ABSTRACT

Title of thesis: UNRAVELING THE LINGUISTIC LOCUS OF STUTTERING FROM CHILDHOOD TO ADULTHOOD: A LONGITUDINAL & POSITIONAL ANALYSIS

Carly Lynn Rosvold, Master of Arts, 2025

Thesis directed by: Nan Bernstein Ratner
Department of Hearing and Speech Sciences

For decades, the locus of stuttering disfluencies has been framed by a binary distinction between function and content words. While this division offered early traction, it now risks constraining deeper inquiry into the structure-sensitive demands that shape fluency development. This study re-examines the linguistic locus of stuttering by integrating grammatical role, utterance position, and developmental trajectory, moving beyond static categories to consider how vulnerabilities unfold across an utterance. It challenges the assumption that lexical class alone governs fluency patterns, proposing instead that the interaction between grammatical role and distributed planning demands more precisely account for persistence and recovery. Longitudinal analyses suggest that while function word stuttering remains stable early on, the interaction between finer-grained grammatical role and position patterns offers stronger differentiation between trajectories. By tracing these pressures, this work lays the foundation for a more dynamically structured model of early stuttering development.

UNRAVELING THE LINGUISTIC LOCUS OF STUTTERING FROM CHILDHOOD TO ADULTHOOD: A LONGITUDINAL & POSITIONAL ANALYSIS

by

Carly Lynn Rosvold

Thesis submitted to the Faculty of the Graduate School of
the University of Maryland, College Park, in partial
fulfillment of the requirements for the degree of
Master of Arts
2025

Advisory Committee:

Dr. Nan Bernstein Ratner, Chair
Dr. Rochelle Newman
Dr. Shelley Brundage

Table of Contents

List of Tables

List of Figures

Introduction

Why do we keep returning to certain questions in stuttering research, even after decades of study? Some might argue that the field has already mapped out the patterns of disfluency with sufficient clarity, but history suggests otherwise. Foundational ideas remain open to scrutiny, not because they are necessarily incorrect, but because they are often incomplete. The long-standing distinction between content and function words as a framework for understanding stuttering loci has shaped much of our knowledge, yet closer examination reveals a series of methodological inconsistencies and theoretical gaps that challenge its explanatory power.

At its core, this study is an attempt to refine, rather than reject, what we think we know about where stuttering occurs and why. The content-function distinction has provided a useful heuristic for decades, but does it truly capture the mechanisms underlying fluency breakdowns? Are we looking at linguistic categories in the right way, or have we been constrained by arbitrary classifications? And crucially, have methodological choices, such as how content-function ratios are computed, contributed to inconsistencies in our understanding of how the loci of disfluencies evolve over time and if so, why?

This thesis takes a critical approach to these questions by re-examining stuttering through three essential lenses: concept, methodology, and population.

1. **Conceptual Boundaries:** Are we over-reliant on a binary distinction?

   Stuttering research has long categorized words into either content or function, but this binary classification may obscure more meaningful linguistic distinctions. What if this framework has limited our ability to uncover subtle influences on fluency? This work explores whether an over-reliance on these broad categories has masked more nuanced patterns in disfluency.

2. **Methodological Challenges**: How do we define and measure content-function ratios? A major limitation in this area of research is the lack of standardization in how stuttering frequency is quantified. Should we measure disfluencies relative to all words spoken or only stuttered words? Should calculations include all utterances or just those containing disfluencies? Without clear methodological consensus, comparisons across studies become tenuous. Addressing these inconsistencies is essential for refining our understanding of the developmental trajectory of stuttering.
3. **Population-Specific Patterns**: Are we missing key differences between groups? Much of what we know about the content-function shift is based on cross-sectional comparisons between children and adults who stutter. However, these studies often fail to distinguish between children who will recover and those who will persist, effectively collapsing two distinct developmental paths into one. As a result, previous findings may be more reflective of group composition than an actual shift in stuttering patterns over time. This study seeks to address this gap by tracking how disfluencies evolve within the same individuals rather than drawing inferences from static snapshots of different age groups.

By integrating these perspectives, this thesis challenges the assumption that the shift from function-word to content-word stuttering is a straightforward developmental transition. Instead, it questions whether this pattern may be a byproduct of methodological and theoretical constraints rather than an inherent feature of stuttering itself. Through a longitudinal, positional, and grammatical lens, this study seeks to refine our understanding of how and why stuttering patterns change over time, moving beyond oversimplified distinctions toward a more precise and empirically grounded framework.

Background

Where within utterances do people stutter more frequently and what increases the likelihood for specific words to be stuttered on? These questions about what impacts the variability of stuttering events happened to be some of the earliest questions posed by researchers in the field of speech-language pathology. Almost 100 years ago, numerous studies attempted to document how linguistic, social-cognitive, and emotional factors influence the likelihood of stuttering (Johnson & Brown, 1935; Brown, 1937; Brown, 1938; Brown & Moren, 1942; Brown, 1945). Brown's initial studies on what he termed the "phonetic factors" of words (1938) identified a pattern in which adults tend to stutter more on consonant-initial words compared to those beginning with vowels, a finding that has been consistently noted in subsequent studies (Hahn, 1942; Wingate, 1979; Au-Yeung et al., 1998; Howell et al., 1999). However, the specific sounds that led to an increased likelihood of stuttering were found to be highly idiosyncratic across adults who stutter (AWS). Later studies, including those by Quarrington and Douglass (1960) and Soderberg (1962), attempted to examine word-initial stuttering patterns by ranking the relative difficulty of word-initial segments. However, unlike Brown's work, which relied on spontaneous speech, these studies used structured reading tasks and analyzed orthographic letters rather than phonemic transcriptions. The resulting methodological differences, along with inconsistent findings, limit the ability to draw firm conclusions about universal patterns in phoneme-level stuttering.

Brown next turned his attention to the "grammatical factors" of words (Brown, 1937). In their seminal studies, Brown and colleagues found that for AWS, content words were more likely to be stuttered than function words (Brown, 1937); words earlier in utterances are more likely to be stuttered than words later in sentences (Brown, 1938); and longer words are more likely to be stuttered than shorter words (Brown & Moren, 1942). Since the work of Brown and colleagues, numerous studies have confirmed this relationship between each of Brown's four features (i.e., word initial phoneme, word grammatical function, word position within a sentence, word length) and the predictability of stuttering events (Hahn, 1942; Wingate, 1979; Au-Yeung et al., 1998; Howell et al., 1999; Au-Yeung et al., 2003; Dworzynski et al., 2004; Warner et al., 2023). However, establishing to what extent these factors are independent of each other is another issue. Could it be the case that content words produce more stuttering than function words largely because content words tend to be longer or because many function words begin with vowels, at least in English? Taylor (1966) addressed this question regarding the impact of consonant-vowel differences, position, and length on stuttering. He discovered that each factor independently affected stuttering, with their influence diminishing in the order listed. Notably, the impact of consonant-vowel differences was approximately twice as significant as that of position and about seven times greater than that of length.

While foundational studies have consistently highlighted that specific linguistic features, such as word-initial phonemes, grammatical functions, word position within sentences, and word length, increase the probability of stuttering events in AWS, researchers have also observed that, in children who stutter (CWS), stuttering tends to occur disproportionately on function words, such as articles, pronouns, and auxiliary verbs. As children who stutter develop, stuttering tends to shift from function words to content words. This pattern has been interpreted as a developmental trajectory tied to linguistic maturation (Bloodstein, 1974; Bloodstein & Gantwerk, 1967; Bernstein, 1981; Yairi & Ambrose, 2005).

Using different data, and from different perspectives on the linguistic characteristics of stuttering events, Bloodstein and Bernstein converged on the idea that early stuttering patterns, though seemingly anomalous when compared to adult distributions, can be explained by positing that stuttering originates in an early stage of syntactic fragmentation. The essence of this assumption is that children repeat the initial word of grammatical constituents (e.g., noun phrases, verb phrases, prepositional phrases) because attempting the unit as a whole is too difficult given their current formulation capacity. This repetition allows time for the child to plan and then execute the remainder of the constituent. Therefore, early in language development, the challenge for children lies in constructing and integrating syntactic units, leading to disfluencies predominantly on function words, which often occur at the beginning of these units in English (heads of grammatical phrases; Bloodstein, 1960).

However, as linguistic competence develops, the challenge for the child transitions to the complex task of retrieving and producing content-rich words, which becomes a significant trigger for stuttering in adults (Silverman, 2004). This idea of developmental changes in speech planning and execution as children mature into adults, is based on the broader view of stuttering as a dynamic condition that evolves with an individual's linguistic and cognitive growth. This would later inspire Peter Howell and colleagues to propose a more integrated theoretical explanation.

According to Howell, the shift in stuttering from primarily occurring on function words in children to content words in adults can be attributed to a strategy used by children to manage their still-developing linguistic and cognitive abilities. This strategy involves using function words as "linguistic filler" that allows them more time to plan the articulation of the subsequent, more complex content words by inserting disfluencies prior to or occurring on function words. His rationale for the need-based delay strategy in stuttering proposes that the complete speech plan for a content word is often absent due to slow planning speed, more complex semantic content, phonetic composition, and greater length, when compared to function words. Howell's Execution and Planning Model (EXPLAN), suggests that stuttering is most likely in the gap between the completion of one speech plan and the readiness of the next (Howell, 2004). Therefore, children who stutter use function words to delay and essentially buy time to plan the articulation of the following content words since function words are typically shorter and simpler, serving as the most practical delay strategy. However, in Howell's perspective, as children grow and their linguistic and cognitive processes mature, their reliance on function words to buffer speech planning decreases and is gradually replaced by some other process that triggers persistent stuttering on content words as stuttering becomes persistent. While Howell made no attempts to define this "other process" within the original theoretical bounds of the EXPLAN model, the redistribution of the location of disfluencies that reflect profiles seen in adults is frequently attributed to the influence of cognitive and affective reactions to the lived experiences of stuttering, leading to anticipation fueled by memories of past events or self-imposed expectations of difficulty (for a review of the affective, behavioral and cognitive consequences (ABCs) of stuttering, see Bloodstein et al., 2021; Bernstein Ratner & Brundage, 2024).

While the EXPLAN model emphasizes the temporal mismatch between linguistic planning and motor execution, it is also important to consider how other speech production theories shed light on the cognitive processes involved in stuttering. One of the most influential is the Covert Repair Hypothesis (CRH; Kolk & Postma, 1997; Postma & Kolk, 1993), which builds upon Levelt's model of normal speech production (Levelt et al., 1999; Roelofs, 2014; later elaborated as the WEAVER++ model). In Levelt's framework, speech begins with a communicative intention, followed by lexical selection and phonological encoding, before culminating in articulation. The CRH proposes that disfluencies arise when the internal monitor detects errors, real or anticipated, during the formulation phase. In response, the speaker initiates a covert correction before articulation, resulting in blocks, repetitions, or prolongations. Function words, due to their rapid processing demands and frequent occurrence at utterance onsets, are particularly susceptible to such breakdowns. This, in turn, offers an explanation for why children often stutter more on function words during early speech development, when linguistic formulation processes are still maturing.

Although CRH has primarily been used to explain adult stuttering, some researchers have extended its predictions to children. However, the results have been mixed. Bakhtiar et al. (2007) found no clear differences in error detection between children who stutter and those who do not, while Yaruss and Conture (1996) reported associations between slips of the tongue and disfluencies in children. These findings suggest that the internal monitoring system may function differently in children than in adults. Particularly during utterance initiation, the combination of high planning demand and still-developing encoding mechanisms may lead the monitor to flag false errors, increasing the likelihood of disfluency. This developmental perspective is consistent with observations that young children stutter most often on function words, which are typically short, occur at the beginning of sentences, and must be integrated rapidly into emerging utterances.

Beyond the EXPLAN and Covert Repair Hypothesis frameworks, several additional models of speech production provide insight into the mechanisms that may underlie fluency breakdowns in early childhood. Levelt's WEAVER++ model, for example, proposes a modular architecture in which speech proceeds from conceptualization to formulation and then articulation, with disruptions at the level of lexical selection or phonological encoding potentially giving rise to stuttering-like disfluencies. In particular, this model emphasizes the role of incremental planning, a process that is inherently sensitive to syntactic structure and utterance position, an emphasis that aligns with the focus of the present study. Similarly, Dell's Spreading Activation Model proposes cascading activation between semantic, lexical, and phonological levels, where competition or delay at any layer can ripple through and delay articulation (Dell, 1986). These models highlight how fluency disruptions may not stem from a single source (e.g., word difficulty or timing mismatches) but instead emerge from distributed, structure-sensitive bottlenecks within the speech planning system. More recent models of language processing, such as the dual-stream framework proposed by Hickok and Poeppel (2007), further reinforce the complexity of speech production, describing dorsal and ventral pathways responsible for mapping sounds to motor representations and meaning, respectively. These accounts share an important implication: that fluency may be disrupted not simply due to word-level properties such as lexical category, but due to where and how these words are embedded within an utterance. Incorporating insights from these broader models strengthens the rationale for analyzing disfluencies in relation to positional and grammatical factors, beyond the conventional binary content/function word distinction.

Despite the popularity of EXPLAN and CRH (which are models specific to stuttering), other speech production theories have received less attention in the ongoing discussions of the content-function word shift. This may reflect a field-wide shift in focus toward affective or cognitive predictors of recovery, rather than structural or linguistic ones. It is also true that the content-function distinction has persisted in the stuttering literature for much longer than it has been used to explain typical speech production. In fact, a literature search of the terms yields primarily research conducted in stuttering, rather than any other population of speakers.

There are probably good reasons for the waning popularity of the content/function distinction in speech-language encoding research. Even in English, many “content” words are inflected with closed-class grammatical elements as they make their way from lexical retrieval to encoding to production. In other languages that inflect closed class elements such as articles, with linguistic properties of their corresponding nouns and in highly agglutinative languages, the distinction between content and function is extremely difficult to make at the “word” level.

While models such as EXPLAN and CRH offer different explanations for why function words may be particularly prone to stuttering in early development, both rely on assumptions about how “function” and “content” words are defined and measured. Yet across the literature, no standardized approach exists for calculating disfluency by word class, and the distinction between content and function words, often treated as intuitive, varies depending on the linguistic framework or analytic tools employed.

This variability becomes especially evident when comparing methodologies across studies. Brown (1945), for instance, categorized words grammatically in scripted reading passages and calculated disfluency rates as the number of stutters divided by the total number of words in each class. This approach suggested greater stuttering on content words, but it reflected patterns in written, not spontaneous, speech.

Later studies, particularly those involving children, turned to spontaneous language samples. Bloodstein and Gantwerk (1967) compared stuttered words in each class relative to their frequency in speech and found that certain function words, like pronouns and conjunctions, were stuttered more often than expected. Wingate (1979) reported a similar pattern in adults. Both studies raised the possibility that stuttering patterns may be influenced by a word’s position in the utterance, not just its lexical class.

In response to these complexities, researchers proposed alternative analytic units. Au-Yeung, Howell, and Pilgrim (1998) introduced the concept of phonological words—content words and their surrounding function words, to better reflect the planning demands of speech production. They found that disfluencies were more likely to occur on function words preceding a content word, suggesting that structural context, rather than word class alone, may offer a more meaningful lens for understanding where and why disfluencies occur.

However, the intersection of word class and utterance *position* remains particularly underexplored in recent years, despite its relevance during the early stages of speech development. Utterance-initial words, often function words, may carry heightened planning demands due to their role in utterance initiation. These structural pressures deserve renewed attention together with existing speech production models.

This is the exact question posed by Buhr and Zebrowski (2009), who argued that it is not the function words themselves but their position within a sentence that relates to stuttering. They specified that the spatial position of function words within a linguistic unit, rather than the function words alone, is the key factor. However, this finding is largely constrained in its interpretation, due to the fact that the utterance level positional values were limited to initial word or non-initial word classifications alone. Earlier, Bloodstein and Gantwerk (1967) had observed that the rate of stuttering events on function words appeared to be largely due to the effect of the high frequency of stuttering on the first word of the sentence. This suggests that the propensity for children to stutter on function words may in fact be due to their common placement at the beginning of sentences in English, where stuttering is most likely to occur (Brown, 1938; Richels et al., 2010). Thus, the pressure to initiate speech fluently, especially in children who are still mastering linguistic formulation, may result in higher disfluency rates at the start of utterances This argument gains further weight when considering Warner et al. (2023), who found that word position within a sentence did not serve as a significant predictor of stuttering in adults. This finding contrasts with observations in children, where sentence-initial words, particularly function words, are more likely to trigger disfluencies. This divergence points to the potential role of developmental differences in how stuttering manifests, suggesting that positional factors may be more crucial in younger populations.

Given this potential confound, it is useful to turn to cross-linguistic research to disentangle the influence of word class from that of within-utterance position. If the positional theory is valid, one would expect to see increased child stuttering on content words in languages where content words more frequently occur in sentence-initial position. Conversely, if the delay strategy hypothesis is correct, children should continue to stutter predominantly on function words, even in languages where utterances begin with content words, since function words could still serve as planning buffers during speech initiation.

The majority of these replication studies have been carried out in European languages such as Spanish and German, which share many linguistic features with English (Dworzynski et al., 2003; Howell & Au-Yeung, 2007; Gkalitsiou et al., 2017). Notably, these languages are non-agglutinative and prepositional (Greenberg, 1963; Chang, 1982; Gaussier & Cancedda, 2001; Bakker & Hekking, 2012; Inaba & Tokizaki, 2018). Furthermore, all three languages utilize definite and indefinite articles before nouns (Eng & O'Connor, 2000). These commonalities support the speculation that in these languages, function words are more likely to appear at the beginning of utterances than content words. However, preliminary findings by Choi and colleagues suggest that the preference for stuttering on function versus content words appears to be influenced by specific linguistic structures, while stuttering on the initial words of utterances is a language-nonspecific feature (Choi et al., 2020). The findings from this study revealed notable differences in the locus of stuttering events in English and Korean, with Korean-speaking children exhibiting a higher frequency of stuttering on content words as opposed to function words. However, both groups showed a greater propensity for stuttering to occur at the beginning of utterances, aligning with prior research that suggests a higher likelihood of stuttering at the start of sentences or speech units across languages.

Synthesis - Bridging Historical Insights with Future Directions

In light of the existing research, is the content/function word shift from children who stutter (CWS) to adults who stutter (AWS) truly a settled matter? Although numerous explanations for this distinction have been proposed over the past two decades, it is clear that the limitations embedded in much of the prior research have significantly constrained how these findings should be interpreted. Methodological inconsistencies, lack of longitudinal designs, and the treatment of CWS as a homogeneous group have all contributed to an oversimplified understanding of disfluency patterns. Yet despite these conceptual "red flags," the notion of a developmental shift from function to content word stuttering remains widely accepted, entrenched in both clinical practice and theoretical discussion. Rewriting this narrative will be challenging, but it is a necessary step toward a more accurate and nuanced account of stuttering development.

Extensive work documenting this age-related shift in the location of disfluencies in people who stutter has predominantly relied on cross-sectional designs that compare groups of CWS to AWS. However, this analytic approach is insensitive to the high remission rate of early stuttering, of whom ~80% will recover spontaneously within the age range of 2-6 years of age (Bloodstein et al., 2021). All prior work reporting on this phenomenon has failed to disambiguate these two distinct groups of CWS, leading to the potential over-representation of recovered CWS in child cohorts. Prior treatment of CWS as a homogenous group has overshadowed possible distinctions in disfluencies that the 20% of persistent stutters evince at early onset. Thus, the "shift" in the locus of stuttering-like disfluencies (SLDs) from CWS to AWS may be an illusion mediated by the transitory nature of early stuttering. This approach has resulted in the loss of any unique characteristics that differentiate between children who recover and those who do not. However, this is not the only concern that clouds the interpretation of past research. Ambiguity about the computational method for obtaining content/function ratios is an additional concern that emerges from a literature review. A clearer understanding of whether stuttering patterns stem from the linguistic properties of content and function words themselves, or from their typical placement within utterances, is essential. Disentangling the effects of word class and utterance position may help determine whether stuttering reflects the cognitive and linguistic demands of semantic and lexical processing, or whether it is more simply a function of structural position within the utterance. Understanding this positional factor in addition to the recovery status of children over time would significantly refine our theoretical models of stuttering and provide a novel contribution towards the refinement of linguistic influences on stuttering. If children who persist and recover from early stuttering differ in these patterns, we might be able to add this information to the growing literature on predictors of stuttering persistence.

The Current Study

Past research has led to the repeated assumption that a "shift" occurs from stuttering on closed-class function words in children to stuttering on open-class content words as an adult. Cross-sectional work has found evidence of this transition from a prevalence of stuttering on function words to more frequent stuttering on content words starting around age nine (Au-Yeung et al., 1998). Although this pattern is widely accepted in the field of fluency, upon closer re-examination, its supporting research appears greatly outdated, lacking appropriate updating based on the advancements in methodology and theory made in the past fifteen years of fluency research, as well as psycholinguistic research in typical language formulation. In the field of stuttering, outdated methodology is most evident in the use of cross-sectional designs to compare profiles between adults and children, and in the treatment of children who stutter (CWS) as a homogeneous group rather than subdividing them by recovery status, potentially leading to an over-representation of recovered CWS in child cohorts. These more glaring gaps in research are not the only concerns that convolute the interpretation of past research. While it has been speculated that stuttering on function words is related to stuttering at the beginning of an utterance (Bloodstein & Grossman, 1981), the question of whether stuttering on function words is driven by an artifact of stuttering at the beginning of an utterance, triggered by aspects of sentence planning, and not strictly by word classification, has not been empirically assessed.

Therefore, the purpose of the present investigation is to answer the following questions:

1. Does the distribution of disfluencies at early onset distinguish children who persist from those who recover?
2. Does the distribution of disfluencies in children who stutter change over time, and does this change differ based on stuttering persistence and recovery status?
3. How do different computational methods for calculating content and function word ratios impact the reliability and validity of measurements regarding shifts in the locus of stuttering?
4. Is the observed prevalence of stuttering on function words at utterance onset better explained by their positional occurrence within a phrase, rather than by their binary

linguistic category, and is this pattern similarly reflected in children who persist and those who recover?

Methods

Participants

This study utilized longitudinal data from the Illinois International Stuttering Research Project (IISRP) Corpus (Yairi & Ambrose, 2005), one of the few available datasets designed to track the natural history of stuttering in early childhood. The corpus, accessed through the FluencyBank database (MacWhinney & Bernstein Ratner, 2020), includes repeated language samples from children who stutter, elicited through spontaneous parent-child play sessions. These recordings span up to eight time points per child, making them well-suited for examining changes in speech patterns over time. This project focused on three time points corresponding to Sample 1 (Year 1), Sample 3 (Year 2), and Sample 5 (Year 3), yielding 158 usable samples from 57 children. These included 12 children classified as children who stutter persistently (CWSP) and 45 classified as children who stutter and recover (CWSR). Participants ranged in age from 23 to 90 months across all timepoints. Within Sample 1, CWSP had a mean age of 38.70 months (SD = 8.39), and CWSR had a mean age of 38.34 months (SD = 8.54). The inclusion of both groups across these three timepoints enabled longitudinal tracking of disfluency patterns relative to both recovery status and linguistic development.

Measures and Coding

Two primary kinds of linguistic analysis were conducted to evaluate stuttering locus: (1) word class-based metrics (i.e., content vs. function words), and (2) positional metrics within utterances. To operationalize these metrics, we relied on automated analyses from the CLAN toolkit. The FluCalc utility (Bernstein Ratner & MacWhinney, 2018) was used to calculate stuttering ratios using the MOR and POST tags produced by the CLAN morphological parser (see Figure 1). FluCalc computes stuttering frequency with respect to various fluency indicators (e.g., part-word repetitions, prolongations, blocks), and categorizes disfluent words into content and function classes based on their syntactic role.

Because there is no standardized way to calculate stuttering frequency by word class, studies have varied widely in both their numerator and denominator choices. Some have calculated stuttering rates as a proportion of stuttered function words to total function words spoken (e.g., Brown, 1945), while others have examined the proportion of stuttered words that fall into each category (e.g., Wingate, 1979). Still others, such as Bloodstein and Gantwerk (1967), used estimates of word frequency across classes as their basis for comparison. These methodological differences affect the interpretation of results and complicate cross-study comparisons.

To directly address these issues, we implemented both tendency and occurrence metrics in our analyses (see Figure 2 for a schematic comparison of these metrics). This dual approach allowed us to explore how stuttering distributes across word classes using complementary perspectives. In Studies 1 and 2, we extracted content and function word ratios from full transcripts, enabling us to calculate the general tendency of stuttering across grammatical categories. In Study 3, we focused only on utterances containing stuttering-like disfluencies (SLDs), from which we derived two metrics:

- **Tendency**, defined as the proportion of all SLDs that occurred on content or function words (similar to Wingate, 1979).

- **Occurrence**, defined as the proportion of content or function words that were stuttered out of the total number of each word type spoken (following the logic of Brown, 1945, and Bloodstein & Gantwerk, 1967).

Other disfluency types were excluded from quantitative analyses unless specified (see Table 1 for coding conventions). While a tendency-based metric has been more commonly referenced in the literature, the occurrence metric offers finer-grained insight into how stuttering patterns relate to total word use, potentially revealing distributional differences that the tendency metric alone may obscure.

For Study 4, positional analysis was conducted using the FREQPOS utility within the CLAN system, which enabled the identification of each word's location within an utterance (initial, medial, or final). Word position was analyzed at the utterance level because utterances provided a consistently identifiable, complete unit of analysis in spontaneous child speech. While clauses represent important syntactic units, naturalistic speech in young children often features incomplete, embedded, or fragmentary clauses that can be difficult to segment reliably. Fluency-coded data were extracted at the word level and merged with positional information from FREQPOS and morphological classifications obtained via the MLU utility. These morphological labels were subsequently collapsed into broader syntactic categories (e.g., nouns, verbs, prepositions, pronouns). This framework allowed us to examine whether the likelihood of stuttering on a given word was better explained by its grammatical class, its position within the utterance, or an interaction between the two.

Statistical Analyses

To address the research questions posed, proportion-based metrics of word class ratios were arcsine square root-transformed in Studies 1 through 3 to stabilize variance (Warton & Hui, 2011). In Study 1, a 2×2 repeated measures ANOVA was used to examine the effects of word class and group at onset. Study 2 employed a linear mixed effects model (LMM) with group, time, and word type as fixed effects, and participant as a random effect. Study 3 utilized a linear model to assess main and interaction effects of measurement type (tendency vs. occurrence), time, and group membership. In Study 4, Firth's Penalized Logistic Regression was selected to reduce bias in parameter estimation due to the small sample size of the CWSP group and to address potential separation in categorical data when predicting group membership. A log-linear model was also used to examine the interaction between grammatical class and utterance position without assuming a dependent variable, and chi-square tests with Cramer's V were used to assess effect sizes and validate associations among categorical variables.

Results

Study 1: Word Class Ratios for SLDs at Stuttering Onset

Research Question 1: Do children who persist (CWSP) and those who recover (CWSR) differ in their tendency to stutter on content vs. function words at onset?

To address whether stuttering tendencies at onset differ between children who recover and those who persist, a 2×2 repeated measures ANOVA was conducted with word type (function vs. content) as a within-subjects factor and group (CWSR vs. CWSP) as a between-subjects factor. On average, both CWSP and CWSR exhibited higher stuttering ratios on function words than on content words, with no apparent group-level differences (see Figure 3).

A 2×2 repeated measures ANOVA was conducted with word type (function vs. content) as a within-subjects factor and group (CWSP vs. CWSR) as a between-subjects factor (see Table 2). The analysis revealed a significant main effect of word type ($F(1,49) = 106.99$, $p < 0.0001$, ges = 0.675), indicating that children stuttered significantly more on function words than content words at onset. This represents a large effect size, with word class accounting for a substantial proportion of variance in disfluency patterns. There was no significant main effect of group ($F(1,49) = 0.0032$, $p = 0.9553$, ges < 0.001), nor a significant interaction between word type and group ($F(1,49) = 0.0105$, $p = 0.9189$, ges < 0.001), suggesting that both CWSP and CWSR demonstrated a similar tendency to stutter more on function words at onset.

To further explore the robust word type effect, a post-hoc paired t-test was conducted across all participants. The results confirmed significantly higher stuttering ratios on function words compared to content words ($p < 0.0001$, Bonferroni corrected), reinforcing function words as a consistent locus of disfluency in early childhood stuttering.

Study 2: Stability of Word Class Disfluency Over Time

Research Question 2: Are word class disfluency patterns stable over time, and do they differ by recovery status?

To examine whether stuttering patterns across word types changed over time or differed by recovery status, a linear mixed effects model was conducted with word type, time, and group as fixed effects. Descriptive statistics indicated that function words were stuttered more frequently than content words across all time points in both groups (see Table 3). These results are summarized in Table 3, which presents the mean stuttering ratios and standard errors for content and function words by group and time point.

The linear mixed effects model revealed no significant main effect of time ($p = 0.622$), indicating that stuttering rates remained stable over the three-year period. Word type, however, had a significant main effect ($p < 0.0001$), confirming that function words were stuttered significantly more than content words at all time points. There was no significant main effect of group ($p = 0.397$), nor any significant interactions, including time × group, time × word type, or time × word type × group. These findings suggest that both CWSP and CWSR followed similar developmental trajectories with respect to the location of disfluencies and that the function word bias remained stable over time.

The model explained 63.8% of the variance in stuttering ratios through fixed effects ($R^2m = 0.638$), while the contribution of random effects was negligible, resulting in an undefined conditional R^2 (R^2c). This suggests that individual differences beyond group membership, time, and word type contributed little to explaining variation in stuttering patterns.

Although function word ratios appeared descriptively higher in CWSR than CWSP by Sample 5, this difference was not statistically significant when evaluated with a Welch's independent t-test ($t(8.93) = -1.013$, $p = 0.338$). The associated effect size, Cohen's d, was also small and did not meet the threshold for practical significance, indicating that the observed difference may reflect sample variability rather than a systematic effect.

Study 3: Tendency vs. Occurrence Metrics

Research Question 3: Do different computational approaches (tendency vs. occurrence) yield different insights about stuttering profiles across groups and time?

This analysis examined whether different computational approaches, tendency (proportion of SLDs on each word class) versus occurrence (proportion of stuttered words relative to total words of that class), could distinguish between children who stutter persistently (CWSP) and those who recover (CWSR).For example, a tendency metric would calculate the percentage of all SLDs that were function words (e.g., 20 function word SLDs out of 30 total SLDs = 66.7%), whereas an occurrence metric would compute the percentage of all function words that were stuttered (e.g., 20 function word SLDs out of 200 total function words spoken = 10%). These metrics reflect different conceptualizations of stuttering distribution, whether disfluencies tend to occur more on one word class (tendency) versus how often a given word class is affected overall (occurrence). This distinction proved meaningful in the data. A significant main effect of ratio type ($p < 2e-16$) confirmed that function word tendency was the strongest and most consistent predictor of stuttering across all samples. Function words consistently exhibited higher stuttering ratios than content words, regardless of group or time point.

No significant main effect of group was observed ($p = 0.67$), and no interactions emerged between group and measurement type or between sample and measurement type. These results indicate that neither computational approach was sensitive to group differences (see Table 4). In addition, no significant effect of sample (i.e., time) was found ($p > 0.05$) (see Figure 5), and stuttering ratios remained stable across the three-year period (see Figure 4).

Together, these results suggest that while function word tendency remains a robust marker of early childhood stuttering patterns, occurrence metrics do not offer additional predictive power for distinguishing recovery status. Measurement approach alone, whether based on distribution (tendency) or likelihood (occurrence), appears insufficient as a standalone diagnostic indicator.

Study 4: Disfluency by Position and Word Class

Research Question 4: Do utterance position and grammatical word class interact to predict disfluency patterns across groups?

General differences in utterance complexity were assessed prior to modeling to determine whether broad differences in utterance complexity might account for group differences. Structures analyzed included complement clauses (COMP), coordinated clauses (COORD), subjects (SUBJ), objects (OBJ), predicate complements (PRED), auxiliary verbs (AUX), conjunctions (CONJ), infinitives (INF), and root clauses (ROOT). A series of Mann-Whitney U tests revealed no statistically significant differences between groups across any syntactic category (all $p > .05$). Effect sizes were uniformly small ($|r_x| < 0.22$), indicating negligible group differentiation even among categories like PRED and AUX, which showed the largest, but non-significant, trends. These findings suggest that long-term stuttering status is not strongly associated with the production of more or less grammatically complex or embedded structures, at least as indexed by these measures at onset.

In addition to analyzing specific syntactic structures, we assessed general utterance complexity using mean length of utterance (MLU) and verbs per utterance. Children who stutter persistently (CWSP) produced slightly shorter utterances (MLU = 4.42) compared to children who recovered (CWSR; MLU = 4.54), although the difference was minimal. Verbs per utterance were identical between groups (CWSP = 0.85; CWSR = 0.85), suggesting comparable clausal density. Together, these findings reinforce the conclusion that eventual recovery status is not strongly associated with producing longer or syntactically more complex utterances at onset. Instead, other factors such as real-time utterance integration and positional demands may play a greater role in differentiating persistence from recovery, a focus of subsequent modeling.

Having ruled out general syntactic complexity, the analysis next examined whether specific positional and grammatical factors predicted persistence. This analysis investigated whether utterance position and grammatical word class, independently or in combination, predicted differences in stuttering patterns between children who stutter persistently (CWSP) and those who recover (CWSR), while controlling for the total number of stuttering-like disfluencies (SLDs) each child produced.

Results from Firth's Penalized Logistic Regression indicated that stuttering on final words, compared to stuttering on initial words, was more likely to be associated with persistent stuttering (CWSP). Specifically, stutters on final words ($\beta = 0.87$, $p = 0.0006$) were significantly more predictive of CWSP than stutters on initial words, which served as the reference category (see Table 5). A parallel Chi-square test confirmed a significant association between group and position ($X^2(2) = 10.48$, $p = 0.005$), though the effect size was small (Cramer's V = 0.078), suggesting that position alone is not a strong predictor of persistence.

Grammatical word class on its own did not significantly differentiate CWSP and CWSR ($X^2(9) = 12.3$, $p = 0.1969$; Cramer's V = 0.085). However, the Firth model identified adverbs ($p = 0.018$), as well as a trend for verbs ($p = 0.099$) and nouns ($p = 0.105$), as showing marginal or significant predictive value. These findings suggest that while grammatical category alone has limited explanatory power, specific categories may still distinguish persistent stuttering profiles.

The most robust finding emerged from the interaction between utterance position and grammatical class. A Monte Carlo Fisher's Exact Test revealed a highly significant relationship ($p < 2.2e-16$), with a much larger effect size (Cramer's V = 0.314) than either group × position or group × word class, indicating that disfluency placement is shaped more by grammatical structure than by group membership alone.

Standardized residual analysis identified several word-position combinations that were significantly over- or underrepresented in disfluent utterances (see Figure 8). Both CWSP and CWSR showed a strong tendency to stutter on final nouns, initial pronouns, and medial prepositions. CWSP were particularly more likely to stutter on final nouns (Residual = 7.46), initial pronouns (Residual = 2.46), and medial prepositions (Residual = 3.15). CWSR exhibited similar patterns, final nouns (Residual = 8.06), initial pronouns (Residual = 4.63), and medial prepositions (Residual = 5.15), but also showed a significant tendency to stutter on medial verbs (Residual = 3.25). Notably, CWSR exhibited significantly fewer disfluencies on medial pronouns than expected (Residual = -5.08), a pattern not observed in CWSP. This positional sensitivity, particularly for function word classes like pronouns and prepositions, supports the importance of considering word class in combination with utterance structure when analyzing early stuttering patterns (see Figure 6 and Figure 7).

Discussion

This study provides a comprehensive re-evaluation of the long-held assumption that childhood stuttering is primarily characterized by disfluencies on function words and that these patterns gradually shift to content word disfluencies as the child matures. Across multiple analyses, our findings reveal that function word stuttering is a robust and stable feature of early childhood stuttering, observed in both children who persist in stuttering (CWSP) and those who recover (CWSR). Importantly, this pattern remained consistent across a three-year span post-onset, independent of how disfluency was measured.

Study 1 demonstrated that word type, rather than recovery status, is the primary determinant of stuttering locus at onset. Both groups stuttered significantly more on function words than on content words, with no group differences or interactions, reinforcing the idea that function word stuttering is a general feature of early disfluency. Study 2 confirmed the stability of this pattern over time: stuttering ratios on function words remained significantly higher than on content words across all time points and groups. There were no significant changes related to time or group, supporting the interpretation that function word dominance in stuttering is consistent and not developmentally contingent on recovery.

Study 3 further clarified the reliability of different computational metrics. Specifically, we compared tendency metrics, the proportion of stuttering-like disfluencies (SLDs) that occurred on function or content words, with occurrence metrics, which calculate the proportion of function or content words that were stuttered out of all words of that same class (e.g., stuttered function words divided by total function words).Tendency metrics emerged as the most stable and robust, demonstrating consistency across both full transcripts and subsets restricted to disfluent utterances. While occurrence ratios offer a novel lens by estimating the likelihood of disfluency relative to the total number of words within a grammatical class within a speech sample, they did not differ significantly by group or across time points. Nonetheless, occurrence metrics carry conceptual appeal, as they offer a potentially more direct indicator of linguistic vulnerability by measuring how often a given word type (e.g., function or content) is stuttered relative to how often it is produced overall. This perspective may provide insight into the susceptibility of specific lexical classes to breakdowns in fluency. Given their theoretical promise, these metrics may warrant further exploration in studies with larger and more diverse samples. Importantly, we recommend that future work clearly define how stuttering frequency is computed and strive for consistency across studies, as variation in denominator choice, whether based on total words, disfluent utterances, or other units, can substantially affect interpretation and cross-study comparisons.

Study 4 examined the role of grammatical class and utterance position in stuttering behavior. While overall associations between group and position (Cramer's V = 0.078) or word class (Cramer's V = 0.085) were statistically weak, the refined logistic regression model, controlling for overall SLD frequency, revealed several grammatical and positional features that significantly predicted persistent stuttering. Children who persisted in stuttering (CWSP) were more likely than their recovered peers to produce disfluencies in final positions, and on specific grammatical categories including adverbs. These effects were not uniform across all categories but emerged consistently in the modeling results, suggesting that stuttering persistence may be associated with breakdowns during sentence-internal formulation, particularly in morpho syntactically demanding contexts.

Crucially, the interaction between position and grammatical category yielded a substantially stronger effect (Cramer's V = 0.314), suggesting that the structure of an utterance, rather than group membership alone, exerts a more robust influence on the locus of stuttering. Function words, particularly pronouns and prepositions, demonstrated strong positional dependencies, with specific word-position combinations showing distinct patterns between CWSP and CWSR. Notably, medial pronouns were significantly underrepresented in the disfluencies of CWSR but not CWSP, suggesting that successful linguistic integration in mid-utterance may differentiate recovery trajectories. This aligns with prior work emphasizing the role of sentence planning and syntactic embedding in stuttering (e.g., Bloodstein & Grossman, 1981; Howell, 2004). The reduced stuttering on medial pronouns in CWSR may reflect greater fluency during syntactic transitions, whereas CWSP may experience continued vulnerability in assembling sentential constituents.

What do these subtleties reveal about potentially meaningful differences emerging in how persistent and recovered children construct their utterances? At utterance onset, CWSP overwhelmingly stuttered on narrow, temporal scaffolds (e.g., *then, now*), whereas CWSR employed a more diverse array of scaffolds, including spatial adverbs (e.g., *here, there, up*) and modal markers (e.g., *maybe, first*). Notably, it was the functional type and planning role of initial adverbs that varied between the groups. This pattern suggests that CWSR may initiate speech with richer, more elaborate message planning that incorporates multiple dimensions of the intended utterance from the outset. A parallel contrast appears at utterance closure. In this instance, CWSP were more likely to stutter on semantically heavy, morphologically complex nouns, including compounds (e.g., *backflip, dollhouse*) and proper names (e.g., *Ninja Turtles, Christopher*), whereas CWSR often closed utterances with semantically lighter, more frequent nouns (e.g., *snake, hoop, head*) or spatial referents (e.g., *here, there*).

These differences raise the possibility that CWSP face greater strain both in initiating and finalizing utterances, perhaps due to narrower scaffolding strategies early and heavier semantic or structural payloads at closure. In contrast, children who recover may distribute planning load more evenly across the utterance, stabilizing fluency through broader early message construction and lighter closure integration. This emerging distinction in planning architecture raises the possibility that differences in early stuttering trajectories may reflect a trade-off in planning strategies, with CWSR distributing cognitive resources more evenly across the utterance and persistent children facing cumulative load at closure. CWSR may allocate greater effort toward early message scaffolding, integrating spatial, temporal, and modal information, which could ease later formulation demands and support more stable utterance endings. In contrast, CWSP may initiate speech with narrower scaffolds, leaving more linguistic and structural integration to occur under pressure later in the utterance.

To situate these findings within a broader developmental context, it is important to consider what is known about children's sentence assembly processes and their sensitivity to structural complexity. Early work by McKee, McDaniel, and Snedeker (1998) demonstrated that toddlers are capable of producing syntactically complex, multi-clause utterances, such as restrictive relative clauses, indicating that fundamental grammatical competence for embedding emerges surprisingly early in development. However, subsequent findings by McDaniel, McKee, and Garrett (2010) reveal that despite this early syntactic availability, children's fluency remains fragile. Disfluencies, particularly pauses and restarts, cluster disproportionately at clause boundaries, exposing the high real-time cognitive demands of assembling multi-phrasal utterances. These studies suggest that sentence-level fluency disruptions do not stem from gaps in grammatical knowledge, but from the developmental challenge of dynamically coordinating message formulation, syntactic assembly, and motor execution across unfolding utterances, a challenge that may underlie persistent stuttering trajectories. Within this context, the present study's findings, that persistent stuttering is associated with disruptions at multiple points of utterance structure, support the view that distributed planning vulnerabilities, not isolated lexical retrieval failures, play a central role in early fluency disorders.

These findings offer a more nuanced perspective on existing models of fluency development. According to the Covert Repair Hypothesis (CRH), disfluencies emerge from internal error detection during pre-articulatory planning, particularly at early morphosyntactic scaffolding points. The observation that CWSP exhibited elevated disfluencies on initial adverbs aligns broadly with this prediction, suggesting vulnerabilities during early stages of utterance formulation. However, given that young children's metalinguistic and speech monitoring capacities are still developing (Clark, 1993; Smith & Zelaznik, 2004; Postma, 2000), it is likely that covert repair operations in early childhood differ qualitatively from those proposed for mature speakers, reflecting more fragility in planning under linguistic load rather than targeted error correction.

In contrast, the EXPLAN model proposes that disfluencies arise when linguistic planning lags behind motor execution, prompting speakers to "buy time" before producing complex material. While CWSP showed elevated disfluencies on semantically and morphologically dense closure nouns, these disruptions typically occurred on the complex word itself, rather than on adjacent material, a pattern not fully consistent with EXPLAN's prediction of anticipatory timing disruptions.

Taken together, these results suggest that while aspects of both CRH and EXPLAN capture elements of early fluency vulnerability, neither model fully accounts for the distributed and position-sensitive patterns observed here. Instead, what we may perceive in early stuttering patterns may reflect a broader sensitivity to dynamic integration demands across the utterance, rather than isolated repair cycles or localized planning-execution mismatches alone.

The partial alignment between existing models and the present findings suggests that the issue may lie not only in the assumptions of individual theories, but in the broader linguistic frameworks that have shaped stuttering research. In particular, the longstanding division between content and function words, a once valuable approximation, which has increasingly constrained how we conceptualize the structural dynamics of disfluency.

Despite its long-standing role in stuttering research, the binary distinction between content and function words has increasingly lost traction in contemporary psycholinguistics.

Modern models of language production emphasize gradations in lexical accessibility, morphosyntactic complexity, and information structure, dimensions that dismantle the simplistic division of grammatical categories that has rooted itself in the discourse on early stuttering. Yet much of stuttering research remains tethered to this outdated framing, anchored to assumptions that cannot account for how utterances are planned and executed in real time. Revisiting and operationalizing word class distinctions in the present study is not merely a history lesson, but a necessary step toward realigning fluency research with current linguistic theory.

Moving beyond the traditional content/function distinction is not simply a matter of redrawing boundaries; it shifts the focus from knowing where stuttering occurs in broad grammatical terms to understanding how children navigate the linguistic structures that lead there. Much like the difference between marking cities on a map and tracing the actual routes between them, refining the structural and positional pressures that shape disfluency brings us closer to the underlying dynamics of fluency vulnerability.

Although the findings here are fine-grained, they are necessary for refining how fluency vulnerability is conceptualized. Identifying structural and positional pressures that shape disfluency patterns sharpens our understanding of where early breakdowns arise, a critical step for developing more precise predictors of persistence risk, especially given the challenges posed by spontaneous recovery. While this work sharpens our understanding of early fluency vulnerabilities, future research will need to test whether these patterns hold across languages and developmental stages, strengthening the case that structural pressures, not just language-specific properties, shape stuttering trajectories. Moreover, future analyses will need to move beyond utterance-level measures, incorporating finer-grained syntactic and prosodic structures to more fully capture the dynamics of planning and breakdown. The present study represents a necessary first step toward that goal, contributing to the broader effort of rebuilding models of fluency development on more detailed, empirically grounded foundations.

In summary, this work offers a novel integration of linguistic theory, computational methodology, and longitudinal data to re-examine foundational assumptions about the nature of early stuttering. Rather than reinforcing the traditional view that early disfluencies are broadly function word driven, the findings highlight a more dynamic interplay between grammatical structure and utterance position, pressures that shape where fluency breakdown is most likely to occur. By demonstrating that specific word-position combinations, not simply lexical class, systematically differentiate children who recover from those who persist, this study refines our understanding of the distributed planning vulnerabilities that underlie early stuttering.

These findings, while rooted in fluency development, reflect a broader tension between the ways we are confined by the structures we impose and the realities we seek to understand. As the field moves beyond older distinctions like content and function words, it must rethink what its categories are meant to capture, and what they risk obscuring. Like all heuristic labels, categories are tools we create, not truths we uncover. Perhaps it's time we expect more from them.

Tables

Stuttering-like disfluencies (SLDs)	**Code**	**Example**	**Notes**
prolongation	:	s:paghetti	Place 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 :	↫r:-r:-r↫rabbit	Lengthening of the "r" segment
word repetition	[/]	dog [/] dog	Further distinctions are done inside FluCalc
Typical Disfluencies (TDs)	**Code**	**Example**	**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 1: CHAT Disfluency Conventions

Note. Only SLDs (e.g., blocks, prolongations, and repetitions) were included in stuttering word class ratio calculations. Other disfluency types were excluded from quantitative analyses unless specified.

Table 2.

Repeated Measures ANOVA: Effects of Group and Word Type on Stuttering Ratios

Effect	df_n	df^d	SS_n	SS^d	F	p	p < .05	GES
Intercept	1	49	58.917	0.131	22,065.950	<.001	*	0.956
Group	1	49	0.000	0.131	0.003	.955		0.000
Word Type	1	49	5.583	2.557	106.990	<.001	*	0.675
Group × Word Type	1	49	0.001	2.557	0.011	.919		0.000

Table 2: Repeated Measures ANOVA examining the effects of Group (CWSP vs. CWSR) and Word Type (Function vs. Content) on stuttering ratios

Table 3.			
Fixed Effects from Linear Mixed Effects Model (RQ2)			
Predictor	Estimate	Std. Error	t value
(Intercept)	0.53	0.07	7.19
Sample (3)	−0.08	0.10	−0.80
Sample (5)	−0.06	0.11	−0.55
Word Type (Function)	0.47	0.10	4.48
Group (CWSR)	0.03	0.09	0.40
Sample (3) × Word Type (Function)	0.14	0.14	0.98
Sample (5) × Word Type (Function)	0.01	0.16	0.04
Sample (3) × Group (CWSR)	−0.06	0.12	−0.53
Sample (5) × Group (CWSR)	−0.04	0.13	−0.31
Word Type (Function) × Group (CWSR)	−0.06	0.12	−0.48
Sample (3) × Word Type (Function) × Group (CWSR)	0.09	0.17	0.55
Sample (5) × Word Type (Function) × Group (CWSR)	0.18	0.18	1.01

Table 3: Fixed Effects from a Linear Mixed Effects Model Predicting Stuttering Ratios by Sample, Word Type, and Group

Table 4.

Fixed Effects from Linear Model (RQ3)

Predictor	Estimate	Std. Error	t value	p	Sig.
(Intercept)	0.24	0.06	3.81	2e-04	***
Group (CWSR)	0.03	0.07	0.42	0.675	
Sample (3)	–0.04	0.08	–0.52	0.606	
Sample (5)	–0.09	0.10	–0.97	0.332	
Content Tendency	0.24	0.09	2.67	0.008	**
Function Occurrence	0.24	0.09	2.67	0.008	**
Function Tendency	0.77	0.09	8.66	<.001	***

Table 4: Fixed effects from linear model predicting arcsine-transformed stuttering ratios by Group (CWSP, CWSR), Sample (1, 3, 5), and Ratio Type (Tendency vs. Occurrence across Function and Content Words)

Table 5.							
Firth Logistic Regression Predicting Persistent Stuttering (Covariate: Total SLDs)							
Predictor	Estimate	SE	95% CI Lower	95% CI Upper	p value	OR	Sig.
(Intercept)	−3.622	0.455	−4.593	−2.782	0.0000	0.03	***
Medial Position	0.174	0.137	−0.097	0.444	0.2068	1.19	
Final Position	0.871	0.248	0.379	1.354	0.0006	2.39	***
Adverb	1.116	0.494	0.184	2.149	0.0181	3.05	*
Aux/Mod	0.735	0.542	−0.305	1.848	0.1685	2.08	
Conjunction	0.115	0.609	−1.094	1.337	0.8514	1.12	
Determiner	0.447	0.476	−0.445	1.449	0.3364	1.56	
Noun	0.711	0.459	−0.142	1.685	0.1051	2.04	
Other	0.600	0.458	−0.252	1.572	0.1743	1.82	
Preposition	0.733	0.498	−0.208	1.772	0.1298	2.08	
Pronoun	0.604	0.449	−0.228	1.563	0.1610	1.83	
Verb	0.727	0.461	−0.131	1.706	0.0995	2.07	
Total SLDs	0.030	0.002	0.026	0.033	0.0000	1.03	***

Table 5: Firth's Logistic Regression Predicting Persistent Stuttering from Utterance Position and Grammatical Class

Figures

```
*INV:   what did Big_Bird do ? •
%mor:   pro:int|what mod|do&PAST n:prop|Big_Bird v|do ?
%gra:   1|4|OBJ 2|4|AUX 3|4|SUBJ 4|0|ROOT 5|4|PUNCT
*INV:   what did Big_Bird do ? •
%mor:   pro:int|what mod|do&PAST n:prop|Big_Bird v|do ?
%gra:   1|4|OBJ 2|4|AUX 3|4|SUBJ 4|0|ROOT 5|4|PUNCT
*CHI:   take [/] take [/] take nothing [/] nothing to [/] to [/] to school . •
%mor:   v|take pro:indef|nothing prep|to n|school .
%gra:   1|0|ROOT 2|1|OBJ 3|2|NJCT 4|3|POBJ 5|1|PUNCT
*CHI:   <and then> [//] and ↫big-big↫Big_Bird ↫a-a-a↫ate . •
%mor:   coord|and n:prop|Big_Bird v|eat&PAST .
%gra:   1|3|LINK 2|3|SUBJ 3|0|ROOT 4|3|PUNCT
*CHI:   ↫big-big↫Big_Bird take him [/] him [/] him friends picture [*] . •
%mor:   n:prop|Big_Bird v|take pro:obj|him n|friend-PL n|picture .
%gra:   1|2|SUBJ 2|0|ROOT 3|2|OBJ2 4|5|MOD 5|2|OBJ 6|2|PUNCT
*INV:   Big_Bird took his friends picture ? •
%mor:   n:prop|Big_Bird v|take&PAST det:poss|his n|friend-PL n|picture ?
%gra:   1|2|SUBJ 2|0|ROOT 3|5|DET 4|5|MOD 5|2|OBJ 6|2|PUNCT
*CHI:   yes . •
%mor:   co|yes .
%gra:   1|0|INCROOT 2|1|PUNCT
```

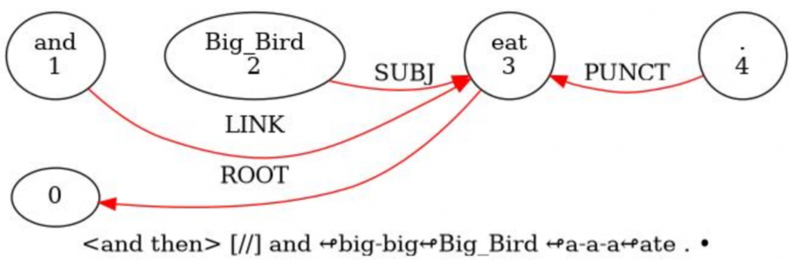

Figure 1: CHAT Transcription

Note. Includes the %mor lines for morphological analysis and the %gra lines for grammatical dependency analysis.

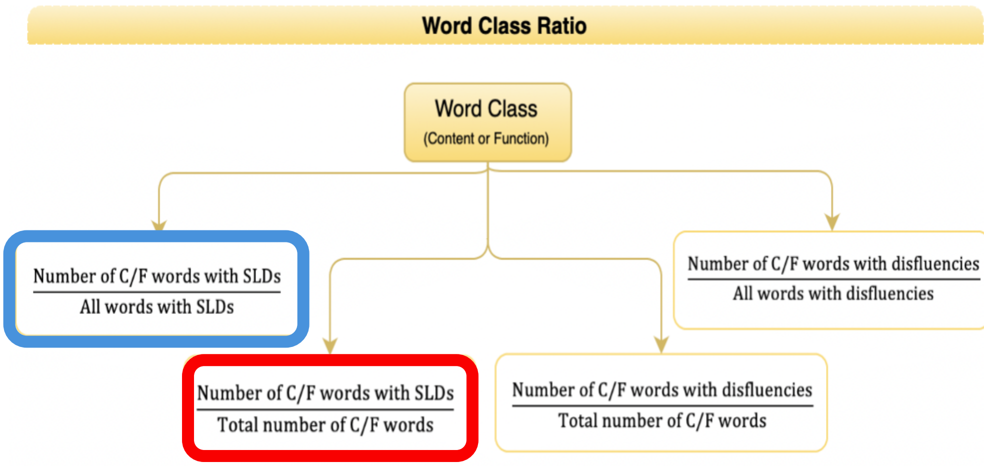

Figure 2: Schematic of Metric Definitions for Stuttering Ratio Calculations

Note. Tendency metrics reflect the proportion of stuttering-like disfluencies (SLDs) occurring on a given word class relative to all SLDs. Occurrence metrics reflect the proportion of words within a grammatical class that were stuttered. This figure illustrates how content and function word ratios were derived in each metric. Use of the +e4 switch provides computations of both the distribution of SLDs that denote type of stutterer (blue) and likelihood of stuttering (red).

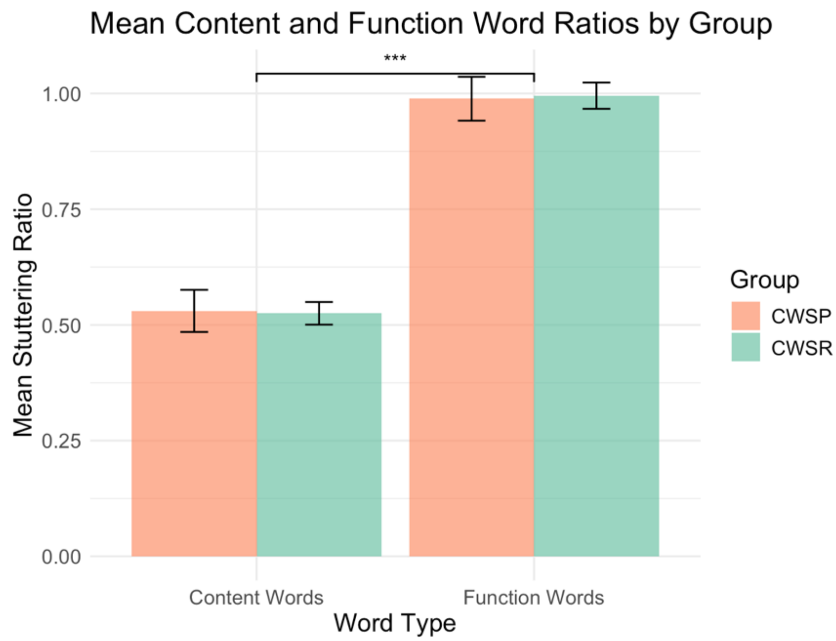

Figure 3: Mean Content and Function Word Stuttering Ratios at Onset by Group

Note. Bars represent mean stuttering ratios on content and function words for children who recovered (CWSR) and children who persisted (CWSP) at the onset of stuttering. Both groups showed higher stuttering ratios on function words.

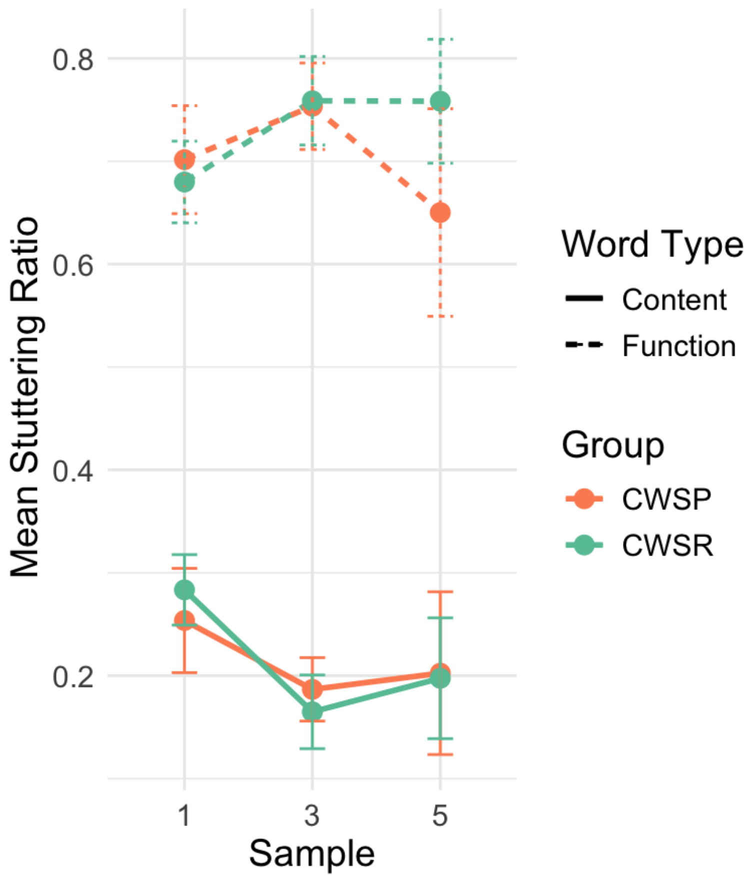

Figure 4: Mean Content and Function Word Stuttering Ratios by Sample

Note. Mean stuttering ratios are shown for content and function words across three time points (Samples 1, 3, and 5), collapsed across groups. Function word ratios remained consistently higher than content word ratios over time.

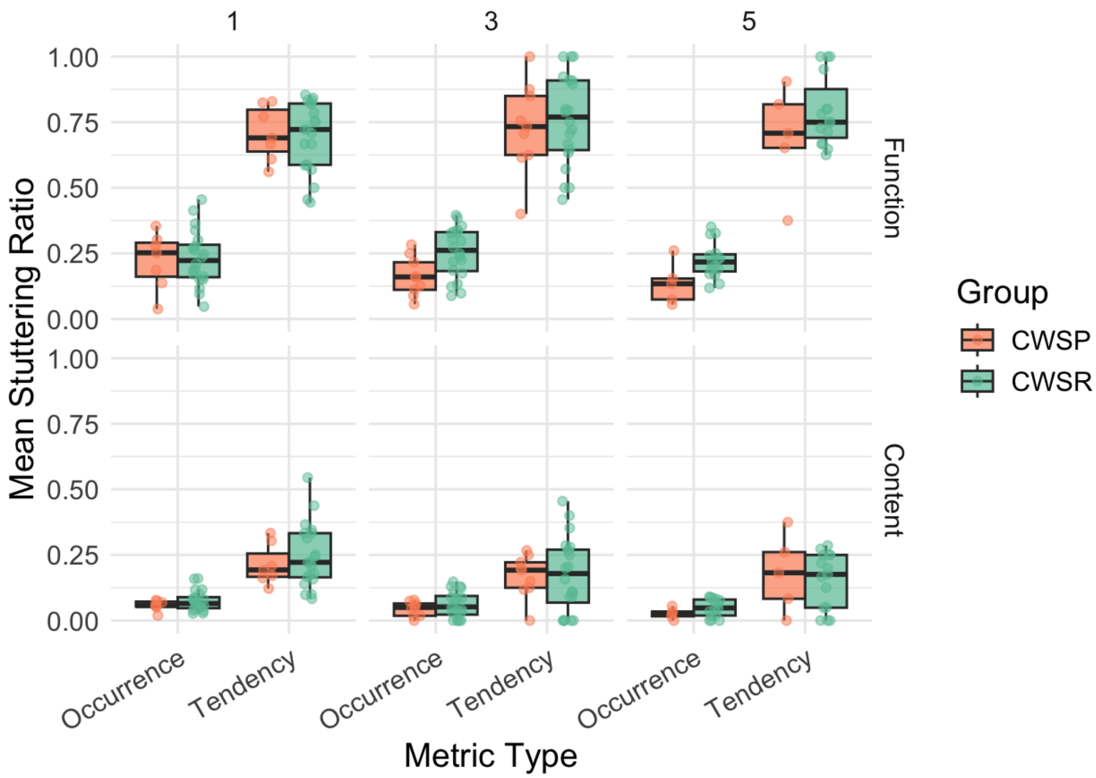

Figure 5: Mean Content and Function Word Stuttering Ratios by Sample and Metric Type

Note. Mean stuttering ratios are plotted separately for content and function words across three time points (Samples 1, 3, and 5), with metrics representing either the tendency (proportion of SLDs on a given word type) or occurrence (SLDs relative to total words of that type).

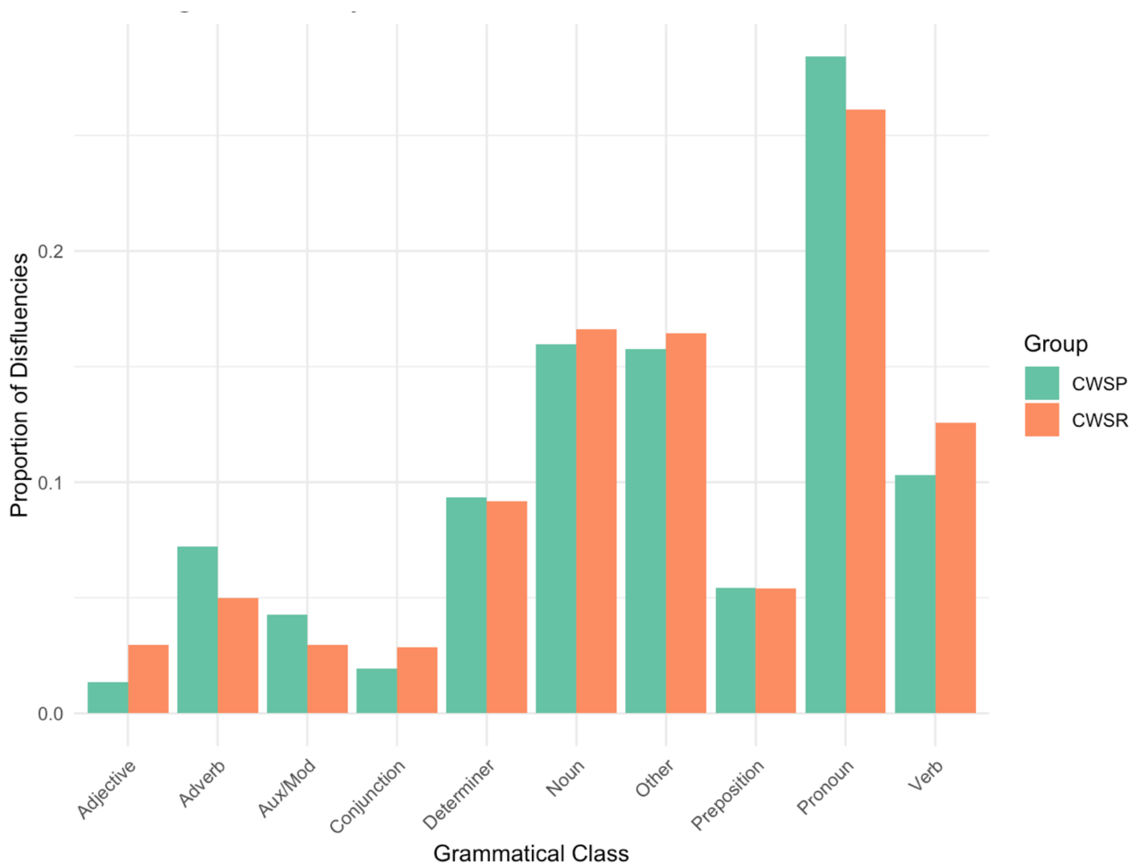

Figure 6: Stuttering Likelihood by Grammatical Word Class in CWSP and CWSR

Note. Bars represent the proportion of disfluencies attributed to each grammatical class, separated by group. Higher rates of stuttering were observed on pronouns, adverbs, and auxiliaries/modals, particularly among children who persisted in stuttering (CWSP). This figure presents raw proportions (unadjusted for total stuttering frequency). Differences shown reflect group-level distributions and do not control for individual variability in total SLDs.

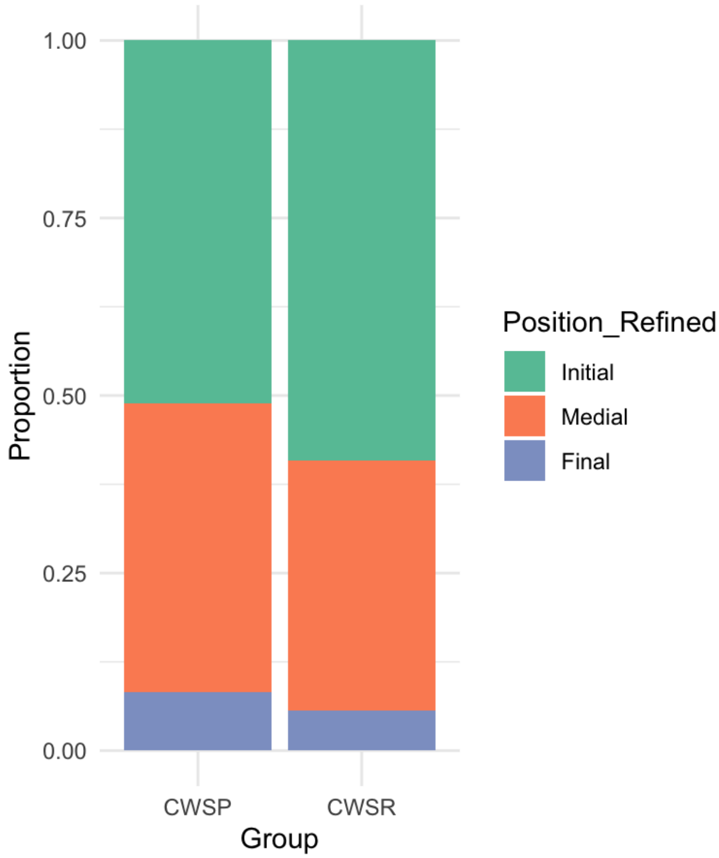

Figure 7: Proportion of Disfluencies Across Utterance Positions in CWSP and CWSR

Note. Each bar represents the proportion of total disfluencies occurring in initial, medial, or final utterance positions, separated by group. CWSR exhibited a higher proportion of disfluencies in initial position, while CWSP showed relatively more disfluencies in medial and final positions. Proportions shown reflect raw counts and are not adjusted for total SLDs per child.

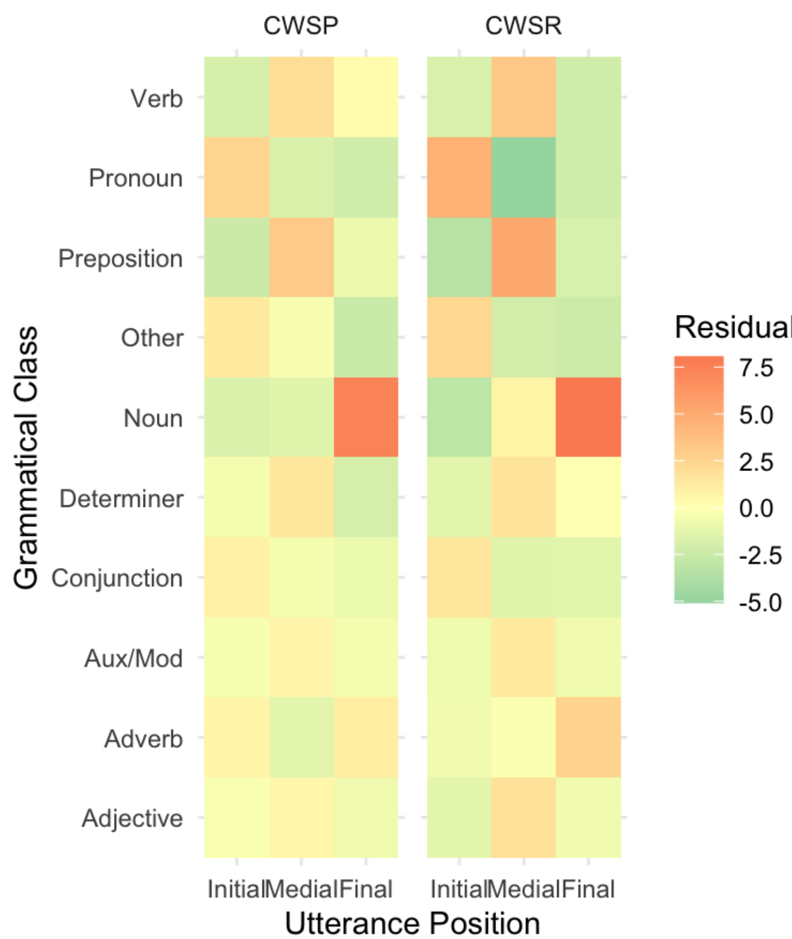

Figure 8: Standardized Residuals for Stuttering by Utterance Position and Grammatical Class in CWSP and CWSR

Note. Residuals represent the difference between observed and expected frequencies of stuttered words in each position × word class cell, based on a reduced log-linear model. Positive values indicate overrepresentation; negative values indicate underrepresentation. Standardized residuals are not corrected for individual differences in overall stuttering frequency. These results should be interpreted as descriptive rather than inferential. Positive values indicate overrepresentation; negative values indicate underrepresentation.

References:

Au-Yeung, J., Howell, P., & Pilgrim, L. (1998). Phonological words and stuttering on function words. *Journal of Speech, Language, and Hearing Research, 41*(5), 1019-1030.

Au-Yeung, J., Vallejo-Giraldo, C., & Howell, P. (2003). A model of how stuttering develops in children. *Advances in Speech-Language Pathology, 5*(1), 1-9.

Bakhtiar, M., Packman, A., & Onslow, M. (2007). The effects of self-monitoring on stuttering frequency. *Journal of Speech, Language, and Hearing Research, 50*(1), 61-72.

Bernstein Ratner, N., & MacWhinney, B. (2018). FluencyBank: Child and adult speech disfluency data. *Journal of Speech, Language, and Hearing Research, 61*(6), 1273-1283.

Bernstein Ratner, N., & Brundage, S. (2024). *The ABCs of Stuttering*. Plural Publishing.

Bloodstein, O. (1960). Phases in young stutterers' disfluency behavior. *Journal of Speech and Hearing Research, 3*(3), 14-20.

Bloodstein, O. (1974). *A Handbook on Stuttering*. National Easter Seal Society.

Bloodstein, O., & Gantwerk, B. F. (1967). Grammatical function in relation to stuttering in young children. *Journal of Speech and Hearing Research, 10*(4), 785-789.

Bloodstein, O., Bernstein Ratner, N., & Brundage, S. (2021). *A Handbook on Stuttering* (7th ed.). Plural Publishing.

Brown, S. (1937). The influence of grammatical function on the incidence of stuttering. *Journal of Speech Disorders, 2*(3), 207-215.

Brown, S. (1938). Stuttering with reference to word length. *Journal of Speech Disorders, 3*(3), 207-215.

Brown, S., & Moren, A. (1942). Studies in the incidence of stuttering. *Journal of Speech and Hearing Disorders, 7*(3), 161-165.

Choi, D., Conture, E., Walden, T., & Jones, R. (2020). Cross-linguistic differences in stuttering loci between English and Korean children. *Journal of Fluency Disorders, 64*, 105755.

Clark, E. V. (1993). *The lexicon in acquisition* (No. 65). Cambridge University Press.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological review, 93*(3), 283.

de Sonneville-Koedoot, C., Stolk, E., Rietveld, T., & Franken, M. C. (2015). Direct versus indirect treatment for preschool children who stutter: The RESTART randomized trial. *Journal of Speech, Language, and Hearing Research, 58*(3), 287-299.

Dworzynski, K., Howell, P., & Au-Yeung, J. (2003). Stuttering on function words in children who stutter: An experimental study. *Journal of Fluency Disorders, 28*(2), 79-98.

Dworzynski, K., Howell, P., & Natke, U. (2004). Predicting stuttering from linguistic factors for German speakers in two age groups. *Journal of Fluency Disorders, 29*(2), 127-136.

Hahn, E. (1942). Some phonetic characteristics of stuttering. *Journal of Speech and Hearing Research, 7*(4), 387-394.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews neuroscience, 8*(5), 393-402.

Howell, P. (2004). The EXPLAN theory of fluency control applied to the diagnosis of stuttering. In B. Maassen, & P. H. H. M. van Lieshout (Eds.), *Speech Motor Control in Normal and Disordered Speech* (pp. 287-311). Oxford University Press.

Howell, P., & Au-Yeung, J. (2007). Stuttering, speech rate, and linguistic factors. *Journal of Speech, Language, and Hearing Research, 50*(1), 123-136.

Johnson, W., & Brown, S. (1935). Speech and stuttering: A study of stuttering and word difficulty. *Journal of Speech Disorders, 1*(1), 1-8.

Kolk, H., & Postma, A. (1997). Stuttering as a covert repair phenomenon. *Speech Communication, 21*(2), 131-138.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1-38.

MacWhinney, B., & Bernstein Ratner, N. (Eds.). (2020). *FluencyBank IISRP*. TalkBank. https://fluency.talkbank.org

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.

McKee, C., McDaniel, D., & Snedeker, J. (1998). Relatives children say. *Journal of Psycholinguistic research, 27*, 573-596.

Quarrington, B., & Douglass, K. (1960). Phonetic difficulty and stuttering. *Journal of Speech and Hearing Research, 3*(1), 42-52.

Postma, A. (2000). Detection of errors during speech production: A review of speech monitoring models. *Cognition, 77*(2), 97-132.

Smith, A., & Zelaznik, H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental psychobiology, 45*(1), 22-33.

Soderberg, G. A. (1962). Stuttering and linguistic complexity. *Journal of Speech and Hearing Research, 5*(3), 194-199.

Taylor, W. (1966). The effect of consonant-vowel differences on stuttering. *Journal of Speech and Hearing Research, 9*(1), 50-58.

Warner, K., Packman, A., Onslow, M., & Harrison, E. (2023). Factors predicting stuttering in adults: A linguistic perspective. *Journal of Fluency Disorders, 63*, 105720.

Wingate, M. E. (1979). The sound of stuttering: A study of phonetic and grammatical factors.

Journal of Speech and Hearing Research, 22(4), 763-773.

Yairi, E., & Ambrose, N. G. (2004). *Early Childhood Stuttering*. PRO-ED, Inc. 8700 Shoal Creek Blvd, Austin, TX 78757.

Yaruss, J. S., & Conture, E. G. (1996). Stuttering and phonological encoding errors: An investigation of slips of the tongue in children who stutter. *Journal of Speech and Hearing Research, 39*(4), 800-815.