

Research Article

Sentence repetition test for measurement of grammatical development in Farsi speaking children

Fatemeh Hasanati¹ - Zahra Agharasouli¹ - Behrouz Mahmoudi Bakhtiyari² – Mohammad Kamali³

1- Department of Speechtherapy, Faculty of Rehabilitation, Tehran University of Medical Sciences, Iran

2- Department of Dramatic Literature, Faculty of Art, University of Tehran, Iran

3- Department of Rehabilitaion Management, Faculty of Rehabilitation, Tehran University of Medical Sciences, Iran

Received: 29 August 2010, accepted: 22 November 2010

Abstract

Background and Aim: valid identification, prevention, and treatment of language disorders are a high priority for the speech and language professionals. One method for studying language development is sentence repetition that is faster to implement and analysis than other procedures. The aim of this project was constructing sentence repetition test as a quick measure of grammatical potency in 2.5 to 4 year old children.

Methods: Sentences appropriate for 2.5 to 4 year old children were selected during several stages by speech and language pathologist and linguists. The validity of sentences was assessed by professional masters in this theme. Subsequently, 41 sentences including those with 80% high validity were selected as the test sentences. Appropriate pictures were also provided with sentences. The test was administrated to 72 children in 3 groups (2.5-3, 3-3.5, and 3.5-4 year olds, gender matched). The reliability was administered with a test-retest design across a 2 weeks interval.

Results: Content validity Index for this test was 80%. "Test-retest reliability" was used for reliability of this test. The Interclass correlation coefficient for this test was 0.95 and standard error measurement was 7.45. The average of scores for sentence repetition, between groups was significant ($p < 0.001$, $p < 0.001$, $p = 0.014$).

Conclusion: This sentence repetition test has the appropriate validity and reliability as well as the capability of proper and quick assessment (screening) of grammatical development in 2.5 to 4 year old Persian speaking children.

Keywords: Sentence repetition test, grammar of language, children, Persian language

آزمون تکرار جمله برای بررسی ساختار دستور زبان کودکان فارسی‌زبان

فاطمه حسناتی^۱ - زهرا آقارسولی^۱ - بهروز محمودی بختیاری^۲ - محمد کمالی^۳

۱- گروه گفتاردرمانی، دانشکده توانبخشی، دانشگاه علوم پزشکی تهران، تهران، ایران

۲- گروه هنرهای نمایشی، دانشکده هنرهای زیبا، دانشگاه تهران، تهران، ایران

۳- گروه مدیریت توانبخشی، دانشکده توانبخشی، دانشگاه علوم پزشکی تهران، تهران، ایران

چکیده

زمینه و هدف: مطالعه زبان و شناسایی دقیق و سریع اختلالات و درمان آن از اولویت بالایی برای متخصصان برخوردار است. یکی از روش‌های بررسی سریع ساختار زبانی، روش تکرار جمله است که جهت اجرا و تجزیه و تحلیل، سریع‌تر از دیگر روش‌ها است. هدف از این پژوهش ساخت آزمون تکرار جمله برای اندازه‌گیری ساختار دستور زبان کودکان ۲/۵ تا ۴ ساله و تعیین روایی و پایایی آن بود.

روش بررسی: در چند مرحله، با کمک آسیب‌شناس گفتار و زبان و زبان‌شناس، جملاتی مخصوص کودکان ۲/۵ تا ۴ ساله تهیه شد. با نظرسنجی از ۱۰ نفر از استادان، روایی محتوایی جملات بررسی شد. از بین این جملات ۴۱ جمله که دارای روایی بالایی ۸۰ درصد بودند به‌عنوان جملات آزمون انتخاب شد. سپس تصاویر مناسب با جملات تهیه گردید و دوباره نظرسنجی انجام شد. ۷۲ کودک (دختر و پسر) در سه گروه سنی شامل ۲/۵ تا ۳، ۳ تا ۳/۵، ۳/۵ تا ۴ سال، در این آزمون شرکت کردند. پایایی به‌روش آزمون-بازآزمون به فاصله ۲ هفته برای همان ۷۲ کودک اجرا شد.

یافته‌ها: شاخص روایی محتوایی جملات نهایی آزمون ۸۰ درصد به‌دست آمد. میزان تکرارپذیری نسبی کل آزمون ۰/۹۵ و خطای معیار اندازه‌گیری آزمون برابر با ۷/۴۵ به‌دست آمد. همچنین تفاوت بین میانگین‌های امتیاز کسب شده از جملات آزمون بین گروه‌های سنی مورد مطالعه معنی‌دار بود ($p < 0.001$ ، $p < 0.001$).

نتیجه‌گیری: آزمون تکرار جمله روشی سریع برای اندازه‌گیری ساختار دستور زبان کودکان ۲/۵ تا ۴ ساله فارسی‌زبان می‌باشد. همچنین آزمون از لحاظ محتوایی، روا بوده و پایایی بسیار بالایی دارد.

واژگان کلیدی: آزمون تکرار جمله، دستور زبان، کودکان، زبان فارسی

(دریافت مقاله: ۸۹/۶/۷، پذیرش: ۸۹/۹/۱)

مقدمه

قرار می‌دهد. این روش علی‌رغم داشتن جنبه‌های مثبت دارای مشکلاتی است. برای مثال، برای محاسبه میانگین طول گفته و یا اشتباهات دستوری گفتار خودانگیزته لازم است از هر کودک حداقل ۵۰ تا ۱۰۰ گفته جمع‌آوری شود. برای ضبط این میزان گفته، به‌ویژه در مورد کودکان کوچک‌تر، به وقت زیادی نیاز است. از طرفی، برای واج‌نگاری در تجزیه و تحلیل ۵۰ گفته، باید حداقل یک تا سه ساعت زمان صرف شود. برای تعیین مرز بین گفته‌ها لازم است به تعداد دفعات زیاد صدای ضبط شده شنیده شود. برای

به‌منظور شناخت چگونگی فراگیری زبان و اختلالات مربوط به آن لازم است زبان در مراحل اولیه تحول مورد بررسی قرار گیرد، زیرا با شناسایی زود هنگام اختلالات زبانی و تعیین شدت و چگونگی آن می‌توان برنامه درمانی مناسبی را تنظیم و اجرا کرد. اکثر مطالعات انجام شده درباره مراحل اولیه فراگیری زبان، بر پایه مشاهده گفتار خودانگیزته کودک در موقعیت‌های طبیعی بوده است (۲۰۱ و ۲۰۳). Dockrell (۲۰۰۳) خاطر نشان می‌کند که این نوع بررسی، اطلاعات مهمی را در اختیار محققان و درمانگران

و سپس جملات طولانی‌تر و با ساختار پیچیده‌تر ارائه می‌شود (۴). نمونه دیگر آزمون غربالگری که از مهارت تکرار استفاده کرده است آزمون غربالگری نحو نورث وسترن (Northwestern Syntax Screening Test: NSST) (۱۹۸۰) است که یکی از بزرگ‌ترین ابزارهای غربالگری برای بررسی نحو زبان است. این آزمون غربالگری برای بررسی زبان بیانی و درکی طراحی شده و برای کودکان ۳ تا ۷ سال و ۱۱ ماه مناسب است. این آزمون با استفاده از روش آزمون‌باز آزمون، امتیاز متوسطی را کسب کرده است (۵۴۲/۰ برای بخش دریافتی و ۶۹۷/۰ برای بخش بیانی) (۶). آزمون غربالگری تحول دستور زبان و واج‌شناسی (Grammar and Phonology Screening test: GAPS) (۲۰۰۶) برای غربالگری سریع اختلالات زبان طراحی شده و مخصوص کودکان پیش‌دبستانی و سال‌های ابتدایی دبستان است. با این آزمون می‌توان مشخص نمود که آیا کودک مهارت لازم در زمینه دستور و واج‌شناسی را جهت فراگیری خواندن کسب کرده است یا خیر. این آزمون شامل ۱۱ جمله و ۸ کلمه بی‌معنی (ناکلمه) است. در این آزمون از مهارت تکرار مستقیم استفاده شده است و توانایی شناسایی اختلالات خواندن و آسیب‌های زبانی را دارد و اجرای آن ۱۰ دقیقه طول می‌کشد. ۶۶۸ کودک ۳/۴ تا ۶/۶ ساله برای هنجارسازی این آزمون مورد بررسی قرار گرفتند. آزمون GAPS در تشخیص آن دسته از کودکانی که نیاز به ارزیابی دقیق‌تر و عمیق‌تر دارند مؤثر است (۷). ابزارهای رسمی و خلاصه که انجام آنها به ۱۰ دقیقه یا کمتر زمان نیاز دارند دیدگاهی قابل قبول و استاندارد جهت غربالگری تأخیر گفتار و زبان ارائه می‌دهند، هرچند غربالگری با این ابزارها قبل از انجام درمان‌های مناسب باید با ابزارهای تشخیصی بیشتر همراه شوند (۸). تا زمان شروع این مطالعه، در ایران هیچ‌گونه تکلیف یا آزمونی برای بررسی روند تحول ساختار نحوی گفتار کودکان به‌دست نیامد. از این رو، این پژوهش به‌منظور آزمون تکرار جمله برای بررسی ساختار دستور زبان کودکان فارسی‌زبان انجام گردید.

روش بررسی

پرهیز از این مشکلات لازم است از روش‌هایی استفاده شود که اجرای آن‌ها آسان‌تر باشد و بتوان با آن‌ها اختلالات گفتار و زبان را به‌سرعت شناسایی کرد (۲). آزمون تکرار جمله این هدف را برآورده می‌کند. تحقیقات نشان داده است تکلیف تکرار جمله، یکی از راه‌های غیرمستقیم برای بررسی قواعد نحوی، معناشناختی و واج‌شناختی گفتار است (۳). همچنین، هنگامی که محدودیت زمانی وجود دارد این روش راهکار مفیدی برای درمانگر است تا در مدت زمان کوتاه توانایی‌های کودک را از راه تکلیف تکرار شناسایی کند. روش دیگر، همراهی تکلیف تکرار با سایر روش‌های جمع‌آوری اطلاعات زبانی است تا یک بازنگری متقابل جهت اعتبار اطلاعات بین این روش‌ها ایجاد شود (۴). از فواید دیگر استفاده از تکلیف تکرار سادگی آن در جهت اجرا و تجزیه و تحلیل نتایج نسبت به سایر روش‌ها است (۵). با استفاده از این روش می‌توان جنبه‌های دستوری و واج‌شناختی زبان را بررسی کرد. Brown (۱۹۵۷) بیان می‌کند که تکلیف تکرار، روشی با روایی و پایایی بالا جهت ارزیابی دانش دستور زبان کودکان است (۵).

Caselli و Devescovi (۲۰۰۶) از طریق روش تکرار جمله تحول اولیه دستور زبان را در کودکان پیش‌دبستانی ۲-۴ ساله بررسی کردند. این بررسی نشان داد که کودکان دو ساله جملات را به‌صورت تلگرافی تکرار می‌کنند. میانگین طول گفته در تکلیف تکرار جمله کودکان ۲ و ۲/۵ ساله حدود ۲ تا ۳ کلمه بود ولی میزان استفاده از حروف اضافه، حروف تعریف و صفت‌ها به‌طور معنی‌داری کم بود. بعد از سن ۳ سالگی حذف کلمات نقشی (عملکردی) آزاد عملاً مشاهده نشد. به‌طور کلی، نتایج Caselli و Devescovi (۲۰۰۶) نشان داد که روش تکرار جمله، برای بررسی ساختار دستوری کودکان اعتبار داشته و توانایی تمیز بین گروه‌های سنی مختلف مورد مطالعه را دارد (۳).

مجموعه استخراج شده زبان کارو (Carrow Elicited Language Inventory: CELI) یک نمونه از آزمون‌های زبانی است که از مهارت تکرار استفاده کرده است. در آن آزمون از فرد خواسته می‌شود که ۵۱ گفته را بعد از گوینده تکرار کند. آزمون با یک استراحت شروع می‌شود و در آن ابتدا جملات با ساختار ساده

نباید باشد، مشخص شد. نسبت روایی محتوایی برای هر مقوله از فرمول زیر محاسبه شد (۱۱):

$$CVR = \frac{ne - \frac{N}{2}}{\frac{N}{2}}$$

که در آن ne برابر تعداد گزینه‌های ضروری است که متخصصان به هر مؤلفه دادند. N تعداد کل افراد ارزیابی‌کننده است. میزان روایی محتوایی بین ۱ و -۱ بوده و هرچه میزان به‌دست آمده به یک نزدیک‌تر باشد نشان می‌دهد که آزمون از روایی بیشتری برخوردار است (۱۱). برای انجام روایی، ابتدا جملات بدون تصویر در اختیار افراد قرار داده شد و نظرات آنها به‌صورت امتیاز ضروری است «۲»، بد نیست «۱» و نباید باشد «۰» بررسی شد. در نهایت از بین ۷۸ جمله، ۴۱ جمله که امتیاز بالای ۰/۶ را کسب کردند انتخاب شدند. ۲۰ جمله امتیاز ۰/۸ و ۱۰ جمله امتیاز ۰/۶ و ۱۱ جمله امتیاز ۱ کسب کردند و بقیه از آزمون حذف شدند. $polit$ و همکاران (۲۰۰۷) با مقایسه روش‌های مختلف تعیین روایی محتوایی نشان دادند که محاسبه شاخص روایی محتوایی (Content Validity Index: CVI)، نسبت به دیگر روش‌های موجود، از مزایای بیشتری برخوردار است که سادگی و قابل فهم بودن این روش از مهم‌ترین مزایای آن محسوب می‌شود (۱۲). شاخص روایی محتوایی از مجموع میانگین‌های CVI به‌دست می‌آید. با این روش CVI جملات انتخاب شده برای آزمون ۸۰ درصد می‌شود که روایی مطلوبی است. سپس برای این جملات تصاویر مناسب طراحی و با استفاده از نرم‌افزار فتوشاپ CS4 رنگ‌آمیزی و ویرایش شد تصاویر آزمون تا حد امکان ساده بودند و مفهوم اصلی جملات را می‌رساندند و از نظر نوع طراحی و اندازه یکسان بودند. ۲ جمله به‌عنوان جملات راهنما و ۳۹ جمله به‌عنوان جملات آزمون انتخاب شدند. سپس جملات به‌وسیله نرم‌افزار Jet Audio با فورمت ۴۴۱۰۰ در یک اتاق آکوستیک ضبط گردید. در نهایت جملات همراه تصاویرشان با نرم‌افزار پاور پوینت تنظیم شدند. پس از تنظیم آزمون از ۵ نفر از افرادی که روایی را انجام داده بودند خواسته شد تا در مورد آزمون همراه با تصاویر به‌صورت

ابتدا در چند مرحله با کمک زبان‌شناس و آسیب‌شناس گفتار و زبان ۷۸ جمله مخصوص کودکان ۲/۵ تا ۴ سال طراحی شد. برای تهیه این جملات از کتاب‌های مناسب این سنین و چند منبع تحقیقی (جليله‌وند ۱۳۷۴، امیری شوکی ۱۳۷۸، مدرس‌زاده ۱۳۸۷) استفاده شد. جملات دارای ویژگی‌های زیر هستند:

۱. جملات به‌ترتیب روند رشد از ساده به پیچیده و از دو کلمه‌ای تا چند کلمه‌ای ساخته شدند؛
 ۲. جملات شامل کلمات محتوایی (اسم، فعل) و کلمات نقشی (حروف اضافه و ربط، نشانه جمع و نشانه زمان افعال) هستند.
 ۳. از کلماتی استفاده شده است که در روند رشد زودتر یاد گرفته می‌شوند و مطابق تنوع اجتماعی-اقتصادی و فرهنگی هستند. همچنین کلمات از ساختار واجی ساده برخوردارند، بنابراین احتمال شکست در آزمون به‌خاطر ناآشنا بودن کلمات یا عدم توانایی تلفظ آنها به کمترین حد رسید (۷)؛
 ۴. از کلمات هم‌خانواده و مشابه استفاده شده است تا از تأثیر تنوع واژگانی بر پاسخ کودک جلوگیری شود (۳)؛
 ۵. ترتیب قرارگیری کلمات در جملات به شکل فاعل، مفعول و فعل به‌همراه متمم هر کدام است؛
 ۶. هر جمله همراه با تصویر ارائه می‌شود؛
 ۷. جملات از نوع مثبت (affirmative) و خبری (declarative) هستند (۹)؛
 ۸. جملات به زبان فارسی رسمی و بدون استفاده از صورت محاوره‌ای ارائه می‌شوند، بنابراین تأثیر تنوع لهجه و فرهنگ به کمترین میزان ممکن می‌رسد (۱۰)؛
- پس از ساخت جملات آزمون تکرار جمله، جهت بررسی روایی محتوایی آن، از ۱۰ نفر از استادان آسیب‌شناس گفتار و زبان (گفتاردرمانگر) و زبان‌شناس نظرسنجی شد. روایی استفاده شده برای این آزمون از نوع محتوایی بود. نسبت روایی محتوایی توسط lawshe (۱۹۷۵) مطرح شده است. در بررسی نسبت روایی محتوایی (Content Validity Ratio: CVR) پاسخ متخصص به هر مؤلفه به‌صورت یکی از سه گزینه: ضروری است، بد نیست و

هفته بعد، از تمام کودکان آزمون مجدد به عمل آمد. روش امتیازدهی آزمون به این ترتیب بود که به تکرار کامل امتیاز ۲ و به عدم تکرار کامل امتیاز ۱ داده می‌شد. با توجه به اینکه تعداد کل جملات ۳۹ جمله است در صورتی که فردی به کل جملات آزمون پاسخ صحیح بدهد یعنی کل جملات را به طور کامل تکرار کند، مجموع امتیاز آزمون ۴۸۴ خواهد شد. امتیازدهی این آزمون شامل دو بخش کلی و دو بخش فرعی است:

۱. امتیاز بخش کلمات محتوایی (شامل اسم‌ها و فعل‌ها) ۱۵۶ است.
۲. امتیاز بخش کلمات نقشی (شامل حروف اضافه، حروف ربط، نشانه جمع و زمان فعل‌ها) ۱۷۲ است.
۳. امتیاز بخش نحو (رعایت ترتیب کلمات در جمله) ۷۸ است.
۴. امتیاز بخش مطابقت فعل و فاعل (بخشی از صرف زبان فارسی) ۷۸ است.

پایایی آزمون با بررسی ضریب همبستگی بین میانگین‌های آزمون-بازآزمون، خطای معیار اندازه‌گیری و ضریب همبستگی پیرسون مورد بررسی قرار گرفت. میانگین امتیازهای گروه‌های سنی مختلف با استفاده از ANONA یک‌طرفه و آزمون توکی تعقیبی مقایسه گردید.

یافته‌ها

نتایج نشان می‌دهد که با افزایش سن بر میانگین امتیازهای کسب شده توسط کودکان در آزمون-بازآزمون افزوده می‌شود. جدول ۱ میانگین امتیاز و انحراف معیار را در دو بار انجام آزمون نشان می‌دهد.

همچنین میانگین امتیازها در دو بار انجام آزمون، ضریب همبستگی بسیار بالایی دارد که از لحاظ آماری معنی‌دار است ($p=0/000$). مقدار ضریب همبستگی (Intraclass Coefficient: ICC) بین میانگین‌های آزمون-بازآزمون و خطای معیار اندازه‌گیری (Standard Error Measurement: SEM) و ضریب همبستگی پیرسون کل آزمون نیز ۰/۹۵ و ۷/۴۵ و ۰/۹۲

مناسب / نامناسب نظر بدهند که هر ۵ نفر به ۳۸ جمله امتیاز مناسب و ۳ جمله باقیمانده را ۴ نفر امتیاز مناسب و ۱ نفر امتیاز نامناسب دادند. در پایان اشکالات موجود در تصاویر و صدا تصحیح و آزمون برای اجرا آماده شد. پایایی آزمون از طریق روش آزمون-بازآزمون انجام شد. ۷۲ کودک، شامل ۳۶ دختر و ۳۶ پسر، در سه گروه سنی ۳-۲/۵ (گروه ۱)، ۳-۳/۵ (گروه ۲) و ۴-۳/۵ (گروه ۳) که به روش تصادفی از مهدکودک‌های مناطقی از شهر تهران، در محدوده بین منطقه راه آهن و بزرگراه همت، که از نظر فرهنگی و اقتصادی به عنوان مناطق متوسط شناخته می‌شوند، انتخاب شده بودند در این آزمون شرکت کردند. در این مطالعه انتخاب نمونه‌ها با استفاده از اطلاعات موجود در پرونده بهداشتی کودک در مهدکودک و اطلاعات دریافتی از مربیان انجام شد. برای کسب اجازه از والدین یا مربیان، توضیحات لازم در مورد پژوهش و اهداف و روش انجام آن ارائه شد و فرم رضایت‌نامه توسط مسئول مهدکودک امضاء گردید. سپس آزمون مورد نظر برای هر کودک به طور جداگانه اجرا گردید و برای جلوگیری از حواس‌پرتی کودک سعی شد آزمون در اتاقی ساکت انجام شود. آزمونگر ابتدا با کودک رابطه دوستانه برقرار و برای آشنا ساختن وی با روش کار، ۲ جمله آزمایشی ارائه می‌کرد. راهنمای انجام آزمون به این گونه است که آزمونگر در کنار کودک نشسته و می‌گوید من جمله‌ای برای تو پخش می‌کنم و سپس تصویر آن را به تو نشان می‌دهم؛ همان جمله را دقیقاً بعد از دیدن تصویر بگو. سپس آزمونگر جمله را پخش می‌کند و بلافاصله تصویر آن را به کودک نشان می‌دهد و می‌گوید حالا تکرار کن. آزمونگر چند ثانیه صبر می‌کند. اگر کودک جواب ندهد دوباره جمله را پخش می‌کند، بدون اینکه تصویر را بردارد. هنگامی که آزمونگر مطمئن شد کودک روش کار را فهمیده است آزمون را ادامه می‌دهد و ابتدا جمله را پخش می‌کند و بلافاصله تصویر مربوطه را ارائه می‌کند. در مرحله پایایی محدوده زمانی خاصی جهت پاسخ‌گویی مشخص نشد. جملات فقط یک بار پخش شد و هیچ نوع راهنمایی در حین انجام آزمون ارائه نگردید. صدای کودکان با استفاده از نرم‌افزار jet Audio جهت تجزیه و تحلیل ضبط گردید. برای انجام پایایی، دو

جدول ۱- شاخص‌های آماری کل آزمون تکرار جمله

p	میانگین کل (انحراف معیار)		تعداد	گروه‌های سنی
	آزمون مجدد	آزمون		
<۰/۰۰۰	۴۲۹/۰۸(۳۰/۸۹)	۴۰۷/۰۸(۳۳/۳۳)	۲۴	۱
<۰/۰۰۰	۴۷۵/۰۸(۳۱/۱۴)	۴۶۶/۴۲(۳۹/۹۴)	۲۴	۲
<۰/۰۰۰	۵۰۹/۵۰(۲۷/۶۴)	۴۹۶/۰۸(۳۳/۲۹)	۲۴	۳
<۰/۰۰۰	۴۷۱/۲۲(۴۴/۴۰)	۴۵۶/۵۳(۵۱/۲۲)	۷۲	کل

تصاویر مربوطه است. این آزمون شامل بخش کلمات محتوایی شامل فعل، اسم، صفت، و کلمات نقشی شامل حروف تعریف و حروف ربط است. در این آزمون از کلمات هم‌خانواده و مشابه و کلماتی که در روند رشد، زودتر یاد گرفته می‌شوند استفاده شده است (۳).

در روند به‌کارگیری آزمون، برای تشخیص یک اختلال نیز، بررسی تکرارپذیری آزمون برای تعیین اعتبار، به‌ویژه زمانی که از آن برای تعیین راهکارهای درمان بیماران استفاده می‌شود بسیار مهم است (۱۴). هرچه مقدار ICC به عدد ۱ نزدیک باشد، نشان‌دهنده تکرارپذیری بالاتر آزمون است و بیانگر آن است که میانگین دو بار انجام آزمون تفاوت معنی‌داری با یکدیگر ندارد (۱۱). در پژوهش حاضر برای بیان درجه تکرارپذیری از تعریف munro (۲۰۰۵) در تفسیر ضریب همبستگی استفاده شد. ضریب همبستگی صفر تا ۰/۲۵ به‌عنوان ارتباط اندک، ۰/۲۶ تا ۰/۴۹ به‌عنوان ارتباط ضعیف، ۰/۵۰ تا ۰/۶۹ به‌عنوان ارتباط متوسط، ۰/۷ تا ۰/۸۹ به‌عنوان ارتباط بالا و ۰/۹ تا ۱/۰۰ به‌عنوان ارتباط بسیار بالا در نظر گرفته شد (۱۵). در این آزمون امتیاز تکرارپذیری برای کل آزمون ۹۵ درصد و ضریب همبستگی آن ۰/۹۲ به‌دست آمد، که نشان می‌دهد آزمون از تکرارپذیری بسیار بالا به‌عبارت دیگر از پایایی بسیار بالایی برخوردار است.

خطای معیار اندازه‌گیری (SEM) یک روش آماری است که میزان خطا را تخمین می‌زند و ارتباط مستقیم با ضریب روایی

به‌دست آمد. جدول ۲ نتایج بررسی پایایی را نشان می‌دهد. با استفاده از روش ANOVA یک‌طرفه و آزمون توکی تعقیبی، میانگین نمرات گروه‌های سنی مختلف مقایسه شد. تفاوت میانگین‌ها با ضریب اطمینان ۹۵٪ بین گروه‌های سنی به این شرح است که گروه سنی ۱ و ۲ $p=۰/۰۰۰$ ، گروه سنی ۱ و ۳ $p=۰/۰۰۰$ ، گروه سنی ۲ و ۳ $p=۰/۰۱۴$. بنابراین تفاوت میانگین‌ها بین همه گروه‌های سنی معنی‌دار است.

بحث

تعیین روایی محتوایی یکی از شرط‌های اصلی هر آزمون جدیدی است. از روش‌های مقبول اخیر در این زمینه، محاسبه شاخص روایی محتوایی (CVI) است (۱۳). براساس مطالعات قبلی که با استفاده از نظرخواهی ۳ تا ۱۰ کارشناس انجام شده است، CVI بزرگ‌تر از ۰/۷۵ یا ۰/۸ به‌عنوان مقدار قابل قبول در نظر گرفته می‌شود (۱۲). در این آزمون نیز از شاخص روایی محتوایی استفاده شده است. به این صورت که از ۱۰ نفر از آسیب‌شناسان گفتار و زبان و زبان‌شناس درباره ساختار و محتوای جملات نظرسنجی به‌عمل آمد و جملاتی که امتیاز بیش از ۸۰ درصد را کسب کردند به‌عنوان جملات آزمون انتخاب شدند که عدد به‌دست آمده برای روایی نشان می‌دهد آزمون از روایی قابل قبولی برخوردار است آزمون تکرار جمله Caselli و Devescovi (۲۰۰۶) نیز همانند آزمون حاضر شامل تعدادی جمله ساده به‌همراه

جدول ۲- نتایج بررسی پایایی آزمون تکرار جمله به روش بررسی ضریب همبستگی و خطای معیار اندازه‌گیری (تعداد ۷۲ نفر)

بخش‌ها	ضریب همبستگی پیرسون	همبستگی بین میانگین‌های آزمون بازآزمون	خطای معیار اندازه‌گیری	p
کلمات محتوایی	۰/۹۱	۰/۹۴	۴/۳۵	۰/۰۰۰
کلمات نقشی	۰/۸۸	۰/۹۴	۵/۰۵	۰/۰۰۰
کل آزمون (۳۹ جمله)	۰/۹۲	۰/۹۵	۷/۴۵	۰/۰۰۰

NSST (۱۹۸۰) است که یکی از بزرگ‌ترین ابزارهای غربالگری برای بررسی نحو است. یکی از روش‌هایی که جهت بررسی پایایی این آزمون استفاده شد آزمون-بازآزمون است که در آن این آزمون پایایی متوسطی را کسب کرده است (۶). در پژوهش حاضر نتیجه تکرارپذیری به‌دست آمده از آزمون-بازآزمون ۹۵٪ به‌دست آمد که تکرارپذیری بسیار بالایی را نشان می‌دهد. یکی دیگر از آزمون‌هایی که قبلاً در مورد آن بحث کردیم GAPS است که برای غربالگری سریع و ساده موقعیت زبانی کودکان طراحی شده است. جهت بررسی پایایی این آزمون نیز، از آزمون-بازآزمون استفاده شد (۷).

مطالعات Caselli و Devescovi در سال ۲۰۰۶ با استفاده از روش تکرار جمله نشان می‌دهد که تکرار جمله در کودکان سنین پایین تلگرافی است. تعداد تکرار جملات در سنین پایین کمتر بوده و اغلب جملات ساده را تکرار می‌کنند. همچنین این نتایج و داده‌ها با نتایج مطالعات پیشین که با بررسی گفتار خودانگیزخته صورت گرفته است همخوانی دارد (۳). Devescovi و Caselli بیان کردند که کودکان ایتالیایی در سن ۳ سالگی شروع به یادگیری بیشتر بخش‌های گرامری می‌کنند که این نتایج شبیه الگوی تولید در گفتار خودانگیزخته است (۳). نتایج بررسی Brown (۱۹۶۳) نیز تأیید می‌کند که در مراحل اولیه تحول زبان حضور کلمات محتوایی بیشتر از کلمات نقشی است (۳). آزمون حاضر نیز نتایج بالا را تأیید می‌کند. نتایج آزمون حاضر نشان می‌دهد که با این روش می‌توان دریافت که روند تحول زبان با افزایش سن

آزمون و انحراف معیار امتیازهای آزمون دارد. SEM برای هر امتیاز دلالت بر بی‌ثباتی و تغییرپذیری دارد که انتظار می‌رود در محدوده امتیاز آزمون دیده شود. به عبارت دیگر SEM دلالت بر این دارد که اگر آزمون با همان شرایط یکسان دوباره تکرار شود امتیازهای کودک چه مقدار می‌تواند تغییر کند. هنگامی که آزمون برای کودک انجام شد امتیازها به‌دست آمده امتیاز واقعی کودک را تخمین می‌زند که شامل مقداری خطاست. به همین علت SEM به‌دست آمده از آزمون کمک می‌کند میزان تفاوت احتمالی امتیاز کودک از امتیاز واقعی مشخص شود. هنگامی که بخواهیم آزمون را برای یک فرد دوباره تکرار کنیم SEM یک محدوده یا درجه اطمینان ایجاد می‌کند (۷). میزان SEM در دو بار تکرار آزمون ۷/۴۵ بوده است. با توجه به اینکه انحراف معیار کل آزمون ۵۱/۲۲ بوده و در میزان معیار خطای اندازه‌گیری مؤثر است، می‌توان نتیجه گرفت که خطای معیار اندازه‌گیری نسبت به میزان انحراف معیار در محدوده مطلوبی است. با این همه باید به خاطر داشته باشیم که این امتیازها فقط مقداری اطلاعات درباره مهارت‌های زبانی کودک فراهم می‌کند که باید با تاریخچه و اطلاعات از منابع مختلف مثل خانواده همراه باشد. همچنین برای تعیین این که آیا آزمون حاضر قادر به تمایزگذاری بین گروه‌های هنجار و گروه مبتلا به اختلال است یا خیر، می‌توان در پژوهش‌های بعدی این آزمون را در گروه‌های دچار انواع اختلالات مورد بررسی قرار داد و نتایج به‌دست آمده را با گروه هنجار مقایسه کرد (۷). در این راستا یکی از آزمون‌های غربالگری که از مهارت تقلید استفاده کرده است

می‌شوند ولی بعضی اعتقاد دارند این دانش وجود دارد (۱۶).

نتیجه‌گیری

نتایج نشان می‌دهد که این ابزار از نظر محتوا و تناسب با زبان فارسی، روایی لازم برای بررسی کودکان ۲/۵ تا ۴ ساله را دارد و سطح عملکرد کودکان به میزان بسیار بالایی در دو مرتبه یکسان است به عبارت دیگر آزمون از قابلیت تکرار بسیار بالایی برخوردار است. همچنین ضریب همبستگی آزمون در دو بار تکرار از لحاظ آماری معنی‌دار است.

سپاسگزاری

نویسندگان مراتب تشکر و سپاس خود را از مسئولان سازمان بهزیستی تهران، به‌ویژه جناب آقای خلیل‌زاده، مدیران مهدکودک‌هایی که در این پژوهش همکاری کردند و کلیه اساتیدی که در روند انجام روایی این آزمون شرکت داشتند، اعلام می‌دارد.

REFERENCES

- Batse E, Bretherton I, Snyder L. From first word to grammar: individual differences and dissociable mechanism. 1st ed. Cambridge: Cambridge University Press; 1988.
- Clark EV. First language acquisition. 1st ed. Cambridge: Cambridge University Press; 2003.
- Devescovi A, Caselli MC. Sentence repetition as a measure of early grammatical development in Italian. *Int J Lang Commun Disord*. 2007;42(2):187-208.
- Tomblin JB, Morris HL, Spriestersbach DC. Diagnosis in speech-language pathology. 2nd ed. San Diego, California: singular publisher group; 2000.
- Crain S, Thornton R. Investigations in universal grammar: a guide to experiments on the acquisition of syntax and semantics. 2nd ed. Cambridge, MA: MIT Press; 2000.
- Ratusnik DL, Klee TM, Ratusnik CM. northwestern syntax screening test: a short form. *J Speech Hear Disord*. 1980;45(2):200-8.
- Gardner H, Froud K, McClelland A, van der Lely HK. Development of the grammar and phonology screening (GAPS) test to assess key markers of specific language and literacy difficulties in young children. *Int J Lang Commun Disord*. 2006;41(5):513-40.
- US Preventive Services Task Force. Screening for speech and language delay in preschool children: recommendation statement. *Pediatrics*. 2006;117(2):497-501.
- Hoff E. Language development. 3rd ed. Belmont, CA: Thompson learning, wadsworth; 2005.
- Leikin M, Ibrahim R, Eviatar Z, Sapir S. Listening with an accent: speech perception in

- a second language by late bilinguals. *J Psycholinguist Res.* 2009;38(5):447-57.
11. Shultz kS, Whitney DJ. *Measurement theory in action, case study and exercise.* 1st ed. Thousand Oaks, CA: Sage publication; 2005.
 12. Polit DF, Beck CT, Owen SV. Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Res Nurs Health.* 2007;30(4):459-67.
 13. Kim Y, Park J, Lee H, Bang H, Park HJ. Content validity of an acupuncture sensation questionnaire. *J Altern Complement Med.* 2008;14(8):957-63.
 14. Hicks GE, Fritz JM, Delitto A, Mishock J. Interrater reliability of clinical examination measures for identification of lumbar segmental instability. *Arch Phys Med Rehabil.* 2003;84(12):1858-64.
 15. Mathur S, Eng JJ, MacIntyre DL. Reliability of surface EMG during sustained contractions of the quadriceps. *J Electromyogr Kinesiol.* 2005;15(1):102-10.
 16. Bishop D, Mogford K. *Language development in exceptional circumstances.* 1st ed. Hove: psychology press Ltd, Taylor & Francis; 1993.