PROSIDING
SENALA
Seminar Nasional
Linguistik Indonesia



Available online:
https://senala.upnjatim.ac.id
E-ISSN:
3123-335X
Prodi Linguistik Indonesia
Universitas Pembangunan Nasional "Veteran" Jawa Timur
Jl. Rungkut Madya, Gn. Anyar, Kec. Gn. Anyar, Surabaya, Jawa Timur 60294

Submitted: 07/10/2026 Accepted: 01/11/2026 Published: 05/11/2026

Research Article

Indonesian Tongue Twisters as Potential Speech Practice Materials for Children with Suspected Dysarthria: An Exploratory Structural Analysis of Ten Stimuli

Mixghan Norman Antono*[1], Aning Ayu Mustikawati[2]
[1]*Universitas Sebelas Maret, Surakarta, Indonesia,*
[2]*Universitas PGRI Madiun, Madiun, Indonesia*
*Correspondence Author, E-mail: mixghan.norman@student.uns.ac.id

Abstract

Background: Indonesian tongue twister materials may differ in length, target concentration, and sequence organization, even when presented as comparable articulation tasks. These differences require description before speaker-based testing. Purpose: To characterize the textual structure of ten Indonesian stimuli and distinguish single-target repetition from paired-target switching for exploratory speech-material design. Methods: A descriptive document-based study analyzed the complete set of ten author-proposed stimuli in an existing manuscript. Word tokens, alphabetic letters, consonant letters, target-letter counts, and target concentration were computed using explicit rules. Switching was examined in three paired-target items, with a sensitivity analysis that preserved comma boundaries. The unit of analysis was written material, not a child or speech recording. Results: The corpus contained 38 word tokens, 166 letters, 88 consonant-letter occurrences, and 64 target-letter occurrences. Pooled target concentration was 72.7%, with item values of 58.3–100.0%. Seven items repeated one target and three contained paired targets. Filtered target streams yielded switching rates of 20.0%, 50.0%, and 71.4%; excluding cross-comma links reduced them to 0.0%, 40.0%, and 66.7%, respectively. Conclusion: Target concentration, sequence switching, and length describe different properties and should be reported separately. The findings support transparent stimulus specification; speech difficulty, developmental suitability, and clinical effectiveness require independent testing.

Keywords: consonant patterns, document analysis, Indonesian tongue twisters, speech material design, stimulus analysis

INTRODUCTION

A tongue twister is not defined adequately by the repeated appearance of one letter. An utterance may concentrate a single articulatory target, alternate competing targets, or combine repetition with different word lengths and consonant contexts. These properties are relevant to stimulus construction because two items assigned the same general label may place different demands on production. The present study examines such differences in a small constructed set of Indonesian materials.

Psycholinguistic experiments connect phonological competition with articulatory realization. Goldrick and Blumstein (2006) and McMillan and Corley (2010) show why the arrangement of competing segments matters beyond their frequency. Acheson and MacDonald (2009) also connect tongue twister effects with serial ordering and verbal working memory. These accounts motivate separate descriptions of target concentration, sequence organization, and length. They do not justify estimating actual speaking difficulty from spelling alone. Material design has particular relevance when practice is considered for children with suspected motor speech difficulties. However, a linguistic description cannot establish clinical suitability. Allison and Hustad (2018) describe heterogeneous pediatric dysarthria profiles, and Korkalainen et al. (2023b) report variation in the strength of pediatric motor speech intervention evidence. Neither source validates the ten Indonesian items studied here. Dysarthria provides a possible future application context, whereas the present research object is the written stimulus set.

Recent research makes the distinction between repeated letters and well-specified production materials more concrete. Loakman et al. (2025) developed phonologically informed English tongue twister generation with both automatic and human evaluation. Their work treats phonological overlap, semantic consistency, and grammaticality as separable requirements. This provides a useful comparison for an AI-assisted stimulus set: producing plausible sentences is only the construction stage, whereas evaluating what those sentences contain requires explicit measures. The current ten-item corpus is much smaller and has neither their phoneme-based representation nor their human evaluation.

Evidence from experimentally elicited errors also highlights the importance of coding position and similarity. Vidović Zorić (2025) manipulated the location and phonetic-feature similarity of target segments in Croatian tongue twisters and examined their substitutions. The study supports treating target position and segment relationships as explicit experimental variables. It does not establish the difficulty of the Indonesian examples here. Instead, it identifies an important boundary of this study: letter counts and filtered target streams provide an initial inventory that does not yet represent syllable position, feature distance, or actual pronunciation.

The study asks three questions. First, how do the ten stimuli differ in word length, consonant-letter composition, and target concentration? Second, which items instantiate single-target repetition and which contain paired-target switching? Third, how does treating commas as analytical boundaries alter the switching description? The contribution is a reproducible item-level analysis, rather than a treatment model or an intervention-effect claim.

THEORETICAL FRAMEWORK

In a cascading account, competing sound representations can influence how a speech plan is realized. This makes sequence order an important design variable. A high frequency of one target is different from frequent changes between two targets. A sentence containing both g and k may retain long homogeneous runs, whereas a sentence containing r and l may alternate repeatedly. The present operationalization measures that distinction in written representations; it does not directly observe competition in a speaker.

Orthography and phonetics are related but distinct levels of analysis. EYD V identifies 21 Indonesian consonant letters and recognizes kh, ng, ny, and sy as combinations representing single consonant sounds (Badan Pengembangan dan Pembinaan Bahasa, n.d.-a, n.d.-b). Accordingly, this study reports consonant-letter occurrences rather than phoneme counts. It does not infer syllable counts from vowel letters or assign observed IPA

transcriptions to unrecorded speech. Target labels are written as letters, while articulatory interpretations remain hypotheses for later pronunciation-based study.

A second distinction concerns representation and performance. Kember et al. (2017) used tongue twisters to elicit errors in adults with dysarthria; such speaker data are fundamentally different from a text inventory. The current analysis can identify what a material contains and how its designated targets are ordered. It cannot determine whether a child understands it, can produce it comfortably, improves through practice, or transfers gains to communication.

Recent Evidence on Planning, Learning, and Lexical Structure

The cascading interpretation is supported by evidence beyond experimentally generated tongue twisters. Alderete et al. (2021) examined phonetic traces in spontaneous adult speech errors and compared them with experimentally elicited errors. Their findings connect competition during planning with aspects of articulatory realization. The present document analysis measures neither phonetic traces nor planning activity; its connection to that account is a rationale for describing stimulus order before testing production.

Developmental evidence further limits direct transfer from adult experiments. Warker et al. (2025) examined children's production and perception of experimentally constrained tongue twisters. Children's errors reflected the constraints of what they produced and, to some extent, what they heard, with a transfer pattern differing from earlier adult findings. A child's previous linguistic exposure and the way a task is presented may therefore matter alongside the visible target sequence. These variables are absent from the current corpus analysis and should be measured in later production studies.

Lexical structure operates at a broader scale as well. Siew and Tan (2026) analyzed word-substitution errors and identified relationships with macro- and meso-level structure in a phonological network. Their unit of analysis is the relationship among lexical word forms, rather than a target-letter count within a sentence. The study motivates considering lexical neighborhoods in an expanded design, while making clear that the present switching index cannot stand in for lexical-network competition.

Length, Phonotactic Structure, and Lexical Familiarity

Length, target concentration, and phonotactic probability must be distinguished. Meylan and Griffiths (2024) modeled relationships among word forms, production effort, and listener recognition, using phonotactic probability rather than treating length as the only relevant property. Their probability-based analysis relies on information about phonemic sequences across a lexicon. In contrast, D in this article is a within-item proportion, and the filtered switching rate is a deterministic summary of a designated target set. Neither quantity estimates the probability of a sound sequence in Indonesian.

Lexical familiarity also requires its own evidence. Su et al. (2023) obtained human familiarity ratings for a large set of simplified Chinese words, illustrating how familiarity can be measured rather than assumed from spelling. Green et al. (2025) compared crowdsourced and AI-generated age-of-acquisition estimates for English vocabulary, distinguishing oral-language acquisition from reading and writing acquisition and evaluating AI estimates against human data. Neither resource provides Indonesian norms for the present items. Their methodological relevance is that plausible AI output and short word length do not, by themselves, validate vocabulary familiarity or developmental appropriateness.

Boundaries and Task Context

Speech boundaries have measurable acoustic properties that a punctuation mark alone cannot supply. Blum et al. (2024) reported evidence of word-initial consonant lengthening across a diverse language sample. Their recordings and time-aligned phonetic data concern the realization of word boundaries, not a rule that every written comma creates a pause. The comma-sensitive analysis in this study is consequently a representational sensitivity check, not an acoustic or prosodic finding.

Task context is another independent source of variation. Meng et al. (2025) examined Chinese tongue twister reading under different background-speech conditions and primary task demands, with eye movements and reading measures. The observed influence of task demands makes it important to specify what a speaker or reader is asked to do with a stimulus. Their reading experiment does not test pediatric dysarthria therapy, but it provides a reason to separate a sentence's written form from the behavioral task built around it.

METHODS

Design and Corpus Provenance

The study used descriptive quantitative document analysis with qualitative interpretation of stimulus structure. The corpus was the complete set of ten author-proposed examples in Table 3 of the manuscript A critical review of tongue twisters and a proposed practice model for children with suspected dysarthria, frozen on 16 October 2026. The stimuli had been developed during AI-assisted editorial revision of an earlier article. They constitute a constructed internal corpus, not naturally occurring speech, a public collection of conventional tongue twisters, or previously validated clinical materials. The source document and its SHA-256 identifier are recorded in the supplementary data package.

All ten examples were retained, assigned identifiers S01–S10 in their original order, and analyzed without changing their wording or punctuation. This is a census of that bounded set, not a random sample of Indonesian tongue twisters. No additional examples were created to increase the sample or to obtain a desired distribution. Target letters were fixed from the existing target annotations before numerical analysis: t, d, k, g/k, s, f, b/p, m, r/l, and n, respectively. They were not selected by searching for whichever letters produced the highest concentration.

Units and Coding Rules

Each written item was the primary unit. For counting, text was lowercased; contiguous strings of a–z were treated as word tokens; spaces and punctuation were excluded from letter counts. Repeated words and proper names were retained. W denotes word tokens, L alphabetic letters, C consonant-letter occurrences, and T occurrences of the designated target set. Consonant letters were those listed by EYD V. None of the words contained kh, ng, ny, or sy; the analysis nevertheless remains letter-based because pronunciation was not observed.

Target concentration was defined as $D = 100 \times T/C$. Its denominator is all consonant-letter occurrences, not all letters or words. Word coverage was the number of word tokens containing at least one designated target, divided by W. A worked example is S01: Tata titip topi Tuti contains nine consonant letters, of which seven are t, giving D = 77.8%. The remaining two consonant letters are p. Counts include targets in every word position; this is not an onset-only analysis.

For items with two designated targets, a filtered target stream was formed by retaining only those target letters in reading order. S04 therefore becomes gggkkk. A switch was a change between adjacent entries of this filtered stream. The switching percentage was 100 × A/(T − 1), where A is the number of changes. These entries need not be adjacent in the original text: vowels and other consonants intervene. Consequently, the measure is a descriptive index of target ordering, not a count of direct articulatory transitions or an established clinical complexity scale. For single-target items, zero switching follows from the coding definition and is not a behavioral result.

Boundary Sensitivity and Verification

The primary stream ignored punctuation. A sensitivity analysis then split an item at commas and removed links that crossed those boundaries. Switches and opportunities were counted within each resulting segment, with the denominator equal to the sum of max(Tsegment − 1, 0). Punctuation did not establish a measured pause; this analysis only examined the dependence of the textual result on a plausible boundary convention. Full streams and boundary-preserved calculations are supplied for every item.

Counts were generated computationally and checked using a second formulation based on per-word letter-frequency totals. Word and target totals were also inspected item by item. This is a consistency audit of deterministic calculations, not independent human coding, inter-rater reliability, or validation of the target labels. AI assisted editorial development, extraction, analysis implementation, and manuscript drafting. No independent linguist or clinical panel reviewed the corpus in this study.

Analysis was descriptive: totals, ranges, means, medians, item-level percentages, and exact switching numerators and denominators. Pooled concentration was calculated from summed T and C; the mean item concentration was calculated separately. No hypothesis tests, confidence intervals for a population, or correlations were used because the small, deliberately constructed set does not support population inference. Length and concentration were not combined into an unvalidated difficulty score.

Role of the Supporting Literature

The references supply theoretical and methodological context; they are not the empirical dataset. A focused literature update on 18 October 2026 prioritized relevant journal publications from 2021–2026 while retaining earlier foundational studies. Publication details and accessible abstracts or article sections were checked through journal, publisher, PubMed, and institutional records. This contextual update was not an exhaustive systematic review, and the resulting bibliography was not treated as a sample of intervention outcomes. No new stimulus, participant, or behavioral observation was added through the literature update.

RESULTS

Corpus Composition and Target Concentration

The ten stimuli comprised 38 word tokens and 36 distinct lowercased word forms. There were 166 alphabetic letters, including 88 consonant-letter occurrences and 64 designated target occurrences. Item length ranged from three to six words, with a mean of 3.8 and a median of 4.0. Five items had four words, four had three words, and one had six words. Consonant-letter counts ranged from six to twelve per item.

Table 1. Complete corpus and item-level counts

ID	Stimulus	Target letters	W	C	T	D %
S01	Tata titip topi Tuti.	t	4	9	7	77.8
S02	Dodi duduk dekat Dede.	d	4	10	7	70.0
S03	Kaki Kiki kaku.	k	3	6	6	100.0
S04	Gita gigit, Kiki ikut.	g/k	4	9	6	66.7
S05	Sari susun susu.	s	3	7	5	71.4
S06	Fafa foto Fifi.	f	3	6	5	83.3
S07	Budi bawa bola, Pipi bawa pita.	b/p	6	12	7	58.3
S08	Mama minum madu.	m	3	7	5	71.4
S09	Rara lari, Lala rileks.	r/l	4	10	8	80.0
S10	Nina menanam, Nani menanak.	n	4	12	8	66.7

W = word tokens; C = consonant-letter occurrences; T = designated target-letter occurrences; D = 100 × T/C. Alphabetic-letter counts L are S01 17, S02 18, S03 12, S04 17, S05 13, S06 12, S07 24, S08 13, S09 18, and S10 22. Letters are counted as written, not as phonetic segments.

Pooled target concentration was 64/88, or 72.7%. The unweighted mean across items was 74.6%, and the median was 71.4%. These summaries differ because the pooled value gives more weight to items with more consonant letters. S03 had the highest concentration at 100.0%: all six consonant letters in Kaki Kiki kaku were k. S07 had the lowest at 58.3%, with seven b/p occurrences among twelve consonant letters. Every word token contained at least one designated target, giving 100.0% target-word coverage for each item.

Seven items had one designated target and three had paired targets. This classification concerns the target annotation, not the entire consonant inventory. For example, the single-target S05 also contains r and n, and S10 contains m and k. S03 was the only item whose consonant-letter inventory contained just its designated target. All other items contained at least one nondesignated consonant letter.

Switching in the Paired Target Items

Table 2. Target ordering under two boundary conventions

ID	Targets	Filtered stream	A / links	Rate %	Within comma segments	Rate %
S04	g/k	gggkkk	1/5	20.0	0/4	0.0
S07	b/p	bbbppbp	3/6	50.0	2/5	40.0
S09	r/l	rrlrllrl	5/7	71.4	4/6	66.7

The primary stream excludes all nontarget letters and punctuation. The sensitivity analysis excludes links crossing commas. Rates use their own displayed denominators, so a change in percentage is not interpreted as a change in speech behavior. S04 contained both target letters but organized them into two homogeneous runs, ggg and kkk. Only one of five links switched target, giving 20.0%. S07 produced bbbppbp, with three switches across six links, or 50.0%. S09 produced rrlrllrl, with five switches across seven links, or 71.4%. Thus, simply assigning two targets did not yield a common switching profile.

Respecting comma boundaries changed all three profiles. In S04, the only switch crossed the comma, leaving zero switches across four within-segment links. S07 retained two switches across five links, or 40.0%. S09 retained four across six links, or 66.7%. The relative order was unchanged, but S04 no longer contained any within-segment target switching. The result demonstrates sensitivity to segmentation rather than an observed advantage or disadvantage of pausing.

Length, Concentration, and Order Describe Different Properties

Length, target concentration, and target order distinguished the stimuli in different ways (Figure 1). Length represented the number of word tokens, concentration represented the proportion of consonant-letter occurrences assigned to the designated target set, and switching represented changes between the two designated targets in a filtered stream. Although all three measures described the same written items, their denominators and units differed. A larger value on one measure therefore did not imply a larger value on another.

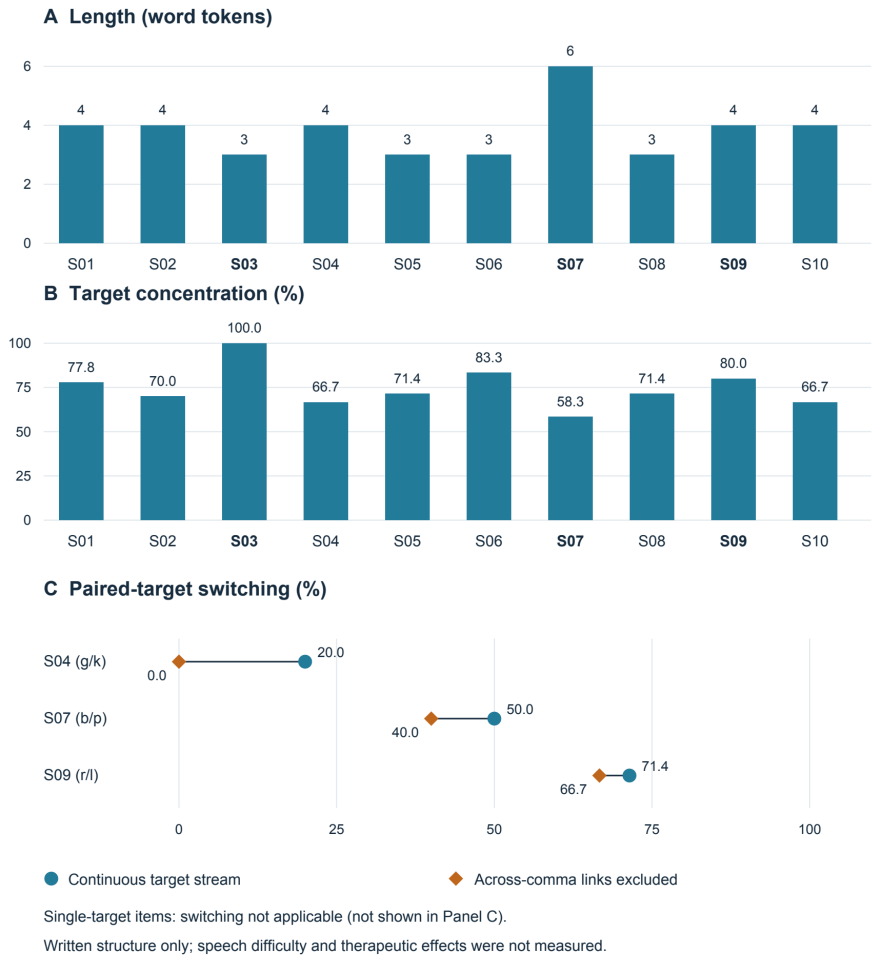


Figure 1. Length, target concentration, and target switching across the ten written stimuli

Panel A shows word-token counts; Panel B shows target concentration, calculated as 100 × designated target-letter occurrences / consonant-letter occurrences. Panel C shows switching percentages for the three paired-target items: blue circles represent the continuous filtered stream, and orange diamonds represent the analysis excluding links across commas. Single-target items are not plotted in Panel C because the paired-target switching measure is not used for those items; their omission does not indicate an absence of all consonant changes in the original sentence. Bold stimulus labels in Panels A and B identify S03, S07, and S09, which illustrate the different properties captured by the measures. Source: the constructed corpus and calculations reported in Tables 1 and 2. The figure describes written structure rather than observed speech performance or therapeutic effects.

The comparison between S03 and S07 illustrates why length and concentration should be interpreted separately. S03 contained three words and six consonant-letter occurrences, all of which were the designated target k, yielding 100.0% concentration. S07 contained six words and twelve consonant-letter occurrences, including seven designated b/p targets, yielding 58.3% concentration. S07 therefore contained twice as many words and twice as many consonant letters as S03, but only one additional target-letter occurrence. Its greater length reflected additional nontarget consonants as well as target material. The longer

stimulus consequently had the lower target concentration, rather than providing a proportionally denser target sequence.

Target ordering added information that neither length nor concentration supplied. S09 contained four words, eight r/l target occurrences, and 80.0% concentration. It also had the highest switching percentage among the three paired-target items: five changes across seven links, or 71.4%. S04 likewise contained four words, but its six g/k targets formed the grouped stream gggkkk, with only one switch across five links (20.0%). Thus, equal word counts and the presence of two designated targets did not produce equivalent ordering patterns. The intermediate profile of S07, with three switches across six links (50.0%), further showed that the longest stimulus did not have the greatest target alternation.

A second comparison separates the number of target occurrences from the composition of the target set. S09 and S10 both contained four words and eight designated target-letter occurrences. However, S09 distributed those occurrences across r and l, whereas S10 designated only n. Their concentrations also differed: 80.0% for S09 and 66.7% for S10, because the corresponding consonant-letter totals were ten and twelve. Equal word counts and equal target counts therefore did not imply equal concentration or a common paired-target structure. As indicated in Panel C, assigning S10 a zero paired-target switching score would obscure this difference; the measure was not applicable to its single-target annotation.

Panel C also makes the boundary sensitivity visible. Excluding links across commas reduced the switching percentages from 20.0% to 0.0% in S04, from 50.0% to 40.0% in S07, and from 71.4% to 66.7% in S09. S04 consequently changed from one observed transition in the continuous target stream to no transitions within its comma-delimited segments. The relative ordering of the three paired items remained S09 > S07 > S04 under both conventions. This stability concerns the descriptive ordering of the written items, while the difference between conventions shows how segmentation affects the calculated index. It does not establish that commas were realized as pauses during speech.

Taken together, Figure 1 supports describing each item through a combination of measures rather than arranging the corpus on a single difficulty scale. S03 exemplified high concentration within a short single-target item; S07 exemplified greater length with lower concentration and intermediate paired-target switching; and S09 exemplified more frequent paired-target switching within a four-word item. These are complementary structural profiles, not observed levels of articulatory difficulty. Whether any profile is more suitable for speech practice in a child with suspected dysarthria remains a question for a separate study using actual productions, specified speaking conditions, and clinically relevant outcomes.

DISCUSSION

Repeated Targets and Competing Sequences

The first contribution is to distinguish dense repetition from paired-target organization within material previously grouped under tongue twisters. The seven single-target items provide repeated written opportunities for a designated consonant, whereas the three paired items offer different ordering patterns. S03 illustrates why high target concentration alone does not establish a strong competing-sequence task. Conversely, the presence of g and k in S04 does not imply frequent alternation because their occurrences are grouped.

This distinction is consistent with using sequence arrangement as a design variable in psycholinguistic research. However, the present target streams differ from the acoustic and articulatory observations in Goldrick and Blumstein (2006) and McMillan and Corley (2010). The numerical descriptions do not replicate their experimental conditions and cannot be used

to import their effect sizes or predict a child's errors. The appropriate inference is that item labels should be accompanied by explicit structural descriptions.

The recent literature also shows why the target inventory should be extended cautiously. Vidović Zorić (2025) examined syllable position and phonetic-feature similarity, whereas the present analysis pools targets across positions. Warker et al. (2025) investigated learning under experimentally manipulated phonotactic constraints, whereas repetition and switching here are properties of fixed texts. These differences suggest concrete next steps: record the location of each target, distinguish the intervening segmental context, and compare matched orderings in an actual production task. They do not justify renaming the current descriptive percentages as observed phonological competition.

Loakman et al. (2025) offer a complementary computational comparison by evaluating generated tongue twisters with phonologically motivated measures and human judgments. The current item-level counts can support data inspection, but a higher target proportion is not equivalent to a better tongue twister. Semantic coherence, naturalness, pronunciation, and comprehensibility would need separate evaluation. In particular, a stimulus with fewer target letters could still create more demanding relationships among competing phonemes, and that possibility remains untested here.

Punctuation and Analytical Decisions

The boundary analysis adds a second contribution. S04 appears to contain a g-to-k switch when all target occurrences are concatenated, yet that switch lies across the written comma. A researcher presenting the item as separate short segments would be implementing a different target-order task from a researcher asking for continuous production. The same distinction reduced the count in S07 and S09 without eliminating their internal switches. Future protocols should therefore specify segmentation and presentation, rather than provide only a sentence and a general instruction to repeat it.

A comma is not an acoustic pause, and the filtered stream discards intervening segments. These limitations prevent a direct interpretation in terms of movement distance, coarticulation, timing, or prosody. They also explain why the analysis reports exact text operations instead of naming the switching percentage an articulatory burden score. Its usefulness is transparency: another analyst can reproduce it or replace it with a phoneme- or recording-based measure.

The boundary result is therefore best understood as a check on what the index preserves and discards. Blum et al. (2024) measured consonant timing at boundaries using acoustic material, whereas the present analysis only excludes a link in a filtered written stream. Meng et al. (2025) varied task demands and auditory context, whereas the present comparison uses the same text under two counting conventions. A future study would need independent pause or boundary annotations and a standardized presentation procedure before linking the comma-sensitive percentages to speech production.

Implications for Speech Material Development

For materials intended for later pediatric study, the corpus can be organized along separate dimensions. Single-target items could be used to investigate repetition, while paired items could be compared under blocked and more alternating arrangements. This is a proposal for experimental control, not a recommended therapeutic progression. A meaningful comparison would match vocabulary familiarity, word length, target opportunities, speaking rate, and presentation while varying target order.

Clinical interpretation must remain separate from textual findings. The heterogeneity described by Allison and Hustad (2018) makes it inappropriate to choose a universal target or progression from these counts. The broad uncertainty in pediatric intervention evidence reported by Korkalainen et al. (2023b) also cannot be resolved by a material inventory. Before use with children with suspected dysarthria, the corpus would need linguistic review, comprehension checks, professionally guided feasibility testing, and assessment of comfort and functional outcomes. No such validation was performed here.

The f example contains no designated v target, and the m example is ordinary written speech material rather than a humming prescription. Neither choice establishes developmental appropriateness or resonance benefit. Proper names and words such as rileks and menanak were retained to preserve the source corpus, not because their familiarity was measured. An expanded study should assess lexical frequency and comprehension independently rather than assume that short spelling makes a word easy.

Requirements for Later Pediatric Validation

Any extension to children with suspected motor speech disorders should begin with appropriate characterization of the population. Iuzzini-Seigel et al. (2022) provide a tutorial for differentiating childhood apraxia of speech and dysarthria through speech features and a structured assessment procedure. Kuschmann et al. (2023) document the assessment and treatment practices of German-speaking clinicians, including a need for further access to dysarthria-specific approaches. These sources explain why a written sequence cannot diagnose a disorder or identify a universal therapeutic target.

Intervention evidence must also be separated by the treatment actually studied. Korkalainen et al. (2023b) synthesize pediatric cerebral palsy motor speech intervention studies with heterogeneous evidence strength. In a separate randomized trial, Korkalainen et al. (2023a) compared Rapid Syllable Transition Treatment with usual care in 14 children with cerebral palsy and dysarthria. That trial reported benefits on speech-related outcomes for its specific intervention. It does not establish the effectiveness of the ten tongue twister stimuli, and its participants and treatment cannot be treated as validation data for the present corpus.

The choice of a later outcome measure should reflect communication beyond repetition. Levy et al. (2021) studied a dual-focus intervention using narrative, acoustic, and participation outcomes. Moya-Galé et al. (2021) examined narrative intelligibility and naturalness in French-speaking children with dysarthria. These studies demonstrate possible domains for an independent clinical evaluation; they do not show that a higher letter concentration or switching percentage produces a corresponding improvement in any of those domains.

Age and measurement conditions require equal attention. Hustad et al. (2021) describe developmental intelligibility trajectories for single-word and multiword productions. Schölderle et al. (2021) distinguish intelligibility, articulation rate, fluency, and communicative efficiency in typically developing children. Mahr et al. (2025) examine the interaction of speaking rate and intelligibility while considering development and utterance length. Together, these studies support specifying age, task, and outcome rather than equating a shorter sentence or a slower requested rate with clearer speech. No normative threshold from these language-specific studies is assigned to an Indonesian item here.

Contextual ratings should likewise not be substituted uncritically for listener transcription. Soriano et al. (2021) examined the relationship between parent-rated Intelligibility in Context Scale scores and transcription intelligibility in children with cerebral palsy. The relevance for a future study is to state whose understanding is being measured and under which conditions.

The current text-only metrics provide no parent rating, listener judgment, or estimate of communicative participation.

Lexical Validation and Construction Bias

Familiarity and age of acquisition remain unmeasured even though every word in this set contains a designated target. Su et al. (2023) provide an example of directly collected familiarity ratings, while Green et al. (2025) explicitly compare AI estimates with human judgments and distinguish oral and print acquisition. An analogous Indonesian validation would need language- and age-appropriate evidence, including for proper names and less frequent words. The present article does not import Chinese or English word ratings into its corpus.

Meylan and Griffiths (2024) and Siew and Tan (2026) also place the current within-item measures in perspective: lexical sequence probability and network relationships require a reference lexicon or a larger annotated corpus. Repeating one target in a deliberately constructed sentence may raise D without telling us how expected, familiar, or confusable the word forms are. Future corpus expansion should therefore document external sources and lexical controls, rather than increase the number of examples through generation alone.

Limitations and Further Research

The corpus is small, deliberately constructed, and derived from an AI-assisted manuscript revision. Its target distribution reflects prior design decisions, including the intended distinction between single and paired targets. The analysis identifies consequences of those choices; it does not discover naturally occurring population frequencies or demonstrate that these sentences are established Indonesian tongue twisters. The results generalize only to the ten displayed items.

The measured level is orthographic. There were no participants, audio recordings, pronunciation judgments, phonetic transcriptions, speech errors, intelligibility scores, or learning measures. Target streams omit nontarget consonants and vowels, which can affect speech production. The denominator also varies across items and target-set sizes, so percentages should be read with their exact counts. Sensitivity to commas shows that even a deterministic index depends on explicit analytical conventions.

Further work should expand the corpus using documented external sources, compare independent linguistic annotations, establish pronunciation-based segmentations, and collect speaker data. A controlled production study could then test whether the textual dimensions predict errors, effort, or rate after accounting for familiarity. Any study involving children would require the appropriate ethical procedures and clinical oversight. The present study supplies an inspectable starting dataset for that work.

CONCLUSION

The ten-item corpus contained 38 word tokens and 64 designated target-letter occurrences among 88 consonant letters. Seven items repeated one target, while three paired-target items differed in switching and sensitivity to comma boundaries. Length, target concentration, and target order captured distinct properties and did not establish a difficulty hierarchy. This document-based analysis contributes reproducible stimulus descriptions for exploratory material design. Pronunciation, developmental suitability, intelligibility, and therapeutic benefit remain questions for independent speaker-based research.

DATA AVAILABILITY

All ten stimuli and item-level results are displayed in this article. The accompanying supplementary package provides the corpus CSV, per-word counts, exact analysis rules, source-document fingerprint, and a reproducible Python analysis. The source manuscript is an unpublished internal document; its stimulus text is reproduced in Table 1.

RESEARCH ETHICS AND AI ASSISTANCE

The study analyzed constructed written materials and collected no human-participant or personal clinical data. No ethics approval number or exemption decision is claimed. AI assistance was used in prior stimulus development and in the present analytical and editorial workflow. Deterministic counting checks do not replace independent linguistic validation.

REFERENCES

Acheson, D. J., & MacDonald, M. C. (2009). Twisting tongues and memories: Explorations of the relationship between language production and verbal working memory. *Journal of Memory and Language, 60*(3), 329–350. https://doi.org/10.1016/j.jml.2008.12.002

Alderete, J., Baese-Berk, M., Leung, K., & Goldrick, M. (2021). Cascading activation in phonological planning and articulation: Evidence from spontaneous speech errors. *Cognition, 210*, Article 104577. https://doi.org/10.1016/j.cognition.2020.104577

Allison, K. M., & Hustad, K. C. (2018). Data-driven classification of dysarthria profiles in children with cerebral palsy. *Journal of Speech, Language, and Hearing Research, 61*(12), 2837–2853. https://doi.org/10.1044/2018_JSLHR-S-17-0356

Badan Pengembangan dan Pembinaan Bahasa. (n.d.-a). *Gabungan huruf konsonan*. Ejaan Bahasa Indonesia yang Disempurnakan Edisi Kelima. Retrieved October 16, 2026, from https://ejaan.kemendikdasmen.go.id/eyd/penggunaan-huruf/gabungan-huruf-konsonan/

Badan Pengembangan dan Pembinaan Bahasa. (n.d.-b). *Huruf konsonan*. Ejaan Bahasa Indonesia yang Disempurnakan Edisi Kelima. Retrieved October 16, 2026, from https://ejaan.kemendikdasmen.go.id/eyd/penggunaan-huruf/huruf-konsonan/

Blum, F., Paschen, L., Forkel, R., Fuchs, S., & Seifart, F. (2024). Consonant lengthening marks the beginning of words across a diverse sample of languages. *Nature Human Behaviour, 8*, 2127–2138. https://doi.org/10.1038/s41562-024-01988-4

Goldrick, M., & Blumstein, S. E. (2006). Cascading activation from phonological planning to articulatory processes: Evidence from tongue twisters. *Language and Cognitive Processes, 21*(6), 649–683. https://doi.org/10.1080/01690960500181332

Green, C., Kong, A. P.-H., Brysbaert, M., & Keogh, K. (2025). Crowdsourced and AI-generated age-of-acquisition (AoA) norms for vocabulary in print: Extending the Kuperman et al. (2012) norms. *Behavior Research Methods, 57*, Article 304. https://doi.org/10.3758/s13428-025-02843-8

Hustad, K. C., Mahr, T. J., Natzke, P., & Rathouz, P. J. (2021). Speech development between 30 and 119 months in typical children I: Intelligibility growth curves for single-word and multiword productions. *Journal of Speech, Language, and Hearing Research, 64*(10), 3707–3719. https://doi.org/10.1044/2021_JSLHR-21-00142

Iuzzini-Seigel, J., Allison, K. M., & Stoeckel, R. (2022). A tool for differential diagnosis of childhood apraxia of speech and dysarthria in children: A tutorial. *Language, Speech, and Hearing Services in Schools, 53*(4), 926–946. https://doi.org/10.1044/2022_LSHSS-21-00164

Kember, H., Connaghan, K. P., & Patel, R. (2017). Inducing speech errors in dysarthria using tongue twisters. *International Journal of Language & Communication Disorders, 52*(4), 469–478. https://doi.org/10.1111/1460-6984.12285

Korkalainen, J., McCabe, P., Smidt, A., & Morgan, C. (2023a). The effectiveness of Rapid Syllable Transition Treatment in improving communication in children with cerebral palsy: A randomized controlled trial. *Developmental Neurorehabilitation, 26*(5), 309–319. https://doi.org/10.1080/17518423.2023.2218485

Korkalainen, J., McCabe, P., Smidt, A., & Morgan, C. (2023b). Motor speech interventions for children with cerebral palsy: A systematic review. *Journal of Speech, Language, and Hearing Research, 66*(1), 110–125. https://doi.org/10.1044/2022_JSLHR-22-00375

Kuschmann, A., Schölderle, T., & Haas, E. (2023). Clinical practice in childhood dysarthria: An online survey of German-speaking speech-language pathologists. *American Journal of Speech-Language Pathology, 32*(6), 2802–2826. https://doi.org/10.1044/2023_AJSLP-23-00076

Levy, E. S., Chang, Y. M., Hwang, K., & McAuliffe, M. J. (2021). Perceptual and acoustic effects of dual-focus speech treatment in children with dysarthria. *Journal of Speech, Language, and Hearing Research, 64*(6S), 2301–2316. https://doi.org/10.1044/2020_JSLHR-20-00301

Loakman, T., Tang, C., & Lin, C. (2025). Train and constrain: Phonologically informed tongue twister generation from topics and paraphrases. *Computational Linguistics, 51*(2), 415–466. https://doi.org/10.1162/coli_a_00544

Mahr, T. J., Rathouz, P. J., & Hustad, K. C. (2025). Speech development between 30 and 119 months in typical children III: Interaction between speaking rate and intelligibility. *Journal of Speech, Language, and Hearing Research, 68*(1), 79–90. https://doi.org/10.1044/2024_JSLHR-24-00356

McMillan, C. T., & Corley, M. (2010). Cascading influences on the production of speech: Evidence from articulation. *Cognition, 117*(3), 243–260. https://doi.org/10.1016/j.cognition.2010.08.019

Meng, Z., Yan, G., Marsh, J. E., & Liversedge, S. P. (2025). Primary task demands modulate background speech disruption during reading of Chinese tongue twisters: An eye-tracking study. *Journal of Cognitive Psychology, 37*(6), 598–615. https://doi.org/10.1080/20445911.2025.2457774

Meylan, S. C., & Griffiths, T. L. (2024). Word forms reflect trade-offs between speaker effort and robust listener recognition. *Cognitive Science, 48*(7), Article e13478. https://doi.org/10.1111/cogs.13478

Moya-Galé, G., Keller, B., Escorial, S., & Levy, E. S. (2021). Speech treatment effects on narrative intelligibility in French-speaking children with dysarthria. *Journal of Speech, Language, and Hearing Research, 64*(6S), 2154–2168. https://doi.org/10.1044/2020_JSLHR-20-00258

Schölderle, T., Haas, E., Baumeister, S., & Ziegler, W. (2021). Intelligibility, articulation rate, fluency, and communicative efficiency in typically developing children. *Journal of Speech, Language, and Hearing Research, 64*(7), 2575–2585. https://doi.org/10.1044/2021_JSLHR-20-00640

Siew, C. S. Q., & Tan, J. F. W. Z. (2026). Production and perception errors from speech error corpora reflect macro- and meso-level structure of the phonological language network. *Language and Speech, 69*(2), 334–355. https://doi.org/10.1177/00238309251330878

Soriano, J. U., Olivieri, A., & Hustad, K. C. (2021). Utility of the Intelligibility in Context Scale for predicting speech intelligibility of children with cerebral palsy. *Brain Sciences, 11*(11), Article 1540. https://doi.org/10.3390/brainsci11111540

Su, Y., Li, Y., & Li, H. (2023). Familiarity ratings for 24,325 simplified Chinese words. *Behavior Research Methods, 55*, 1496–1509. https://doi.org/10.3758/s13428-022-01878-5

Vidović Zorić, A. (2025). The influence of phonological similarity and segment position within a word or syllable on speech production: Evidence from speech errors. *Suvremena lingvistika, 51*(99), 61–91. https://doi.org/10.22210/suvlin.2025.099.04

Warker, J., Richtsmeier, P., & Durkee, J. (2025). Children's speech errors reflect phonotactics that they say and hear. *Language Learning and Development, 21*(2), 103–117. https://doi.org/10.1080/15475441.2024.2372246