Article:

Masso, S., McLeod, S., Baker, E. et al. (1 more author) (2016) Polysyllable productions in preschool children with speech sound disorders: Error categories and the Framework of Polysyllable Maturity. International Journal of Speech-Language Pathology, 18 (3). pp. 272-287. ISSN 1754-9507

https://doi.org/10.3109/17549507.2016.1168483

Polysyllable productions in preschool children with speech sound disorders: Error categories and the Framework of Polysyllable Maturity

Sarah Masso

Sharynne McLeod

Charles Sturt University, Bathurst, Australia

Elise Baker

The University of Sydney, Australia

Jane McCormack

Charles Sturt University, Albury, Australia

Correspondence to:

Sarah Masso

Charles Sturt University

Locked Bag 450, Silverwater NSW 2128, Australia

Tel: +61-2-9752 9031

Fax: +61-2-6338 4417

Email: smasso@csu.edu.au

Keywords: speech impairment, phonology, assessment

Abstract

Purpose: Children with speech sound disorders (SSD) find polysyllables difficult; however, routine sampling and measurement of speech accuracy are insufficient to describe polysyllable accuracy and maturity. This study had two aims: (1) compare two speech production tasks and, (2) describe polysyllable errors within the Framework of Polysyllable Maturity. Method: Ninety-three preschool children with SSD from the Sound Start Study (4;0-5;5 years) completed the Polysyllable Preschool Test (POP; Baker, 2013) and the Diagnostic Evaluation of Articulation and Phonology (DEAP-Phonology; Dodd et al., 2002). Result: Vowel accuracy was significantly different between the POP and the DEAP-Phonology. Polysyllables were analysed using the seven Word-level Analysis of Polysyllables (WAP) error categories: (1) substitution of consonants or vowels (97.8% of children demonstrated common use), (2) deletion of syllables, consonants or vowels (65.6%), (3) distortion of consonants or vowels (0.0%), (4) addition of consonants or vowels (0.0%), (5) alteration of phonotactics (77.4%), (6) alteration of timing (63.4%), and (7) assimilation or alteration of sequence (0.0%). The Framework of Polysyllable Maturity described five levels of maturity based on children's errors. Conclusions: Polysyllable productions of preschool children with SSD can be analysed and categorised using the WAP, and interpreted using the Framework of Polysyllable Maturity.

Keywords: speech sound disorders, speech impairment, phonology, assessment, polysyllables

Polysyllable productions in preschool children with speech sound disorders: Error categories and the Framework of Polysyllable Maturity

INTRODUCTION

Children's productions of polysyllables are associated with phonological awareness and literacy abilities at school-age (Larrivee & Catts, 1999), and may assist with the identification of preschool children with speech sound disorders (SSD) who may be at risk of future literacy difficulties (Preston et al., 2013). Before further research into the association between polysyllable accuracy and literacy development can be conducted with preschool children with SSD, three issues need to be investigated: (1) task/s appropriate for sampling polysyllable productions of preschool children, (2) methods used to analyse and categorise error productions, and, (3) a framework to interpret the errors and polysyllable maturity of preschool children with SSD. This study addresses those needs through an investigation of the polysyllable productions by preschool children with phonologically-based SSD of unknown origin.

Speech sound disorders are common communication difficulties in children. Children may have unintelligible speech based on perception, phonological and/or production-based difficulties (International Expert Panel on Multilingual Children's Speech, 2012). The focus of the current research is preschool children with phonologically-based SSD. That is, children who demonstrate delayed phonological development in the absence of overt motor speech difficulties. The speech of preschool children with phonologically-based SSD reflects pattern-based errors reminisent of younger children with typically developing speech and language (Broomfield & Dodd; 2004; Dodd, 1995; Shriberg et al., 2005). These errors may involve systemic substitutions such as replacing one class of phoneme for another, or structural simplifications such as deleting sounds or syllables in particular word positions. It is thought that substitution and deletion error patterns reflect poorly specified underlying phonological representations (Anthony, Williams, Aghara, et al., 2010). It is the lack of specified underlying phonological representations which has been linked to poor phonological awareness skills and literacy outcomes (Anthony et al., 2011; Elbro, 1996; Fowler, 1991; Fowler & Swainson, 2004; Swan & Goswami, 1997). Given that between 30% and 77% of school-aged children with SSD have literacy difficulties (Anthony et al., 2011), there is a need to be able to better identify the preschool children with SSD most at risk of future literacy difficulties, before they start formal literacy instruction. One way to identify these preschool children could be to sample and closely analyse their productions of polysyllables. Polysyllables are words of three or more syllables and are also known as multisyllabic words (Kehoe & Stoel-Gammon, 1997; Mason, Bérubé, Bernhardt & Stemberger, 2015) or polysyllabic words (Gozzard, Baker & McCabe, 2006; James, 2006; James, van Doorn & McLeod, 2008). Polysyllables are thought to be more taxing than monosyllabic words as they contain more information to be perceived, stored in underlying representations and produced. Evidence to support this idea comes from research on typical speech acquisition. When compared to monosyllables, polysyllables take a longer time to acquire (James et al., 2008) and, when sampled, may assist with the identification of children with different sub-classifications of SSD (Vick et al., 2014). What follows is an overview of literature on typically developing children' acquisition of polysyllables, and consideration of how polysyllables might be sampled, analysed and measured.

Typical acquisition of polysyllabic words

Polysyllable acquisition in typically developing children has been studied using three main measures: phoneme accuracy, syllable accuracy, and stress accuracy (James, 2006; James et al., 2008). James (2006) proposed that the least mature realisation of polysyllables, as demonstrated by young children, primarily contain the stressed syllable, with the correct vowel (i.e. children frequently delete the less-salient weak syllables). As children grow older and their production of polysyllables becomes more mature, children gradually develop phonemic accuracy of the consonants within the stressed syllables (James et al., 2008). During these early polysyllable productions, non-final weak syllables are frequently omitted (e.g. *banana* [nana] for /bəˈnanə/ and *elephant* [ˈɛfənt] for /ˈɛləfənt/) as children develop vowels in context (also known as syntagmatic productions) (James et al., 2008). From the age of 2;4, typically developing children produce more mature polysyllables by including all the syllables (including weak and strong syllables) and demonstrate correct stress accuracy although phonemic accuracy may be reduced (e.g. *elephant* [ˈɛwəfənt] for /ˈɛləfənt/). In children aged four years and over, James (2006) observed that phonemic accuracy was realised to the detriment of stress accuracy (e.g. *elephant* [ˈɛlifənt] for /ˈɛləfənt/). James (2006) hypothesised that the shift in polysyllable accuracy as children grew older occurred because stored phonological representations became more specified. Children with more specified phonological representations for segmental information remained more true to the segmental information at the expense of stress accuracy. Thus, James' work with children who have typically developing speech and language suggests that a word-level analysis of polysyllable productions, based on unique categories exploring phoneme accuracy, syllable accuracy, and stress accuracy, may provide insight into the nature of the errors and the maturity of the phonological systems of children with phonologically-based SSD. Considering that a word-level analysis appears to be important, it is necessary to determine how these words might be sampled.

Assessment of polysyllabic words

Single word picture-naming tasks are commonly used by speech-language pathologists (SLPs) to assess children's speech (McLeod & Baker, 2014). An important component of any sampling task is the capacity to measure a range of possible phonological errors across a sample of word positions, in a relatively short period of time (Eisenberg &

Hitchcock, 2010; Kirk & Vigeland, 2015). In an analysis of 22 picture-naming tests, James (2006) highlighted that commonly-used speech assessments infrequently sampled polysyllables.

Consequently, it has been suggested that routine speech sampling tasks be supplemented with polysyllable tasks (Baker & Munro, 2011; James et al., 2008). However, it has yet to be empirically established that such supplementation is needed to more comprehensively describe the nature of children's phonological difficulties, and in particular, their production of polysyllables. Recent evidence suggests that children with SSD may be classified in to subgroups based on speech production accuracy across a number of disyllabic word and polysyllable tasks (Vick et al., 2014). Thus, research is needed to compare the information yielded from a routine speech sampling task containing relatively few polysyllables with a polysyllable-specific sampling task in preschool children with phonologically-based SSD of unknown origin. It would be expected that the children's polysyllable productions would be less accurate given that children with typically developing speech and language demonstrate systemic (segmental) and structural phonological processes in polysyllables well beyond the ages previously reported based on mono- and disyllabic word stimuli (James, Ferguson & Butcher, 2016). In addition, it has been reported that children with typically developing speech and language demonstrate lower consonant and vowel accuracy when saying polysyllables compared with monosyllabic and disyllabic words (James, 2001; Vance, Stackhouse & Wells, 2005). Sampling tasks aside, consideration also needs to be given to how a sample of polysyllables produced by children with phonologically-based SSD could best be analysed.

Analysis and measurement of polysyllables

A common method for analysing and measuring speech production accuracy in children with SSD is the calculation of percentage of consonants correct (PCC) (Shriberg, 1993; Shriberg et al., 2010), or computing a standard score based on a raw PCC value (e.g. Dodd & Crosbie, 2005). Over the years, PCC has been used widely as a standard objective measure to demonstrate baseline speech accuracy and improvements in speech accuracy following intervention. PCC has also been associated with perceptual ratings of children's speech intelligibility (McLeod, Harrison & McCormack, 2012) and children's early literacy skills (Larrivee and Catts, 1999). Specifically, Larrivee and Catts, (1999) reported that a PCC based on the production of polysyllabic real and nonwords (referred to as a MULTI-PCC measure) accounted for the greatest independent variance in reading outcomes at the end of year 1 in 30 children with phonologically-based SSD. Although more helpful than a broad measure of PCC, it could be argued that a polysyllabic PCC score alone does not yield meaningful information about the different types of errors in children's productions. Using a PCC measure, all consonants are considered equal, regardless of phonotactic context, phonetic complexity, or age of acquisition (Preston, Ramsdell, Oller, Edwards, & Tobin, 2011).

Qualitative differences in speech accuracy were considered in a longitudinal analysis of children's speech and literacy skills at five points in time (preschool to Grade 4) in Sweden (Magnusson & Nauclér, 1993; Magnusson & Nauclér, 1999; Nauclér & Magnusson, 1990). The children in these studies (mean age in preschool = 6;4, mean age at the end of Grade 1 = 7;11) were assessed for "phonological deviance" (Magnusson & Nauclér, 1993, p. 97) based on a repetition task of "long and unfamiliar words" (Nauclér & Magnusson, 1999, p. 174). The outcomes of these studies suggested that children who demonstrated phoneme errors in context (syntagmatic errors) may be less phonologically aware than children who demonstrate phoneme errors in isolation (paradigmatic speech errors) (Magnusson & Nauclér, 1990). The preliminary results of this longitudinal study highlight the need for further investigation in to categories of polysyllable errors, and polysyllable maturity, of children with SSD.

A more comprehensive method for analysing children's productions of polysyllables was reported by Mason et al. (2015). They studied the polysyllable productions of 10 typically developing Canadian English-speaking children aged 5-years and eight French-speaking children with protracted phonological development (i.e. SSD) age 3- to 4-years. Using the basic tenets of nonlinear phonology, Mason et al. classified the children's productions of six polysyllables as either matches or non-matches with the adult pronunciation. They then categorised and tallied the matches and non-matches on five different measures of a word (including foot structure, stress pattern, syllable shape, timing units and segments) to create a multisyllabic word metric with and without vowels. They also calculated phonological mean length of utterance (pMLU) and PCC. They discovered that the multisyllabic word metric (with and without vowels) was more valuable than PCC and pMLU as it more finely differentiated phonological accuracy across the six polysyllabic words. The multisyllabic word metric numerically captured more information about segmental and prosodic aspects of the words. The preliminary findings of Mason and colleagues (2015) support the need for more comprehensive sampling, analysis and measurement of a wider sample of polysyllables with a larger sample of English-speaking preschool children with SSD. Additionally, their findings indicate a need for more detailed descriptions of the types of polysyllable errors produced by preschool children with SSD. They also raise the need to consider the frequency of those errors.

Frequency of errors on polysyllabic words

Although the occurrence of an error can be quite easy to measure (e.g. an error can be present or absent), interpretation of the frequency of occurrence can be challenging. To date, there is no agreed upon criterion for determining what is and is not a problem when reporting the frequency of occurrence of an error for preschool children with SSD (Kirk & Vigeland, 2015). For instance, in studies which report the presence of phonological processes in children with typically developing speech and language, speech errors have been considered “present” based on: one instance of the use of a phonological process (Hodson & Paden, 1981), at least five instances of a process (Dodd et al., 2003), or on 20% of opportunities to use the process (Haelsig & Madison, 1986; Roberts, Burchinal & Footo, 1990). In a recent study by James, Ferguson and Butcher (2016) with typically developing children aged 3-, 4-, 5-, 6- and 7-years, a phonological process was considered to be in use at each year age if: (1) 20% of the children produced the process and (2) the median percent process occurrence was greater than 5%. While such relatively low criterions may be useful for studying typical acquisition and differential diagnosis (i.e. determining if a child does or does not have a SSD), it has the potential to conceal differences between common and less common errors. One way of revealing differences in error frequency, particularly within a clinical sample of children with SSD, is to set a higher criterion. This approach is recommended when wanting to identify the dominant errors in the speech of children with SSD requiring intervention. For example, Hodson and Paden (1991) recommend a criterion of 40% occurrence of an error pattern within an adequate sample, such as a minimum of 10 opportunities for a specific error pattern (Hodson & Paden, 1991).

Study aims

In light of the extant literature highlighting the gradual maturity of polysyllables in typically developing children, and the potentially important role that polysyllables might play in identifying preschool children most at risk of future literacy difficulties, there is a need to study the polysyllable productions of preschool children with SSD. Specifically, there is a need to consider how polysyllables might best be sampled and how polysyllable errors produced by preschool children with SSD could be described with respect to their level of maturity. Thus, the first aim of this study is to determine whether there is a significant difference between consonant, vowel, and total phoneme accuracy between two different speech sampling tasks: a routine speech sampling task, and a polysyllable-specific sampling task. The second aim is to describe and categorise the polysyllable errors demonstrated by preschool children with SSD using the Word-level Analysis of Polysyllables (WAP; Masso, 2016a; see Supplementary Appendix A), then describe the errors within the Framework of Polysyllable Maturity (Masso, 2016b). In light of the literature attesting to the fact that polysyllables can contain segmental (consonant and vowel) and structural errors (word shape, length and stress), the WAP consists of seven error categories: (1) substitution of consonants or vowels, (2) deletion of syllables, consonants or vowels, (3) distortion of consonants or vowels, (4) addition of consonants or vowels, (5) alteration of phonotactics, (6) alteration of timing, and (7) assimilation and/or alteration of sequence. Categories 1, 2, 3, and 4 were based on the SODA analysis (substitution, omission, distortion, addition) by van Riper (1939) and categories 1, 2, 5, 6, and 7 were informed by James' (2006) work on polysyllable acquisition. Further information about the WAP is provided in the method.

METHOD

Context and participant recruitment

The children described in this study were participants in the Sound Start Study (McLeod, Baker, McCormack, Wren & Roulstone, 2013-2015) conducted in early childhood centres across Sydney, Australia. The Sound Start Study was a cluster randomised controlled trial exploring the effectiveness of a computer-based, first-phase phonological intervention program for preschool children. The Sound Start Study included six stages of data collection. The current study reports on data collected during stages 2 and 3 for years 1 and 2 of the Sound Start Study (see Figure 1). Stages 2 and 3 were designed to identify suitable participants for the intervention that occurred in stage 4.

Cluster sampling, a hierarchical sampling method, was implemented to select participants for the study. A total of 54 early childhood centres were invited to participate in years 1 and 2 of the Sound Start Study to represent a range of socioeconomic areas based on the Australian Bureau of Statistics, Index of Relative Socio-economic Advantage and Disadvantage (IRSAD, ABS, 2008) and 33 centres agreed to participate. The director/principal of these centres provided written consent for the involvement of their centre. A total of 1353[1] 4- to 5-year-old children were enrolled at the participating centres. During stage 1, screening questionnaires were distributed to the parents of all 4- to 5-year-old children and 852 (62.9%) questionnaires were returned. Of these families, the parents of 782 children (91.7%) provided consent for their children's teacher to complete a similar questionnaire and 729 (93.2%) of the teacher questionnaires were returned.

Eligibility for participation in stage 2 required that children's parents or teachers reported concerns about their speech sound development based on the Parents' Evaluation of Developmental Status (PEDS; Glascoe, 2000). Children were excluded from further participation if the parent and/or teacher reported that the child had a persistent hearing loss, cleft lip and/or palate or a developmental delay. Children were also excluded if parents reported that the child's English language proficiency was less than their home language proficiency. Of the 852 children in stage 1, 655 were ineligible: 623 children did not meet the inclusion criteria for further participation (as outlined above), 24 parents did not provide consent for further assessment, one child did not provide assent for further assessment, two children were excluded due to a parent-reported prior diagnosis of childhood apraxia of speech (CAS), and five were unable to participate for other reasons (e.g. moving out of area).

---

1 Demographic data was obtained from 32 of the 33 participating sites. A total of 1340 children attended 32 participating sites. Demographic data was only collected from participants at the final site based on questionnaire returns (n=13).

A total of 197 children completed assessments in stage 2. Eligibility for participation in stage 3 required that children had delayed phonological development (less than one standard deviation from the typical mean) characterised by the presence of typical, but ongoing phonological error patterns or processes with or without the presence of additional, atypical speech sound errors. Children's speech production eligibility was determined based on the results of their Diagnostic Evaluation of Articulation and Phonology – Phonology subtest (DEAP-Phonology, Dodd, Hua, Crosbie, Holm & Ozanne, 2002) and analysis using the PROPH+ module of Computerized Profiling v 9.7.0 (Long et al., 2008). Children were excluded from stage 3 if they presented with articulation errors only (e.g. children who only had a lisp), had features consistent with CAS, or if they fell below 2 standard deviations from the mean as defined on the Preschool Test of Nonverbal Intelligence (PTONI; Ehrler & McGhee, 2008). Children who met all inclusion criteria, and did not meet any exclusionary criteria, were invited to participate in stage 3 and 97 children were identified as eligible. The parents of two of these children withdrew their consent for ongoing participation in the research. Consequently, stage 3 assessments were undertaken by 95 children. This study reports the results of the 93 children who completed the Polysyllable Preschool Test (Baker, 2013) in stage 3 of the Sound Start Study (two children were unable to complete this task). The children reported from this study came from 29 sites (mean number of participants at each centre = 3.21; range = 1-10).

Insert Figure 1 here

Participants

Demographic characteristics. The 93 children in this study were aged between 4;0-5;5 (*M* = 55.6 months, *SD* = 4.3 months). There were 58 males (62.4%) and 35 females

(37.6%). Participants' IRSAD deciles[2] ranged from 1 to 10 with a mean decile of 5.72 (*SD* = 3.3). Postcode and IRSAD information was not available for five participants. The parents of 88 children (94.6%) reported that English was the only language spoken by the child in the home whereas the parents of five children (5.4%) reported that a language other than English was spoken at home. All children had an identified phonologically-based SSD of unknown origin.

Developmental characteristics. Receptive and expressive language, hearing, oromotor, nonverbal intelligence, and receptive vocabulary were assessed and characteristics are detailed in Appendix A. The participants demonstrated a mean receptive vocabulary score based on the Peabody Picture Vocabulary Test – Fourth Edition (Dunn & Dunn, 2007) of 93.9 (*SD* = 14.6) and mean non-verbal intelligence of 100.3 (*SD* = 17.3) based on the PTONI (Ehrler & McGhee, 2008).

Parent reported case history information. The parents were invited to complete a questionnaire during stage 2. The parents of 86 (92.5%) children returned questionnaires. Two (2.4%) children were reported to have a history of hearing difficulties and six (7.1%) children had previously failed a hearing test. Positive family histories of speech, language, literacy or hearing difficulties were identified: five (6.5%) had mothers, eight (10.5%) had fathers, 14 (18.4%) had sisters, 12 (16.2%) had brothers, and nine (12.2%) had cousins with a history of speech, language, literacy or hearing difficulties.

Instruments

Children's speech sound production skills were assessed using the Phonology subtest from the Diagnostic Evaluation of Articulation and Phonology (DEAP-Phonology, Dodd et al., 2002). The DEAP-Phonology contains a total of 50 words; 27 x one-syllable words, 18 x

[2] Within the IRSAD coding system, postcodes throughout Australia are allocated a decile between 1-10 to represent the social and economic context of residents within that postcode. Areas with a low index score indicate an area of least advantage whereas areas with a higher index score represent areas of most advantage.

two-syllable words, four x three-syllable words, and one x four-syllable word. The DEAP-Phonology allows for phonological assessment of children aged 3;0-6;11 and is used by SLPs to assess children's production of consonants, vowels, and use of phonological processes. The DEAP-Phonology provides normative data for PCC as well as percent vowels correct (PVC), and percent phonemes correct (PPC).

Children's productions of polysyllables were measured using the Polysyllable Preschool Test (POP, Baker, 2013). The POP is a single-word picture-naming task that contains a total of 30 words: 20 x three-syllable words, eight x four-syllable words, and 2 x five-syllable words. The POP word list represents a variety of weak and strong onset stress patterns at each word length. The stimuli for the POP were presented via Powerpoint™ slides on a computer screen with one- or two- stimuli pictures per slide.

Procedure

Approval to undertake this study was gained from the Charles Sturt University Ethics in Human Research Committee (approval number 2013/070). Consent was gained from the parents and assent was gained from the children prior to undertaking the assessments. Speech production data, including the DEAP-Phonology and the POP were collected during stage 2 and stage 3 respectively (see Figure 1). Spontaneous production of the target words during the DEAP and the POP was encouraged; however, if prompting was required, cues were provided in the following order: (1) semantic, (2) binary choice, and (3) direct imitation. Each assessment during stage 2 and 3 (see Appendix A) was conducted in a quiet room within their early childhood centre by one of two speech-language pathologists. Assessments at stage 2 were completed in 30-45 minutes and assessments at stage 3 were completed in 45-60 minutes. All assessments were completed, and scored, as described in the assessment manual, relevant journal article or, in terms of the POP, following training with the author of the test (Baker, 2013). All assessment sessions were recorded on a Panasonic HC-V700 video camera with external Hahnel Mk100 uni-directional microphone and speech production tasks were also recorded using Zoom H1 audio recorders.

Online broad phonetic transcription was completed for all DEAP-Phonology (Dodd et al., 2002) samples and checked by the assessing speech-language pathologist at the end of the assessment session. Transcription of the POP (Baker, 2013) was completed by the first author from de-identified audio recordings using Phon v 1.6.2 (Rose et al., 2006). Transcriptions were checked with the audio recording as many times as necessary using Sennheiser HD429 stereo headphones.

Reliability

Inter- and intra-judge point-by-point reliability was completed for the DEAP-Phonology with the two speech-language pathologists involved in data collection. Reliability for the POP was completed by the first and third authors. Reliability for each measure was based on a randomly selected 10% sample of the recorded data. Based on 2104 data points of the DEAP-Phonology, intra-judge agreement for broad phonetic transcription was 89.4% and inter-judge agreement was 88.7%. In cases of mismatches, the decision of the speech-language pathologist who completed the initial transcription was upheld. Based on 1966 data points of the POP, intra-judge agreement for broad phonetic transcription was 91.3% and inter-judge agreement was 87.3%. In cases of mismatches, the decision of the first author was upheld. In a study of the sources of variance which may affect reliability in phonetic transcription, Shriberg and Lof (1991) described “acceptable agreement” as >85% (p. 255).

Data analysis

Aim 1. Analysis of the DEAP-Phonology speech samples was undertaken using the PROPH+ module of Computerised Profiling v 9.7.0 (Long et al., 2008). The DEAP-Phonology samples were analysed to identify PCC, PVC, and PPC. PPC for the DEAP samples was manually calculated based on PCC and PVC raw scores. A review of all

PROPH+ output files did not highlight any words that were not analysed due to errors in processing. Manual PCC, PVC, and PPC calculations of the POP (Baker, 2013) were completed using consistent scoring rules to those used within the PROPH+ software where any phoneme omission, deletion, substitution or distortion was recorded as incorrect. The number of correct consonants, vowels, and total phonemes were then calculated as a percentage of the target phonemes in the sample. The participants' PCC, PVC, and PPC as measured using the DEAP-Phonology and the POP samples were entered in to SPSS version 21.0 (IBM, 2012). Descriptive statistics and frequency distributions for all variables were extracted.

Aim 2. Analyses of children's polysyllable productions on the POP (Baker, 2013) were completed manually by the first author using the WAP (Masso, 2016a; see Supplementary Appendix A). The WAP is an item-based analysis tool design to capture categories and sub-categories or error in children's productions of polysyllables. The error categories were based on the frequently-identified polysyllable errors present in the speech production of children with typically developing speech and language (James, 2006) and consideration of speech errors present in children with SSD of unknown origin (i.e. substitution errors and distortion errors).

The main categories and sub-categories on the WAP are as follows: (1) Substitution of consonants or vowels: including number of consonants and vowels substituted (e.g. calculator /ˈkælkjəleɪtə/ → [ˈtætəleɪtə] was coded as 2 instances of consonant substitution), (2) Deletion of syllables, consonants or vowels: including the presence, and number of occurrences, of syllable, vowel and consonant deletion in each word position (e.g. butterfly /ˈbʌtəflaɪ/ → [ˈbʌflaɪ] was coded as 1 instance of within word syllable initial consonant deletion, and 1 instance of within word vowel deletion. The deletion of the syllable was also coded a 1 instance of within word weak syllable deletion), (3) Distortion of consonants or vowels: including number of vowels distorted (including vowels of excessive length) and consonants distorted (including misarticulations such as lateralisation) (e.g. tyrannosaurus /təˈɹænəsɔɹəs/ → [təˈɹænəɬɔːɹəɬ] was coded as 2 instances of consonant distortion), (4) Addition of consonants or vowels: including number of consonants and vowels added to the target word (e.g. echidna /əˈkɪdnə/ → [kəˈkɪdnə] was coded as 1 instance of consonant addition), (5) Alteration of phonotactics: including alterations to word length (number of syllables) and shape (total phonemes present) (e.g. helicopter /hɛləˈkɒptə/ → [ˈhʌk.tə] was coded as 5 of 9 phonemes present), (6) Alteration of timing: including inaccuracy of stress, weak- and strong- onset inaccuracy, and/or presence of syllable segregation (e.g. butterfly /ˈbʌtəflaɪ/ → [ˈbʌ.ˈtʌ.ˈflaɪ] was coded as incorrect stress and syllable segregation). The definition of syllable segregation was consistent with that used by Murray, McCabe, Heard, and Ballard (2015). That is, syllable segregation was noted when there were "noticeable gaps between syllables" (p. 47), (7) Assimilation and/or alteration of sequence: included the presence of assimilation and/or alteration of segments within the target word. (e.g. tyrannosaurus /təˈɹænəsɔɹəs/ → [ˈtænəɹæsɔɹəs] was coded as 1 instance of alteration of sequence).

Initial WAP analysis involved the binary identification of all seven main categories as being present or absent for each word of the target word list. Those who demonstrated no presence of the main category were defined as having "no use", presence on 1-11 words as "less common" use, and presence on 12 or more words (40% of total words produced; cf. Hodson, 2006) as "common" use. This relatively high criterion of 40% was in keeping with Hodson and Paden (1991) and was used to meet the need for a criterion that identified and separated errors that were common from errors that were less common. Note, that the presence of different WAP categories of error were not mutually exclusive.

RESULT

All 93 participants completed the DEAP-Phonology (Dodd et al., 2002): 92 produced all 50 words and 1 child produced 49 words. Therefore a total of 4649 tokens from the DEAP-Phonology were analysed. The 93 participants were required to produce 30 polysyllables from the POP (Baker, 2013) and while 70 produced all 30 words, 13 produced 29 words, and 10 produced between 25-28 words. Therefore a total of 2749 tokens from the POP were analysed.

Percentage of consonants, vowels, and total phonemes correct

The first aim was to determine whether there was a significant difference between PCC, PVC, and PPC between a routine measure of speech production and a measure of polysyllable production. The PCCs across the DEAP and the POP ranged from 29.1- 86.5 and 28.1-85.9 respectively. A similar pattern was observed based on the PVC across the DEAP and the POP (range = 67.9-100; 61.8-99.0) and the PPC across the DEAP and the POP (range = 47.3-90.9; 43.9-88.7). Wilcoxon signed-rank tests were conducted to determine whether there was a significant difference between PCC, PVC, and PPC based on the two speech production assessment tasks. Wilcoxon signed-rank tests were selected because the PCC, PVC, and PPC from the DEAP-Phonology was not normally distributed, with PCC skewness of -0.97 (*SE* = 0.25) and kurtosis of 0.94 (*SE* = 0.50), PVC skewness of -1.22 (*SE* = 0.25) and kurtosis of 1.88 (*SE*= 0.50), and PPC skewness of -1.06 (*SE* = 0.25) and kurtosis of 1.27 (*SE*= 0.50). The results of the Wilcoxon signed-rank test are presented in Table I. Table I demonstrates that there was no significant difference between PCC, Z = -0.28, $p = 0.78$ and PPC, Z = -1.70, $p = 0.09$. However, a statistically significant difference in PVC was identified between the two speech production tasks, $Z = -7.39$, $p < 0.00$. The participants demonstrated lower vowel accuracy on the POP compared to the DEAP-Phonology.

Insert Table I here

Analysis of seven categories of polysyllable error

The second aim was to categorise polysyllable errors made by children with phonologically-based SSD and describe children's polysyllable maturity considering these errors. The participants' productions of 30 words from the POP (Baker, 2013) were analysed using the WAP (Masso, 2016a) to identify the presence of seven categories of error. The occurrence of each error category is outlined in Table II and was measured as present or not present in each word. The most commonly occurring error category across the 30 words was substitutions (*M*= 20.19, *SD* = 4.06). In order, the next most commonly occurring error categories were: alteration of phonotactics (*M*= 16.33, *SD* = 5.33), deletions (*M*= 15.18, *SD* = 5.58), alteration of timing (*M*= 13.26, *SD* = 5.29), assimilation and/or alteration of sequence (*M*= 3.73, *SD* = 2.29), additions (*M*= 1.97, *SD* = 1.5), and distortions (*M*= 0.83, *SD* = 1.38). Within the analysis structure of the WAP, each word was analysed based on the presence or absence of the main category and then, the occurrence of each subcategory for any present main category.

Insert Table II here

Category 1: Substitution of consonants or vowels (SUB)

Main category analysis. Substitution of consonants or vowels was a frequent pattern across the participants, with 91/93 (97.8%) participants demonstrating common use of this pattern. The mean occurrence of the substitution of consonants or vowels across the participants' productions of 30 words was 20.19 (*SD* = 4.06, range = 10-28). The top two words that most frequently included the substitution of consonants or vowels were: *tyrannosaurus* and *thermometer* (see Table II).

Subcategory analyses. Substitution of consonants or vowels was further analysed in two different ways by considering the number of consonant substitutions (mean = 25.8) and vowel substitutions (mean = 9.46).

Category 2: Deletion of syllables, consonants or vowels (DEL)

Main category analysis. Deletion of syllables, consonants or vowels was a frequent pattern across the participants, with 61/93 (65.6%) demonstrating common use of this pattern. Deletion of syllables, consonants or vowels was initially measured as a binary variable (present or not present in each word). The mean occurrence of the deletion of syllables, consonants or vowels across the participants' productions of 30 words was 15.18 (*SD* = 5.58, range = 5-28). The top two words that most frequently included the deletion of syllables, consonants or vowels were: *computer* and *barbeque* (see Table II).

Subcategory analyses. Deletion of syllables, consonants or vowels was further analysed in 20 different ways by considering the number and type of syllables deleted (six subcategories), consonants deleted (six subcategories) and vowels deleted (eight subcategories) (see Table III). Consonants were deleted more frequently than vowels with syllable position and stress playing a role in the frequency of deletion. Non-final weak syllables were most frequently deleted. This was reflected in the frequency of word-initial (mean = 2.43) and within word (mean = 2.15) weak syllable deletion as well as word-initial (mean = 3.27) and within word (mean = 2.84) weak vowel deletions. Syllable-initial consonants of weak syllables were the most frequently deleted consonants (mean = 9.12).

Category 3: Distortion of consonants or vowels (DIST)

Main category analysis. Distortion of consonants or vowels was an infrequent pattern across the participants, with none of the participants demonstrating common use of this pattern. The mean occurrence of the distortion of consonants or vowels across the participants' productions of 30 words was 0.83 (*SD* =1.38, range = 0-7). The top two words that most frequently included the distortion of consonants or vowels were: *koala* and *animals* (see Table II).

Subcategory analyses. Distortion of consonants or vowels was further analysed in two different ways by considering the number of distortions of vowels (more frequent) and consonants (less frequent) specifically (see Table III).

Category 4: Addition of consonants or vowels (ADD)

Main category analysis. Addition of consonants or vowels was an infrequent pattern across the participants, with none of the participants demonstrating common use of this pattern. Addition of consonants or vowels was initially measured as a binary variable (present or not present in each word). The mean occurrence of the addition of consonants or vowels across the participants' productions of 30 words was 1.97 (*SD* = 1.57, range = 0-8). The top two words which most frequently included the addition of consonants or vowels were: *hippopotamus* and *spaghetti* (see Table II).

Subcategory analyses. Addition of consonants or vowels was further analysed in two different ways by specifically considering the number of consonants added (more frequent) and the number of vowels added (less frequent) (see Table III).

Insert Table III here

Category 5: Alteration of phonotactics (AP)

Main category analysis. Alteration of phonotactics was a frequent pattern across the participants, with 72/93 (77.4%) demonstrating common use of this pattern. Alteration of phonotactics was initially measured as a binary variable (present or not present in each word). The mean occurrence of the alteration of phonotactics across the participants' productions of 30 words was 16.33 (*SD* = 5.33, range = 5-28). The top two words that most frequently included the alteration of phonotactics were: *kangaroo* and *caterpillar* (see Table II).

Subcategory analyses. Alteration of phonotactics was further analysed in two ways by considering the length and shape of the words (see Table III). Overall, children's word shape was less accurate than word length.

Category 6: Alteration of timing (AT)

Main category analysis. Alteration of timing was a frequent pattern across the participants, with 59/93 (63.4%) demonstrating common use of this pattern. Alteration of timing was initially measured as a binary variable (present or not present in each word). The mean occurrence of the alteration of timing across the participants' productions of 30 words was 13.26 (*SD* = 5.29, range = 1-23). The top two words that most frequently included the alteration of timing were: *hippopotamus* and *computer* (see Table II).

Subcategory analyses. Alteration of timing was further analysed in four ways by considering the stress accuracy, weak onset accuracy, strong onset accuracy, and syllable segregation of the words (see Table III). Overall, children produced stress incorrectly more frequently on weak-onset syllables (mean = 4.97) compared to strong-onset syllables (mean =2.54).

Category 7: Assimilation and/or alteration of sequence (ASEQ)

Main category analysis. Assimilation and/or alteration of sequence was an infrequent pattern across the participants, with none of the participants demonstrating common use of this pattern. Assimilation and/or alteration of sequence was measured as a binary variable (present or not present in each word). The mean occurrence of assimilation and/or alteration of sequence across the participants' productions of 30 words was 3.73 (*SD* = 2.29, range = 0-10). The top two words that most frequently included this category were: *hippopotamus* and *vegetables* (see Table II).

Subcategory analysis. The category of assimilation and/or alteration of sequence (ASEQ) was further analysed in two ways by considering the presence of assimilation and the presence of an alteration of the sequence of segments within the word (see Table III). Both subcategories were represented occasionally with assimilation present more frequently (mean = 2.12) than alteration of sequence (mean = 0.86) (see Table III).

The Framework of Polysyllable Maturity

The Framework of Polysyllable Maturity (Masso, 2016b; see Table IV) describes five levels of polysyllable maturity for the children in this sample. Substitutions were not included as an element of the Framework due to the high prevalence of substitution errors among the participants and the need for a parsimonious scaffold to describe polysyllable maturity in children with phonologically-based SSD. The inclusion of substitutions as a component did not add meaningful variance to the levels as all the children at each level demonstrate frequent substitution errors. The Framework accounts for all children (n = 91) who demonstrated a high frequency of substitution errors (that is, common use of substitutions). The Framework documents five levels of polysyllable maturity based on the presence and/or absence of the remaining three most frequent error categories demonstrated by children with phonologically-based SSD: deletions (with specific interest in the divide between the presence of vowel deletion and consonant deletion), alteration of phonotactics, and alteration of timing. Within the Framework, children were deemed to demonstrate that a category was present when they demonstrated common use of that category. That is, they demonstrated the use of the category on at least 40% of the sampled words. Levels were determined based on previous evidence exploring phonotactic and stress accuracy during the typical development of polysyllables (James, 2006). In Level A, 17/91 (18.7%) children presented with frequent deletion, phonotactic, and timing errors. In Level B, 37/91 (40.7%) children demonstrated phonotactic and timing errors. In Level C 17/91 (18.7%) children demonstrated phonotactic errors and fewer timing errors. The five (5.5%) children in Level D presented with more accurate phonotactics but ongoing alterations of timing. In Level E, 15 (16.5%) children did not demonstrate deletion, phonotactic or timing errors and only demonstrated substitution errors. Thus, 71/91 (78.0%) children who presented with common substitution errors also presented with frequent errors in phonotactic structure.

Insert Table IV here

DISCUSSION

This study aimed to explore the polysyllable productions of preschool children with phonologically-based SSD of unknown origin through a comparison of speech accuracy on a routine measure of speech production (DEAP; Dodd & Dodd, 2002) and a polysyllable task

(POP; Baker, 2013) as well as a comprehensive analysis of the categories of error present in children's speech. Results from this study indicate that the use of PCC alone as a measure of speech accuracy, particularly when based on a routine speech sampling task with few polysyllables, may mask qualitative differences between children in a heterogeneous sample of children with phonologically-based SSD of unknown origin. The results will be discussed in the following sections with regard to each of the research aims and within the context of the broader literature.

Consonant, vowel and total phoneme accuracy

The analysis presented in this study highlights that children have lower vowel accuracy when saying polysyllables compared to a routine speech sampling task which includes few polysyllables. This finding suggests that vowel accuracy may be underestimated when using picture naming tasks that predominantly consist of mono- and disyllables. Thus, this finding demonstrates the contribution of polysyllable sampling tasks to tax children's phonological system, particularly for vowels, beyond what is possible based on routine speech sampling tasks. Therefore, polysyllables could be considered an essential component of any assessment battery to explore the range of possible phonological errors in preschool children with SSD. This confirms and extends the work of previous authors who have advocated for the use of polysyllables in assessment (Baker & Munro, 2011; Gozzard et al., 2006; James, 2006; James et al., 2008, Kehoe & Steol-Gammon, 1997; Mason et al., 2015; Young, 1991). In addition to the importance of sampling polysyllables, the results also highlight the importance of polysyllable measures moving beyond PCC, PVC, and PPC alone to account for qualitative difference in children with SSD.

Children's PCC was not significantly different between the DEAP-Phonology (Dodd et al., 2002) and the POP (Baker, 2013) in this study. A lack of significance is surprising, particularly because previous reports suggested that polysyllables are more prone to error than mono- and di-syllables (James, 2006). Thus, based on these PCC or PPC data alone, we run the risk of concluding that there is no significant advantage in completing a polysyllable sampling task with children who have SSD. However, the significant difference in vowel accuracy between the DEAP-Phonology and the POP demonstrates the value of polysyllable sampling with children who have SSD of unknown origin.

It is widely agreed that children's paradigmatic production of vowels (that is, vowels produced in isolation) is achieved by approximately 3 years of age (Allen & Hawkins, 1980; Smit, Hand, Freilinger, Bernthal, & Bird, 1990). Whereas, children's syntagmatic productions (that is, vowels produced in context) continue to develop beyond the preschool years, particularly in non-final weak syllables (James et al., 2008; James et al., 2016). The prominent vowel errors demonstrated by our sample of children were to those vowels within non-final weak syllables. This highlights the need for polysyllables to be sampled in assessment to allow for an adequate number of opportunities for children to produce, or attempt, non-final weak syllables. Kirk and Vigeland (2015) suggested that a minimum of four opportunities for each error pattern should be sampled when assessing the phonological capabilities of children. Through polysyllable sampling, there are an increased number of possible occurrences of non-final weak vowel and/or total syllable deletion, as well as atypical structural errors (e.g. deletion of stressed vowels).

Seven categories of error are present in the polysyllable productions of preschool children with SSD of unknown origin

In this study, we demonstrated a comprehensive method of analysis of children's productions of polysyllables (using the WAP; Masso, 2016a) that was developed after consideration of categories of error identified in children with typically developing speech and language (James, 2006) in addition to two categories of error (distortions and substitutions) which frequently occur in the speech of children with SSD (van Riper, 1939).

In the following sections, the errors made by children in the POP will be discussed in order of frequency. Errors are discussed in terms of the observations of this analyses compared with previous literature based on children's speech accuracy and ongoing considerations regarding how each category may reflect the specification of children's underlying phonological representations.

Using the WAP, the most frequent category of polysyllable error present within this sample of children was substitution of consonants or vowels. This is not surprising due to the frequency of substitution errors in typically developing preschool children (Dodd, Holm, Hua & Crosbie, 2003; Hodson & Paden, 1981; James et al., 2016) and children with SSD (McLeod, Harrison, McAllister, & McCormack, 2013; Shriberg et al., 2010) as well as the number of systemic phonological processes that affect segmental accuracy (e.g. fronting and stopping). The high frequency of substitution errors in this sample does further highlight that children with phonologically-based SSDs are likely to have poor underlying phonological representations of polysyllables which is supported by the previous observations of Anthony and colleagues (2011). The high frequency of this category highlights the importance of ongoing consideration for the substitution errors demonstrated in children's speech, whether speech is sampled in mono-, di- or polysyllables.

The second most frequent category present within this sample was alteration of phonotactics (i.e. addition or deletion of singleton consonants, consonant clusters or vowels). This work extends previous research by Mason et al. (2015) about the phonotactic accuracy on a sample of polysyllables produced by children with SSD. The high frequency of this category reflects the significant number of deletion errors demonstrated by the children in this sample and thus, may also be a measure of the accuracy/inaccuracy of children's phonological representations due to deletions. Shriberg et al. (2005) theorised that deletion errors may be indicative of weak or missing phonological representations for the target consonant. The use of a phonotactic measure of polysyllables may be useful to monitor overall improvement in word length and shape, in addition to the reduction of deletion errors following intervention. The frequency of this category of error highlights the importance of word-level analysis for children with SSD (e.g. using the WAP, Masso, 2016a, or the method of polysyllable measurement grounded in non-linear phonological theory presented by Mason et al., 2015) in addition to analyses based on phonological processes.

The third most frequent error category demonstrated was the deletion of syllables, consonants or vowels. The current research quantifies the high frequency of polysyllable deletion errors in children with SSD of unknown origin (65.6% of all children demonstrating common use). This high rate of deletion errors is consistent with previous research reporting deletion frequency of preschool-aged children with SSD (e.g. Rvachew et al., 2007). As mentioned previously, children's deletion of syllables, consonants or vowels in polysyllables may imply poorly specified underlying representations (Anthony et al., 2011). Thus, the children in this sample who demonstrated deletions, particularly of whole syllables, may have more poorly specified stored representations than those who did not demonstrate deletions. These data also support the notion that children's phonological system may be more taxed during the production of polysyllables and may manifest as a larger number of deletion and substitution errors in children with poorer phonological representations. Thus, children who present with fewer polysyllable errors, particularly fewer non-final weak syllable deletions, have more mature polysyllable productions than those children with frequent non-final weak syllable deletions. These data lay the foundation for further exploration of the maturity of children's polysyllable productions and the possible relationship between polysyllable maturity, phonological processing and emergent literacy skills.

Alteration of timing was the fourth most frequent error category and implies the ongoing challenges that children with SSD face achieving stress accuracy in polysyllables. Until now, very little was known about the accuracy of stress within the polysyllable productions of preschool children with SSD. Interestingly, the sample presented in this study demonstrated more prominent stress inaccuracy than the clinical population of younger children with SSD (mean age of 3;10 years) reported by Vick and colleagues (2014). The sample of younger children demonstrated stress inaccuracy of 29.5% on a sample of five repetitions of six disyllabic words. The stress inaccuracy reported by Vick et al. (2014) represents markedly less stress inaccuracy than the 45.2% demonstrated by this sample although this discrepancy is likely due to the increased stress complexity of polysyllables over disyllables, and the larger speech sampling task used in the current study. Interestingly, children with typically developing speech and language who demonstrate more phonotactic accuracy (as measured by fewer consonant and vowel deletions) may temporarily achieve reduced stress accuracy as their polysyllable productions mature (James, 2006; James et al., 2008).

The three other error categories that were identified within the POP speech samples occurred infrequently: addition of consonants or vowels, distortion of consonants or vowels, and assimilation and/or alteration of sequence. The infrequent or less common occurrence of vowel and consonant additions (e.g. epenthesis) and alterations in sequence (e.g. metathesis) are in line with previously reported infrequency of these error patterns in children with typically developing speech and language (James et al., 2016; James, 2006; McLeod, van Doorn & Reed, 2001) and children with SSD (McLeod, van Doorn & Reed, 1997). Similarly, the infrequent demonstration of distortion of consonants or vowels was also unsurprising as distortions are more characteristic of children with articulation-based SSD (Shriberg et al., 2010).

The Framework of Polysyllable Maturity highlights qualitative difference in children with SSD

Substitution errors featured prominently in the polysyllable productions of children with phonologically-based SSD. Ninety-one of the 93 (97.8%) children in this sample presented with frequent substitution errors. The Framework (Table IV) was developed to describe, interpret, and classify the polysyllable productions of these children. The Framework highlights the qualitative differences that exist within a sample of children with phonologically-based SSD when polysyllable maturity is considered based on deletion errors, alterations of phonotactics, and alterations of timing. Substitution errors can mask analysis and clinical judgement about children's polysyllables when considering broader phonological development. Clinically, comprehensive word-level analysis of polysyllables may assist with ongoing monitoring of polysyllable maturity and generalisation of intervention targets across to non-treatment linguistic elements.

Polysyllable acquisition described by James (2006), and James and colleagues (2008), accounted for changes to the frequency of deletion errors, alterations of phonotactics, and alterations of timing, but did not explicitly account for substitution errors (beyond those expected in typical development and those which demonstrate the preservation of phonotactics). Thus, the Framework can be used to classify polysyllable accuracy observed in children with typically developing speech and language (James, 2006; James et al., 2008) as well as for children with SSD.

When considering the children in this study, a large proportion of children described within the Framework (71/91 children; 78.0%) demonstrated common errors in phonotactic structure. This high frequency of phonotactics errors highlights the wide-spread inadequacy in the polysyllable productions of children with SSD and the possible insight that polysyllables may provide in understanding children's phonological representations. For example, that children in Level A, who demonstrate the least mature realisation of polysyllables (in addition to frequent substitutions), may have more poorly specified phonological representations than those represented in the later levels of polysyllable maturity. Thus, if stored phonological representations form part of the foundations for children's phonological awareness and emergent literacy development, these children with the poorest specified polysyllables (those in Level A) may be at most risk of emergent literacy difficulties. This idea requires more investigation and evidence-based support.

Future research

Based on this preliminary work, further research using the WAP and the Framework is indicated to determine whether children's polysyllable maturity may also be linked to long-term speech prognosis, severity of disorder, response to treatment or considerations for treatment planning. More work is required before polysyllables could be used as a primary diagnostic tool for preschool children with suspected SSD. The recent findings of James et al. (2016) suggest that children with typically developing speech and language present with frequent cluster reduction and substitution errors (i.e. fricative simplification) within polysyllables up until at least the age of seven years old. Thus, in preschool, children with typically developing speech and language also demonstrate high frequency of errors when saying polysyllables (James et al., 2016). The findings of the current investigation pave the way for future research regarding the diagnostic value of polysyllable sampling and analysis in identifying preschool children with poor polysyllable maturity and those who may be most at-risk of future literacy difficulties. Further research should examine the clinical profiles of preschool children with respect to their level of polysyllable maturity and emergent literacy abilities in preschool as well as children's literacy acquisition during the school years. In light of previous research (e.g. Preston et al., 2013; Rvachew et al., 2007), future research should also consider children's atypical speech errors when exploring links between polysyllable accuracy and literacy development. The data analysed using the WAP (Masso, 2016a) was only based on one single-word assessment task at one point in time. It may also be valuable to complete the WAP on children's polysyllable samples across a number of time-points to explore the development of polysyllable accuracy and maturity in preschool children with SSD and later development when children reach school age. It would also be valuable to compare children's abilities to produce polysyllables in single-word versus connected speech contexts and changes in their polysyllable accuracy between sampling contexts, over time.

Limitations

There are three primary limitations to this study. First, many of the children in our sample performed poorly on the oromuscular function task. Eighty-eight (96.7%) participants demonstrated oromotor function outside the typical range based on the protocol described by Robbins and Klee (1987). This poor performance is likely due to the co-ordinated speech movement component of the test (including the repetition of real words) that accounted for most of the variance in the original reported sample (Robbins & Klee, 1987) and targets speech production accuracy. Children with delayed phonological development, as represented in our sample, perform poorly on single-word repetition tasks and this is reflected in their oromuscular function scores. Future studies should consider the bias of the Robbins and Klee (1987) protocol for children with phonologically-based difficulties in the absence of motor planning or motor control difficulties.

A further limitation of the current method of analysis is the implementation of a consistent criterion for the presence of each main error category. In this case, a threshold of 40% was selected to indicate when children demonstrated the common use of an error (or category of error). This consistent criterion, used across all categories regardless of expected frequency of errors, may reduce the sensitivity of the WAP to identify children who demonstrate certain types of low-frequency errors (e.g. metathesis).

The polysyllable sampling task (POP; Baker, 2013), and analysis method (WAP; Masso, 2016a), have not been tested in any other contexts to date. The use polysyllable sampling tasks with more robust item selection (i.e. reaching a threshold item discrimination score as described by James, 2001) and tested content validity may produce varying results to those presented here. Similarly, without a normative sample of the POP, it is difficult to determine the threshold at which errors present on this test may indicate clinically meaningful data. Finally, the, deletion errors (Category 2, Deletion of syllables, consonants or vowels), addition errors (Category 4, Addition of consonants or vowels) and phonotactic errors (Category 5, Alteration of phonotactics) occurrences are not mutually exclusive. That is, through the deletion of syllables and/or the addition or deletion of consonants and vowels, children change the phonotactic structure of polysyllables. It was reasoned that identification of all three categories would be helpful in discerning whether individual children demonstrate poor phonotactic structure in general, or, if they have more specific issues involving addition or deletion. Further research would be needed to determine if and/or how each of these three measures offer unique insights into children's productions of polysyllables.

Conclusion

The current research has clinical implications for the routine assessment and analysis of the speech production abilities of preschool children with SSD of unknown origin. These data highlight the usefulness of polysyllables when assessing children with SSD to gain a broader understanding of children's speech production accuracy in different speech contexts. Through comprehensive polysyllable analysis of children's polysyllable errors (e.g. using the WAP), greater insight into the maturity of children's phonological systems beyond segmental accuracy can be gained and interpreted using the Framework.

Different methods of sampling and analysing speech production result in different conclusions about the severity of children's speech difficulties and the importance of different types of errors within children's speech. This study confirms the recommendations of previous authors that polysyllables be included in assessments of children's speech, because they are clinically valuable to tax children's phonological systems. Measures of PCC, PVC, and PPC alone may mask broader phonological differences between preschool children with SSD. Further, greater insight into children's polysyllable errors and level of polysyllable maturity can be gained through more in-depth qualitative analysis of error types.

ACKNOWLEDGMENTS

This research was supported by Australian Research Council Discovery Grant (DP130102545). The first author acknowledges support from an Australian Postgraduate Award scholarship from the Department of Education. The authors thank Yvonne Wren, Sue Roulstone, Kate Crowe, Charlotte Howland, and Felicity McKellar who provided support with the design of the study, data collection, and entry.

REFERENCES

Allen, G., & Hawkins, S. (1980). Phonological rhythm: definition and development. In G. Yeni-Komshian, J. Kavanagh, & C. Ferfuson (Eds.), *Child Phonology, Volume 1* (pp. 227-256). New York, NY: Academic Press.

Anthony, J. L., Aghara, R. G., Dunkelberger, M. J., Anthony, T. I., Williams, J. M., & Zhanga, Z. (2011). What factors place children with speech sound disorders at risk for reading problems? *American Journal of Speech-Language Pathology, 20*, 146–160.

Anthony, J. L., Williams, J. M., Aghara, R. G., Dunkelberger, M., Novak, B., & Mukherjee, A. (2010). Assessment of individual differences in phonological representation. *Reading and Writing: An Interdisciplinary Journal, 23*(8), 969-994.

Australian Bureau of Statistics (2008). *An introduction to Socio-Economic Indexes for Areas (SEIFA) 2006*. Canberra, Australia: Australian Bureau of Statistics.

Baker, E. (2013). *Polysyllable Preschool Test.* Sydney, Australia: Author.

Baker, E., & Munro, N. (2011). An overview of resources for assessing toddlers' productions of polysyllables. *ACQuiring Knowledge in Speech, Language, and Hearing, 13*(2), 58-62.

Broomfield, J. & Dodd, B. (2004). The nature of referred subtypes of primary speech disability. *Child Language Teaching and Therapy, 20*, 135-151.

Dodd, B. (1995). *Differential diagnosis and treatment of children with speech disorder*. London, UK: Whurr.

Dodd, B., & Crosbie, S. (2005). A procedure for classification of speech disorders. In B. Dodd (Ed.), *Differential diagnosis and treatment of children with speech disorder* (pp. 140-162). London, UK: Whurr.

Dodd, B., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: A normative study of British English-speaking children. *Clinical Linguistics and Phonetics, 17*(8), 617-643.

Dodd, B., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2002). *Diagnostic Evaluation of Articulation and Phonology (DEAP)*. London, UK: Psychological Corporation.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4)*. Bloomington, MN: Pearson Assessment.

Ehrler, D. J., & McGhee, R. L. (2008). *Primary Test of Nonverbal Intelligence*. Austin, TX: Pro-Ed.

Eisenberg, S. L., & Hitchcock, E. R. (2010). Using standardized tests to inventory consonant and vowel production: A comparison of 11 tests of articulation and phonology. *Language, Speech, and Hearing Services in Schools, 41*(4), 488-503.

Elbro, C. (1996). Early linguistic abilities and reading development: A review and a hypothesis. *Read Writing, 8*(6), 453-485.

Fowler, A. (1991). How early phonological development might set the stage for phoneme awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Lieberman* (pp. 97–117). Hillsdale, NJ: Erlbaum.

Fowler, A. E., & Swainson, B. (2004). Relationships of naming skills to reading, memory, and receptive vocabulary: Evidence for imprecise phonological representations of words by poor readers. *Annals of Dyslexia*, *54* (2), 247-280.

Glascoe, F. P. (2000). *Parents Evaluation of Developmental Status: Authorized Australian Version*. Parkville, Victoria: Centre for Community Child Health.

Goldman, R., & Fristoe, M. (2000). *The Goldman-Fristoe Test of Articulation – Second edition.* Minneapolis, MN: Pearson Assessments.

Gozzard, H., Baker, E., & McCabe, P. (2006). Children's productions of polysyllabic words. *ACQuiring Knowledge in Speech, Language and Hearing, 8*(3), 113-116.

Haelsig, P.C., & Madison, C. L. (1986). A study of phonological processes exhibited by 3-, 4-, and 5-year-old children. *Language, Speech, and Hearing Services in Schools, 17,* 107-114.

Hodson, B. (2006). Identifying phonological patterns and projecting remediation cycles: Expediting intelligibilty gains of a 7 year old Australian child. *Advances in Speech-Language Pathology, 8*(3), 257-264.

Hodson, B. W., & Paden, E. P. (1981). Phonological processes which characterize unintelligible and intelligible speech in early childhood. *Journal of Speech and Hearing Disorders, 46*(4), 369-373.

Hodson, B. W., & Paden, E. P. (1991). *Targeting intelligible speech: A phonological approach to remediation.* Austin, TX: Pro-Ed.

International Expert Panel on Multilingual Children's Speech. (2012). *Multilingual children with speech sound disorders: Position paper*. Retrieved from: http://www.csu.edu.au/research/multilingual-speech/position-paper.

James, D. G. H. (2001). Use of phonological processes in Australian children ages 2 to 7;11 years. *International Journal of Speech-Language Pathology, 3*(2), 109-127.

James, D. G. H. (2006). *Hippopotamus is so hard to say: Children's acquisition of polysyllabic words* (Doctoral thesis, University of Sydney, Sydney Australia). Retrieved from http://hdl.handle.net/2123/1638

James, D. G. H., Ferguson, W. A., & Butcher, A. (2016, early online). Assessing children's speech using picture-naming: The influence of differing phonological variables on some speech outcomes. *International Journal of Speech-Language Pathology*, doi: 10.3109/17549507.2015.1101159.

James, D. G. H., van Doorn, J., & McLeod, S. (2008). The contribution of polysyllabic words in clinical decision making about children's speech. *Clinical Linguistics and Phonetics, 22*(4-5), 345-353.

Kehoe, M., & Stoel-Gammon, C. (1997). The acquisition of prosodic structure: An investigation of current accounts of children's prosodic development. *Language, 73*(1), 113-144.

Kirk, C., & Vigeland, L. (2015). Content coverage of single-word tests used to assess common phonological error patterns. *Language, Speech, and Hearing Services in Schools*, *46*, 14-29.

Larrivee, L. S., & Catts, H. W. (1999). Early reading achievement in children with expressive phonological disorders. *American Journal of Speech-Language Pathology, 8*(2), 118-128.

Long, S. H., Fey, M. E., & Channell, R. W. (2008). *Computerized Profiling (PROPH).* Retrieved from www.computerizedprofiling.org

Magnusson, E., & Nauclér, K. (1990). Reading and spelling in language-disordered children-linguisitic and metalinguistic prerequisites: Report on a longitudinal study. *Clinical Linguistics and Phonetics*, *44*(1), 49-61.

Magnusson, E., & Nauclér, K. (1993). The development of linguistic awareness in language-disordered children. *First Language*, *13*, 93-111.

Mason, G., Bérubé, D., Bernhardt, B. M., & Stemberger, J. (2015). Evaluation of multisyllabic word production in Canadian English- or French-speaking children within a non-linear phonological framework, *Clinical Linguistics and Phonetics*, *29*(8-10), 666-685.

Masso, S. (2016a). *Word-level Analysis of Polysyllables.* Sydney, Australia: Author.

Masso, S. (2016b). *Framework of Polysyllabic Maturity.* Sydney, Australia: Author.

McLeod, S., & Baker, E. (2014). Speech-language pathologists' practices regarding assessment, analysis, target selection, intervention, and service delivery for children with speech sound disorders. *Clinical Linguistics and Phonetics*, *28*(7-8), 508-531.

McLeod, S., Baker, E. M., McCormack, J. M., Wren, Y. E. & Roulstone, S. E. (2013-2015). *A sound start: Innovative technology to promote speech and pre-literacy skills in at-risk preschoolers.* Australian Research Council (ARC) Discovery Grant (DP130102545).

McLeod, S., Harrison, L., J., McAllister, L., & McCormack, J. (2013). Speech sound disorders in a community study of preschool children. *American Journal of Speech-Language Pathology*, *22*(3), 503-522.

McLeod, S., Harrison, L. J., & McCormack, J. (2012). Intelligibility in Context Scale: Validity and reliability of a subjective rating measure. *Journal of Speech, Language, and Hearing Research*, *55*, 648–656.

McLeod, S., van Doorn, J., & Reed, V. A. (2001). Normal acquisition of consonant clusters. *American Journal of Speech-Language Pathology*, *10*(2), 99-110.

McLeod, S., van Doorn, J., & Reed, V. A. (1997). Realizations of consonant clusters by children with phonological impairment. *Clinical Linguistics and Phonetics, 11*, 85-113.

Murray, E., McCabe, P., Heard, R., Ballard, K. (2015). Differential diagnosis of children with suspected childhood apraxia of speech. *Journal of Speech, Language, and Hearing Research*, *58*(1), 43-60.

Nauclér, K., & Magnusson, E. (1999). Reading and writing development: A longitudinal study from pre-school to adolescence: Status report. *Working Papers, 47*, 169-180.

Preston, J., Hull, M., & Edwards, M. L. (2013). Preschool speech error patterns predict articulation and phonological awareness outcomes in children with histories of speech sound disorders. *American Journal of Speech-Language Pathology*, *22*, 173-184.

Preston, J. L., Ramsdell, H. L., Oller, D. K., Edwards, M. L., Tobin, S. J. (2011). Developing a weighted measure of speech sound accuracy. *Journal of Speech, Language, and Hearing Research, 54*, 1-18.

Roberts, J. E., Burchinal, M., & Footo, M. M. (1990). Phonological process decline from 2;6 to 8 years. *Journal of Communication Disorders, 23*, 205-217.

Robbins, J., & Klee, T. (1987). Clinical assessment of oropharyngeal motor development in young children. *Journal of Speech and Hearing Disorders, 52*(3), 271-277.

Rose, Y., MacWhinney, B., Byrne, R., Hedlund, G., Maddocks, K., O’Brien, P., & Wareham, T. (2006). Introducing Phon: A software solution for the study of phonological acquisition. In D. Bamman, T. Magnitskaia & C. Zaller (Eds.), *Proceedings of the 30th Annual Boston University Conference on Language Development*, (pp. 489-500). Somerville, MA: Cascadilla Press.

Rvachew, S., Chiang, P.-Y., & Evans, N. (2007). Characteristics of speech errors produced by children with and without delayed phonological awareness skills. *Language, Speech, and Hearing Services in Schools, 38*, 60-71.

Shriberg, L. D. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech and Hearing Research, 36*(1), 105-140.

Shriberg, L. D., Fourakis, M., Hall, S. D., Karlsson, H. B., Lohmeier, H. L., McSweeny, J. L., . . . Wilson, D. L. (2010). Extensions to the Speech Disorders Classification System (SDCS). *Clinical Linguistics and Phonetics, 24*(10), 795-824.

Shriberg, L. D., Lewis, B. A., Tomblin, J. B., McSweeny, J. L., Karlsson, H. B., & Scheer, A. R. (2005). Toward diagnostic and phenotype markers for genetically transmitted speech delay. *Journal of Speech, Language, and Hearing Research, 48*(4), 834-852.

Shriberg, L. D., & Lof, G. L. (1991). Reliability studies in broad and narrow phonetic transcription. *Clinical Linguistics and Phonetics, 5*(3), 225-279.

Smit, A., Hand, L., Freilinger, J., Bernthal, J., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*, 779-798.

Swan, D., & Goswami, U. (1997). Picture naming deficits in developmental dyslexia: The phonological representations hypothesis. *Brain and Language, 56*, 334-353.

Vance, M., Stackhouse, J., & Wells, B. (2005). Speech-production skills in children aged 3-7 years. *International Journal of Language and Communication Disorders, 40*(1), 29-48.

Van Riper, C. (1939). *Speech correction: Principles and methods*. Englewood Cliffs, NJ: Prentice-Hall.

Vick, J. C., Campbell, T. F., Shriberg, L. D., Green, J. R., Truemper, K., Rusiewicz, H. L., & Moore, C. A. (2014). Data-driven subclassification of speech sound disorders in preschool children. *Journal of Speech, Language, and Hearing Research, 57*(6), 2033-2050.

Young, E. C. (1991). Analysis of young children's ability to produce multisyllabic English nouns. *Clinical Linguistics and Phonetics, 5*(4), 297-316.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2013). *Preschool Language Scales, Fifth Edition (screening) - Australian and New Zealand Language Adapted Edition (PLS-5)*. Sydney, Australia: Pearson.

Appendix A. Characteristics of the children based on developmental assessments (*n* = 93)

Aspect of communication	**Assessment task**	**Subscale**	**Stage**	**Mean (*SD*)**	**Range**	**WNL *n* (%)**	**Not WNL *n* (%)**	**Valid data**
Expressive and Receptive Language	Preschool Language Scales – Screening test (PLS-5; Zimmerman et al., 2013)	Language total raw score (out of a possible 6)	S2	2.8 (1.6)	0-6	35[a] (38.5%)	56 (61.5%)	91
Receptive Vocabulary	Peabody Picture Vocabulary Test – fourth edition (PPVT, Dunn & Dunn, 2007)	Form A – standard score (possible standard score between 20-160)	S3	93.9 (14.6)	45-123	88[b] (94.6%)	5 (5.4%)	93
Hearing	Pure-tone audiometry (1.5, 1, 2, 4kHz at 40dB)		S2/3	-	-	91 (97.8%)	2 (2.2%)	93
Oromotor skills	Robbins & Klee (1987)	Structure (out of a total possible of 24)	S2	22.3 (1.9)	13-24	74 (81.3%)	17 (18.7%)	91
		Function (out of a total possible of 112)	S2	91.7 (11.0)	29-110	3 (3.3%)	88 (96.7%)	91
Nonverbal intelligence	Primary Test of Nonverbal Intelligence (PTONI, Ehrler & McGhee, 2008)	Nonverbal index (possible standard score between 46-149)	S2	100.3 (17.3)	70-146	93 (100.0%)	0 (0.0%)	93

Note. S2 = Stage 2 of the Sound Start Study; S3 = Stage 3 of the Sound Start Study; WNL = Within normal limits; [a] Achieved the pass criteria for age – total raw score of 4 or more (4-year-olds) or 5 or more (5-year-olds), [b] Achieved a standard score of 70 or above (within 2 SD of mean)

Table I.

Percentage of Consonants, Vowels, and Phonemes (total) Correct on Two Speech Assessments (n = 93) and results of the Wilcoxon signed-rank test.

Measure	DEAP			POP			Z	Sig. (2-tailed)
	Mean	Median	Range	Mean	Median	Range		
Percentage of consonants correct (PCC)	64.8	66.2	29.1-86.5	64.6	64.1	28.1-85.9	-0.28	0.78
Percentage of vowels correct (PVC)	91.1	92.4	67.9-100	82.9	83.8	61.8-99.0	-7.39	**0.00***
Percentage of phonemes correct (PPC)	74.2	75.0	47.3-90.9	72.7	73.5	43.9-88.7	-1.70	0.09

Note. DEAP = Diagnostic Evaluation of Articulation and Phonology – Phonology subtest (Dodd et al., 2002); POP = Polysyllable Preschool Test (Baker, 2013); * significant $p<.001$

Table II.

*Mean Number of Words Displaying an Error Per Category (*n *= 93), and the Number of Participants Showing the Presence of Each Error Category on the Polysyllable Preschool Test.*

Category	Mean number of words displaying an error (max. words = 30)			Number of children showing the presence of each error category			Top 2 words affected
	Mean	*SD*	Range	No use (%)[a]	Less common use 1-11 (%)[b]	Common use 12(%)[c]	
1. Substitution of consonants or vowels	20.19	4.06	10-28	0 (0.0%)	2 (2.2%)	91 (97.8%)	*tyrannosaurus, thermometer*
2. Deletion of consonants or vowels	15.18	5.58	5-28	2 (2.2%)	22 (23.7%)	61 (65.6%)	*computer, barbeque*
3. Distortion of consonants or vowels	0.83	1.38	0-7	53 (57.0%)	40 (43.0%)	0 (0.0%)	*koala, animals*
4. Addition of consonants or vowels	1.97	1.57	0-8	80 (86.0%)	13 (14.0%)	0 (0.0%)	*hippopotamus, spaghetti*
5. Alteration of phonotactics	16.33	5.33	5-28	0 (0.0%)	15 (16.1%)	72 (77.4%)	*kangaroo, caterpillar*
6. Alteration of timing	13.26	5.29	1-23	0 (0.0%)	31 (33.3%)	59 (63.4%)	*hippopotamus, computer*
7. Assimilation and/or alteration of sequence	3.73	2.29	0-10	3 (3.2%)	90 (96.8%)	0 (0.0%)	*hippopotamus, vegetables*

Note. [a] The number of participants (and percentage of total) who demonstrated the no use of the error category, [b] The number of participants (and percentage of total) who demonstrated less common use of the error category (presence on 1-11 words), [c] The number of participants who demonstrated common use of the error category (presence on at least 12 words; at least 40% of words).

Table III.

Mean Occurrence, SD and Range Of Errors on the Polysyllable Preschool Test for Each WAP Subcategory across All Participants (N = 93), the Number of Participants Demonstrating the Error At Least Once and the Top Two Words Affected.

Main category and subcategory	Mean Occurrence	*SD*	Range	Maximum opportunities[a]	Participants[b]	Top 2 words affected
1. Substitution of consonants or vowels (SUB)						
a) Substitution of consonant/s	25.8	8.99	8-54	128	93 (100%)	*spaghetti, tyrannosaurus*
b) Substitution of vowel/s	9.46	4.50	0-20	102	92 (98.9%)	*calculator, cauliflower*
2. Deletion of syllables, consonants or vowels (DEL)						
Total syllable deletion						
a) Word-initial stressed syllables[c]	0.29	0.56	0-2	18	22 (23.7%)	*avocado, escalator*
b) Word-initial weak syllables[c]	2.43	2.57	0-11	12	67 (72.0%)	*thermometer, hippopotamus*
c) Within word stressed syllables[c]	0.18	0.49	0-3	21	14 (15.1%)	*thermometer, tyrannosaurus*
d) Within word weak syllables[c]	2.15	2.50	0-11	21	61 (65.6%)	*hippopotamus, animals*
e) Word-final stressed syllables[c]	0.04	0.20	0-1	13	4 (4.3%)	*spaghetti, potato*
f) Word-final weak syllables[c]	0.27	0.59	0-3	17	19 (20.4%)	*thermometer, cauliflower*
Consonant deletion						
g) Presence of consonant deletion[d]	14.2	5.54	4-28	30	93 (100%)	*computer, calculator*
h) Total number of consonants deleted	18.8	8.50	5-41	128	93 (100%)	*computer, calculator*
i) Syllable-initial consonant/s of stressed syllables[c]	4.18	2.46	0-13	52	90 (96.8%)	*barbeque, computer*
j) Syllable-initial consonant/s of weak syllables[c]	9.12	5.20	2-26	47	93 (100%)	*calculator, ambulance*
k) Syllable-final consonant/s of stressed syllables[c]	3.61	1.89	0-8	15	92 (98.9%)	*vegetables, calculator*
l) Syllable-final consonants of weak syllables[c]	2.08	1.60	0-8	12	81 (87.1%)	*computer, elephant*
Vowel deletion						
m) Presence of deletion of vowels[d]	6.69	5.19	0-20	30	91 (97.8%)	*tyrannosaurus, thermometer*
n) Total number of vowels deleted	7.24	5.70	0-22	102	91 (97.8%)	*tyrannosaurus, thermometer*
o) Word-initial stressed vowel[c]	0.38	0.67	0-3	18	27 (29.0%)	*avocado, hippopotamus*
p) Word-initial weak vowel[c]	3.27	2.85	0-11	12	78 (83.9%)	*tyrannosaurus, thermometer*
q) Within word stressed vowel/s[c]	0.24	0.56	0-3	21	17 (18.3%)	*tyrannosaurus, thermometer*
r) Within word weak vowel/s[c]	2.84	2.96	0-12	21	73 (78.5%)	*animals, hippopotamus*
s) Word-final stressed vowel[c]	0.09	0.28	0-1	13	8 (8.6%)	*koala, barbeque*

t) Word-final weak vowel[c]	0.43	0.68	0-3	17	32 (34.3%)	*cauliflower, thermometer*
3. Distortion of consonants or vowels (DIST)						
a) Distortion of consonants	0.22	0.76	0-5	128	9 (9.7%)	*echidna, ambulance*
b) Distortion of vowels	0.66	1.14	0-6	102	35 (37.6%)	*koala, barbeque*
4. Addition of consonants or vowels (ADD)						
a) Addition of consonants	1.7	1.50	0-7	-	71 (76.3%)	*spaghetti, echidna*
b) Addition of vowels	0.7	1.0	0-4	-	42 (45.2%)	*echidna, thermometer*
5. Alteration of phonotactics (AP)						
a) Length incorrect	7.88	5.64	0-22	102	91 (97.8%)	*medicine, tyrannosaurus*
b) Shape incorrect	25.8	13.7	6-60	230	93 (100%)	*calculator, ambulance*
6. Alteration of timing(AT)						
a) Stress incorrect	12.2	5.14	1-23	30	93 (100%)	*thermometer, echidna*
b) Weak onset incorrect[c]	4.97	2.79	1-11	12	91 (97.8%)	*tyrannosaurus, echidna*
c) Strong onset incorrect[c]	2.54	2.01	0-8	18	84 (90.3%)	*calculator, butterfly*
d) Syllable segregation	3.75	3.07	0-13	30	83 (89.2%)	*calculator, hippopotamus*
7. Assimilation and/or Alteration of sequence (ASEQ)						
a) Assimilation	2.12	1.98	0-11	30	77 (82.8%)	*hippopotamus, vegetables*
b) Alteration of sequence	0.86	1.11	0-4	30	43 (16.1%)	*spaghetti, escalator*

Note. [a]Maximum number of possible occurrences of each category (except ADD) based on the POP word list, there was no maximum number of opportunities for the main category of 'addition of consonants or vowels (ADD)', [b]Number of participants demonstrating the pattern on at least one occasion, [c]This variable was only present for words which included this linguistic element, hence the lower number of opportunities for this error type, [d]In the main category of 'deletion of syllables, consonants or vowels (DEL)' consonant/vowel deletion was identified as both present or absent per word, as well as the number of across all vowels/consonants in the POP word list.

Table IV.

The Framework of Polysyllable Maturity (Masso, 2016b) applied to Children in the Present Study who Demonstrated Substitution Errors (SUB) on at Least 40% of Words (n = 91)

Maturity	Description	DEL		AP	AT	Number of children
		Vowel	Cons			
Least ⇕	A. Vowel and consonant deletions, alterations of phonotactics, and alterations of timing.	+	+	+	+	17 (18.7%)
	B. Alterations of phonotactics due to consonant deletions or other (e.g. additions) although some vowel deletions may be present. Alterations of timing are ongoing.	-	+/-	+	+	37 (40.7%)
	C. Fewer instances of alterations of timing. Alterations of phonotactics due to consonant deletions or other (e.g. additions) are ongoing.	-	+/-	+	-	17 (18.7%)
	D. Fewer instances of alterations of phonotactics. Alterations of timing present.	-	-	-	+	5 (5.5%)
Most	E. Fewer instances of deletions, alterations of phonotactics or alterations of timing.	-	-	-	-	15 (16.5%)

Note. + indicates the presence of the feature in at least 40% of words, - indicates the absence of the feature in at least 40% of words, +/- indicates the presence or absence of the feature, DEL = Deletion of syllables, consonants or vowels, AP = Alteration of phonotactics, AT = Alteration of timing (stress). Progression of maturity based on James, 2006.

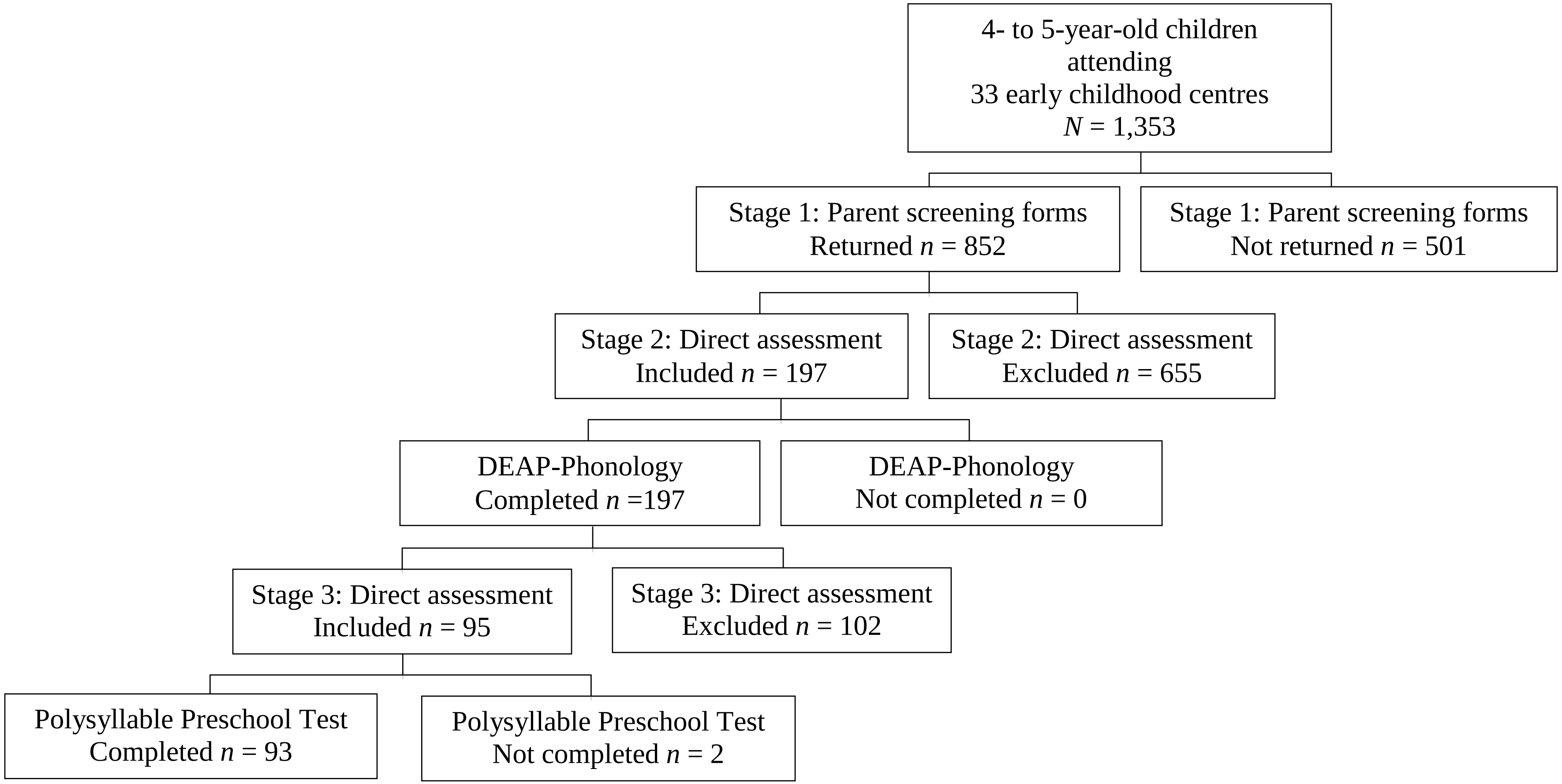
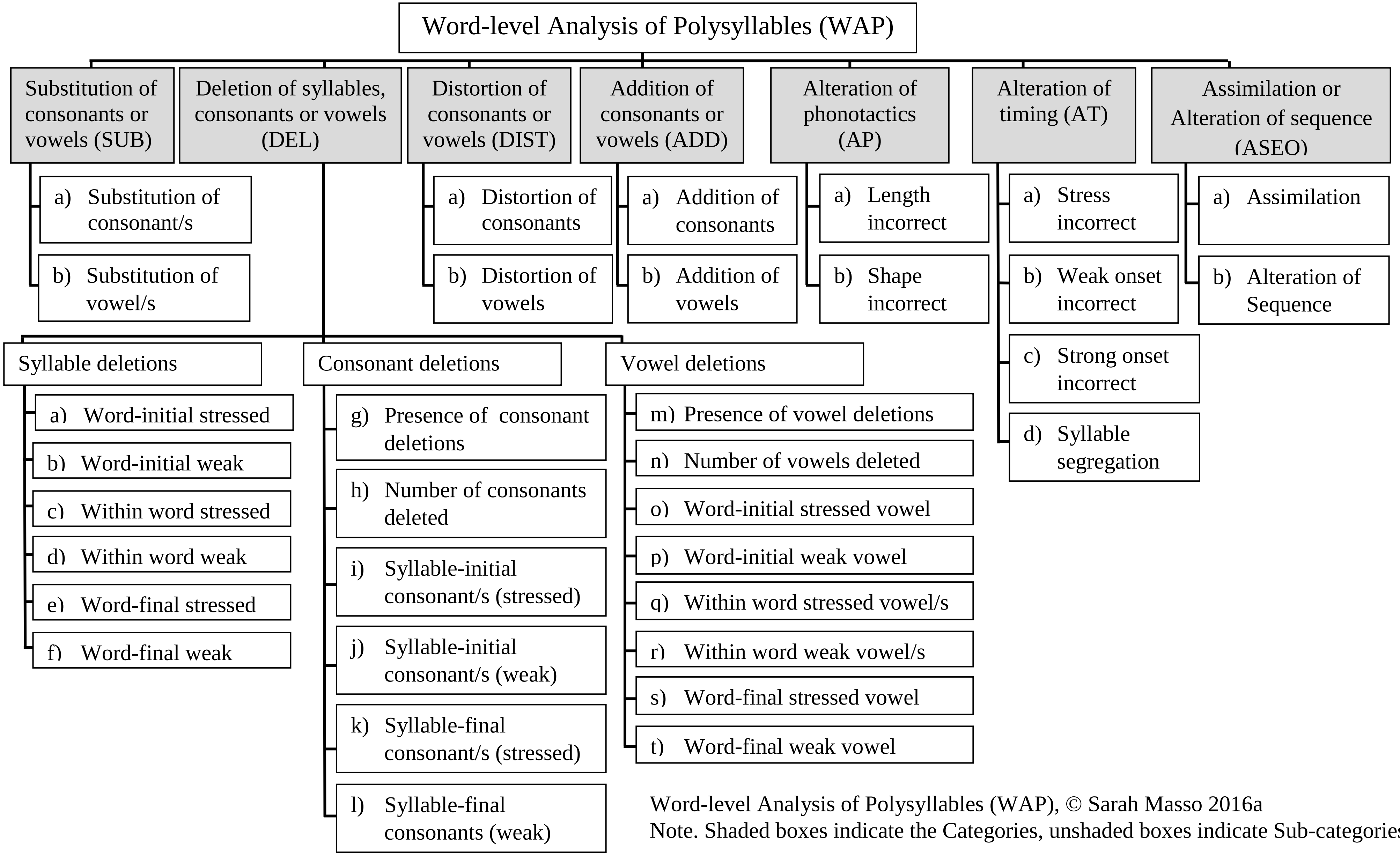


Figure 1. Derivation of participants from years 1 and 2 of the Sound Start Study used in the present investigation. Stage 1 refers to the first stage of the Sound Start Study. Stage 2 refers to the second stage of the Sound Start Study (where only children whose parents and/or teachers reported concerns about their speech sound development were directly screened). Stage 3 refers to the third stage of the Sound Start Study (where only children with delayed phonological development and typical non-verbal intelligence were assessed).

Supplementary Appendix A



Word-level Analysis of Polysyllables (WAP), © Sarah Masso 2016a
Note. Shaded boxes indicate the Categories, unshaded boxes indicate Sub-categories

Word-level Analysis of Polysyllables (WAP)

Item Analysis: Definition and Description of Each Category considered for each word

Category and Subcategories	**Definition**	**Binary or numeric analysis**	**Example**
1. Substitution of consonants or vowels (SUB)	Did this child substitute any vowels or consonants in this word? (if *yes*, complete subcategory analysis)	Yes/no (binary)	*calculator* /'kælkjəleɪtə/ → ['tælkjəleɪtə] coded as "yes"
a) Substitution of consonant/s	Number of consonants substituted	Numeric	*calculator* /'kælkjəleɪtə/ → ['tæpəleɪtə] coded as 2 consonant substitutions
b) Substitution of vowel/s	Number vowels substituted	Numeric	*calculator* /'kælkjəleɪtə/ → ['kælkjuleɪtə] coded as 1 vowel substitution
2. **Deletion of syllables, consonants or vowels (DEL)**	Did this child delete any syllables, consonants or vowels or in this word? (if *yes*, complete subcategory analysis)	Yes/no (binary)	*butterfly* /'bʌtəflaɪ/ → ['bʌflaɪ] coded as "yes"
Syllable deletions			
a) Word-initial stressed syllables [a]	Number of word-initial stressed syllables deleted (max = 1)	Numeric	*butterfly* /'bʌtəflaɪ/ → [tə'faɪ] coded as 1 word-initial stressed syllable deletion
b) Word-initial weak syllables [a]	Number of word-initial weak syllables deleted (max = 1)	Numeric	*pyjamas* /pə'ʤaməz/ → ['ʤaməz] coded as 1 word-initial weak syllable deletion
c) Within word stressed syllables [a]	Number of within word stressed syllables deleted	Numeric	*tyrannosaurus* /tə'ɹænəsɔɹəs/ → [tə'nəsɔɹəs] coded as 1 within word stressed syllable deletion
d) Within word weak syllables [a]	Number of within word weak syllables deleted	Numeric	*hippopotamus* /hɪpəpɒtəməs/ → [hɪ'pɒməs] coded as 2 within word weak syllable deletions
e) Word-final stressed syllables [a]	Number of word-final stressed syllables deleted (max = 1)	Numeric	*butterfly* /'bʌtəflaɪ/ → ['bʌtə] coded as 1 word-final stressed syllable deletion
f) Word-final weak syllables [a]	Number of word-final weak syllables deleted (max = 1)	Numeric	*pyjamas*/pə'ʤaməz/ → [pə'ʤa] coded as 1 word-final weak syllable deletion
Consonant deletions			
g) Presence of consonant deletion	Did this child delete any consonants in this word?	Yes/no (binary)	*butterfly* /'bʌtəflaɪ/ → ['bʌflaɪ] coded as "yes"
h) Total number of consonants deleted	Number of consonants deleted	Numeric	*butterfly* /'bʌtəflaɪ/ → ['bʌtə] coded as 2 consonant deletions
i) Syllable-initial consonant/s of stressed syllables [a]	Number of syllable-initial consonant/s of stressed syllables deleted	Numeric	*butterfly* /'bʌtəflaɪ/ → ['ʌ.tə.faɪ] coded as 2 syllable-initial consonant deletion (stressed syllable)

j) Syllable-initial consonant/s of weak syllables [a]	Number of syllable-initial consonant/s of weak syllables deleted	Numeric	*butterfly* /'bʌtəflaɪ/ → ['bʌflaɪ] coded as 1 syllable initial consonant deletion (weak syllable)
k) Syllable-final consonant/s of stressed syllables [a]	Number of syllable-final consonant/s of stressed syllables deleted	Numeric	*calculator* /'kælkjəleɪtə/ → ['kækjəleɪtə] coded as 1 syllable-final consonant deletion (stressed syllable)
l) Syllable-final consonants of weak syllables [a]	Number of syllable-final consonants of weak syllables deleted	Numeric	*pyjamas*/pə'ʤaməz/ → [pə'ʤamə] coded as 1 syllable-final consonant deletion (weak syllable)
Vowel deletions			
m) Presence of deletion of vowels	Did this child delete any vowels in this word?	Yes/no (binary)	*butterfly* /'bʌtəflaɪ/ → ['bʌflaɪ] coded as "yes"
n) Deletion of vowels	Number vowels deleted	Numeric	*butterfly* /bʌtəflaɪ/ → [bʌflaɪ] coded as 1 vowel deletion
o) Deletion of word-initial stressed vowel[a]	Number of word-initial stressed vowels deleted (max = 1)	Numeric	*butterfly* /'bʌtəflaɪ/ → [tə. 'flaɪ] coded as 1 word-initial vowel deletion (stressed vowel)
p) Deletion of word-initial weak vowel[a]	Number of word-initial weak syllables deleted (max = 1)	Numeric	*echidna* /ə'kɪdnə/ → ['kɪdnə] coded as 1 word-intial vowel deletion (weak vowel)
q) Deletion of within word stressed vowel/s[a]	Number of within word stressed vowel/s deleted	Numeric	*echidna* /ə'kɪdnə/ → [ənə] coded as 1 within word vowel deletion (stressed vowel)
r) Deletion of within word weak vowel/s[a]	Number of within word weak vowel/s deleted	Numeric	*butterfly* /'bʌtəflaɪ/ → ['bʌflaɪ] coded as 1 within word vowel deletion (weak vowel)
s) Deletion of word-final stressed vowel[a]	Number of word-final stressed vowels deleted (max = 1)	Numeric	*butterfly* /'bʌtəflaɪ/ → ['bʌ.tə.] coded as 1 word-final vowel deletion (stressed vowel)
t) Deletion of word-final weak vowel[a]	Number of word-final weak vowels deleted (max = 1)	Numeric	*echidna* /ə'kɪdnə/ → [ə'kɪd] coded as 1 word-final vowel deletion (weak vowel)
3. Distortion of consonants or vowels (DIST)	Did this child distort any vowels or consonants in this word? (if *yes*, complete subcategory analysis)	Yes/no (binary)	*tyrannosaurus* /tə'ɹænəsɔɹəs/ → [tə'ɹænəsˡɔːɹəsˡ] coded as "yes"
a) Distortion of consonants	Number of consonants distorted	Numeric	*tyrannosaurus* /tə'ɹænəsɔɹəs/ → [tə'ɹænəɬɔːɹəɬ] coded as 2 consonant distortions
b) Distortion of vowels	Number of vowels distorted	Numeric	*tyrannosaurus* /tə'ɹænəsɔɹəs/ → [tə'ɹænəɬɔːɹəɬ] coded as 1 vowel distortion
4. Addition of consonants or vowels (ADD)	Did this child add any vowels or consonants in this word? (if *yes*, complete subcategory analysis)	Yes/no (binary)	*echidna* /ə'kɪdnə/ → [ə'kɪdənə] coded as "yes"
a) Addition of consonants	Number of consonants added	Numeric	*echidna* /ə'kɪdnə/ → [kə'kɪdnə] coded as 1 consonant addition
b) Addition of vowels	Number of vowels added	Numeric	*echidna* /ə'kɪdnə/ → [ə'kɪdənə] coded as 1 vowel addition

5. **Alteration of phonotactics (AP)**	Did this child alter the phonotactics of this word? (if *yes*, complete subcategory analysis)	Yes/no (binary)	*helicopter* /hɛlə'kɒptə/ → [hʌ'kɒptə] coded as "yes"
c) Length[b]	Number of syllables present	Numeric	*helicopter* /hɛlə'kɒptə/ → ['hʌk.tə] coded as 2 syllables
d) Shape[b]	Total number of phonemes present	Numeric	*helicopter* /hɛlə'kɒptə/ → ['hʌk.tə] coded as 5 phonemes
6. Alteration of timing (AT)	Did this child alter the timing of the word? (if *yes*, complete subcategory analysis)	Yes/no (binary)	*butterfly* /'bʌtəflaɪ/ → ['bʌ.tə.flaɪ] and *echidna* /ə'kɪdnə/ → [ə'kɪdənə] are both coded as "yes"
a) Stress incorrect	Was the stress incorrect?	Numeric	*butterfly* /'bʌtəflaɪ/ → ['bʌtɜflaɪ] coded as 1 (yes)
b) Weak onset incorrect	Was the stress of the initial syllable (weak) incorrect?	Numeric	*echidna* /ə'kɪdnə/ → ['kɪdnə] coded as 1 (yes)
c) Strong onset incorrect	Was the stress of the initial syllable (strong) incorrect?	Numeric	*butterfly* /'bʌtəflaɪ/ → ['bʌtəflaɪ] coded as 0 (no)
d) Syllable segregation	Did this child segregate the syllables within this word?	Numeric	*butterfly* /'bʌtəflaɪ/ → ['bʌ.tʌ.flaɪ]) coded as 1 (yes)
7. Assimilation and/or Alteration of sequence (ASEQ)	Was assimilation or an alteration of sequence present within this word? (if *yes*, complete subcategory analysis)	Yes/no (binary)	*tyrannosaurus* /tə'ɹænəsɔɹəs/ → ['tænəɹæsɔɹəs] coded as "yes"
a) Assimilation	Was assimilation present?	Numeric	*calculator* /'kælkjəleɪtə/ → ['tætəleɪtə] coded as 1 to indicate the presence of assimilation
b) Alteration of sequence	Did this child alter the sequence of segments?	Numeric	*tyrannosaurus* /tə'ɹænəsɔɹəs/ → ['tænəɹæsɔɹəs] coded as 1 to indicate the presence of an alteration of the sequence

Note. This analysis tool can be used for any polysyllable word list. [a] This variable was only present for words which included this linguistic element (e.g., a variable indicating the deletion of a word-initial stressed vowel was only included for words which contained a word-initial stressed vowel), [b]Number of syllables (length) or phonemes (shape) were calculated based on the number actually produced in each word. If the number of syllables or phonemes was more than the intended target, the additional length was calculated as a negative of the total (e.g., *Echidna* /ə'kɪdnə/ → [ə'kɪdənə] was coded as length of 5 (the possible number of phonemes was 6; minus 1 for the additional syllable). The diacritic /'/ is used to notate primary stress. The creation of categories 1, 2, 3, and 4 were informed by van Riper (1939) and categories 1, 2, 5, 6, and 7 were informed by James (2006).