ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

# Predicting Which Children Will Normalize Without Intervention for Speech Sound Disorders

**Carol Kit Sum To,[a] Sharynne McLeod,[b] Ka Lam Sam,[a] and Thomas Law[c]**

[a] The University of Hong Kong, Hong Kong SAR, China [b] Charles Sturt University, Bathurst, New South Wales, Australia [c] The Chinese University of Hong Kong, Hong Kong SAR, China

ARTICLE INFO

Article History:
Received August 16, 2021
Revision received December 9, 2021
Accepted January 16, 2022

Editor-in-Chief: Bharath Chandrasekaran
Editor: Stephanie Borrie

https://doi.org/10.1044/2022_JSLHR-21-00444

ABSTRACT

**Purpose:** The speech of some children does not follow a typical normalization trajectory, and they develop speech sound disorders (SSD). This study investigated predictive correlates of speech sound normalization in children who were at risk of SSD.
**Method:** A prospective population cohort study of 845 Cantonese-speaking preschoolers was conducted over 2.5 years to examine (a) children who resolved nonadult realizations of consonants (normalized) and (b) those who had persisting speech sound difficulties (did not normalize). From these 845, a sample of 82 participants characterized as having SSD (1.25 *SD*s below the mean in a standardized speech assessment, with a delay in initial consonant acquisition or with one or more atypical errors) was followed for 2 years at 6-month intervals or until the completion of their initial consonant inventory. Data from 43 children who did not receive speech-language pathology services were analyzed with survival analysis to model time to normalization while controlling for covariates. The target event (outcome) was the completion of their initial consonant inventory.
**Results:** Under the no-intervention condition, the estimated median time to normalization was 6.59 years of age. Children who were more likely to normalize or normalized in a shorter time were stimulable to all errors and more intelligible as rated by caregivers using the Intelligibility in Context Scale. Those who showed atypical error patterns did not necessarily take longer to normalize. Similarly, expressive language ability was not significantly associated with speech normalization.
**Conclusions:** Stimulability and intelligibility were more useful prognostic factors of speech normalization when compared to (a)typicality of error patterns and expressive language ability. Children with low intelligibility and poor stimulability should be prioritized for speech-language pathology services given that their speech errors are less likely to resolve naturally.

Children's speech development begins with vegetative sounds (e.g., reflexive cries and burps), followed by more voluntary sound production (e.g., cooing and laughing), and then vocal play that approaches speechlike sounds. Babbling, which is characterized by true consonants and vowels in syllable, emerges before the production of first words (Mitchell & Kent, 1990). Next, children develop their phonological system by adding speech sounds in their inventory. The expansion of their repertoire appears to happen in a time-dependent manner whereby children acquire particular sounds at a certain age level (McLeod & Crowe, 2018; Smit et al., 1990; To et al., 2013). Speech development may appear to be orderly if mastery is simply defined as the age at which the correct production of phonemes in all word positions and contexts is achieved by the majority of children. If orderliness also implies a fixed order of chronological development of sounds, this is clearly not the case. Like other aspects of human

---

Correspondence to Carol Kit Sum To: tokitsum@hku.hk. ***Disclosure:*** *Carol Kit Sum To is a co-author of the Hong Kong Cantonese Articulation Test, used in this study. The other authors have declared that no other competing financial or nonfinancial interests existed at the time of publication.*

development (e.g., motor skill development), individual developmental progression in speech sound development is variable, and the order and timing of emergence of milestones is not necessarily obligatory (McLeod & Crowe, 2018). Some children may skip milestones and produce others first, regress to earlier ones, or master several milestones at the same time, but eventually, they attain the mature status without experiencing negative impact in daily life. Individual differences exist during the course of speech and language development (Bates et al., 1995; McLeod & Hewett, 2008; McLeod et al., 2001; Nelson, 1981; Shore, 1994; Tomblin et al., 2014). Not all children learning the same language acquire individual sounds and words in the same manner. Some may begin with a set of sounds and add new sounds according to their own timetable and pattern. Some children add sounds "cautiously" by working on a set of phonetically related sounds and then expand the inventory systematically (Ferguson, 1979; Vihman et al., 1986). Some other children may be "risk takers" in that they try new sounds of different types and show a changing sound repertoire during early years of speech sound acquisition (Ferguson, 1979; Vihman et al., 1986). Some children may be "slow" in speech sound mastery at the beginning but "catch up" later (Paul & Jennings, 1992). Various causal mechanisms give rise to individual differences. Perception, cognition, learning, social understanding, and culture all come to the play (Davis & Bedore, 2013). For example, children's consonant accuracy is closely related to the phonotactic probability and neighborhood density of the child's language (Storkel et al., 2010). In a corpus study, Zamuner et al. (2004) reported that toddlers produced a coda consonant like /–n/ more accurately if that consonant is in a high-probability environment, as in /nin/, than a relatively low-probability environment, as in /von/. In an experimental study, after controlling for the frequency of within-word consonant sequences in nonwords, Munson (2001) found that both 3- and 8-year-old children were faster, more accurate, and less variable when producing high English frequency sequences (e.g., /st/) compared with low-frequency sequences (e.g., /fp/).

Variation in the timing of speech sound acquisition also may be reflected in differences in the prevalence of children with speech sound disorders (SSD) at different age cohorts. Shriberg et al. (1999) reported a 3.8% prevalence rate of SSD in children at the age of 6 years in the United States while Wren et al. (2016) reported a similar rate of 3.6% in a slightly older sample of 8-year-olds in the United Kingdom. Subsequently, Campbell et al. (2003) used the same diagnostic criteria as that of Shriberg et al. (1999) and reported a higher prevalence estimate at the age of 3 years of 15.6% (100/639). The lower prevalence rate of SSD in older cohorts may suggest that a subgroup of children underwent the process of normalization and resolved their speech problem naturally (Shriberg et al., 1994). More direct evidence about variation in timing and normalization without intervention (natural recovery) comes from observational studies. Bralley and Stoudt (1977) followed 60 children who had at least one articulation error in Grade 1 until they were in Grade 5. They indicated that of the 60 children with misarticulations, 78% resolved naturally without direct intervention. Related findings were reported by Shriberg et al. (1994), who investigated the association of speech normalization with many speech-related variables (e.g., phonological patterns, phoneme acquisition, and oral motor skills) and risk factors (e.g., demographic information and psychosocial variables). Out of the 54 children with SSD investigated, 10 showed normalization within 1 year of diagnosis, and among these 10 children, four did not receive any intervention during the course of the study.

When children continue to experience difficulties in speech sound production during school years, they may develop persistent speech problems. SSD is a broad term including phonological and motor speech (articulation) disorders. Different classification systems have been proposed to characterize subgrouping of these disorders. Based on etiology and demographic factors, Shriberg (1993) proposed the Speech Disorders Classification System categorizing SSD into seven subtypes: speech delay–genetic, speech delay–otitis media with effusion, speech delay–apraxia, speech delay–dysarthria, speech delay–developmental psychosocial involvement, and two kinds of speech errors limited to speech sound distortion. These subtypes have been evolving with research findings and may have implications on the prognosis of later speech and language outcomes. However, empirical studies on the long-term outcomes of these subtypes have yet to be developed. McLeod and Baker (2017) proposed five types of SSD grouped in two categories, phonology (phonological impairment and inconsistent speech disorder) and motor speech (articulation impairment, childhood apraxia of speech, and dysarthria), and outlined assessment and intervention approaches for each. Stackhouse and Wells (1997) put forward a psycholinguistic framework to classify SSD in terms of information processing. The framework specifies breakdowns in the speech processing system of an individual child and provides explanations to individual differences in the patterns observed in SSD, as well as the literacy problems (Pascoe et al., 2006). The validity framework has been supported by an intervention study devised based on the speech processing system (Pascoe et al., 2005). The framework provides insights into what makes the surface manifestation of a child with SSD more severe or resistant to change. Speech difficulties often are associated with long-term negative outcomes on communication, interpersonal interactions, and academic achievement (McCormack et al., 2009).

## Risk Factors of SSD

Clinicians often encounter the following question from parents, teachers, and doctors: "Can children outgrow speech errors without intervention?" There are two potential outcomes for children with SSD: normalization and long-term (persistent) SSD (Roulstone et al., 2009; Wren et al., 2016). This study aimed to study four consistently reported risk factors of SSD related to children's speech and language profile, namely, low stimulability, intelligibility, presence of atypical errors, and expressive language difficulties, and how these factors related to children's time to normalization. The contribution of these risk factors can improve our knowledge of what speech-language characteristics may be more likely related to pathology underlying SSD rather than normal individual differences in children's schedule and timing of development.

### Stimulability

Stimulability refers to a child's ability to modify speech production errors when stimulated by a clinician with models and cues (Glaspey & Stoel-Gammon, 2007; Lof, 1996; Powell & Miccio, 1996). It has been adopted as a routine assessment procedure in clinical speech pathology since it was first introduced by Lee Edward Travis in 1931 (Travis, 1931). The actual procedures evolved substantially and vary across clinicians' practice (e.g., the number of presented stimulations, stimuli in isolation vs. at word level, or in nonsense syllables). Regardless of these variations, stimulability performance offers insights regarding children's potential of producing a sound under the most supportive circumstances and their actual performance in day-to-day situations (Dinnsen & Elbert, 1984; Glaspey & Stoel-Gammon, 2007). Previously, researchers had uncovered the prediction power of stimulability regarding the potential for improvement in speech sound production with and without treatment (e.g., Carter & Buck, 1958; Diedrich, 1983; Powell et al., 1991). Of children in the first and second grades who did not receive treatment, those with high stimulability scores demonstrated significant improvement, implying speech errors can be resolved naturally (Carter & Buck, 1958; Sommers et al., 1967).

### Intelligibility

Speech intelligibility refers to the degree to which the listener understands what the speaker says and has been described as the most practical single index to apply in evaluating oral communication competence (Subtelny, 1977). Clinicians have used this as an important measure to determine the presence of SSD and the need for intervention (McLeod et al., 2020; Williams et al., 2021), implying that children with poor intelligibility are not likely to resolve their errors naturally and intervention is necessary. There are a number of ways to measure intelligibility (Kent et al., 1994). McLeod et al. (2012) developed the Intelligibility in Context Scale (ICS), a parent rating scale to estimate children's intelligibility with various communication partners. The ICS has been translated into 60+ different languages and validated in 14 languages (McLeod, 2020). The traditional Chinese translation of the ICS (ICS-TC) has been normed and validated on the Cantonese population (Kok & To, 2019; Ng et al., 2014).

### Type of Speech Errors

As young children start to learn their ambient language, predictable patterns of sound errors are observed in their production due to motor and perceptual restrictions (Stampe, 1969). These expected "typical" patterns of speech sound errors have been well documented. For example, typical substitutions and syllable structure errors are observed in most young children with or without SSD (Shriberg & Kwiatkowski, 1980), though these errors are more frequently observed in the speech of children with SSD over a longer time (e.g., Ingram, 1976). When children are exposed more to the mature form in the ambient language, these patterns can be gradually suppressed so that their word productions match the mature form (Stampe, 1969). Atypical errors refer to substitutions, syllable structure errors, and distortions that are not generally found in typical phonological development. These error patterns are thought to be associated with poor speech normalization. Leonard (1973) recommended that children who exhibit atypical error patterns be given priority for intervention as their speech sound systems are not likely to normalize naturally. Leahy and Dodd (1987) supported this claim with their case study of a 3-year-old child who used predominantly atypical patterns. This child showed little progress in the number and types of error patterns used when no intervention was provided but a decrease in the use of these patterns and an increase in consonant accuracy upon intervention. More recently, Dodd et al. (2018) also found that children with fewer atypical errors were more likely to resolve their errors, regardless of whether intervention was available or not.

### Expressive Language Ability

Children with SSD have been shown to exhibit concomitant language difficulties (e.g., Baker & Cantwell, 1987; Flipsen, 2003; Preston & Edwards, 2009; Shriberg & Austin, 1998; Shriberg et al., 1986, 1999). In a large-scale population study, Shriberg et al. (1999) reported that about 11%–15% of children with persisting speech delay showed specific language impairment (i.e., developmental language disorder [DLD]) and that comorbidity of speech delay and language impairment was 1.3%. Comorbidity of SSD and language impairment may imply a higher risk for later reading and spelling problems in school years than isolated SSD alone (Hayiou-Thomas et al., 2017; Lewis et al., 2018; Pennington & Bishop, 2009). A recent longitudinal study explored the long-term outcomes of preschool SSD (Lewis et al., 2019). In this study, children with early SSD were followed up in their school years and adolescence. The results, in general, indicated that children with better early speech and language skills at preschool and school years are less likely to have persistent SSD during adolescence and that language difficulties at school years appeared to be associated with incomplete speech normalization. However, there is a sizeable group of children showing isolated SSD (Lewis et al., 2000).

## Natural History Studies

Correlation analyses and regression analyses previously have been employed to explore the role of risk factors on normalization. The fundamental question that these statistical analyses address is whether a predictor factor (e.g., types of phonological patterns) is related to the occurrence of an outcome (i.e., normalization). These techniques do not address questions about the time that it takes for normalization to occur. The fact is that children normalize not only in association with certain characteristics or demographic background but also at different points in time. It means that questions such as when normalization is most likely to occur and whether children who resolve early differ from those who resolve late or will not resolve before a certain time have been left unanswered until now. Oberklaid et al. (2002) pointed out that current evidence regarding predictor variables is not strong enough to guide speech and language screening in routine surveillance system for children in Australia. This policy statement has acknowledged the limited amount of high-level evidence by calling for an increase in longitudinal studies to provide more robust evidence for prognosis and risk factors. Among different longitudinal study designs, natural history studies offer a good way of examining the outcomes of a disease/disorder in relation to the nature or symptoms of the individuals.

In related fields, natural history can be defined as the evolution of a disease/disorder in the absence of intervention (Fletcher et al., 1988). In the field of speech-language pathology, there are very few natural history studies of speech and language disorders (Law et al., 2000) with some notable exceptions. Roulstone et al. (2009) investigated the natural history of SSD within a case–control study with two groups of participants, one with SSD (cases) and an alike group of participants who did not have SSD (controls). In their study, 741 children diagnosed with SSD (cases) were compared with a group of children with typical speech ability (controls) in a longitudinal study covering 2, 5, and 8 years of age. They found that the trajectory of speech sound acquisition in children who exhibited persisting problems at 8 years of age showed clear differences in speech sound error rates when compared with the controls who resolved their speech errors either naturally or with intervention. In a subsequent study, Wren et al. (2013) highlighted that useful tasks for identifying older children with atypical speech were speech performance on a connected speech task and a nonword repetition as well as identification of phonological patterns on a single-word task. Related results on measures that differentiated transient and longer term problems were replicated in McIntosh and Dodd (2008), who followed 10 children aged 25–35 months for a 12-month period. All children were assessed at three time points using norm-referenced articulation tests. Percentage of consonants correct (PCC) at baseline did not predict the children's final speech outcome 1 year later. On the other hand, their type of error patterns appeared to be more sensitive: two of the children who showed atypical errors at 2 years of age continued to exhibit the similar errors at 3 years of age.

Natural history studies, like other types of longitudinal studies, are often plagued with dropout bias. Participants who withdraw in the middle of the study or those who do not show normalization even by the end of the study period have to be excluded from subsequent analysis. This results in the loss of information, thereby reducing the power of the tests and introducing bias into the analyses. Therefore, Gruber (1999a) made use of the technique of survival analysis to study the natural history of SSD.

## Survival Analysis

Survival analysis is a technique originally developed by biostatisticians to determine how long a patient with a particular disease/disorder may be expected to survive. It can address questions about time that cannot be addressed by simple correlation, linear and logistic regression, or other longitudinal procedures such as growth curve analysis. The time scale for an individual participant usually begins at the start of the study. However, participants can also enter the study later, and their entry time is marked as time zero in survival analysis. As mentioned above, participants who drop out in the middle of the study or do not show the target outcome (e.g., normalization) by the end of the study period have to be excluded from the final analysis in traditional correlational studies. In survival analysis, we still can include these participants. These two types of lost cases are considered as the same category statistically and are termed *censored* when participants did not normalize during the testing period or withdrew from the study.

There are two basic functions in survival analysis. *Hazard functions* describe the conditional probability that a child will normalize during a discrete time period given that that child has exhibited errors until that time interval. Examining the hazard functions can inform us the time of greatest risk for the occurrence of an event or the time when the event is more likely to occurrence. *Survival functions* describe the unconditional probability of continuously having speech errors beyond time $t$ for a randomly selected child, that is, the accumulated period-by-period risks of event occurrence. By definition, the survival function will never increase and will decrease over time. The use of these two functions enables examination of situations where normalization probabilities change with time $t$. This is the advantage of survival analysis—supporting evaluation of data throughout the study period rather than confining the analysis to cross-sectional examination of the final event, resulting in findings of clinical usefulness (i.e., estimates of the probability that children will show speech normalization over a given time period). Estimates of the hazard or survival functions for different groups of participants can be compared to examine whether differences exist in the probabilities of normalization (Gruber, 1999b).

Gruber (1999a) successfully provided an estimate of the probability of normalization at a certain time, given the presence of speech difficulties up to that time. However, the results from the study were still not sufficient to reliably determine the likelihood of which growth path (i.e., natural normalization or prolonged normalization) an individual child with a particular profile may undergo. Gruber (1999a) also pointed out that the clinical application of his findings may be limited by "misestimation" and "misspecification" biases arising from the model used (p. 458). The first type of bias may be due to the nonproportional hazards in children with different ages, whereby older children may be less likely to undergo normalization than younger children. An additional factor to age is the number of initial consonants in the inventory relative to other children of the same age. Children with a larger existing consonant inventory could achieve normalization faster than those with more missing consonants. In this case, an interaction term of age and the number of initial consonants at the initial time point can be included to represent the effect. The second type of bias may be due to the assumption made about the exponential nature of the survival time distribution, which may not be accurate. The above issues can be addressed or alleviated by using the technique of the Cox proportional hazard models, which makes no assumption about the survival–time distribution. Cox proportional hazard models allowed the estimation of the effect of more than one covariate. It means that the possible effect of various potential risk factors was taken into account when calculating the probability estimates of normalization.

In this article, speech normalization is defined as completion of speech sound inventory, which is the expected "target event" in the survival analysis. During the course of children's speech sound development, we can observe (a) whether or not children continue to have unresolved speech errors and (b) the time of completed development (normalization) by those who can complete the phonemic inventory expected for their language. We can estimate from this information the probability that a child with certain characteristic in the sample will show speech normalization for each point in time during the study period. For example, one of the primary goals of this survival analysis is to compare the normalization probability functions for individual participants classified as having developmental phonological errors only versus atypical errors.

## Clinical Relevance: Access to Speech-Language Pathology Service and Waiting Lists

Understanding whether children outgrow speech errors without intervention is of clinical importance. Throughout the world, there are fewer speech-language pathologists (SLPs) than the number of children who need their services (Mulcair et al., 2018). Families, referring professionals (doctors and teachers), and SLPs frequently wonder whether they should "watch and wait" or refer young children for speech-language pathology services (Morgan et al., 2017). In many countries, there are long waiting lists for services (McGill et al., 2021). Long waiting lists for access to speech-language pathology services can impact children and families, SLPs, and society (McGill et al., 2020). Some children do not receive speech-language pathology services at all but may be able to resolve the errors at a later time point. However, some children continue to show a large number of errors even when they receive belated therapy. Some children may show minimal or no progress given the low dose of intervention provided (Law & Conti-Ramsden, 2000). For example, children with childhood apraxia of speech require a high dose of intensive treatment in order to make observable progress (Preston et al., 2018). Delay in services or insufficient intervention frequency may lead to poor speech outcomes impacting children's education, social development, and occupational prospects (Glogowska et al., 2000; McCormack et al., 2009; McLeod et al., 2019).

## This Study

This study presents the findings of a naturalistic cohort study investigating speech normalization and normalization rates in Cantonese-speaking preschool children in Hong Kong SAR, China, who may be at risk of SSD. All children born in Hong Kong are eligible for the Developmental Surveillance Scheme, a public service provided by the Department of Health in Hong Kong (Family Health Service, 2021). Community nurses conduct surveillance interviews and observation on children's development areas including motor, communication, social behaviors, self-care, vision, and hearing. These surveillance interviews are scheduled when the child is at 6, 12, and 18 months of age, which also coincides with their vaccination schedule to facilitate access and compliance. Children with developmental concerns, including speech and language needs, identified at these interviews will either continue to be monitored at subsequent visits or be referred for a detailed developmental assessment by relevant professionals. Children with speech and language concerns identified at the surveillance interviews are referred to the Child Assessment Service (2021), also a public service under the Department of Health, for a comprehensive speech and language assessment. Following the assessment, according to the level of needs, children are further referred to appropriate speech-language pathology services (termed *speech therapy*) to access intervention. These speech-language pathology services usually provide intervention to children until they reach school age. For school-age children, they are then managed by the SLP (termed *speech and language therapist*) based at their primary schools (Education Bureau, 2021). The public speech-language pathology service often has an extensive waiting list—a situation that has extended for at least a decade and impacts intervention intensity (To et al., 2012). Children with speech and language needs may also access speech-language pathology services through private sectors even though private speech-language pathology service is costly and rarely covered by insurance.

The objectives of this study were to quantify speech normalization rates at 2.5-year follow-up and to investigate predictors of time to normalization. Given the prior evidence demonstrating that children with SSD who show persistent speech difficulties (i.e., less likely to resolve) demonstrate lower stimulability, lower intelligibility, and atypical errors, we hypothesized that being nonstimulable, having low speech intelligibility, and having atypical errors may increase the risk of longer speech normalization time. With mixed findings about comorbidity of SSD and DLD, it was not clear if low expressive language ability may present as a significant risk factor of SSD.

## Method

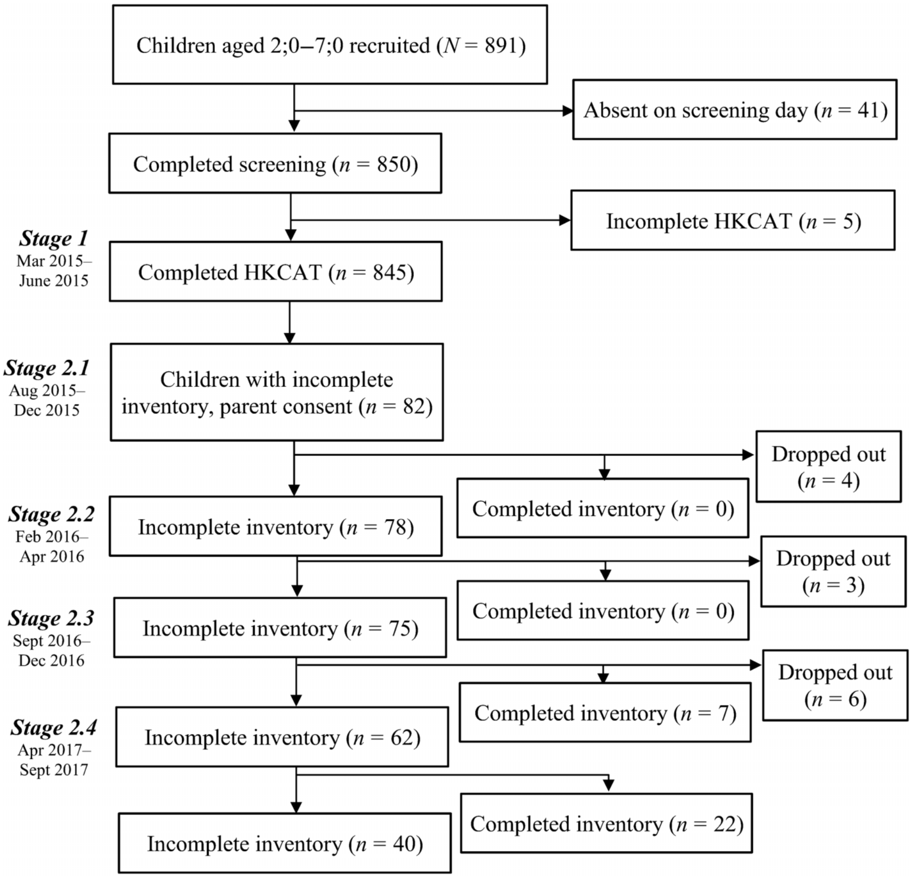
A prospective cohort study was conducted to examine (a) children who resolved nonadult realizations of speech sounds (i.e., had normalized production of speech sounds) and (b) those who had persisting speech sound difficulties (did not normalize) over 2.5 years.

### Participants

The participants of this study were selected using a two-stage sampling procedure (see Figure 1). In the first stage, 891 children were recruited using a stratified random sampling procedure with the stratifying parameters of age, sex, and region. Eight age groups with intervals of 6 months were included. The youngest group comprised children aged below 3 years (2;4–2;11 [years; months]), and the oldest group comprised children aged above 6 years (6;0–6;9). Kindergartens and preschools in the three main regions of Hong Kong (Hong Kong Island, Kowloon, and the New Territories) were invited to participate. At least three kindergartens were selected randomly based on the kindergarten list of the Education Bureau. All the children studying in each kindergarten were invited to participate. This sample had been reported in a previous study examining the cutoffs for the speech screening tool, the ICS-TC, which is a seven-item parent report instrument evaluating children's speech intelligibility (Kok & To, 2019). The ICS-TC had been validated on Cantonese-speaking population in Hong Kong (Kok & To, 2019; Ng et al., 2014). Of the 891 children recruited, 41 children were absent on the day of the assessment. The total samples therefore consisted of 850 children aged 2;4–6;9 with an approximately equal number of boys and girls. The parents of all child participants provided written consent for their own and their child's participation.

In Stage 1, the initial screening sessions, 845 of all these 850 children completed a series of direct assessments, and parental questionnaires were administered, including (a) the Hong Kong Cantonese Articulation Test (HKCAT; Cheung et al., 2006), a standardized norm-referenced single-word speech assessment tool eliciting all speech sounds in Cantonese via naming 41 pictures; (b) the ICS-TC (Kok & To, 2019); (c) a simplified version of the oro-mechanism examination (OME); and (d) the Hong Kong Cantonese Receptive Vocabulary Test (HKCRVT; Cheung et al., 1997), a standardized language assessment tool normed on Cantonese-speaking children. Undergraduate students in the Speech and Hearing Sciences Program at The University of Hong Kong were recruited as testers in the data collection process. All testers were familiar with the tests as they had learned and/or used them in the speech and hearing sciences curriculum. One week before data collection, the testers received 1 day of training, highlighting the administration of these tests and the recording procedures. Testers were asked to record data online for the OME and the HKCRVT. The direct assessment process took place in the children's preschools under the supervision of the first author. All the assessment processes were audio- and video-recorded. Parents were asked to complete a questionnaire regarding their language and demographic background and the ICS-TC. The completed forms were returned to the class teachers who collected all the documents for the researchers.

**Figure 1.** Flow diagram for the 891 children who participated in the research. HKCAT = Hong Kong Cantonese Articulation Test.



Children from Stage 1 were selected to participate in Stage 2 (follow-up longitudinal study) if (a) their standardized scores in the HKCAT fell under −1.25 *SD*, or they could not pronounce the initial consonants that were expected at their age, or they had one or more atypical errors (e.g., dentalization of alveolar fricatives and affricates; To et al., 2013) including distortions (i.e., a realization that is recognized as the target phoneme but perceived as acoustically different and unacceptable); (b) they were not receiving intervention at the point of screening; (c) they did not show any significant deficits during the OME; and (d) they had not been diagnosed with other neurodevelopmental disorders such as autism or attention-deficit/hyperactivity disorder. Parents of these children were contacted by phone, starting from the oldest age group, followed by the younger age groups. Parents received an explanation about the research aim, the nature of the study (as an observational study and not an intervention study), their commitment as a research participant (visiting the university for data collection 4 times at most), and participant compensation fees of about U.S. $25 for each visit. A total of 82 children aged 2;3–6;2 at Stage 1 agreed to participate in Stage 2.

## Procedure

About 4–6 months after the Stage 1 screening, the 82 children were invited to The University of Hong Kong for the initial assessment of Stage 2 (Stage 2.1). Follow-up interviews were conducted 6, 12, and 18 months after Stage 2.1. Participants who were not interviewed when due were interviewed within 2 weeks whenever possible. Each participating family was paid U.S. $25 after each visit. Follow-up consisted of direct assessment of the child conducted at approximately 6-month intervals. These assessments sought to identify the speech sound production and language abilities. For this purpose, the HKCAT and the Hong Kong Cantonese version of the Reynell Developmental Language Scales (RDLS-C; Hong Kong Society for Child Health and Development, 1987) were administered. For children who had reached the age of 7;0 at any point of assessment, the RDLS-C was replaced by the Narrative Test (To et al., 2010) of the Hong Kong Cantonese Oral Language Assessment Scale (T'sou et al., 2006). At the last point of assessment, caregivers were queried if intervention was received during the study and the content of the therapy.

## Data Analysis

Speech samples collected using the HKCAT were transcribed based on narrow transcription and input into the PHON system (Version 3.3.0; Hedlund & Rose, 2020) by speech-language pathology students in the Speech and Hearing Sciences Program at The University of Hong Kong. PHON is a free phonological analysis software available from https://www.phon.ca/phon-manual/getting_started.html. Distortions were counted as errors and were marked with diacritics in PHON. Some frequently used diacritics includes marking sounds that were dentalized̪, lateralized$^1$, palatalizedʲ, and nasalized˜. When the type of distortion could not be classified, the realization was marked with a vertical tilde̾ in PHON.

All of the transcriptions were cross-checked by a research assistant with a strong phonetic background. Mismatches were resolved through discussion between the research assistant and the corresponding students. All the transcriptions in the PHON system were then reviewed by the first author who made the final decision regarding the transcription. All errors produced by each individual child were identified within the PHON system. The types and tokens of the errors were summarized in a Microsoft Excel spreadsheet. Standardized scores of the HKCAT in Stage 1 were computed.

## Survival Analysis

Survival analysis was used to analyze time to normalization (i.e., time to event). Event time (i.e., the survival time) is the difference between the initial time, when no one showed a completed inventory (normalization) and all participants are considered as having a potential to normalize the speech (i.e., at risk in survival analysis) and the time at which the child showed a completed inventory (normalization; if it does occur). In other words, the initial time was birth, when all individuals had the potential to normalize and thus all children were in the potential set at that time. When children normalized their speech by completing the consonant inventory, they were no longer in the potential set.

To complete a survival analysis, a binary outcome (target event) is required. The outcome of interest was word-initial consonant inventory (as opposed to inventory of final consonants) because word-initial consonants are more frequent in Cantonese than word-final consonants (To et al., 2013), and children with SSD have "significantly lower consonant accuracy in word-initial position than within-word or word-final positions" (McLeod & Masso, 2019, p. 71). Therefore, completion of the initial consonant inventory (complete/incomplete), as evaluated by the HKCAT, was the target event. Accuracy was determined based on relational analysis, that is, comparing each child's realization against the adult form. In the HKCAT, there are at least three items for each of the initial consonants (except the consonants of /kʷ, kʷʰ/, /ŋ/, or zero initial, which are relatively infrequent in daily speech) in the contexts of front, mid, and back vowels. Operationally, the child had to achieve more than 75% accuracy of each of the 18 initial consonants in Cantonese in order to be defined as having a completed initial consonant inventory (i.e., normalization). To illustrate, if children only produced one initial consonant inaccurately while other consonants were correct, they were considered to have an incomplete mastery of initial consonant. In other words, the possible events were either complete mastery of all Cantonese initial consonants or incomplete mastery of the initial consonants.

PCC was not adopted as a measure to define the target event because it is not binary in nature. PCC serves as a proxy of children's speech sound production ability relative to other children of the same age. However, it is possible that a child may show an acceptable PCC (i.e., > 85%) but still misarticulate one or a small set of phonemes that may also affect intelligibility. In comparison, the target event of completion of the initial consonant inventory enabled consideration of every individual phoneme.

Observed time to normalization (in years) was evaluated using nonparametric Kaplan–Meier curves (Kaplan & Meier, 1958). Kaplan–Meier curves plot the probability of normalization over time while taking account of censored (i.e., missing) data points. The Kaplan–Meier product limit was used to estimate the cumulative probability of normalization (i.e., survival function) with the factors of error (a) typicality, stimulability, passing the expressive language test, and passing the mean ICS-TC cutoff. Two control covariates, (a) an interaction term of size of consonant inventory and age and (b) sex, were also included to control for their possible confounding influence. A multivariate Cox regression model was developed with these variables.

### Error Atypicality/Typicality

This variable was dichotomously defined in this study. All speech errors produced by each child were classified into typical and atypical error patterns. Typical error patterns refer to patterns produced by 5% or more of the children in the general population who speak Cantonese as their native language who are aged 2;6 or above. Atypical error patterns are those used by fewer than 5% of the children (To et al., 2013; see Table 1). Patterns other than those in Table 1 were coded as atypical. For example, backing was regarded as a developmental pattern because it was exhibited by approximately 5%–10% of children when 3 years old, whereas initial consonant deletion was considered an atypical process for Cantonese-speaking children because it was rare (cf. English). Distortions were also considered a type of atypical errors, which occurs when a

Table 1. Developmental errors coded in this study.

| Target | | Typical realizations (produced by > 5% of children) |
|---|---|---|
| /pʰ/ | → | [p] |
| /t/ | → | [k] |
| /tʰ/ | → | [t, kʰ, tsʰ] |
| /k/ | → | [t] |
| /kʰ/ | → | [t, tʰ, k] |
| /kʷ/ | → | [p, t, w] |
| /kʷʰ/ | → | [pʰ, k, kʰ, kʷ] |
| /ts/ | → | [t] |
| /tsʰ/ | → | [t, tʰ, ts, s̪] |
| /f/ | → | [p] |
| /s/ | → | [t, ts, tsʰ, s̪] |
| /h/ | → | [ʔ] |
| /l/ | → | [ʔ, n, j] |

target phoneme is replaced by a sound that is slightly to severely off the target or by a sound that is not present in the language.

Error patterns were counted as a pattern only if the same error was noted twice or more on the HKCAT. If children demonstrated the use of any atypical error patterns, their status for this variable was coded as *atypical*. If children did not show any atypical error patterns and only typical developmental errors, their status was coded as *typical*.

### Stimulability

This variable was also dichotomously defined for each child in the study as *stimulable* or *nonstimulable*. If children were stimulable for all misarticulated sounds in nonsense syllables of /Ca/, and /Ci/ or /Cy/, they were considered stimulable. On the other hand, if children were stimulable for part of the sounds or nonstimulable for all the sounds in the above conditions of nonsense syllables, they were considered nonstimulable.

### Intelligibility

Each child's mean ICS-TC scores were computed and checked against the stated cutoffs with reference to the child's corresponding age groups (Kok & To, 2019). This variable was dichotomously defined as having the mean scores higher or lower than the stated cutoff. For example, the cutoffs for the age groups of 4;0–4;5 and 5;6–5;11 are 3.64 and 4.79, respectively.

### Expressive Language Ability

All participants completed a standardized language assessment, the RDLS-C, during Stage 2.1. The child's performance on the standardized expressive language test was examined as a potential contributing factor of speech normalization. This variable is dichotomously defined as having standardized scores of higher or lower than −1.0 *SD* on the expressive scale.

### Sex

Male sex has been repeatedly found to be a risk factor for SSD (Hyde & Linn, 1988; McLeod & Baker, 2017). This variable was included to control for its effect in the final model.

### Interaction of Consonant Inventory Size and Age at Stage 1

For the binary outcome of normalization, the number of initial consonants at one time point would have an impact on the probability of normalization; the larger the consonant inventory, the more likely a child can achieve normalization. As maturity plays an important role in acquiring initial consonants, age was also a potential factor associated with normalization. However, it was also possible that older children may be less likely to undergo normalization than younger children (Gruber, 1999a). In this case, the concurrent contribution of the age and the size of the consonant inventory could be captured by an interaction term of age and number of initial consonants at Stage 1. This interaction term was included in the final multivariate model as a continuous covariate.

## Results

Among the 82 participants who were followed up at Stage 2.1, the number of participants who were lost to follow-up at Stage 2.2 was four; at Stage 2.3, it was three, and at Stage 2.4, it was six (see Figure 1). No children showed a completed initial consonant inventory at Stages 2.1 and 2.2, whereas seven completed their initial consonant inventory at Stage 2.3 and 22 children completed their initial consonant inventory at Stage 2.4. The median period of estimated time to normalization of all the 82 children was 6.82 years. At Stage 2.4, parents were asked about whether their child had received speech-language pathology services. Among the 82 participants, 39 of the caregivers reported that their child began speech-language pathology services during the course of the study. Most of the caregivers were not able to recall the exact onset of the services or the content. Table 2 summarizes the background of the 39 children who had started speech-language pathology services and those without this service during the study period. The median period of estimated time to normalization for these 39 children was 7.89 years.

## Study Sample Characteristics

With the aim of examining the natural history of SSD, the following analyses focused on the 43 children whose caregivers confirmed the absence of intervention during the course of the study. The mean age of these 43 study participants was 4.09 years (*SD* = 0.982, range:

**Table 2.** Background of the children with and without speech-language pathology services.

| Variable | Started speech-language pathology services ($n$ = 39) | Without speech-language pathology services ($n$ = 43) | Statistics |
|---|---|---|---|
| $M_{\text{age}}$ in years (range) | 4.23 (3.30–5.73) | 4.09 (2.23–6.20) | $F(1, 81) = 0.617$, $p = .434$ |
| Female:male | 13:26 | 14:29 | $\chi^2 = 0.006$, $p = .941$ |
| Consonant inventory (maximum = 19) | 11.74 (5–17) | 13.33 (6–17) | $F(1, 81) = 6.032$, $p = .016$ |
| Typical:atypical error | 5:34 | 13:25 | $\chi^2 = 8.546$, $p = .003$ |
| Simulable:nonstimulable | 17:21 | 22:21 | $\chi^2 = 0.334$, $p = .564$ |
| Passed:failed expressive language test | 27:12 | 26:17 | $\chi^2 = 0.687$, $p = .407$ |
| Passed:failed mean ICS-TC cutoff | 5:34 | 14:29 | $\chi^2 = 4.476$, $p = .034$ |

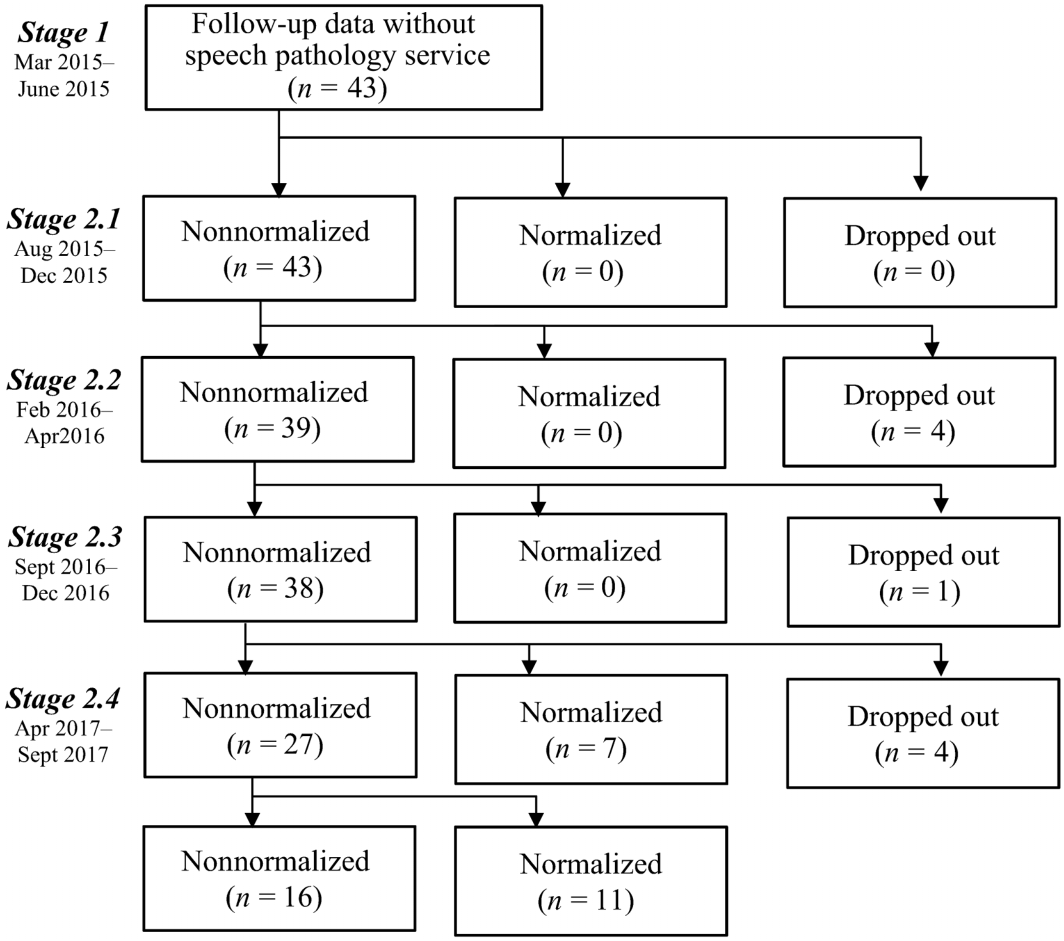
*Note.* ICS-TC = Intelligibility in Context Scale–Traditional Chinese.

2.23–6.20 years), and 32.6% were girls. The sample of study participants without speech-language pathology services and those who started speech-language pathology services were comparable in all demographic and clinical characteristics except for three variables. The study sample had a larger consonant inventory at Stage 1. As a group, the study sample consisted of fewer children with atypical error patterns and more children who passed the mean ICS-TC cutoff (see Figure 2).

## Predicting Normalization

Among the study sample, the number of participants who dropped out from the study without completing the

**Figure 2.** Flow diagram of the study for the 43 children who received no intervention during follow-up.



inventory was four at Stage 2.1, one at Stage 2.2, and four at Stage 2.3 (see Figure 1). No children showed normalization at Stages 2.1 and 2.2, whereas seven children demonstrated normalization at Stage 2.3 and 11 children demonstrated normalization at Stage 2.4. The median period of estimated time to normalization of these 43 children was 6.59 years.

Figures 3a and 3b show the Kaplan–Meier curves for time to normalization compared by stimulability and by passing the mean ICS-TC cutoff. The plot in Figure 3a

**Figure 3.** Kaplan–Meier curves showing the survival functions for children (a) who are stimulable and nonstimulable and (b) who are intelligible and not intelligible (mean ICS-TC). Censored means that those participants did not normalize during the testing period or withdrew from the study. ICS-TC = Intelligibility in Context Scale–Traditional Chinese (McLeod et al., 2012).

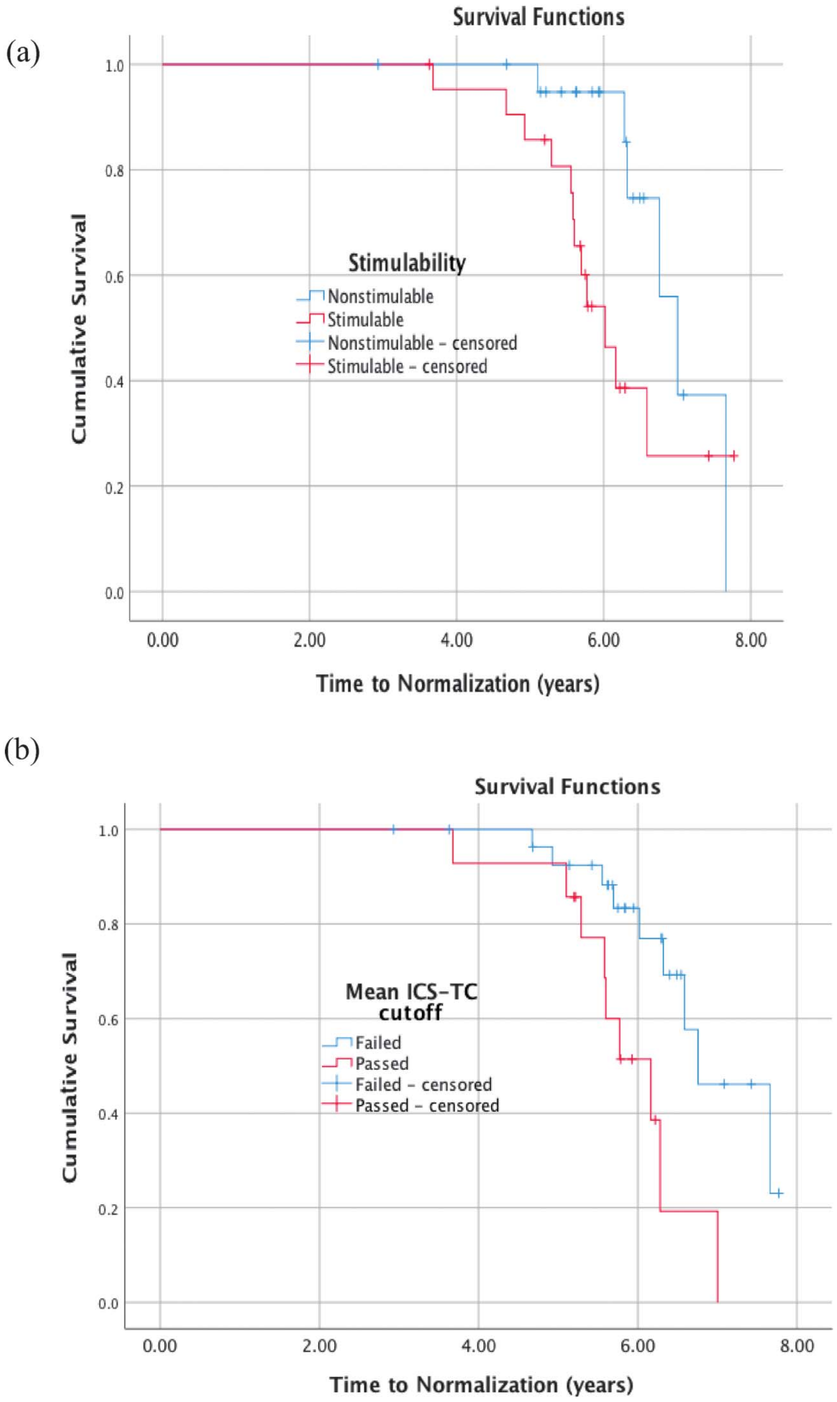


**Table 3.** Median time to normalization of the categorical variables.

| Variable | Subgroup | Median time in years (95% CI) | Log-rank test *p* |
|---|---|---|---|
| Stimulability | Stimulable (*n* = 22) | 6.02 [5.50, 6.54] | .051 |
| | Nonstimulable (*n* = 21) | 7.00 [6.49, 7.52] | |
| Error (a)typicality | Typical (*n* = 18) | 6.76 [5.89, 7.62] | .315 |
| | Atypical (*n* = 25) | 6.32 [5.81, 6.83] | |
| Expressive language test | Passed (*n* = 14) | 6.76 [5.83, 7.68] | .587 |
| | Failed (*n* = 29) | 6.59 [5.53, 7.64] | |
| Mean ICS-TC cutoff | Passed (*n* = 14) | 6.17 [5.36, 6.97] | .008 |
| | Failed (*n* = 29) | 6.76 [5.83, 7.69] | |
| Sex | Male (*n* = 29) | 7.00 [6.22, 7.78] | .017 |
| | Female (*n* = 14) | 6.17 [5.59, 6.74] | |

*Note.* CI = confidence interval; ICS-TC = Intelligibility in Context Scale–Traditional Chinese (McLeod et al., 2012).

illustrated that the cumulative survival proportion (i.e., proportion of children that had not normalized) appeared to be higher in the nonstimulable group compared to the stimulable group. It means that children who were nonstimulable took significantly longer to show normalization than those who were stimulable. Similarly, as in Figure 3b, children who did not pass the ICS-TC (i.e., were unintelligible) took longer to normalize than those who passed the ICS-TC (i.e., were intelligible). The curves were statistically significantly different as determined by the log-rank test. The Kaplan–Meier curves for other variables were not shown given that they were statistically nonsignificant. Table 3 shows the median time to normalization of each of the categorical variables (i.e., the time at which 50% of the participants have reached normalization) and the *p* values of the corresponding log-rank test. With the alpha level of .10, the median time to normalization demonstrated significant group differences in the variables of stimulability, mean ICS-TC, and sex.

## Cox Proportional Hazard Regression Model

Before the calculation of the multivariate hazard ratios, univariate hazard ratios for individual variables when they were considered separately without the variable of time are calculated (see Table 4). With the alpha level of .10, the only statistically significant variables were stimulability and mean ICS-TC as well as the covariates of sex and the interaction term of age and size of the consonant inventory. The final Cox regression analyses were conducted to evaluate whether findings of the variables of interest, including (a)typicality of error patterns, stimulability, expressive language skills, and speech intelligibility, still held after controlling for the covariates of sex and the interaction term of age and size of the initial consonant inventory at Stage 1 (see Table 5).

Controlling for the two covariates, there was no significant effect of error (a)typicality and expressive language ability on time to normalization. It means that atypical error patterns or not passing the expressive language test was not significantly associated with a longer time to normalize. Greater stimulability, however, was associated with a significantly shorter time to normalization (Wald = 6.743, hazard ratio = 0.174, 95% confidence interval [0.047, 0.652], Cox *p* = .009). A hazard ratio of 0.174 implies that participants who were stimulable at Stage 1 were 17.4% more likely to normalize when compared to those who were nonstimulable and tended to have a shorter duration of normalization compared to those who were nonstimulable at Stage 1. Similarly,

**Table 4.** Univariate hazard ratios for potential of time to normalization.

| Variable | Sample estimates *M* | Univariate hazard ratio (95% CI) | *p* |
|---|---|---|---|
| Stimulability | 0.512 | 2.600 [0.963, 7.014] | .059 |
| Error (a)typicality | 0.585 | 1.646 [0.617, 4.392] | .320 |
| Passed expressive language test | 0.390 | 1.102 [0.425, 2.855] | .842 |
| Passed mean ICS-TC cutoff | 0.341 | 3.458 [1.300, 9.193] | .013 |
| Sex | 0.317 | 3.375 [1.182, 9.636] | .023 |
| Consonant Inventory × Age | 57.442 | 0.951 [0.924, 0.979] | .001 |

*Note.* CI = confidence interval; ICS-TC = Intelligibility in Context Scale–Traditional Chinese (McLeod et al., 2012); Consonant Inventory × Age = a continuous variable representing the interaction of size of consonant inventory and age.

**Table 5.** Cox proportional hazards model predicting time to normalization.

| Variable | Hazard ratio | 95% CI | *SE* | Wald | *p* |
|---|---|---|---|---|---|
| (A)typical error patterns | 1.616 | [0.476, 5.486] | 0.624 | 0.592 | .442 |
| Stimulability | 0.174 | [0.047, 0.652] | 0.672 | 6.743 | .009 |
| Passed mean ICS-TC cutoff | 0.229 | [0.068, 0.768] | 0.617 | 5.701 | .017 |
| Passed expressive language test | 0.336 | [0.089, 1.269] | 0.677 | 2.587 | .108 |
| Sex = male | 1.147 | [0.358, 3.672] | 0.594 | 0.053 | .817 |
| Consonant Inventory × Age | 0.911 | [0.863, 0.962] | 0.028 | 11.421 | .001 |

*Note.* CI = confidence interval; SE = standard error; ICS-TC = Intelligibility in Context Scale–Traditional Chinese (McLeod et al., 2012); Consonant Inventory × Age = a continuous variable representing the interaction of size of consonant inventory and age.

participants whose mean ICS-TC passed the stated cutoff (i.e., who were rated as intelligible by their caregivers) were 22.9% more likely to normalize.

## Discussion

This study investigated the probability of speech normalization in children without intervention using the statistical technique of survival analysis, which took into account of censored data points. Time to normalization was evaluated using nonparametric Kaplan–Meier curves. The main findings can be summarized in two points. Children who were more likely to normalize or normalized in a shorter time were more intelligible and were stimulable to all speech sound errors. Children who exhibited atypical error patterns at Stage 1 did not necessarily demonstrate a lower probability in speech normalization as previously suggested (cf. Morgan et al., 2017).

### Error Atypicality/Typicality

It has long been assumed that children exhibiting typical developmental error patterns are those children who are "delayed," as opposed to "disordered" (Dodd, 2014). These children with a profile of speech delay have been described as being more likely to have a shorter normalization time than children exhibiting typical error patterns; thus, intervention is deemed necessary for children with atypical speech errors since self-correction is less likely (Dodd et al., 2018; Morgan et al., 2017). Participants who commenced speech-language pathology services demonstrated a more atypical error profile than those without speech-language pathology services at Stage 1. One possible interpretation is that SLPs may have used speech error types to prioritize cases for intervention. However, no association between error type and time to normalization was found in this study, in either the univariate Cox regression or the final multivariate Cox proportional hazard model where the contribution of other factors was controlled for (i.e., stimulability, expressive language ability, intelligibility, interaction between age and size of the initial consonant inventory, and sex). In other words, children with atypical error patterns may not necessarily take a longer time to normalization than those children with typical developmental errors when no intervention was received, contrary to expectations from previous literature (cf. Morgan et al., 2017). That is, some participants demonstrated atypical error patterns at Stage 1 and then resolved all errors and achieved speech normalization within the study period without any speech-language pathology intervention. This contrasts with what was observed by Morgan et al. (2017), who found that type of error (i.e., "delay or disorder," p. 201) was the only significant predictor in the logistic regression model that regressed error types, sex, and PCC at the age of 4 years against normalization. There may be several reasons for the discrepant findings between this study and that of Morgan et al. First, attrition in the Morgan et al. study (67/160, 41.9%) might have introduced bias that was not taken into account in the final modeling. In comparison, the attrition rate of this study was relatively smaller (15.8% of all 82 children and 20.9% of those without intervention). The attrition sample in the Morgan et al. study was reported to show similar socioeconomic status, sex, family history, and nonverbal IQ as the sample under investigation. However, it was not clear if the children who declined further follow-up (54/160, 33.8%) showed a "delayed" versus a "disordered" profile or less severe SSD at the age of 4 years. The data loss may lead to different results in the final regression model. In addition, it was reported that the "data were limited on whether children were assessed or received therapy, and no detail was provided on what type of therapy was applied" (p. 202). It is possible that children even with a delayed profile may only resolve the speech errors under the intervention condition. Thus, further research is required to test Morgan et al.'s (2017) claim that having a delayed pattern was more likely to resolve when compared to those who made atypical errors who may need more support. A second explanation may be possible differences between atypical errors in English compared with atypical errors in Cantonese. For example, although many atypical errors overlap, backing is considered to be atypical in English, but typical in Cantonese (To et al., 2013). Another explanation for the nonsignificance of the error type in the current model may be due to the inclusion of other potential factors that will be discussed in the following section.

### Stimulability

Stimulability was found to be a useful predictor of speech normalization in the no-intervention condition after controlling for the effect of error type, sex, passing the expressive language test, passing the mean ICS-TC cutoff, and the interaction of age and size of inventory. Stimulability has long been reported as a strong positive prognostic factor of SSD put forward in 1931. Powell and Miccio (1996) argued that if children can imitate misarticulated consonants with stimulation, those consonants are likely to be added to the phonetic inventory even without intervention. Stimulability is therefore considered a dynamic assessment (Bain, 1994; Glaspey & Stoel-Gammon, 2007) that indicates children's potential within their zone of proximal development (Vygotsky, 1978). Being stimulable entails the integrity of the sensory input and generally intact linguistic and motor output system (Powell & Miccio, 1996). Lof (1996) also argued that stimulability may reflect a child's focus (Kwiatkowski & Shriberg, 1993), implying the propensity of a child to focus on the productions and being motivated to change are essential. A number of early studies have provided consistent findings supporting the important role of stimulability in prognosis during the early 1960s (Bain, 1994; Carter & Buck, 1958; Irwin et al., 1966; Milisen, 1954). With the emergence of phonological assessment procedures between the 1970s and 1980s, more research and clinical attention was devoted to the systematic changes of the consonant inventory and rule-based learning of speech acquisition. The articulatory aspects of speech production including stimulability may have been de-emphasized, and the focus on stimulability research shifted from an assessment tool to its contribution to generalization in intervention (Miccio & Elbert, 1996; Powell, 1996; Powell et al., 1991). The current findings reiterate the valuable role of stimulability in the evaluation process. It is worth noting that stimulability was defined as being stimulable for all misarticulated consonant phonemes in at least two different consonant–vowel nonsense syllables rather than in isolation or in real words as in previous studies (Flint & Ingham, 2005; Lof, 1996). The decision to use syllable-level stimuli was based on the fact that syllable appears to be a particularly important unit for speech motor control (Levelt, 1999; Tourville & Guenther, 2011). Not using real word stimuli was to avoid children accessing their existing motor plans of the words that may have been well learned but incorrect. Even with much stimulation, it is difficult to change the form of production. If stimulability aims to explore children's potential or their learnability in the production of speech sounds independent of semantic influence, nonsense syllables are deemed more appropriate. A dichotomous measure of stimulable versus nonstimulable was adopted in this study to represent children's global ability to imitate misarticulated consonants. With this binary differentiation, some children were stimulable to all misarticulated consonants, whereas some were nonstimulable to all consonants and some were stimulable to some consonants. It is possible when the measure is phoneme based and the target outcome is graded rather than binary, the trend may be even more robust.

### Intelligibility

The second useful predictor of normalization was parent-reported intelligibility, specifically passing the mean ICS-TC cutoff. This implies that children who exhibited a mean ICS-TC that was lower than the stated cutoff (less intelligible) were less likely to normalize and took longer time to normalization. The finding that parent-reported intelligibility was associated with children's time to normalization whereas atypical errors were not adds to the worth and convenience of this quick measure, furthering the call that these data be routinely collected by SLPs (Ireland et al., 2020). Parent-reported intelligibility on the ICS has been found to be correlated with PCC in many different languages and countries (McLeod, 2020).

### Expressive Language Ability

Finally, expressive language ability was not significantly associated with speech normalization. This observation was generally in line with the study of Morgan et al. (2017), who reported that core language scores were not a significant predictor of later speech outcomes at the age of 7 years. In other words, passing an expressive language test may not be a protective factor of SSD such that children with age-appropriate expressive language skills may or may not show normalization of their speech errors.

## Limitations and Future Studies

This study focused on children who did not receive intervention over the 2.5-year study period. Although the study started with a large sample, the sample size of the population of interest was relatively small. Those 39 children who started speech-language pathology services during the study period were shown to have significantly more atypical error patterns and a smaller consonant inventory at the initial time point. Future analyses may include a greater range of children with SSD and may also include individuals receiving current standard care, which may affect the manifestation of SSD. Although a rigorous three-layer reliability check was used, transcription reliability measures were not calculated. Additionally, although there is evidence that the amount of infant- and child-directed speech influences children's language (Cartmill et al., 2013; Dilley et al., 2020; Montag et al., 2018; Werker, 2018), there is limited research about the influence on the likelihood of the resolution of SSD. Future research could be conducted to consider the variable of infant- and child-directed speech modifications to engage attention and stimulate speech and language. Finally, although much evidence has substantiated that SSD is associated with later language and literacy issues in English-speaking children, longer term follow-up of these children who resolved errors by self-correction and those with intervention can provide a valuable means for determining which factors are associated with literacy outcomes.

## Clinical Implications and Conclusions

This study revealed that when no intervention was provided for children with SSD, longer term (2.5 years) outcome may be predicted by stimulability and intelligibility, but not by atypical errors or expressive language ability. The median time to normalization was 6.59 years of age. Children with greater stimulability and higher intelligibility were more likely to resolve their errors naturally and took a shorter time to normalize. Atypical error patterns and expressive language ability, however, were not useful prognostic factors of speech normalization. In other words, they did not represent additional risk factors on speech sound acquisition, and children who showed atypical error patterns and/or weaker language ability did not necessarily take longer to normalize. In conclusion, children who present with better stimulability and higher intelligibility are more likely to represent instances of typical developmental variation rather than atypical development. Speech-language services may speed up speech normalization in these children. However, for children who can self-correct their errors (i.e., children who are stimulable) without intervention and for as long as these errors do not impact their intelligibility, prescribing intervention targeting at speech sound may not be necessary. Paul (2000) discussed the impact of ascribing unnecessary intervention to children with circumscribed early language delay by suggesting they use their time developing other talents or interests. The present results suggest that children with low intelligibility and poor stimulability should be prioritized for speech-language pathology services since their speech errors are less likely to resolve naturally. Finally, stimulability testing and a speech intelligibility rating by caregivers are important components of routine clinical assessment for children with potential SSD, and these measures can be used for caseload prioritization.

## Acknowledgments

This research was supported by the General Research Fund, Research Grants Council of Hong Kong SAR, China, awarded to C. K. S. To, S. McLeod, T. Law, K. Y. S. Lee, M. W. L. Ng, and K. Y. Cheung. The authors thank Kathy Lee, Manwa L. Ng, and Kayan Cheung for their support. Special thanks go to Kelvin Cheung, Wing Heen Ng, Daphne Sze, and all student helpers who assisted in data collection and transcription. Last but not least, the authors are grateful to all children and parents who participated in this study.

## References

**Bain, B. A.** (1994). A framework for dynamic assessment in phonology: Stimulability revisited. *Clinics in Communication Disorders, 4*(1), 12–22.

**Baker, L., & Cantwell, D. P.** (1987). A prospective psychiatric follow-up of children with speech/language disorders. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*(4), 546–553. https://doi.org/10.1097/00004583-198707000-00015

**Bates, E., Dale, P. S., & Thal, D.** (1995). Individual differences and their implications for theories of language development. *The Handbook of Child Language, 30,* 96–151.

**Bralley, R. C., & Stoudt, R. J.** (1977). A five-year longitudinal study of development of articulation proficiency in elementary school children. *Language, Speech, and Hearing Services in Schools, 8*(3), 176–180. https://doi.org/10.1044/0161-1461.0803.176

**Campbell, T. F., Dollaghan, C. A., Rockette, H. E., Paradise, J. L., Feldman, H. M., Shriberg, L. D., Sabao, D. L., & Kurs-Lasky, M.** (2003). Risk factors for speech delay of unknown origin in 3-year-old children. *Child Development, 74*(2), 346–357. https://doi.org/10.1111/1467-8624.7402002

**Carter, E. T., & Buck, M.** (1958). Prognostic testing for functional articulation disorders among children in the first grade. *Journal of Speech and Hearing Disorders, 23*(2), 124–133. https://doi.org/10.1044/jshd.2302.124

**Cartmill, E. A., Armstrong, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C.** (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences, 110*(28), 11278–11283. https://doi.org/10.1073/pnas.1309518110

**Cheung, P. S. P., Lee, K. Y. S., & Lee, L. W. T.** (1997). The development of the Cantonese Receptive Vocabulary Test for children aged 2–6 in Hong Kong. *European Journal of Disorders of Communication, 32*(1), 127–138. https://doi.org/10.3109/13682829709021465

**Cheung, P. S. P., Ng, A. K. H., & To, C. K. S.** (2006). *Hong Kong Cantonese Articulation Test*. Language Information Sciences Research Centre, City University of Hong Kong.

**Child Assessment Service.** (2021). *Your clinic visit*. https://www.dhcas.gov.hk/en/assessment_journey.html

**Davis, B. L., & Bedore, L. M.** (2013). *An emergence approach to speech acquisition: Doing and knowing*. Psychology Press. https://doi.org/10.4324/9780203375303

**Diedrich, W. M.** (1983). Stimulability and articulation disorders. *Seminars in Speech and Language, 4*(4), 297–311.

**Dilley, L., Lehet, M., Wieland, E. A., Arjmandi, M. K., Kondaurova, M., Wang, Y., Reed, J., Svirsky, M., Houston,**

**D., & Bergeson, T.** (2020). Individual differences in mothers' spontaneous infant-directed speech predict language attainment in children with cochlear implants. *Journal of Speech, Language, and Hearing Research, 63*(7), 2453–2467. https://doi.org/10.1044/2020_JSLHR-19-00229

**Dinnsen, D. A., & Elbert, M.** (1984). On the relationship between phonology and learning. In M. Elbert, D. A. Dinnsen, & G. Weismer (Eds.), *Phonological theory and the misarticulating child* (ASHA Monographs No. 22, pp. 59–68). American Speech-Language-Hearing Association.

**Dodd, B.** (2014). Differential diagnosis of pediatric speech sound disorder. *Current Developmental Disorders Reports, 1*(3), 189–196. https://doi.org/10.1007/s40474-014-0017-3

**Dodd, B., Ttofari-Eecen, K., Brommeyer, K., Ng, K., Reilly, S., & Morgan, A.** (2018). Delayed and disordered development of articulation and phonology between four and seven years. *Child Language Teaching and Therapy, 34*(2), 87–99. https://doi.org/10.1177/0265659017735958

**Education Bureau.** (2021). *Professional support for students with educational needs from the Education Bureau*. https://www.edb.gov.hk/en/edu-system/special/support/wsa/pro-support/index.html

**Family Health Service.** (2021). *Developmental surveillance scheme*. https://www.fhs.gov.hk/english/health_info/child/13033.html

**Ferguson, C. A.** (1979). Phonology as an individual access system: Some data from language acquisition. In C. J. Fillmore, D. Kempler, & W. S.-Y. Wang (Eds.), *Individual differences in language ability and language behavior* (pp. 189–201). Academic Press. https://doi.org/10.1016/B978-0-12-255950-1.50016-0

**Fletcher, R. H., Fletcher, S. W., & Wagner, E. H.** (1988). *Clinical epidemiology: The essentials*. Williams & Wilkins.

**Flint, C. B., & Ingham, J. C.** (2005). Pretreatment stimulability and percentage of consonants correct as predictors of across-phoneme generalization. *Contemporary Issues in Communication Science and Disorders, 32*(Spring), 53–63. https://doi.org/10.1044/cicsd_32_S_53

**Flipsen, P.** (2003). Articulation rate and speech-sound normalization failure. *Journal of Speech, Language, and Hearing Research, 46*(3), 724–737. https://doi.org/10.1044/1092-4388(2003/058)

**Glaspey, A., & Stoel-Gammon, C.** (2007). A dynamic approach to phonological assessment. *International Journal of Speech-Language Pathology, 9*(4), 286–296. https://doi.org/10.1080/14417040701435418

**Glogowska, M., Roulstone, S., Enderby, P., & Peters, T. J.** (2000). Randomised controlled trial of community based speech and language therapy in preschool children. *BMJ, 321,* 923. https://doi.org/10.1136/bmj.321.7266.923

**Gruber, F. A.** (1999a). Probability estimates and paths to consonant normalization in children with speech delay. *Journal of Speech, Language, and Hearing Research, 42*(2), 448–459. https://doi.org/10.1044/jslhr.4202.448

**Gruber, F. A.** (1999b). Tutorial: Survival analysis—A statistic for clinical, efficacy, and theoretical applications. *Journal of Speech, Language, and Hearing Research, 42*(2), 432–447. https://doi.org/10.1044/jslhr.4202.432

**Hayiou-Thomas, M. E., Carroll, J. M., Leavett, R., Hulme, C., & Snowling, M. J.** (2017). When does speech sound disorder matter for literacy? The role of disordered speech errors, co-occurring language impairment and family risk of dyslexia. *The Journal of Child Psychology and Psychiatry, 58*(2), 197–205. https://doi.org/10.1111/jcpp.12648

**Hedlund, G., & Rose, Y.** (2020). *Phon 3.3.0* [Computer software]. https://phon.ca

**Hong Kong Society for Child Health and Development.** (1987). *Manual of the Reynell Developmental Language Scales, Cantonese (Hong Kong) Version.*

**Hyde, J. S., & Linn, M. C.** (1988). Gender differences in verbal ability: A meta-analysis. *Psychological Bulletin, 104*(1), 53–69. https://doi.org/10.1037/0033-2909.104.1.53

**Ingram, D.** (1976). *Phonological disability in children* (Vol. 2). Elsevier.

**Ireland, M., McLeod, S., Farquharson, K., & Crowe, K.** (2020). Evaluating children in U.S. public schools with speech sound disorders. *Topics in Language Disorders, 40*(4), 326–340. https://doi.org/10.1097/tld.0000000000000226

**Irwin, R. B., West, J. F., & Trombetta, M. A.** (1966). Effectiveness of speech therapy for second grade children with misarticulations: Predictive factors. *Exceptional Children, 32*(7), 471–479. https://doi.org/10.1177/001440296603200705

**Kaplan, E. L., & Meier, P.** (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association, 53*(282), 457–481. https://doi.org/10.1007/978-1-4612-4380-9_25

**Kent, R. D., Miolo, G., & Bloedel, S.** (1994). The intelligibility of children's speech: A review of evaluation procedures. *American Journal of Speech-Language Pathology, 3*(2), 81–95. https://doi.org/10.1044/1058-0360.0302.81

**Kok, E. C. E., & To, C. K.** (2019). Revisiting the cutoff criteria of Intelligibility in Context Scale–Traditional Chinese. *Language, Speech, and Hearing Services in Schools, 50*(4), 629–638. https://doi.org/10.1044/2019_LSHSS-18-0073

**Kwiatkowski, J., & Shriberg, L. D.** (1993). Speech normalization in developmental phonological disorders: A retrospective study of capability-focus theory. *Language, Speech, and Hearing Services in Schools, 24*(1), 10–18. https://doi.org/10.1044/0161-1461.2401.10

**Law, J., Boyle, J., Harris, F., Harkness, A., & Nye, C.** (2000). Prevalence and natural history of primary speech and language delay: Findings from a systematic review of the literature. *International Journal of Language & Communication Disorders, 35*(2), 165–188. https://doi.org/10.1080/136828200247133

**Law, J., & Conti-Ramsden, G.** (2000). Treating children with speech and language impairments: Six hours of therapy is not enough. *BMJ, 321*(7266), 908–909. https://doi.org/10.1136/bmj.321.7266.908

**Leahy, J., & Dodd, B.** (1987). The development of disordered phonology: A case study. *Language and Cognitive Processes, 2*(2), 115–132. https://doi.org/10.1080/01690968708406353

**Leonard, L. B.** (1973). The nature of deviant articulation. *Journal of Speech and Hearing Disorders, 38*(2), 156–161. https://doi.org/10.1044/jshd.3802.156

**Levelt, W. J.** (1999). Producing spoken language: A blueprint of the speaker. In *The neurocognition of language* (pp. 83–122). Oxford University Press.

**Lewis, B. A., Freebairn, L., Tag, J., Benchek, P., Morris, N. J., Iyengar, S. K., Gerry Taylor, H., & Stein, C. M.** (2018). Heritability and longitudinal outcomes of spelling skills in individuals with histories of early speech and language disorders. *Learning and Individual Differences, 65,* 1–11. https://doi.org/10.1016/j.lindif.2018.05.001

**Lewis, B. A., Freebairn, L., Tag, J., Igo, R. P., Jr., Ciesla, A., Iyengar, S. K., Stein, C. M., & Taylor, H. G.** (2019). Differential long-term outcomes for individuals with histories of preschool speech sound disorders. *American Journal of Speech-Language Pathology, 28*(4), 1582–1596. https://doi.org/10.1044/2019_AJSLP-18-0247

**Lewis, B. A., Freebairn, L. A., & Taylor, H. G.** (2000). Follow-up of children with early expressive phonology disorders. *Journal of Learning Disabilities, 33*(5), 433–444. https://doi.org/10.1177/002221940003300504

**Lof, G. L.** (1996). Factors associated with speech-sound stimulability. *Journal of Communication Disorders, 29*(4), 255–278. https://doi.org/10.1016/0021-9924(96)00013-5

**McCormack, J., McLeod, S., McAllister, L., & Harrison, L. J.** (2009). A systematic review of the association between childhood speech impairment and participation across the lifespan. *International Journal of Speech-Language Pathology, 11*(2), 155–170. https://doi.org/10.1080/17549500802676859

**McGill, N., Crowe, K., & McLeod, S.** (2020). "Many wasted months": Stakeholders' perspectives about waiting for speech-language pathology services. *International Journal of Speech-Language Pathology, 22*(3), 313–326. https://doi.org/10.1080/17549507.2020.1747541

**McGill, N., McLeod, S., Crowe, K., Wang, C., & Hopf, S. C.** (2021). Waiting lists and prioritization of children for services: Speech-language pathologists' perspectives. *Journal of Communication Disorders, 91,* 106099. https://doi.org/10.1016/j.jcomdis.2021.106099

**McIntosh, B., & Dodd, B. J.** (2008). Two-year-old's phonological acquisition: Normative data. *International Journal of Speech-Language Pathology, 10*(6), 460–469. https://doi.org/10.1080/17549500802149683

**McLeod, S.** (2020). Intelligibility in Context Scale: Cross-linguistic use, validity, and reliability. *Speech, Language and Hearing, 23*(1), 9–16. https://doi.org/10.1080/2050571X.2020.1718837

**McLeod, S., & Baker, E.** (2017). *Children's speech: An evidence-based approach to assessment and intervention*. Pearson.

**McLeod, S., & Crowe, K.** (2018). Children's consonant acquisition in 27 languages: A cross-linguistic review. *American Journal of Speech-Language Pathology, 27*(4), 1546–1571. https://doi.org/10.1044/2018_AJSLP-17-0100

**McLeod, S., Davis, E., Rohr, K., McGill, N., Miller, K., Roberts, A., Thornton, S., Ahio, N., & Ivory, N.** (2020). Waiting for speech-language pathology services: A randomised controlled trial comparing therapy, advice and device. *International Journal of Speech-Language Pathology, 22*(3), 372–386. https://doi.org/10.1080/17549507.2020.1731600

**McLeod, S., Harrison, L. J., & McCormack, J.** (2012). The Intelligibility in Context Scale: Validity and reliability of a subjective rating measure. *Journal of Speech, Language, and Hearing Research, 55*(2), 648–656. https://doi.org/10.1044/1092-4388(2011/10-0130)

**McLeod, S., Harrison, L. J., & Wang, C.** (2019). A longitudinal population study of literacy and numeracy outcomes for children identified with speech, language, and communication needs in early childhood. *Early Childhood Research Quarterly, 47,* 507–517. https://doi.org/10.1016/j.ecresq.2018.07.004

**McLeod, S., & Hewett, S. R.** (2008). Variability in the production of words containing consonant clusters by typical 2- and 3-year-old children. *Folia Phoniatrica et Logopaedica, 60*(4), 163–172. https://doi.org/10.1159/000127835

**McLeod, S., & Masso, S.** (2019). Screening children's speech: The impact of imitated elicitation and word position. *Language, Speech, and Hearing Services in Schools, 50*(1), 71–82. https://doi.org/10.1044/2018_LSHSS-17-0141

**McLeod, S., van Doorn, J., & Reed, V. A.** (2001). Consonant cluster development in two-year-olds: General trends and individual difference. *Journal of Speech, Language, Hearing Research, 44*(5), 1144–1171. https://doi.org/10.1044/1092-4388(2001/090)

**Miccio, A. W., & Elbert, M.** (1996). Enhancing stimulability: A treatment program. *Journal of Communication Disorders, 29*(4), 335–351. https://doi.org/10.1016/0021-9924(96)00016-0

**Milisen, R.** (1954). A rationale for articulation disorders. *Journal of Speech and Hearing Disorders, 4*(Monograph Suppl), 6–17.

**Mitchell, P. R., & Kent, R. D.** (1990). Phonetic variation in multisyllable babbling. *Journal of Child Language, 17*(2), 247–265. https://doi.org/10.1017/S0305000900013751

**Montag, J. L., Jones, M. N., & Smith, L. B.** (2018). Quantity and diversity: Simulating early word learning environments. *Cognitive Science, 42*(S2), 375–412. https://doi.org/https://doi.org/10.1111/cogs.12592

**Morgan, A., Ttofari Eecen, K., Pezic, A., Brommeyer, K., Mei, C., Eadie, P., Reilly, S., & Dodd, B.** (2017). Who to refer for speech therapy at 4 years of age versus who to "watch and wait." *The Journal of Pediatrics, 185,* 200–204.e1. https://doi.org/10.1016/j.jpeds.2017.02.059

**Mulcair, G., Pietranton, A. A., & Williams, C.** (2018). The International Communication Project: Raising global awareness of communication as a human right. *International Journal of Speech-Language Pathology, 20*(1), 34–38. https://doi.org/10.1080/17549507.2018.1422023

**Munson, B.** (2001). Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language, and Hearing Research, 44*(4), 778–792. https://doi.org/10.1044/1092-4388(2001/061)

**Nelson, K.** (1981). Individual differences in language development: Implications for development and language. *Developmental Psychology, 17*(2), 170–187. https://doi.org/10.1037/0012-1649.17.2.170

**Ng, K. Y. M., To, C. K. S., & Mcleod, S.** (2014). Validation of the Intelligibility in Context Scale as a screening tool for preschoolers in Hong Kong. *Clinical Linguistics & Phonetics, 28*(5), 316–328. https://doi.org/10.3109/02699206.2013.865789

**Oberklaid, F., Wake, M., Harris, C., Hesketh, K., & Wright, M.** (2002). *Child health screening and surveillance: A critical review of the evidence*. National Health and Medical Research Council.

**Pascoe, M., Stackhouse, J., & Wells, B.** (2005). Phonological therapy within a psycholinguistic framework: Promoting change in a child with persisting speech difficulties. *International Journal of Language & Communication Disorders, 40*(2), 189–220. https://doi.org/10.1080/13682820412331290979

**Pascoe, M., Stackhouse, J., & Wells, B.** (2006). *Persisting speech difficulties in children: children's speech and literacy difficulties*. Wiley.

**Paul, R.** (2000). Predicting outcomes of early expressive language delay: Ethical implications. In D. V. M. Bishop & L. B. Leonard (Eds.), *Speech and language impairments in children: Causes, characteristics, intervention and outcome* (pp. 195–209). Psychology Press.

**Paul, R., & Jennings, P.** (1992). Phonological behavior in toddlers with slow expressive language development. *Journal of Speech and Hearing Research, 35*(1), 99–107. https://doi.org/10.1044/jshr.3501.99

**Pennington, B. F., & Bishop, D. V.** (2009). Relations among speech, language, and reading disorders. *Annual Review of Psychology, 60*(1), 283–306. https://doi.org/10.1146/annurev.psych.60.110707.163548

**Powell, T. W.** (1996). Stimulability considerations in the phonological treatment of a child with a persistent disorder of speech-sound production. *Journal of Communication Disorders, 29*(4), 315–333. https://doi.org/10.1016/0021-9924(96)00015-9

**Powell, T. W., Elbert, M., & Dinnsen, D. A.** (1991). Stimulability as a factor in the phonological generalization of misarticulating preschool children. *Journal of Speech and Hearing Research, 34*(6), 1318–1328. https://doi.org/10.1044/jshr.3406.1318

**Powell, T. W., & Miccio, A. W.** (1996). Stimulability: A useful clinical tool. *Journal of Communication Disorders, 29*(4), 237–253. https://doi.org/10.1016/0021-9924(96)00012-3

**Preston, J. L., & Edwards, M. L.** (2009). Speed and accuracy of rapid speech output by adolescents with residual speech sound errors including rhotics. *Clinical Linguistics and Phonetics, 23*(4), 301–318. https://doi.org/10.1080/02699200802680833

**Preston, J. L., McAllister, T., Phillips, E., Boyce, S., Tiede, M., Kim, J. S., & Whalen, D. H.** (2018). Treatment for residual rhotic errors with High- and low-frequency ultrasound visual feedback: A single-case experimental design. *Journal of Speech, Language, and Hearing Research, 61*(8), 1875–1892. https://doi.org/10.1044/2018_JSLHR-S-17-0441

**Roulstone, S., Miller, L. L., Wren, Y., & Peters, T. J.** (2009). The natural history of speech Impairment of 8-year-old children in the Avon Longitudinal Study of Parents and Children: Error rates at 2 and 5 years. *International Journal of Speech-Language Pathology, 11*(5), 381–391. https://doi.org/10.1080/17549500903125111

**Shore, C. M.** (1994). *Individual differences in language development*. Sage.

**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. https://doi.org/10.1044/jshr.3601.105

**Shriberg, L. D., & Austin, D.** (1998). Comorbidity of speech-language disorder: Implications for a phenotype marker for speech delay. In R. Paul (Ed.), *The speech-language connection* (pp. 73–117). Brookes.

**Shriberg, L. D., & Kwiatkowski, J.** (1980). *Natural process analysis (NPA): A procedure for phonological analysis of continuous speech samples*. Wiley.

**Shriberg, L. D., Kwiatkowski, J., Best, S., Terselic-Weber, B., & Hengst, J.** (1986). Characteristics of children with phonologic disorders of unknown origin. *Journal of Speech and Hearing Disorders, 51*(2), 140–161. https://doi.org/10.1044/jshd.5102.140

**Shriberg, L. D., Kwiatkowski, J., & Gruber, F. A.** (1994). Developmental phonological disorders. II: Short-term speech-sound normalization. *Journal of Speech and Hearing Research, 37*(5), 1127–1150. https://doi.org/10.1044/jshr.3705.1127

**Shriberg, L. D., Tomblin, J. B., & McSweeny, J. L.** (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language, and Hearing Research, 42*(6), 1461–1481. https://doi.org/10.1044/jslhr.4206.1461

**Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A.** (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55*(4), 779–798. https://doi.org/10.1044/jshd.5504.779

**Sommers, R. K., Leiss, R. H., Delp, M. A., Gerber, A. J., Fundrella, D., Smith, R. M., Revucky, M. V., Ellis, D., & Haley, V. A.** (1967). Factors related to the effectiveness of articulation therapy for kindergarten, first, and second grade children. *Journal of Speech and Hearing Research, 10*(3), 428–437. https://doi.org/10.1044/jshr.1003.428

**Stackhouse, J., & Wells, B.** (1997). *Children's speech and literacy difficulties. I: A psycholinguistic framework*. Whurr.

**Stampe, D.** (1969). The acquisition of phonetic representation. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society* (Vol. 5, No. 1, pp. 443–454). Chicago Linguistic Society.

**Storkel, H. L., Maekawa, J., & Hoover, J. R.** (2010). Differentiating the effects of phonotactic probability and neighborhood density on vocabulary comprehension and production: A comparison of preschool children with versus without phonological delays. *Journal of Speech, Language, and Hearing Research, 53*(4), 933–949. https://doi.org/10.1044/1092-4388(2009/09-0075)

**Subtelny, J.** (1977). Assessment of speech with implications for training. In F. Bess (Ed.), *Childhood deafness* (pp. 183–194). Grune & Stratton.

**To, C. K. S., Cheung, P. S., & McLeod, S.** (2013). A population study of children's acquisition of Hong Kong Cantonese consonants, vowels, and tones. *Journal of Speech, Language, and Hearing Research, 56*(1), 103–122. https://doi.org/10.1044/1092-4388(2012/11-0080)

**To, C. K. S., Law, T., & Cheung, P. S. P.** (2012). Treatment intensity in everyday clinical management of speech sound disorders in Hong Kong. *International Journal of Speech-Language Pathology, 14*(5), 462–466. https://doi.org/10.3109/17549507.2012.688867

**To, C. K. S., Stokes, S. F., Cheung, H.-T., & T'sou, B.** (2010). Narrative assessment for Cantonese-speaking children. *Journal of Speech, Language, and Hearing Research, 53*(3), 648–669. https://doi.org/10.1044/1092-4388(2009/08-0039)

**Tomblin, J. B., Oleson, J. J., Ambrose, S. E., Walker, E., & Moeller, M. P.** (2014). The influence of hearing aids on the speech and language development of children with hearing loss. *JAMA Otolaryngology-Head & Neck Surgery, 140*(5), 403–409. https://doi.org/10.1001/jamaoto.2014.267

**Tourville, J. A., & Guenther, F. H.** (2011). The DIVA model: A neural theory of speech acquisition and production. *Language and Cognitive Processes, 26*(7), 952–981. https://doi.org/10.1080/01690960903498424

**Travis, L. E.** (1931). *Speech pathology*. Appleton.

**T'sou, B., Lee, T. H.-T., Tung, P., Man, Y., Chan, A., To, C. K. S., & Chan, Y.** (2006). *Hong Kong Cantonese Oral Language Assessment Scale*. Language Information Sciences Research Centre, City University of Hong Kong.

**Vihman, M. M., Ferguson, C. A., & Elbert, M.** (1986). Phonological development from babbling to speech: Common tendencies and individual differences. *Applied Psycholinguistics, 7*(1), 3–40. https://doi.org/10.1017/S0142716400007165

**Vygotsky, L.** (1978). Interaction between learning and development. *Readings on the Development of Children, 23*(3), 34–41.

**Werker, J. F.** (2018). Perceptual beginnings to language acquisition. *Applied Psycholinguistics, 39*(4), 703–728. https://doi.org/10.1017/S0142716418000152

**Williams, A. L., McLeod, S., & McCauley, R. J.** (2021). *Interventions for speech sound disorders in children* (2nd ed.). Brookes.

**Wren, Y., McLeod, S., White, P., Miller, L. L., & Roulstone, S.** (2013). Speech characteristics of 8-year-old children: Findings from a prospective population study. *Journal of Communication Disorders, 46*(1), 53–69. https://doi.org/10.1016/j.jcomdis.2012.08.008

**Wren, Y., Miller, L. L., Peters, T. J., Emond, A., & Roulstone, S.** (2016). Prevalence and predictors of persistent speech sound disorder at eight years old: Findings from a population cohort study. *Journal of Speech, Language, and Hearing Research, 59*(4), 647–673. https://doi.org/10.1044/2015_JSLHR-S-14-0282

**Zamuner, T. S., Gerken, L., & Hammond, M.** (2004). Phonotactic probabilities in young children's speech production. *Journal of Child Language, 31*(3), 515–536. https://doi.org/10.1017/S0305000904006233