detailed nature of such an investigation is beyond the purpose and scope of our present work.

Finally, this study suggests that whether an utterance will later contain a grammatical error may influence disfluency occurring earlier in the utterance, and it suggests that stalls and revisions may both be caused by planning demand and increases in language skill in play-based language contexts, for monolingual English-speaking children who do not present with any obvious language delay or disorder. The generalizability of the findings may be limited to children with similar profiles. The data are also skewed to higher education levels; the maternal education levels for CWS and CWNS from the UMD–CMU corpus

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

are higher than typical parent education in the United States, and this may limit generalizability as well.

Because this study used cross-sectional data, it is also restricted in the conclusions that it can draw about how disfluencies relate to general profiles of language development. Fortunately, the CWS and CWNS from the UMD–CMU corpus are being followed across three yearly visits, and these data will soon be available to interested researchers at http://Fluency.Talkbank.org.

It is also currently unknown how grammatical demand relates to fluency when children are older and produce fewer grammatical errors. Older children who produce a limited number of grammatical errors may be able to avert frank errors via preproduction monitoring but still show evidence that increased grammatical challenge increases disfluency rate. Conversely, as children acquire more robust language systems, grammatical complexity may influence fluency less. Some evidence already suggests the second may be true (contrast Bernstein Ratner & Sih, 1987, and Silverman & Bernstein Ratner, 2002). A related question is whether there is an age/developmental level so young that children are not aware that their utterances are ungrammatical and therefore do not show an association between grammaticality and fluency.

Lastly, our analysis of the order of disfluencies and grammatical errors followed an a priori plan, designed to include the types of disfluencies expected to be associated with grammatical errors occurring in an utterance. If current findings are replicated and the association between revisions and ungrammatical utterances is robust, it may be of interest to include revisions in future analyses looking at the order of disfluency and grammatical errors.

Data Availability Statement

The data sets analyzed during this study are available in the FluencyBank repository, http://Fluency.Talkbank.org.

Acknowledgments

This material is based upon work supported by the National Science Foundation (NSF) under Grant 1449815 awarded to principal investigator (PI) Colin Phillips; Julianne Garbarino was supported as a trainee on this grant. This work was also supported by the National Institute on Deafness and Other Communication Disorders Grant R01DC015494 awarded to Nan Bernstein Ratner and Brian MacWhinney (co-PIs) and NSF Behavioral and Cognitive Sciences Grant 1626300/1626294 awarded to Nan Bernstein Ratner (co-investigator [co-I]) and Brian MacWhinney (co-I). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF or National Institutes of Health. This study was completed by the first author in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the University of Maryland. The authors are grateful to Allison Vance, Elizabeth Kim, Gabrielle Woska, Madison Burns, Kaitlin McCloskey, Rachel Raqueno, and Valerie Hsieh for their assistance with this project. They would also like to thank members of the first author's dissertation committee, Natasha Cabrera, Rochelle Newman, Jared Novick, and Stacy Wagovich, for their feedback.

References

Ambrose, N. G., & Yairi, E. (1994). The development of awareness of stuttering in preschool children. *Journal of Fluency Disorders, 19*(4), 229–245. https://doi.org/10.1016/0094-730X(94)90002-7

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

Anderson, J. (2007). Phonological neighborhood and word frequency effects in the stuttered disfluencies of children who stutter. *Journal of Speech, Language, and Hearing Research, 50* (1), 229–247. https://doi.org/10.1044/1092-4388(2007/018)

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

Bernstein, N. E. (1981). Are there constraints on childhood disfluency? *Journal of Fluency Disorders, 6*(4), 341–350. https://doi.org/10.1016/0094-730X(81)90021-8

Bernstein Ratner, N. (1997). Linguistic and perceptual characteristics of children at stuttering onset. In E. C. Healey & H. F. M. Peters (Eds.), *Proceedings of the Second World Congress on Fluency Disorders* (pp. 3–6). Nijmegen University Press.

Bernstein Ratner, N., Brundage, S. B., & Fromm, D. (2020). *A clinician's complete guide to CLAN and PRAAT*. https://talkbank.org/manuals/Clin-CLAN.pdf

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

Bernstein Ratner, N., & Sih, C. C. (1987). Effects of gradual increases in sentence length and complexity on children's dysfluency. *Journal of Speech and Hearing Disorders, 52*(3), 278–287. https://doi.org/10.1044/jshd.5203.278

Bernstein Ratner, N. & Wijnen, F. (2007). The vicious cycle: Linguistic encoding, self-monitoring and stuttering. In J. Au-Yeung (Ed.), *Proceedings of the Fifth World Congress on Fluency Disorders* (pp 84–90).

Boomer, D. S. (1965). Hesitation and grammatical encoding. *Language and Speech, 8*(3), 148–158. https://doi.org/10.1177/002383096500800302

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

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

Charest, M., & Johnston, J. R. (2011). Processing load in children's language production: A clinically oriented review of research. *Canadian Journal of Speech-Language Pathology and Audiology, 35*(1), 18–31.

Charest, M., & Skoczylas, M. J. (2019). Lexical diversity versus lexical error in the language transcripts of children with developmental language disorder: Different conclusions about lexical ability. *American Journal of Speech-Language Pathology, 28*(3), 1275–1282. https://doi.org/10.1044/2019_AJSLP-18-0143

Choi, D., Sim, H., Park, H., Clark, C. E., & Kim, H. (2020). Loci of stuttering of English- and Korean-speaking children who stutter: Preliminary findings. *Journal of Fluency Disorders, 64,* Article 105762. https://doi.org/10.1016/j.jfludis.2020.105762

Clark, H., & Fox Tree, J. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition, 84*(1), 73–111. https://doi.org/10.1016/S0010-0277(02)00017-3

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test–Revised*. American Guidance Service.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test–Fourth Edition*. Pearson Assessments.

Dunn, M., Flax, J., Sliwinski, M., & Aram, D. (1996). The use of spontaneous language measures as criteria for identifying children with specific language impairment: An attempt to reconcile clinical and research incongruence. *Journal of Speech and Hearing Research, 39*(3), 643–654. https://doi.org/10.1044/jshr.3903.643

Eisenberg, S. L., & Guo, L.-Y. (2016). Using language sample analysis in clinical practice: Measures of grammatical accuracy for identifying language impairment in preschool and school-aged children. *Seminars in Speech and Language, 37*(2), 106–116. https://doi.org/10.1055/s-0036-1580740

Eisenberg, S. L., & Guo, L.-Y. (2018). Percent grammatical responses as a general outcome measure: Initial validity. *Language, Speech, and Hearing Services in Schools, 49*(1), 98–107. https://doi.org/10.1044/2017_LSHSS-16-0070

Eisenberg, S. L., Guo, L.-Y., & Germezia, M. (2012). How grammatical are 3-year-olds? *Language, Speech, and Hearing Services in Schools, 43*(1), 36–52. https://doi.org/10.1044/0161-1461(2011/10-0093)

Ellis Weismer, S., Venker, C. E., Evans, J. L., & Moyle, M. J. (2013). Fast mapping in late-talking toddlers. *Applied Psycholinguistics, 34*(1), 69–89. https://doi.org/10.1017/S0142716411000610

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D. J., Bates, E., Hartung, J., Pethick, S., & Reilly, J. (1993). *The MacArthur Communicative Development Inventory*. Singular Publishing Group.

Fraundorf, S. H., & Watson, D. G. (2014). Alice's adventures in um-derland: Psycholinguistic sources of variation in disfluency production. *Language, Cognition and Neuroscience, 29*(9), 1083–1096. https://doi.org/10.1080/01690965.2013.832785

Gaines, N. D., Runyan, C. M., & Meyers, S. C. (1991). A comparison of young stutterers' fluent versus stuttered utterances on measures of length and complexity. *Journal of Speech and Hearing Research, 34*(1), 37–42. https://doi.org/10.1044/jshr.3401.37

Gardner, M. F. (1990). *Expressive One-Word Picture Vocabulary Test–Revised*. Academic Therapy Publications.

Gerken, L. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language, 30*(4), 431–451. https://doi.org/10.1016/0749-596X(91)90015-C

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. Academic Press.

Gordon, P. A., & Luper, H. L. (1989). Speech disfluencies in nonstutterers. *Journal of Fluency Disorders, 14*(6), 429–445. https://doi.org/10.1016/0094-730X(89)90029-6

Hall, N. E., Wagovich, S. A., & Bernstein Ratner, N. (2007). Language considerations in childhood stuttering. In E. G. Conture & R. F. Curlee (Eds.), *Stuttering and related disorders of fluency* (3rd ed., pp. 153–167). Thieme Medical.

Haynes, W. O., & Hood, S. B. (1978). Disfluency changes in children as a function of the systematic modification of linguistic complexity. *Journal of Communication Disorders, 11*(1), 79–93. https://doi.org/10.1016/0021-9924(78)90055-2

Hollister, J., Van Horne, A. O., & Zebrowski, P. (2017). The relationship between grammatical development and disfluencies in preschool children who stutter and those who recover. *American Journal of Speech-Language Pathology, 26*(1), 44–56. https://doi.org/10.1044/2016_AJSLP-15-0022

Hudson Kam, C. L., & Edwards, N. A. (2008). The use of uh and um by 3- and 4-year-old native English-speaking children: Not quite right but not completely wrong. *First Language, 28*(3), 313–327. https://doi.org/10.1177/0142723708091149

Kleinow, J., & Smith, A. (2000). Influences of length and syntactic complexity on the speech motor stability of the fluent speech of adults who stutter. *Journal of Speech, Language, and Hearing Research, 43*(2), 548–559. https://doi.org/10.1044/jslhr.4302.548

Lee, L. L. (1974). *Developmental sentence analysis*. Northwestern University.

Leech, K. A., Bernstein Ratner, N., Brown, B., & Weber, C. M. (2017). Preliminary evidence that growth in productive language differentiates childhood stuttering persistence and recovery. *Journal of Speech, Language, and Hearing Research, 60*(11), 3097–3109. https://doi.org/10.1044/2017_JSLHR-S-16-0371

Leech, K. A., Bernstein Ratner, N., Brown, B., & Weber, C. M. (2019). Language growth predicts stuttering persistence over and above family history and treatment experience: Response to Marcotte. *Journal of Speech, Language, and Hearing Research, 62*(5), 1371–1372. https://doi.org/10.1044/2019_JSLHR-S-18-0318

Logan, K. J., & Conture, E. G. (1995). Length, grammatical complexity, and rate differences in stuttered and fluent conversational utterances of children who stutter. *Journal of Fluency Disorders, 20*(1), 35–61. https://doi.org/10.1016/0094-730X(94)00008-H

Luckman, C., Wagovich, S. A., Weber, C., Brown, B., Chang, S.-E., Hall, N. E., & Bernstein Ratner, N. (2020). Lexical diversity and lexical skills in children who stutter. *Journal of Fluency Disorders, 63,* 105747–14. https://doi.org/10.1016/j.jfludis.2020.105747

Maclay, H., & Osgood, C. E. (1959). Hesitation phenomena in spontaneous English speech. *WORD, 15*(1), 19–44. https://doi.org/10.1080/00437956.1959.11659682

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.). Lawrence Erlbaum Associates.

MacWhinney, B., & Osser, H. (1977). Verbal planning functions in children's speech. *Child Development, 48*(3), 978–985. https://doi.org/10.2307/1128349

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

Maner, K. J., Smith, A., & Grayson, L. (2000). Influences of utterance length and complexity on speech motor performance in children and adults. *Journal of Speech, Language, and Hearing Research, 43*(2), 560–573. https://doi.org/10.1044/jslhr.4302.560

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions

McDaniel, D., McKee, C., & Garrett, M. F. (2010). Children's sentence planning: Syntactic correlates of fluency variations. *Journal of Child Language, 37*(1), 59–94. https://doi.org/10.1017/S0305000909009507

McGregor, K. K., & Leonard, L. B. (1994). Subject pronoun and article omissions in the speech of children with specific language impairment: A phonological interpretation. *Journal of Speech and Hearing Research, 37*(1), 171–181. https://doi.org/10.1044/jshr.3701.171

McKee, C., McDaniel, D., & Garrett, M. F. (2017). Children's performance abilities: Language production. In E. M. Fernandez & H. S. Cairns (Eds.), *The handbook of psycholinguistics* (pp. 459–480). Wiley-Blackwell. https://doi.org/10.1002/9781118829516.ch22

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

Oetting, J. B., & McDonald, J. L. (2001). Nonmainstream dialect use and specific language impairment. *Journal of Speech, Language, and Hearing Research, 44*(1), 207–223. https://doi.org/10.1044/1092-4388(2001/018)

Oetting, J. B., & Pruitt, S. (2005). Southern African-American English use across groups. *Journal of Multilingual Communication Disorders, 3*(2), 136–144. https://doi.org/10.1080/14769670400027324

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Richels, C. G., Buhr, A., Conture, E. G., & Ntourou, K. (2010). Utterance complexity and stuttering on function words in preschool-age children who stutter. *Journal of Fluency Disorders, 35*(3), 314–331. https://doi.org/10.1016/j.jfludis.2010.06.001

Richels, C. G., & Conture, E. G. (2007). An indirect treatment approach for early intervention for childhood stuttering. In E. G. Conture & R. F. Curlee (Eds.), *Stuttering and related disorders of fluency* (3rd ed., pp. 77–99). Thieme Medical Publishers.

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. (2001). The Leading-Edge: The significance of sentence disruptions in the development of grammar. *Journal of Speech, Language, and Hearing Research, 44*(5), 1131–1143. https://doi.org/10.1044/1092-4388(2001/089)

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)

Semel, E., Wiig, E. H., & Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals Preschool– Second Edition*. Pearson Assessments.

Silverman, S., & Bernstein Ratner, N. B. (2002). Measuring lexical diversity in children who stutter: Application of *vocd*. *Journal of Fluency Disorders, 27*(4), 289–304. https://doi.org/10.1016/S0094-730X(02)00162-6

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

Tichenor, S., Leslie, P., Shaiman, S., & Yaruss, J. S. (2017). Speaker and observer perceptions of physical tension during stuttering. *Folia Phoniatrica et Logopaedica, 69*(4), 180–189. https://doi.org/10.1159/000486032

Tomblin, J. B., Records, N. L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research, 39*(6), 1284–1294. https://doi.org/10.1044/jshr.3906.1284

Tsai, P. T. (2018). Phonological neighborhood effect in spontaneous speech in adults who stutter. *Journal of Fluency Disorders, 58*, 86–93. https://doi.org/10.1016/j.jfludis.2018.08.005

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

Vasiç, N & Wijnen, F. (2005). Stuttering as a monitoring deficit. In R. Hartsuiker, R. Bastiaanse, A. Postma, & F. Wijnen (Eds.), *Phonological encoding and monitoring in normal and pathological speech* (pp. 226–247). Psychology Press.

Wagovich, S. A., & Bernstein Ratner, N. (2007). Frequency of verb use in young children who stutter. *Journal of Fluency Disorders, 32*(2), 79–94. https://doi.org/10.1016/j.jfludis.2007.02.003

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

Wall, M. J., Starkweather, C. W., & Cairns, H. S. (1981). Syntactic influences on stuttering in young child stutterers. *Journal of Fluency Disorders, 6*(4), 283–298. https://doi.org/10.1016/0094-730X(81)90017-6

Watson, J. B., Byrd, C. T., & Carlo, E. J. (2011). Effects of length, complexity, and grammatical correctness on stuttering in Spanish-speaking preschool children. *American Journal of Speech-Language Pathology, 20*(3), 209–220. https://doi.org/10.1044/1058-0360(2011/10-0019)

Wiig, E. H., Secord, W. A., & Semel, E. (1992). *Clinical Evaluation of Language Fundamentals–Preschool*. Psychological Corporation.

Yang, J. S., Rosvold, C., & Bernstein Ratner, N. (2022). Measurement of lexical diversity in children's spoken language: Computational and conceptual considerations. *Frontiers in Psychology, 13*, 905789. https://doi.org/10.3389/fpsyg.2022.905789

Yaruss, J. S. (1999). Utterance length, syntactic complexity, and childhood stuttering. *Journal of Speech, Language, and Hearing Research, 42*(2), 329–344. https://doi.org/10.1044/jslhr.4202.329

Yaruss, J. S. (2004). Documenting individual treatment outcomes in stuttering therapy. *Contemporary Issues in Communication Science and Disorders, 31*(Spring), 49–57. https://doi.org/10.1044/cicsd_31_S_49

Yaruss, J. S., Newman, R. M., & Flora, T. (1999). Language and disfluency in nonstuttering children's conversational speech. *Journal of Fluency Disorders, 24*(3), 185–207. https://doi.org/10.1016/S0094-730X(99)00009-1

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

Zimmerman, I., Steiner, V., & Pond, R. (1992). *Preschool Language Scale–Third Edition*. Psychological Corporation.

Downloaded from: https://pubs.asha.org Carnegie Mellon University on 09/10/2024, Terms of Use: https://pubs.asha.org/pubs/rights_and_permissions