



ТЕСТИРАЊЕ НА ВЕРОДОСТОЈНОСТА НА НЕЗАВИСНИТЕ ФОНОЛОШКИ МЕРКИ КАЈ ГОВОРНОТ НА ДВЕГОДИШНО ДЕТЕ: ПИЛОТ-СТУДИЈА

Кетрин Мари ВИТЛЕР
Шари Ли ДЕВЕНЕИ

Оддел за специјална едукација
и комуникациски нарушувања,
Универзитет на Небраска во Омаха, САД

Примено: 17.03.2016
Прифатено: 29.06.2016

Резиме

Вовед: Во областа на патологијата на говорниот јазик, многумина претпоставуваат дека најчесто употребуваните неформални мерки на звукот на говорот се сигурни. Сепак, недостатокот на научни докази за поддршка на оваа претпоставка е проблематичен. Логопедите често користат звучни анализи на неформалниот говор за утврдување на основните однесувања според кои може да се мери терапевтскиот напредок. Неколкумина истражувачи ја испитале веродостојноста од тест-ретестирањето на неформалните фонолошки мерки при оценување на говорот кај мали деца. Од клинички аспект, податоците поврзани со овие мерки се од клучно значење за олеснување на донесувањето одлуки врз основа на докази а во однос на проценката и третманот на говорниот јазик.

Цели: Целта на оваа студија беше да се идентификува базата на докази во врска со привремената веродостојност на две такви неформални мерки на говорниот звук, фонетскиот инвентар и анализата на зборовната форма кај двегодишни деца.

TEST-RETEST RELIABILITY OF INDEPENDENT PHONOLOGICAL MEASURES OF 2-YEAR-OLD SPEECH: A PILOT STUDY

Katherine Marie WITTLER
Shari Leigh DEVENEY

Department of Special Education
and Communication Disorders,
University of Nebraska at Omaha, USA

Received: 17.03.2016
Accepted: 29.06.2016
Original Article

Abstract

Introduction: Within the field of speech-language pathology, many assume commonly used informal speech sound measures are reliable. However, lack of scientific evidence to support this assumption is problematic. Speech-language pathologists often use informal speech sound analyses for establishing baseline behaviors from which therapeutic progress can be measured. Few researchers have examined the test-retest reliability of informal phonological measures when evaluating the speech productions of young children. Clinically, data regarding these measures are critical for facilitating evidence-based decision making for speech-language assessment and treatment.

Objectives: The aim of the present study was to identify the evidence-base regarding temporal reliability of two such informal speech sound measures, phonetic inventory and word shape analysis, with two-year-old children.

Адреса за кореспонденција:

Шари ДЕВЕНЕИ

Department of Special Education
and Communication Disorders,
University of Nebraska at Omaha,
6005 Dodge Street,
512E Roskens Hall, Omaha, NE 68182; САД
Тел: (402) 554-2993, (402)968-6083
Е-пошта: sdevaney@unomaha.edu

Corresponding address:

Shari DEVENEY

Department of Special Education
and Communication Disorders,
University of Nebraska at Omaha,
6005 Dodge Street,
512E Roskens Hall, Omaha, NE 68182; USA
Phone: (402) 554-2993, (402)968-6083
E-mail: sdevaney@unomaha.edu



Методи: Истражувачите ги испитувале анализите спроведени од примероци на разговорен говор земени со точно по една недела растојание кај три деца на 29 до 33-месечна возраст. Снимените 20-минутни конверзации во форма на игра биле направени додека децата правеле некаква интерактивна активност со нивните мајки. Потоа примероците биле транскрибирани со помош на Меѓународната фонетска азбука (МФА) и биле анализирани со помош на двете неформални мерки наведени погоре.

Резултати: Врз основа на визуелниот преглед на податоците, веродостојноста на тест-ретестираните изведби на почетната согласка и на групи согласки беше нестабилна меѓу двата разговорни примероци. Сепак, фонетските инвентари за последните согласки и анализи на формата на зборовите беа релативно стабилни.

Заклучок: Иако се потребни повеќе податоци, резултатите од оваа студија укажуваат дека академскиот кадар, клиничките едукатори и логопедите треба да бидат претпазливи при толкување на анализите на неформалниот говор засновани на комуникациски примероци базирани на игра кај малите деца.

Клучни зборови: комуникациски нарушувања, фонетски инвентар, фонологија, говор, патологија на говорниот јазик, анализа на зборовната форма.

Вовед

Логопедите се потпираат на формални (на пример, стандардизирани тестирања базирани на критериуми) и неформални (на пример, опсервации, клиничка проценка, нестандардизирани) мерки на проценка за сеопфатна евалуација на развојот на говорот и јазикот кај детето. Овие мерки се користат за утврдување на подобноста на услугите за интервенција овозможени во образовни или клинички установи. Тие исто така нудат описни информации од клучно значење за утврдување на основната форма на изразување кај детето со цел следење на напредокот на терапевтските алатки за говорниот јазик (1). Неформалните алатки за проценка не се наменети за споредба со поголема група (на пример, не се стандардизирани или нормативни). Некои видови неформални мерки за оценување вклучуваат клиничко набљудување, досие на родител, како и анализа на примерок од говорот на детето добиени од набљудувањето за време

Methods: The researchers examined analyses conducted from conversational speech samples taken exactly one week apart for three children 29- to 33-months of age. The videotaped 20-minute play-based conversational samples were completed while the children interacted with their mothers. The samples were then transcribed using the International Phonetic Alphabet (IPA) and analyzed using the two informal measures noted above.

Results: Based on visual inspection of the data, the test-retest reliability of initial consonant and consonant cluster productions was unstable between the two conversational samples. However, phonetic inventories for final consonants and word shape analyses were relatively stable over time.

Conclusion: Although more data is needed, the results of this study indicate that academic faculty, clinical educators, and practicing speech-language pathologists should be cautious when interpreting informal speech sound analyses based on play-based communication samples of young children.

Keywords: communication disorders, phonetic inventory, phonology, speech, speech-language pathology, word shape analysis

Introduction

Speech-language pathologists (SLPs) rely on formal (i.e., standardized, criterion-based testing) and informal (i.e., observations, clinical judgment, non-standardized) assessment measures to comprehensively evaluate a child's speech and language development. These measures are used for determining eligibility for intervention services provided in educational or clinical settings. They also offer descriptive information critical to establishing a child's baseline performance for therapeutic speech-language progress monitoring (1). Informal assessment tools are not intended to be compared to a larger group (i.e., not standardized or norm-referenced). Types of informal assessment measures include clinical observation, parent report, and analysis of a sample of the child's speech obtained from a



на игра или разговор.

Независни фонолошки мерки, еден вид анализа на говорен примерок, се користат за проценка на продукциите на говорни звуци, без споредба со стандардите кај возрастни луѓе, а во суштина на SLP му обезбедуваат информации за она што детето може да произведе, а не што тој / таа не може (2). Постојат повеќе неформални независни фонолошки (на пример, говорен звук) анализи што им се на располагање на SLP, вклучувајќи фонетски инвентари и анализи на зборовната форма. Фонетски инвентар е записник на различните звуци што детето ги користи, дури и на звуци што не се произведени во точната зборовна позиција (2). Еден логопед ќе го заврши индексот на сите звуци при говорот на детето без споредба со зборовните форми кај возрастните со цел да се утврдат говорните звуци присутни во говорот на детето. На пример, ако детето рече: „tat“ наместо „cat“, производството на / t / на почетокот на зборот е снимено во инвентарот на говорните звуци и покрај тоа што обидот на детето да ја каже возрастната зборовна форма „cat“ била неточна. За детето ќе се знае дека може да каже / t / во почетна позиција на зборот во примерот погоре, иако зборот „cat“ беше произведен погрешно. Анализата на зборовната форма е записник на комплексноста на звучните комбинации кои детето ги користи во зборовите што ги кажува. На пример, зборовите со само една согласка и самогласка се помалку сложени од зборовите со согласка, самогласка, а потоа друга согласка. Двете анализи на фонетски инвентари и на зборовната форма се важни во дијагностичкиот процес за утврдување на основната линија на третман. Односно, фонетскиот инвентар му пренесува на SLP кои звуци детето ги произведува и во кои позиции во зборот (иницијална, медијална, финална и / или група согласки), па SLP има идеја за тоа од каде да започне со третман за да се насочи кон звуците што детето сè уште не ги произведува. Анализата на зборовната форма дава информации за нивото на комплексноста на звучните комбинации коишто обично ги користи детето.

Биле спроведени неколку студии што ја испитуваат веродостојноста на тест-ретестираните неформални фонолошки мерки, како фонетскиот инвентар и анализата на зборовната форма при оценување на говорот кај ма-

play observation or conversation.

Independent phonological measures, one type of speech sample analysis, are used to assess speech-sound productions without comparison to an adult standard and essentially provide the SLP with information as regards what the child can produce, rather than what he/she cannot.(2) There are several informal independent phonological (i.e., speech sound) analyses available to SLPs including *phonetic inventories* and *word shape analyses*. Phonetic inventory is a record of different sounds the child uses, even sounds not produced in the correct word position (2). An SLP completes an index of all the speech sounds a child produces without reference to adult word forms in order to determine the speech sounds present in the child's conversational speech. For example, if a child said "tat" in place of "cat," the /t/ production at the beginning of the word is recorded in a speech sound inventory even though the child's attempt of the adult word form "cat" was inaccurate. The child would be credited for production of the /t/ sound in initial position in the above example even though the word "cat" was produced incorrectly. Word shape analysis is a record of the complexity of sound combinations used by the child in the words that they produce. For example, words with only a consonant (C) and vowel (V) sound (e.g., 'CV' sound shape such as in "go" or "she") are less complex than words with a consonant, vowel, and then another consonant sound. Both phonetic inventories and word shape analyses are important in the diagnostic process for establishing a treatment baseline; that is, phonetic inventory conveys to the SLP which sounds the child produce and in which word positions (initial, medial, final, and/or consonant cluster) so the SLP has an idea of where to begin implementing treatment to target the sounds the child does not yet produce. Word shape analysis provides information on the complexity level of sound combinations typically used by the child.

Few studies have been conducted in order to examine the test-retest reliability of informal phonological measures like phonetic inventory and word shape analysis when evaluating the



лите деца. Еден истражувач, Morris (3) ги оценил примероците на говор на десет нормално развиени деца на возраст од 18 до 22 месеци користејќи го фонетскиот инвентар и анализата на зборовната форма. За децата било утврдено дека се со нормален развој преку користење на комбинација од мерки за скрининг на јазик и речник. Откако била утврдена веродостојноста од оваа метода, секој пар мајка - дете учествувале во две 20-минутни сесии во форма на игра, меѓусебно одделени точно една недела. Говорот на децата потоа беше анализиран од видеозаписите користејќи независни фонолошки мерки за анализа. Резултатите од оваа студија покажаа дека веродостојноста на тест-ретестираните анализи спроведени на два различни примероци за говор од едно исто дете а во период од две последователни недели, беа нестабилни и не го претставуваа истиот број или опсег на произведени звуци. Morris открил дека за деца во оваа возрасна група, бројот на продукции на првата согласка се најнестабилна мерка. Продукцијата на финалната согласка и анализата на зборовната форма беше констатирано дека се умерено стабилни во период од една недела, но ниту една не го достигна потребното ниво на значење. Morris истакна дека ако неформалните мерки не се постојани цело време, забележаните подобрувања во текот на терапијата на говорниот јазик не може да претставуваат вистински напредок, туку наместо тоа, може да се сметаат за „производ на нестабилна мерка“ (3). Preston, Ramsdell, Oller, Edwards, и Tobin предложија користење на адаптирани мерки за точноста на говорните звуци и се повикаа на студијата на Morris (3) како причина за тоа. Адаптираните мерки што се користат во нивната студија, *Адаптирани мерки за точноста на говорните звуци* (WSSA), беше независна фонолошка мерка (4). Таа функционира слично како анализата *Процент на точни консонанти* (PCC) во која зборовните продукции на децата се споредуваат со стандардните на возрасни и процентот што е точно произведен се пресметува врз основа на бројот на согласни фонеме (звуци) што детето ги произведува на ист начин како и возрасните (5). Меѓутоа, со методата WSSA, различни видови грешки се бројат поинаку. На пример, необичните грешки и испуштањето на фонемите се бројат повеќе отколку грешките кои вклучуваат вообичаени замени. WSSA беше

speech productions of young children. One researcher, Morris (3) evaluated the speech samples of ten typically developing 18-to-22 month old children using phonetic inventory and word shape analysis. Children were determined to be typically developing through use of a combination of language and vocabulary screening measures. Once eligibility was determined, each mother-child dyad participated in two 20-minute play sessions occurring exactly one week apart. The children's speech was then analyzed from the videotaped samples using independent phonological analysis measures. Results of this study showed that the test-retest reliability of analyses conducted on two different speech samples collected from the same child one week apart were unstable and did not necessarily represent the same number or range of speech sounds produced. Particularly, Morris found that for children in this age range the number of initial consonant productions were the least stable measure over time. Final consonant productions and word shape analyses were determined to be moderately stable over the one-week time span, but neither reached a required level of significance. Morris noted that if the informal measures were not consistent over time, the perceived improvements during speech-language therapy may not represent true progress. Instead, it may be considered as "an artifact of an unstable measure" (3).

Preston, Ramsdell, Oller, Edwards, and Tobin suggested using a weighted measure for speech sound accuracy and referenced the Morris (3) study as a reason for doing so. The weighted measures used in their study, the *Weighted Speech Sound Accuracy Measure* (WSSA), was an independent phonological measure. (4) It functions as the *Percent Consonants Correct* (PCC) analysis, in which children's word productions are compared to the adult standard of production and the percentage produced accurately is calculated based on the number of consonant phonemes (sounds) the child produces in the same way as adults (5). However, with the WSSA, different types of errors are weighted differently. For example, unusual errors and



искористена за да се анализираат примероците на говор кај деца на различна возраст, и тоа и кај деца со нормален развој и кај деца со нарушена продукција на говорни звуци. Резултатите покажаа дека и веродостојноста и валидноста е висока кога WSSA се користи како мерка за анализа. Оваа истражувачка група овозможи дополнително оправдување за употреба на независни фонолошки мерки на анализа, но употребата на метриката на WSSA не била применета на пошироки контексти на говор, вклучувајќи примероци од поврзан говор.

Друго истражување ја поддржа употребата на неформални релативски мерки за фонолошка анализа (на пример, споредување на исказот на детето наспроти зборовната форма на возрасните), бидејќи веродостојноста на тест-ретестирањето беше силна кога примероци од комуникацијата беа собрани од мали деца на училишна возраст, преку примена на интервју (6). Истражувачите собраа примероци на говор од 20 деца од градинка, коишто беа изложени на ризик од поспор развој на говор и/или јазик. Примероците беа земени со растојание од една недела. Потоа, истражувачите ја измерија точноста на говорниот звук преку користење на неформални релативски мерки. Наодите од студијата покажаа силна веродостојност на тест-ретестирањето на неформално релативските анализи пресметани од структурирани примероци на комуникација за оваа возрасна група.

Van Severen, Van Den Berg, Molemans, и Gillis (7) ги проучувале ефектите од големината на разговорниот говор кај примероци од 30 холандски учесници на возраст од 6 до 24 месеци. Тие пресметале фонетски инвентари користејќи метода на повторување (на пример, тие постојано влечеле случајни примероци од разговорниот говор од еден поголем, подолг примерок). Истражувачите заклучиле дека биле пронајдени недоследности за оваа млада возраст и веродостојноста за понатамошни неформални анализи базирани врз примероци од разговорниот говор, во голема мера зависи од должината на вкупниот добиен примерок. Со оглед на претходните наоди на недоследности во врска со употребата на примероци од разговорниот говор на мали деца и веродостојноста на последователните неформални мерки, целта на овој наод беше основана. Оваа истражувачка студија имаше за цел да

phoneme omissions are weighted more heavily than errors involving common substitutions. The WSSA was used to analyze speech samples from a variety of age ranges for both typically developing and disordered speech sound productions. Results indicated that both reliability and validity is high when the WSSA is used as an analysis measure. This research group provided further justification for the use of independent phonological analysis measures, but the use of the WSSA metric has not been applied to an extended variety of speech contexts including connected speech sampling.

Other research has supported the use of informal relational measures for phonological analysis (i.e., comparing the child's utterance to the adult form of the word) because the test-retest reliability was strong when communication samples were collected for young school-aged children through a structured interview (6). Researchers collected speech samples from 20 kindergartners who were at risk for speech and/or language delays. The samples were taken one week apart. Then, the researchers measured speech sound accuracy using informal relational measures.

The study findings indicated strong test-retest reliability for informal relational analyses calculated from structured communication samples for this age group.

Van Severen, Van Den Berg, Molemans, and Gillis (7) studied the effects of conversational speech sample size with 30 Dutch-speaking participants whose ages range from 6- to 24-months. They calculated phonetic inventories using a bootstrapping procedure (i.e., they repeatedly drew random conversational speech samples from a larger, lengthier sample). Researchers concluded that inconsistencies were found for this young age group and the reliability for further informal analyses based from the conversational speech sample depended heavily on the length of the overall sample obtained.

Given previous findings of inconsistencies related to the use of conversational speech samples of young children and reliability of subsequent informal measures, the aim of the present was established. This exploratory

го зголеми делото на Morris (3) преку одредување на тест-ретестирање на веродостојноста на независни фонолошки анализи кај малку постари деца, од 29 до 33 месеци, деца кои се 7 до 15 месеци постари од оние кои беа целна група во студијата на Morris (3). Со оглед на недостатокот на податоци во врска со веродостојноста на тест-ретестираниот фонолошки развој, во оваа област беше предложена пилот-работа од различни причини, вклучувајќи и обезбедување на оправдување за понатамошна истрага преку собирање на прелиминарни податоци, проценка на одржливоста на техниките за анализа на податоци и проценка на потенцијалните различни исходи за идните студии.

Со цел постојано донесување одлуки врз основа на докази во текот на процесот на проценка и третман, SLP треба да имаат информации базирани на докази во врска со веродостојноста на неформалните независни фонолошки анализи собрани од комуникациските примероци на мали деца во различни фази на развојот на јазикот. Специфичната возрастна група беше цел на оваа студија, бидејќи во оваа фаза за време на типичниот јазичен развој, поголемиот дел од децата доживуваат брз раст на речникот што резултира со експресивна „експлозија“ на вокабуларот и почетокот на искази составени од повеќе зборови. Освен тоа, многу мали деца коишто одат на терапии за говор исто така почнуваат со транзиција од искази со еден кон искази со повеќе зборови и веродостојните мерки за документирање на фонолошките добивки во текот на овој процес се од суштинско значење за донесување одлуки, следење на напредокот, како и планирање на третманот. Се осврнавме на следниве истражувачки прашања:

1. Кај двегодишни деца, каква е краткорочната (во рок од една недела) тест-ретестираната веродостојност на фонетскиот инвентар кога се пресметува со користење на 20-минутниот примерок од комуникација?
2. Кај двегодишни деца, каква е краткорочната (во рок од една недела) тест-ретестираната веродостојност на анализата на зборовната форма кога се пресметува со користење на 20-минутниот примерок од комуникација?

study was aimed at extending Morris' work by determining test-retest reliability of independent phonological analyses over time for a slightly older child population, 29- to 33-month olds, and children seven- to 15-months older than the sample used in the Morris study (3). In view of the scarcity of data regarding the test-retest reliability of phonological development, pilot work in this area was proposed for a variety of reasons including providing justification for further investigation through the collection of preliminary data, assessing the viability of data analysis techniques, and estimating potential outcome variability for future full-scale studies.

In order to consistently use evidence-based decision making throughout the assessment and treatment process, SLPs need to have evidence-based information regarding the reliability of informal independent phonological analyses collected from communication samples of young children across differing language development stages. The specific age range was targeted in the present study because at this stage during typical language development the majority of children experience rapid vocabulary growth resulting in an expressive vocabulary 'explosion' and the onset of multiword utterances. Moreover, many young children receiving speech-language services are also beginning to transition from one- to multiword utterances and reliable measures for documenting phonological gains during this process are essential for evaluation decisions, progress monitoring, and treatment planning. The following research questions were addressed:

1. For two-year-old old children, what is the short-term (within one week) test-retest reliability of phonetic inventory when calculated using a 20-minute communication sample?
2. For two-year-old old children, what is the short-term (within one week) test-retest reliability of word shape analysis when calculated using a 20-minute communication sample?



2. Материјали и методи

Сите интеракции со учесниците, регрутирањето, како и процедурите за проектот беа спроведени во согласност со етичките стандарди на Универзитетот во Небраска во Омаха. Оригиналното истражување беше одобрено од страна на ова управно тело пред почетокот на собирање на податоците (ИРБ # 035-14-ЕП).

2.1 Испитаници

Испитаниците беа три деца на возраст од 29 до 33 месеци. За децата беше утврдено дека немаат доцнење во развојот на говорот преку завршување и оценување на две скрининг мерки. Најпрво беше спроведена анализата *Јазична скала за предучилишна возраст – петто издание* (ПЛС-5) во текот на првата од двете експериментални сесии (8). ПЛС-5 е стандардизиран, нормативен инструмент за проценка што најчесто се користи од страна на СЛП во образовни средини за да се оценат рецептивните и експресивните јазични вештини кај децата од предучилишна возраст. Со цел да бидат вклучени во студијата, како деца без застој во јазичните вештини, учесниците требаше да постигнат стандарден резултат ($M = 100$, $SD = 15$) од 85 или погоре за вкупниот развој на јазикот. Тројцата учесници постигнаа резултати над просекот на експресивната комуникација (стандарден опсег на резултат: 119 - 126) и еден учесник постигна резултати над просекот на тестот за аудитивно разбирање (стандарден резултат: 120). Други резултати укажуваа на јазичната функција во рамките на типичните очекувања според возраста. Покрај тоа, *Инвентарот за комуникативен развој на MacArthur Bates – Зборови и реченици* (ИКД) се користи како мерка за скрининг (9). ИКД претставува контролен формулар за родители од 680 зборови, кој е стандардизиран и нормативен инструмент за мерење на експресивниот вокабулар на мало дете. Повторно, ова е мерка којашто рутински се користи од страна на логопедот при оценување на речникот и јазичните вештини на малите деца. Резултатите од или над 25 проценти го квалификуваа детето да учествува во оваа студија. Сите учесници беа монолингвални чиј мајчин е англискиот јазик, а чиишто родители не искажаа загриженост за развојот на говорот или на јазикот, ниту пак за сетилата за слух или вид. Види табела 1 и 2 за описни информации за учесниците.

2. Materials and Methods

All participant interactions, recruitment, and project procedures were conducted in accordance with the ethical standards of the University of Nebraska at Omaha and University of Nebraska Medical Center Institutional Review Board. The original research was approved by this governing body prior to the beginning of data collection (IRB #035-14-EP).

2.1 Participants

Participants included three children between the ages of 29- to 33-months of age. Children were identified as having no known delay in language development through completion and scoring of two screening measures. First, the *Preschool Language Scale - Fifth Edition* (PLS-5) was administered during the first of two experimental sessions (8). The PLS - 5 is a standardized, norm-referenced assessment instrument commonly used by SLPs in educational settings to evaluate the receptive and expressive language skills of preschool-aged children. In order to be included in the study as a child without a language delay, the participants needed to receive a standard score ($M = 100$, $SD = 15$) of 85 or above for total language development. The three participants scored above average on the expressive communication subtest (standard score range: 119-126) and one participant scored above average on the auditory comprehension subtest (standard score: 120). Other scores indicated language function within typical age-level expectations. In addition, the *MacArthur Bates Communicative Development Inventory - Words and Sentences* (CDI) was used as a screening measure (9). The CDI is a 680-word parent checklist that is standardized and norm-referenced to measure a young child's expressive vocabulary. Again, this is a measure routinely used by practicing SLPs when evaluating the vocabulary skills of young children. Scores at or above the 25th percentile qualified a child to participate in the present study. All participants were monolingual native English speakers whose parents reported neither concerns for speech or language development nor hearing or vision abilities. See Tables 1 and 2 for descriptive participant information.



Табела 1. Описни информации за учесниците: стандардизирани мерки /
Table 1. Participant Descriptive Data: Standardized Measures

Опис / Descriptor	Учесници / Participants		
	Учесник/Participant 1	Учесник/Participant 2	Учесник/Participant 3
Возраст/Age	33 месеци / months	30 месеци / months	29 месеци / months
Пол/Gender	машки / male	женски / female	женски / female
PLS-5 експ. ^a / PLS-5 Exp. ^a			
Станд. резултат/ Standard Score	126	119	119
Процент/ Percentile	95	90	90
PLS-5 ауд. ^b / PLS-5 Aud. ^b			
Станд. резултат / Standard Score	120	112	103
Перцентил/ Percentile	91	79	58
CDI/CDI III ^c / CDI/CDI III ^c			
Перцентил/ Percentile	75 (CDI)/65(CDI III)	90 (CDI)/90 (CDI III)	25

^a Предучилишна јазична скала – 5. издание, поттест за експресивна комуникација

^b Предучилишна јазична скала – 5. издание, поттест за аудиоразбирање

^c MacArthur Bates Развој на инвентар за комуникација – зборови и реченици (CDI)/MacArthur Bates Развој на инвентар за комуникација – продолжение (CDI III) спроведено кај деца на возраст од над 30 месеци. Два учесника беа на возраст од 30 месеци или постари (P1 и P2), поради тоа и учесниците покажаа процентен опсег и за CDI и за CDI III.

^a Preschool Language Scale – 5th Edition, Expressive Communication subtest

^b Preschool Language Scale – 5th Edition, Auditory Comprehension subtest

^c MacArthur Bates Communicative Development Inventory - Words and Sentences (CDI)/MacArthur Bates Communicative Development Inventory Extension (CDI III) administered to children over 30 months of age. Two of the participants were 30 months of age or older (P1 and P2), therefore; these participants have a reported percentile rank for both the CDI and CDI III.

Табела 2. Описни информации за учесниците: мерки базирани на примерок /
Table 2. Participant Descriptive Data: Sample-based Measures

Опис / Descriptor	Учесници / Participants		
	Учесник/Participant 1	Учесник/Participant 2	Учесник/Participant 3
ПДИ ^a / MLU ^a	3.32	2.50	1.98
Сесија 1 / Session 1	3.70	2.42	1.66
Сесија 2 / Session 2			
Вкупно зборови/ Total Words			
Сесија 1 / Session 1	402	284	230
Сесија 2 / Session 2	319	573	197
PI ^b			
Почетни согласки /Initial Consonants			
Сесија 1 / Session 1	15	1	4
Сесија 2 / Session 2	14	18	10
Последни согласки / Final Consonants			
Сесија 1 / Session 1	14	12	7
Сесија 2 / Session 2	13	10	7
Групи согласки / Consonant Clusters			
Сесија 1 / Session 1	16	5	4
Сесија 2 / Session 2	12	13	4
Форма на збор ^c / Word Shape ^c			
Сесија 1 / Session 1	8	8	8
Сесија 2 / Session 2	8	8	8

^a Просечна должина на изговор (ПДИ)

^b Фонетски инвентар (ФИ) за продуктивни согласки употребени на определена позиција во најмалку два различни збора во текот на примерокот

^c Форма на збор, Анализа на осум различни целни форми на зборови: V, CV, CVCV, VC, CVC, CCVC, CVCC и CVCVC

^a Mean Length of Utterance (MLU)

^b Phonetic Inventory (PI) for productive consonants, those produced in a particular position across at least two different words during the sample

^c Word Shape Analysis of eight different target word shapes: V, CV, CVCV, VC, CVC, CCVC, CVCC, and CVCVC



2.2 Поставка и процедури

Процесот на собирање податоци беше спроведен и евидентиран во патолошка клиника за говорен јазик во склоп на универзитетскиот кампус. Тука дипломирани студенти по предметот логопедија вообичаено даваа своја проценка и терапевтски услуги на населението во рамките на заедницата. Во клиниката, стандардни мали индивидуални соби за терапија вклучуваа маса на склопување, три до четири столчиња, како и мал кабинет. Освен тоа, секоја соба за терапија беше опремена со систем за далечинско снимање, систем за снимање на интерактивна сесија (ISR), сетран да снима аудио и визуелна документација на клинички интеракции преку видеокамери инсталирани во таванот.

Откако првиот автор ги администрира мерките за скрининг, 20-минутни примероци од разговори беа собрани од секое дете додека тој/таа беше во интеракција со неговите/нејзините родители, со играчки на располагање за игра. Две различни групи на играчки (на пример: фарма, храна, кујна, автомобили со гаража) беа случајно доделени на секој учесник за секоја од двете сесии спроведени со точно една недела растојание. За секој учесник, истиот избор на играчки беше достапен секоја недела. На детето и родителот / негувателот им беше кажано да ги користат играчките, истовремено или еден по еден, врз основа на интересите на детето. Сесиите беа снимени со помош на системот на ISR инсталиран во сите клинички соби. Овие снимки подоцна беа разгледани од страна на истражувачки кадар за транскрипција и анализа. Говорните продукции на секој учесник беа транскрибирани со помош на Меѓународната фонетска азбука (МФА) и анализирани со помош на две одделни анализи на неформален говор, фонетски инвентар и зборовна форма. Процедурите за веродостојност беа усогласени со оние што се користат од страна на Morris со тоа што сите вокализации од секоја сесија беа разгледани и транскрибирани. Транскрипциите беа завршени од страна на првиот автор и втор препишувач. Вторите препишувачи беа студенти од отсекој логопедија, обучени во транскрипција на ИПА. За разлика од Morris (3), истражувачите се обидуваа да постигнат усогласеност фонема по фонема и кај двајцата препишувачи. Иако истражувачите изразиле загриженост во врска со помалата веродостојност на тран-

2.2 Setting and Procedures

All data collection was conducted and recorded at a university campus speech-language pathology clinic. At this site, graduate students in speech-language pathology typically provided assessment and therapeutic services to a community-based clinical population. In the clinic, standard small individual therapy rooms included an adjustable table, three to four chairs, and a small cabinet. In addition, each therapy room was fitted with a remote recording system, the Interactive Session Recorder (ISR) system, set up to record audio and visual documentation of clinical interactions through video cameras secured in the ceiling.

After the first author administered the screening measures, two 20-minute conversational speech samples were collected from each child while he/she was interacting with his/her parent with toys available for play. Two different sets of toys (e.g., farm, grocery, kitchen, cars with garage) were randomly assigned to each participant for each of the two sessions conducted exactly one week apart. For each participant, the same two sets of randomly assigned toys were available each week. The child and caregiver were instructed to utilize the toys either simultaneously or one at a time based on their child's interests. The sessions were recorded using the ISR system present in all clinic rooms which includes adjustable wall/ceiling mounted video and audio recorders. These recordings were later reviewed by research personnel for transcription and analysis. Each participant's speech productions were transcribed using the International Phonetic Alphabet (IPA) and analyzed using two separate informal speech sound analyses; phonetic inventory and word shape. Reliability procedures were aligned with those used by Morris (3) in that all vocalizations from each session were reviewed and transcribed. Transcriptions were completed by the first author and a second transcriber. Second transcribers were students majoring in speech-language pathology trained in IPA transcription. In contrast to Morris, the researchers did attempt to achieve phoneme-by-phoneme agreement between the two



скрипцијата со користење на овој метод, мотивацијата за постигнување на прецизна и точна мерка на производството на секоја фонема беше огромна.

Кога доаѓаше до несогласување помеѓу транскрипцијата на првиот автор и вториот преписувач, а поврзани со одредена фонема, двајцата заедно повторно го прегледуваа записот од неусогласениот наод и врз исказот на детето постигнуваа договор. Доколку не можеше да се постигне договор по три обиди на повторен преглед, исказот не се користеше во крајната анализа. Првичната веродостојност беше само над 62% (38% несогласување); Сепак, по разгледувањето на исказите кај коишто постоеше неусогласеност, преписувачите успеаја да решат 100% од несогласувањата, така што заедно дискутираа за видеозаписите. Ниската веродостојност на транскрипцијата претставува документирана грижа во областа на патологијата на говорниот јазик, особено кога фонетски се транскрибира поврзаниот говор на млади деца чишто ограничени резултати од фонолошкиот развој во говорните примероци се сосема различни од примероците на зрел и добро формиран говор кај возрасните (10). Дури и со попрецизните правила на договор, првичната веродостојност на транскрипцијата во оваа студија беше 11% повисока од онаа на Morris кој забележа 27% почетно несогласување (3). Главната причина за зголемувањето на неусогласеноста е: (1) во оваа студија истражувачите се обидуваат да постигнат усогласеност фонема по фонема и (2) инструментите за снимањето на податоците. Morris користел комбинација на самостојни и далечински микрофони, додека во оваа студија се користат уреди за снимање монтирани на ѕидовите и таванот (3). Иако се користат во тивка, мала соба со минимална бучава во заднината, оддалеченоста на уредот за снимање од звучниот извор може да резултира со понизок квалитет на звукот за подоцнежната транскрипција, што би можело да резултира со зголемување на варијабилноста на транскрипцијата.

Двете неформални анализи на говорните звуци, фонетскиот инвентар и анализата на зборовната форма, беа спроведени за секој 20-минутен примерок на говор на секој учесник. Резултатите биле споредени меѓу првиот и вториот примерок на секое дете преку визуелна анализа.

transcribers. Although, the researchers noted concern about lower transcription reliability using this method, the motivation for achieving a precise and accurate measure of each phonemic production was paramount.

When instances of disagreement occurred between the first author and a second transcriber related to a particular phoneme, together the two reviewed the child's production of the disagreed upon utterance and reached an agreement. If an agreement could not be reached after three review attempts, the utterance was not used in the final analysis. Initial inter-rater reliability was just above 62% (38% disagreement); however, after reviewing utterances with disagreement, transcribers resolved 100% of disagreements by discussing the videos together. Low inter-rater transcription reliability is a documented concern in the field of speech-language pathology, particularly when phonetically transcribing the connected speech of young children whose limited phonological development results in speech samples that are quite different from an adult's sample of mature and well-formed speech production (10). Even with the more precise agreement requirements, the initial transcription reliability of the present study was 11% higher than that of Morris who noted 27% initial disagreement (3). The primary reason for the increase in disagreement is likely two-fold: (1) the present study researchers' attempt to achieve phoneme-by-phoneme agreement and (2) data recording instrumentation. Morris used a combination of freestanding and remote microphones, while the present study used wall/ceiling mounted recording devices (3). Although used in a quiet, small room with little background noise interference, the distance of the recording device from the sound-source may have resulted in lower-quality sound capture for later transcription, which could have resulted in increased transcription variability.

The two informal speech sound analyses, phonetic inventory and word shape analysis, were conducted for each 20-minute speech sample for each participant. Results were compared between each child's first and second sample through visual analysis.



3. Резултати

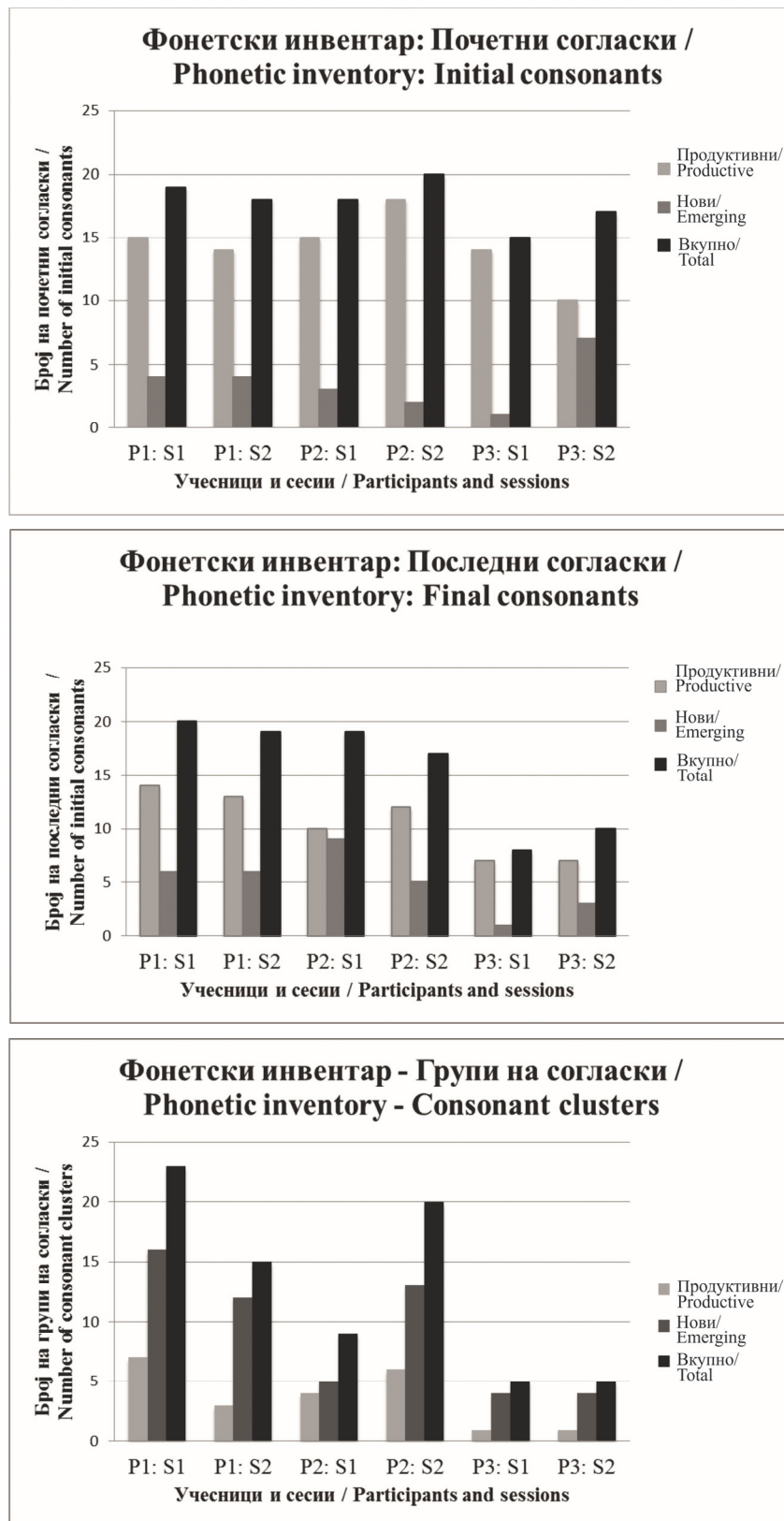
3.1 Фонетски инвентар

При пресметување на фонетските инвентари на учесниците, согласките се сметаат за „продуктивни“ за одреден збор позиција ако детето покаже употреба на согласка во два различни збора за време на примерокот. Согласките се сметаат за „нови“ ако детето ги произведува само во еден збор или позиција во текот на примерок. Визуелната анализа на резултатите за продуктивни, нови и вкупен број на согласки (види слика 1) укажаа на тест-ретестирани недоследности за пресметаните фонетски инвентари. Како недоследност се смета ако постои разлика од три или повеќе согласки во еден целен збор. Оваа точка на исклучување беше избрана бидејќи претставува разлика од нешто повеќе од една стандардна девијација и за почетната (2.75 за сесија 1 и 2.91 сесија 2) и за крајните согласки (1.48 за сесија 1 и 2.05 за сесија 2) во наодите од студијата од Morris (3). Поточно, како што може да се види во табела 2 и слика 1, вториот и третиот учесник (П2, П3 соодветно) произведоа неконзистентен број на продуктивни почетни согласки помеѓу двете сесии (15 во првата сесија, 18 во втората сесија и 4; 10, соодветно), додека првиот учесник (П1) произведе конзистентен број на продуктивни почетни согласки (15; 14). Визуелната анализа на произведените финални согласки кај сите тројца учесници остана релативно конзистентна меѓу двете сесии (14; 13, 12; 10, 7 и 7). Сепак, недоследности беа забележани во бројот на продуктивните согласни групи произведени меѓу две сесии кај двајца од тројцата учесници. П1 и П2 користеле неконзистентен број на продуктивни согласни групи (16, 12, 5; 13), додека П3 покажа употреба на конзистентен број на согласни групи (4 : 4) и во двете сесии.

3. Results

3.1 Phonetic Inventory

When calculating participant phonetic inventories, consonants were considered ‘productive’ for a particular word position if the child demonstrated use of the consonant in two different words during the sample. Consonants were considered ‘emerging’ if the child produced them in only one word or position during the sample. The visual analysis of the results for productive, emerging, and total consonants used (see Figure 1) indicated test-retest inconsistencies for the calculated phonetic inventories. An inconsistency was indicated if a difference of three or more consonant productions in a target word position were present. This cutoff point was selected because it represented a difference of just over one standard deviation for both initial (2.75 for session 1 and 2.91 session 2) and final consonants (1.48 for session 1 and 2.05 for session 2) in the study findings from Morris (3). Specifically, as evident in Table 2 and Figure 1, the second and third participants (P2, P3 respectively) produced an inconsistent number of productive initial consonants between the two sessions (15 in the first session; 18 in the second sessions and 4; 10, respectively), while the first participant (P1) produced a consistent number of productive initial consonants (15; 14). Visual analysis of productive final consonant productions by all three participants remained relatively consistent between the two sessions (14; 13, 12; 10, 7; 7, respectively). However, inconsistencies were noted in the number of productive consonant clusters produced between the two sessions in two of the three participants. P1 and P2 used an inconsistent number of productive consonant clusters (16; 12, 5; 13), while P3 demonstrated the use of a consistent number of consonant clusters (4; 4) across the two sessions.



Слика 1. Фонетски инвентар на почетни согласки, последни согласки и групи согласки по сесија.

Figure 1. Phonetic inventories of initial consonants, final consonants, and consonant clusters by session.



3.2 Анализа на зборовната форма

При пресметување на присуството на формата на целниот збор во примероците од комуникацијата, на учесниците им се признаваше зборовната форма доколку редоследот на звуците беа произведени во најмалку два различни збора во целиот примерок. Истражувачите посебно анализираа осум различни целни зборовни форми, вклучувајќи V, CV, CVCV, VC, CVC, CCVC, CVCC и CVCVC. Како што е прикажано во табела 2, нема суштински разлики помеѓу две сесии за секој учесник. Сите учесници произведоа најмалку два различни збора во секоја од осумте категории на целни зборови. Како резултат на тоа, наодите од анализата на зборовната форма беа конзистентни кај сите учесници низ двете сесии.

Дискусија

Оваа студија претставува ран обид да се започне да се утврди веродостојноста на тест-ретестираните неформални независни фонолошки анализи, фонетски инвентар и анализа на зборовна форма кај двегодишни деца без застој во говорниот јазик. Од клинички аспект, податоците во врска со овие мерки се доста вредни за олеснување на донесувањето одлуки врз основа на докази за дијагноза и интервенција на говорниот јазик. Во оваа студија, резултатите од првото истражување поставуваат прашање дали во врска со веродостојноста на тестирањето на фонетскиот инвентар кај 29 до 33-месечни деца пресметани од 20-минутен комуникациски примерок не се покажаа мешани наоди. Резултатите покажуваат делумна поддршка за претходните заклучоци во врска со неконзистентните мерки на фонетскиот инвентар (3,7). Двајца од тројцата учесници добија неконзистентни фонетски инвентари за почетните звучни и согласни парови во зборообразувањето (на пример, изговарање на две соседни согласки како "sn" во "snake (змија)" или "pl" во "plate (плоча)"), додека два од трите наведени конзистентни профили за продукција на финалниот звук во зборот во текот на две сесии за собирање податоци. Овие наоди покажуваат дека двајца од тројцата учесници произведуваат приближно ист број на продуктивни финални согласки во сесија 1 како што направија во сесија 2, но не го дадоа приближно истиот број на продуктивни почетни согласки или согласни групи меѓу сесиите. Morris покажа дека за деца на возраст од 18 до 24 месеци, продукцијата на финални согласки е постабилна од почетната продукција и таа не

3.2 Word Shape Analysis

When calculating the presence of target word shapes in the communication samples, participants were credited for a word shape if the sequence of sounds were produced in at least two different words across the sample. The researchers specifically analyzed eight different target word shapes including V, CV, CVCV, VC, CVC, CCVC, CVCC, and CVCVC. As shown in Table 2, no substantive differences were found across the two sessions for any participant. All participants produced at least two different words in each of the eight target word shape categories. Consequently, word shape analysis findings were consistent for all participants across the two sessions.

Discussion

The present study represents an early attempt to begin to determine the test-retest reliability of two different informal independent phonological analyses, phonetic inventory and word shape analysis for 2-year-old children with no known language delay. Clinically, data concerning these measures are quite valuable for facilitating evidence-based decision making for speech-language diagnosis and intervention. In the present study, results for the first research question regarding the test-retest reliability of phonetic inventory of 29- to 33-month old children calculated from a 20-minute communication sample indicated mixed findings. The results indicated partial support for previous conclusions regarding inconsistent measures of phonetic inventory (3, 7). Two of the three participants obtained inconsistent phonetic inventories for word-initial sound productions and consonant cluster productions (i.e., production of two adjacent consonant sounds such as "sn" in "snake" or "pl" in "plate"), while two of the three indicated consistent profiles for word-final sound productions over the two data collection sessions. These findings indicated that two of the three participants produced approximately the same number of productive final consonants in Session 1 as they did in Session 2, but did not produce approximately the same number of productive initial consonants or consonant clusters across sessions. Morris found that for 18- to 24-month old children, final sound productions were more stable than initial



ги испита инвентарите на продукцијата на согласни групи (3). Какви било разлики во резултатите забележани меѓу двете студии може да бидат поврзани со фактори како што е малиот примерок на ова истражување и разликите во возраста и општата разговорливост на учесниците за секоја сесија. На пример, „деца кои произведуваат поголем број вкупно зборови, произведуваат поголем број различни зборови, кои во оваа студија може да вклучат и две фонетски варијации на истиот збор“ (3). Во суштина, општата разговорливост на даден учесник би можела да влијае врз стабилноста на примерокот на комуникацијата. Продолжувањето на оваа истражувачка студија, вклучувајќи и поголем број учесници, ќе обезбеди повеќе конечни заклучоци во оваа област.

Второто истражувачко прашање на кое се осврнавме во оваа студија беше во врска со веродостојноста на тестираната анализа на зборовната форма пресметана со користење на 20-минутниот примерок од комуникацијата. Во овој случај, резултатите се во согласност со Morris (3) во тоа што сите учесници покажаа конзистентни зборовни форми во текот на двете сесии. Наодите од оваа студија покажале дека мерката што се користи за анализа на зборовната форма беше повеќе временски веродостојна од фонетскиот инвентар во период од една недела со три деца на возраст од 29 до 33 месеца.

Употребата на 20-минутните примероци од комуникацијата, што беше во согласност со процедурите на Morris, може да биде причина за овие разлики во конзистентноста на мерењата (3). Како што е наведено, во професионалната заедница на патолози за говорен јазик не е воспоставено стандардно време за мерење на комуникацискиот примерок и не се применува конзистентно во анализите (7). Сегашните наоди ја поддржуваат понатамошната истрага за оптимално време за примерок на комуникација употребено за анализа, бидејќи наодите не се конзистентни низ проучените анализи на мерења. На пример, 20-минутен примерок базиран на игра дава доволно време за секој учесник да покаже повеќе начини на употреба на голем број на различни форми на зборови, но не може да им овозможи доволно време за сите да ја прикажат употребата на сите звуци при говорот во нивниот експресивен репертоар. Потребни се повеќе истражувања во овие важни променливи, особено кога тие се однесуваат на анализа на говорот кај мали деца.

productions and she did not examine inventories of consonant cluster productions (3). Any differences in findings noted between the two studies may be related to factors such as the small sample size of the current study and the differences in ages and general talkativeness of child participants for each session. For instance, “children who produced a greater number of total words produced a greater number of different words, which in this study could include up to two phonetic variations of the same word (3).” Essentially, the general talkativeness of a participant could influence the robustness of the communication sample. An extension of the present exploratory study including a greater number of participants would provide more definitive findings in this area.

The second research question addressed in the present study regarding the test-retest reliability of a word shape analysis calculated using the 20-minute communication sample. In this instance, findings were consistent with Morris (3) in that all participants demonstrated consistent word shapes over the two sessions.

The present study findings indicated that measure used for word shape analysis was more temporally reliable than phonetic inventory over a one-week period with three 29- to 33-month old children.

The use of a 20-minute communication sample, which was consistent with Morris’ procedures, may be at cause for these differences in measurement consistencies (3). As noted, a standard communication sample measurement time has not been established and applied consistently in the speech-language pathology professional community (7). The present findings support further investigation into the optimal communication sample time used for analysis since the findings were inconsistent across studied measurement analyses. For instance, the 20-minute play-based sample time allotted adequate time for each participant to demonstrate multiple uses of a number of different word shapes, but may not have allowed adequate time for each to demonstrate the use of all of the speech sounds in their expressive repertoire. More research into these important variables is needed, particularly as they relate to the analysis of speech for young children.



4.1 Образовно и клиничко значење

Бидејќи не постои метод за мерење, формален или неформален, што ќе има целосна согласност со администрациите (на пр. CDI има корелација низ едномесечните тест-ретест администрации која се движи помеѓу 0,61-0,95) (9), наодот дека неколку случаи на разликување во времето за земање примероци директно не обезбедува доказ дека оваа мерка е несоодветна за клиничка употреба. Напротив, кога се користат неформални фонолошки мерки во клинички услови за евалуација, развојот на целта и / или следење на напредокот со текот на времето, лекарите треба да бидат внимателни при толкувањето. Повеќе извори на податоци треба да се земат предвид за време на евалуацијата и при следење на терапевтскиот напредок, така што одлуките и извештаите се базираат на повеќе од една алатка за мерење на конвергенцијата. Покрај тоа, истражувачите повикуваат на претпазливост кај академските и клиничките едукатори во областа на патологија на говорот кога ги советуваат дипломираните студенти да користат неформални независни фонолошки мерки пресметани од примерок на комуникација на некое мало дете. Бидејќи временската рамка веродостојна за фонетски пописи сè уште не е утврдена за деца под три години, треба да бидеме внимателни кога се залагаме за нивната примена кај оваа млада популација. Покрај тоа, сите SLP коишто работат со млада клиентела треба да вложат слична воздржаност при користење на фонетски инвентари за основни или описни информации. Додека овие мерки обезбедуваат описни информации корисни за документирање на терапевтскиот напредок, недоследните индикатори на веродостојноста на тест-ретестирањето укажуваат на тоа дека може да биде потребна употребата на алтернативни методи за проценка или да се преземат повеќе основни мерки за да се обезбеди порепрезентативен примерок за говорните способности на детето. Не сите неформални независни мерки се создадени подеднакво и веродостојноста на тест-ретестирањето се чини дека се разликува кај различни анализи, со тоа што кај анализите на зборовната форма се чини дека е постабилна отколку кај анализите на фонетски инвентари. Конечно, SLP треба да бидат свесни за факторите кои можат да влијаат врз веродостојноста на овие неформални мерки, вклучително и возраста на детето, должината на примерокот, фонетските материјали, големината на корпусот, како и постапката за собирање на податоци и транскрипцијата.

4.1 Educational and clinical significance

Since no measurement method – formal or informal - will have perfect agreement across administrations (e.g., the CDI has a correlation across one-month test-retest administrations that ranges from 0.61-0.95) (9), the finding that a few instances of difference across sampling times does not directly provide evidence that the measure is inappropriate for clinical use. Rather, when informal phonological measures are used in clinical settings for evaluations, goal development, and/or progress monitoring over time, clinicians should use caution in interpretation. Multiple sources of data should be considered during an evaluation and when monitoring therapeutic progress so that decisions and reports are based on more than one converging measurement tool. Additionally, the researchers urge caution among speech-pathology academic and clinical educators when advising graduate students to use informal independent phonological measures calculated from a young child's communication sample. Because the temporal reliable for phonetic inventories has not been established for children under three-years of age, we should use caution when advocating for their use with this young population. In addition, practicing SLPs working with a young clientele should exert similar restraint when using phonetic inventories for baseline or descriptive information. While these measures provide descriptive information helpful for baseline and therapeutic progress documenting, inconclusive test-retest reliability indicators imply that use of other alternative assessment tools may be necessary or multiple baseline measures undertaken to provide a more representative sample of a child's speech production abilities. Not all informal independent measures are created equally and test-retest reliability does seem to vary across analyses with word shape analyses appearing to be more stable over time than phonetic inventories. Finally, SLPs need to be aware and mindful of the factors that may influence the reliability of these informal measures including age of child, length of sample collected, phonetic materials, corpus size, as well as data collection and transcription procedures.

4.2 Ограничувања и идни насоки

Голем број фактори ја ограничуваат можноста за генерализација на резултатите од сегашната студија вклучувајќи ги техниките на мали димензии на примерокот и анализата на податоци. Како резултат на тоа, оваа студија сè уште не нуди јасни клинички упатства за примена на неформални фонолошки мерки со мали деца, ниту пак доволно ја проширува базата на докази за клинички одлуки во областа на патологијата на говорниот јазик. Наместо тоа, оваа студија обезбедува оправдување за понатамошна истрага во таа област и зголемување на свеста за техниките на анализа на податоци кои треба да се земат предвид во идните студии.

Иако овие прелиминарни наоди покажуваат делумна поддршка за резултатите на Morris, проширувањето кон малку постари деца со понапредни јазични вештини, малите димензии на примерокот значително ја ограничува генерализацијата на наодите за општата популација. На пример, поради малата димензија на примерокот, авторите може само визуелно да ги анализираат податоците, наместо да ги тестираат наодите со употреба на статистички методи. Со цел да се направи анализа на нешто повеќе од само споредба на визуелни прикази, и со цел попрецизно да се поддржат наодите на Morris, потребна е поголема димензија на примерокот. Идните студии слични на оваа, кои вклучуваат дополнителни опсези на возраст и над и под оние кои се вклучени во оваа студија, ќе обезбедат информации за неформалната фонолошка анализа за широк спектар на млади деца во различни развојни фази на говорниот јазик. Вклучувањето на деца со застој во говорот би ги олеснило генерализираните наодите за клиничкото население, кои може да функционираат поинаку од децата што се застапени во оваа студија (3).

Техниките на анализа на податоци користени во оваа студија вклучуваа и мерка за употреба на дадена зборовна форма. Резултатите на сите тројца учесници беа на горната граница на оваа мерка. Резултатите што допираат до горната граница предизвикуваат практични проблеми при пресметувањето на веродостојноста, бидејќи тие спречуваат да се измерат променливите или се проценети над одредено ниво, или како во овој случај, надвор од дистрибуцијата на осум различни целни зборовни форми. Конкретната анализа на зборовната форма што се користи во оваа студија беше избрана во обид да се придржува кон мерките што ги користи Morris. Сепак, со оглед на возраста и развојот на говорот кај учесни-

4.2 Limitations and future directions

A number of factors limit the potential for generalization of the present study results including the small sample size and data analysis techniques. Consequently, the present study does not yet provide clear clinical guidelines for use of informal phonological measures with young children nor does it sufficiently expand the evidence base for clinical decision making in speech-language pathology. Rather, the present study provides justification for further investigation in the area and an increased awareness of data analysis technique to be considered in future studies.

Although these preliminary findings indicated partial support for Morris' outcomes extending to slightly older children with more advanced language skills, the small sample size severely restricted the generalization of the findings to the general population. For instance, due to the small sample size, the authors could only visually analyze the data rather than test the findings using statistical methods. In order to conduct analyses beyond visual display comparisons and more accurately account for support of Morris' findings, a larger sample size is needed. Future studies along this vein that include additional age ranges - both above and below those included in the present study - would provide information on informal phonological analyses for a wide range of young children at differing developmental language stages. The inclusion of children with language delays would facilitate generalizable findings to clinical populations who may perform differently than the children represented in the present study. (3).

The data analysis techniques utilized in the present study included a measure of word shape use. All three participants performed at ceiling on this measure. Ceiling-level performances cause practical problems when calculating reliability because they prevent a variable from being measured or estimated above a certain level, or as in this case, beyond the distribution of eight different target word shapes. The particular word shape analysis utilized in the present study was selected in an effort to adhere closely with measured used by Morris. However, given the age and language development status of the Morris participants compared to those of the present participants, a



ците на Morris во споредба со сегашните учесници, може да се користи покомплексна техника за мерење на перформансите. Идните студии кои вклучуваат деца над 24-месечна возраст може да вклучуваат посуптилни мерки за анализа на зборовната форма со цел подобро да се долови различното изразување. Дополнителна забелешка на сегашните наоди вклучени во постапките кои се користат за собирање и транскрибирање на примероците од комуникацијата е, дека, и покрај тоа што беа реплицирани повеќето процедури кои се користат од страна на Morris во оваа студија, постои широк спектар на процедурални разлики кои се користат во областа за каков било број на примероци од комуникацијата и неформалните фонолошки мерки. Во оваа студија, почетната веродостојност беше 62%, прилично ниска кореспонденција со оглед на тоа што 80% веродостојност е често цел во кодирањето на веродостојноста. Ова се покажа како проблем во областа на логопедијата бидејќи фонетската транскрипција на говорните примероци често се користи за да се утврдат соодветните цели за интервенција, како за следење на напредокот со текот на времето (10). Друг придонес за варијабилноста при собирање и транскрипција на примерокот од комуникацијата е изборот на различни сетови на играчки, укажувајќи на потенцијалот за различни фонетски материјали кои би можеле да влијаат врз бројот на зборовите кои децата ги произведуваат и врз нивните фонетски инвентари. Многу SLP различно го користат фонетскиот материјал кој се користи во сесии, во зависност од одговорите на детето на понудениот материјал. Според тоа, иако изборот на играчки може да биде извор на варијабилност во оваа студија, тој служи како еколошки валидна застапеност на промената која постои во клиничката пракса. Конечно, врз фонетскиот инвентар може да влијае големината на корпусот, па така одлуката на истражувачите да се користат целиот транскрибиран примерок, а не само првите 100 збора, исто така би можело да придонесе за варијации, и покрај тоа што оваа процедура била применета подеднакво кај сите учесници. Идните студии може да се осврнат на некои од овие потенцијални извори на варијабилност и да одговорат на потребата за повеќе истражувања за должината на примерокот од говорот и процедурите за фонетска транскрипција кои се користат за клиничкото и нормативното население, со цел да се прочистат процедурите за донесување оценка врз основа на докази во областа на логопедијата.

more complex technique could have been utilized to measure performance. Future studies that include children over 24-months of age could include more nuanced measures of word shape analysis to better capture performance diversity.

An additional caveat to the present findings involved the procedures used to collect and transcribe the communication sample.

Although in the present study, most procedures used by Morris³ were replicated, there exists a wide variety of procedural differences utilized in the field for any number of communication sample collections and informal phonological measures. For instance, low inter-judge reliability when transcribing speech samples of young children occurs because the speech to be transcribed is usually quite different from adult-like forms.¹⁰ In the present study, initial interrater reliability was 62%, a fairly low correspondence considering 80% reliability is often the goal in reliability coding. This has proven to be a concern in the field of speech-language pathology as phonetic transcriptions of speech samples are often used to determine appropriate goals for intervention as well as monitor progress over time (10). Other contributions to variability in collection and transcribing the communication sample include choice of differing toy sets, indicating the potential for different phonetic materials which could affect the number of words children produce and the phonetic inventories elicited. Many practicing SLPs vary the phonetic material used across elicitation sessions depending on child responses to provided material. Therefore, although toy selection may be a source of variability in the present study, it serves as an ecologically-valid representation of the variability that exists in clinical practice. Finally, phonetic inventory may be affected by corpus size, so the researchers' decision to use the entire transcribed sample rather than only the first 100 words may also have contributed to variance even though this procedure was applied consistently across participants. Future studies may address some of these potential sources of variability and address the need for more research into the length of speech sample collected and phonetic transcription procedures used for clinical and normative populations in order to refine evidenced-based assessment procedures for the speech-language pathology profession.



Како заклучок, за мал примерок од двегодишни деца без познат застој во говорот, веродостојноста на тест-ретестирањето на изведувањето на почетната согласка и групите согласки беше нестабилна меѓу два примерока од разговор, добиени со разлика од една недела под речиси идентични услови. Сепак, фонетските инвентари на финалните согласки и анализите на зборовната форма се чинат релативно стабилни во текот на период од една недела. Иако се потребни повеќе податоци, прелиминарните резултати од оваа пилот-студија покажаа дека академскиот стручен кадар, клиничките едукатори и логопеди треба да имаат на ум дека еден наод со неколку случаи на разлики пронајдени кај неформални примероци не значи дека фонолошката мерка сама по себе е несоодветна, туку укажува на тоа дека резултатите кај мали деца треба да се толкуваат внимателно.

Конфликт на интереси

Авторите изјавуваат дека не постои конфликт на интереси.

Референци / References

1. Paul R. Introduction to clinical methods in communication disorders. Third edition. Baltimore, Maryland USA: Brookes; 2014.
2. Bernthal J, Bankson N, Flipsen, P. Articulation and phonological disorders: Speech sound disorders in children. Seventh edition. Boston, Massachusetts USA: Pearson; 2013.
3. Morris, SR. Test-retest reliability of independent measures of phonology in the assessment of toddlers' speech. *Lang Speech Hear Serv Sch* 2009; January;40(1):46–52.
4. Preston JL, Ramsdell, HL, Oller, KD, Edwards, ML, Tobin, SJ. Developing a weighted measure of speech sound accuracy. *J Speech Lang Hear Res*. 2011 February;54(1):1–18.
5. Shriberg LD, Austin D, Lewis BA, McSweeney JL, Wilson DL. The Percentage of Consonants Correct (PCC) Metric Extensions and Reliability Data. *J Speech Lang Hear Res*. 1997 August;40(4):708–722.
6. Heilmann J, DeBrock L, Riley-Tillman TC. Stability of measures from children's interviews: The effects of time, sample length, and topic. *Am J Speech Lang Pathol* 2013; August;22(3):463–475.
7. Van Severen L, Van Den Berg R, Molemans I, Gillis S. Consonant inventories in the spontaneous speech of young children: A bootstrapping procedure. *Clinical Linguistic & Phonetics* 2012; February;26(2):164–187.
8. Zimmerman IL, Steiner VG, Pond RE. *Preschool language scales*. Fifth edition. Boston, Massachusetts USA: Pearson; 2011.
9. Fenson L, Dale PS, Reznick JS, Thal D, Bates E, Hartung JP, Reilly JS. *MacArthur communicative development inventories*. San Diego, California USA: Singular; 1993.
10. Oller DK, Ramsdell HL. A weighted reliability measure for phonetic transcription. *J Speech Lang Hear Res* 2006; December;49(6):1391–1411.

In conclusion, for a small sample of two-year-old children without known language delays, the test-retest reliability of initial consonant and consonant cluster productions was unstable between two conversational samples obtained one week apart under near-identical conditions. However, phonetic inventories for final consonants and word shape analyses appeared relatively stable over the one-week time period. Although more data is needed, the preliminary results of this pilot study indicated that academic faculty, clinical educators, and practicing SLPs should be mindful that while a finding of a few instances of difference across informal phonological measurement samplings does not mean the measure itself is inappropriate, it does indicate that outcomes should be cautiously interpreted for young children.

Conflict of interests

Authors declare no conflict of interests.