

Research Article

Developing a story retelling test for the assessment of language structure in Persian-speaking children

Salimeh Jafari¹, Zahra Agharasouli¹, Yahya Modaresi², Mohammad Kamali³

¹- Department of Speech therapy, School of Rehabilitation, Tehran University of Medical Sciences, Iran

²- Department of Linguistics, Institute for Humanities and Cultural Studies, Tehran, Iran

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

Received: 21 December 2011, accepted: 21 May 2012

Abstract

Background and Aim: Valid identification, prevention, and treatment of language disorders have a high priority for the clinicians. Story retelling is a method for studying language development which is faster than other procedures in implementing and analyzing, and has a reliable scoring system. The aim of this project is to construct story retelling test for assessing language structure in 6-7 year old, Persian children.

Methods: An appropriate story for 6-7 year old children was written by the help of a speech pathologist, a linguist, and a novelist. The validity of the constructed story was checked by 14 experts. Then related pictures were design and their validity was examined. The test was conducted on 72 children (36 girls, 36 boys) of 6-7 years. The reliability was administered with a test-retest design with a two week interval. The reliability of the scoring system was evaluated with inter-rater method.

Results: Content validity index for this test was 89%, interclass correlation was 83%, and standard error of measurement was 2.76. The interclass correlation for scoring system was 93% and Pearson's correlation was 90%. Correlation of the test-retest scores was statistically significant ($p < 0.01$).

Conclusion: This story retelling test has appropriate validity and reliability. It is suitable for quick assessment of language structure in 6-7 year old, Persian children.

Keywords: Test, story retelling, language structure, children, Persian language

ساخت آزمون بازگویی داستان برای ارزیابی ساختار زبان در کودکان فارسی‌زبان

سلیمه جعفری^۱، زهرا آقارسولی^۱، یحیی مدرسی^۲، محمد کمالی^۳

^۱ - گروه گفتار درمانی، دانشکده توانبخشی، دانشگاه علوم پزشکی تهران، ایران

^۲ - گروه زبان‌شناسی، پژوهشگاه علوم انسانی و مطالعات فرهنگی، تهران، ایران

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

چکیده

زمینه و هدف: شناسایی دقیق و سریع اختلالات زبان و درمان آن از اولویت بالایی برای درمانگران برخوردار است. یکی از روش‌های بررسی سریع ساختار زبانی، بازگویی داستان است که برای اجرا و تحلیل، سریع‌تر از دیگر روش‌ها است و سیستم امتیازدهی پایایی دارد. هدف از این پژوهش ساخت و تعیین روایی و پایایی آزمون بازگویی داستان در کودکان شش تا هفت ساله فارسی‌زبان بود.

روش بررسی: ابتدا با کمک گفتاردرمانگر، زبان‌شناس و داستان‌نویس کودک، داستانی مناسب کودکان شش تا هفت ساله نوشته شد. برای بررسی روایی محتوایی داستان، از ۱۴ متخصص، نظرسنجی شد. سپس تصاویر مناسب با داستان تهیه و به صورت لوح فشرده در آمد، سپس دوباره نظرسنجی صورت گرفت. پایایی به روش آزمون-بازآزمون، به فاصله دو هفته روی ۷۲ کودک در گروه سنی شش تا هفت سال شامل ۳۶ دختر، ۳۶ پسر اجرا شد. برای بررسی پایایی مقیاس امتیازدهی از روش پایایی بین ارزیاب استفاده شد.

یافته‌ها: شاخص روایی محتوایی آزمون بازگویی داستان ۸۹ درصد، میزان تکرارپذیری نسبی کل آزمون ۸۳ درصد و خطای معیار اندازه‌گیری آن ۲/۷۶ بود. همچنین تکرارپذیری نسبی مقیاس امتیازدهی ۹۳ درصد و همبستگی پیرسون برای آن ۹۰ درصد به دست آمد. همبستگی امتیازهای دو بار انجام آزمون معنی‌دار بود ($P < 0.01$).

نتیجه‌گیری: آزمون بازگویی داستان، روشی سریع برای ارزیابی ساختارهای زبان کودکان شش تا هفت ساله فارسی‌زبان است. آزمون از لحاظ محتوایی و کیفی، روا بوده و پایایی بالایی دارد.

واژگان کلیدی: آزمون، بازگویی داستان، ساختار زبان، کودکان، فارسی‌زبان

(دریافت مقاله: ۹۰/۹/۳۰، پذیرش: ۹۱/۳/۱)

مقدمه

می‌دهد، ولی این اطلاعات فقط درباره وضعیت کنونی زبان کودک بوده و فاقد هرگونه قابلیت پیش‌بینی درباره چگونگی وضعیت زبان کودک در آینده است (۲). Pena (۲۰۰۶) نیز خاطر نشان می‌کند که برای ارزیابی زبان، به‌طور سنتی، از آزمون‌هایی استفاده می‌شود که توانایی زبانی کودک را فقط در سطح کلمه یا جمله بررسی می‌کنند. این نوع آزمون‌ها نمی‌توانند اختلال واقعی کودک را نشان دهند، زیرا زبان کودک در بافت طبیعی بررسی نمی‌شود (۳). در سال‌های اخیر برای پرهیز از این مشکلات، متخصصان به دنبال

زبان نظام پیچیده‌ای است و برای انجام یک ارزیابی جامع و قابل اعتماد از آن، لازم است نمونه گفتاری کامل و جامعی از آن تهیه شود. این نمونه باید به گونه‌ای باشد که بتوان از طریق آن به بود یا نبود مشکل زبانی پی برد، حتی وضعیت زبانی کودک در آینده را پیش‌بینی کرد (۱). اکثر مطالعاتی که درباره زبان انجام گرفته، بر پایه مشاهده گفتار خودانگیزه کودک در موقعیت‌های طبیعی بوده است. Dockrett (۲۰۰۳) خاطر نشان می‌کند که این نوع بررسی اطلاعات مهمی را در اختیار درمانگران و محققان قرار

(۲۰۰۳) ابزار هنجار داستان‌گویی ادمونتون، (Edmonton Narrative Norm Instrument: ENNI) را بر طبق مدل مؤلفه‌های دستور زبان قصه طراحی کردند (۱۱). Gilman و Pearson (۲۰۰۴) آزمون زبان داستانی (Test of Narrative Language: TNL) را برای ارزیابی درک و بیان مهارت داستان‌گویی طراحی کردند (۱۲). مطالعات انجام شده بیانگر این است که ارزیابی زبان از طریق داستان‌گویی یک روش روا، پایا، حساس و با خطای کمتر نسبت به سایر آزمون‌های هنجار و غیرهنجار برای ارزیابی زبان است. در بررسی آزمون‌های فارسی مشخص شد که تا زمان شروع این مطالعه، هیچ‌گونه تکلیف یا آزمونی به شکل داستانی برای بررسی عناصر زبانی کودکان در ایران تهیه نشده است. بر این اساس، پژوهش حاضر با هدف ساخت آزمونی به عنوان ابزار اندازه‌گیری ساختار زبان در کودکان شش تا هفت ساله انجام شد.

روش بررسی

برای ساخت آزمون بازگویی داستان ابتدا پژوهش‌ها و آزمون‌های مربوط و مشابه در دسترس، مورد بررسی قرار گرفت. در این میان، دو آزمون BST و NAP که آزمون‌های کامل و جامعی هستند و در اکثر پژوهش‌ها با هم استفاده شده‌اند به عنوان مدل انتخاب شد. ابتدا در چند مرحله با کمک زبان‌شناس و گفتاردرمانگر (آسیب‌شناس گفتار و زبان)، داستانی که در برگیرنده اهداف این پژوهش بود طراحی شد. داستان طراحی شده ویژگی‌های زیر را دارد:

- ۱- دارای ساخت هر می است، یعنی نقطه شروع، اوج و پایان دارد؛
- ۲- دارای طرحی ساده است، یعنی دارای موضوعات پیچیده نیست؛
- ۳- شامل کلماتی است که در روند رشد زودتر یاد گرفته می‌شوند و مطابق تنوع اجتماعی-اقتصادی و فرهنگی هستند. همچنین کلمات از ساختار فونولوژیک ساده برخوردارند، بنابراین احتمال شکست در آزمون به خاطر ناآشنا بودن کلمات یا عدم توانایی تلفظ آنها، به کمترین حد می‌رسید؛
- ۴- در انتخاب کلمات داستان از واژگان این سن استفاده شد؛
- ۵- دارای کلمات

روش‌هایی بوده‌اند که از یک طرف زبان کودک را در بافت طبیعی مورد ارزیابی قرار دهد و از طرفی اجرای آنها آسان‌تر باشد و بتوان با آنها اختلالات گفتار و زبان را به سرعت شناسایی کرد. آزمون بازگویی داستان این هدف را برآورده می‌کند، زیرا شکل، فرم و کاربرد زبان را با هم مورد بررسی قرار می‌دهد و اطلاعات بیشتری نسبت به سایر آزمون‌های زبانی، در اختیار متخصصان قرار می‌دهد (۴). از طرفی، این شیوه، رویکردی متنوع و انعطاف‌پذیر است و می‌تواند برای ارزیابی انواع نقایص زبانی به کار رود. همچنین، هنگامی که محدودیت زمانی وجود دارد این روش راهکار مفیدی برای درمانگر است تا در مدت‌زمان کوتاه توانایی کودک را ارزیابی کند و حتی در مواردی می‌تواند شرایط زبانی و مهارت‌های خواندن و نوشتن، مهارت اجتماعی و ارتباطی کودک در سال‌های آینده را نیز پیش‌بینی کند. از طرفی، این تکلیف، سازماندهی شناختی را در ذهن فعال می‌کند که این امر با مشاهدات و آزمون‌های هنجار قابل دسترس نیست. از فواید دیگر استفاده از این تکلیف، سادگی آن برای اجرا و تجزیه و تحلیل نتایج نسبت به سایر روش‌ها است (۶۵). مهارت داستان‌گویی، در خلال طیفی وسیع شامل داستان‌گویی شخصی، خیالی، توصیفی، توضیحی و بازگویی داستان، که هر کدام معایب و مزایای خاص خود را دارند، مورد ارزیابی قرار می‌گیرد. برای ارزیابی ساختارهای زبان کودکان پیش‌دبستانی، در مطالعات گوناگون بیشتر از تکلیف بازگویی داستان استفاده می‌شود (۷). Renfrew (۱۹۷۹) آزمون داستان اتوبوس (Bus story test: BST) را برای غربالگری و تشخیص اختلالات زبانی طراحی کرد. در این آزمون از تکلیف بازگویی داستان استفاده شده است (۸). Mercer Meyer (۲۰۰۰) آزمون قورباغه کجایی؟ (Frog where are you test) را برای ارزیابی توانایی زبانی کودکان طراحی کرد. این آزمون شامل یک مجموعه ۲۴ تایی از تصاویر متوالی است که بیشتر در مطالعات بررسی داستان‌گویی خیالی استفاده می‌شود (۹). Bliss و همکاران (۱۹۹۸) نیمرخ ارزیابی داستان‌گویی (Narrative assessment profile: NAP) را به عنوان ابزاری برای ارزیابی داستان‌گویی در کودکان سن دبستان طراحی کردند (۱۰). Schneider و همکاران

محتوایی (اسم، فعل، صفت) و کلمات عملکردی (حروف اضافه، ربط، شناسه) است؛ ۶- برخی جملات ساختار پایه و پیرو دارند؛ ۷- متناسب با سن کودک طراحی شده است؛ ۸- دارای تنوع در ارجاعات و حروف ربط است؛ ۹- از روابط منطقی و ترتیب زمانی برخوردار است؛ ۱۰- انسجام و پیوستگی متن رعایت شده است.

در مجموع، داستان دارای یک موضوع اصلی، ده جمله پیچیده، هشت نوع حرف ربط، هیجده اطلاعات اصلی و ده اتفاق متوالی است.

پس از نوشتن داستان اصلی فیل بازیگوش و داستان آزمایشی خرگوش شکمو بر اساس معیارهای بالا که هر دو دارای ساخت مشابه بودند، داستان برای بررسی روایی صورتی در اختیار دو داستان‌نویس کودک قرار گرفت. پس از تأیید روایی صورتی (facial validity)، برای بررسی روایی محتوایی نیز از ۱۴ نفر از اساتید گفتاردرمانگر (آسیب‌شناس گفتار و زبان) و زبان‌شناس نظرخواهی شد.

ابتدا روایی داستان با نسبت روایی محتوایی (Content Validity Ratio: CVR) که توسط Lawshe (۱۹۷۵) مطرح شده است مورد بررسی قرار گرفت. میزان روایی محتوایی بین مثبت یک و منفی یک (۱ و -۱) بود که هر چه میزان به دست آمده به یک نزدیک‌تر باشد، نشان می‌دهد که آزمون از روایی بالاتری برخوردار است (۱۳). برای انجام روایی، ابتدا داستان‌ها بدون تصویر در اختیار متخصصان قرار داده شد و نظرات آنها به صورت امتیاز ضروری است دو، مفید است ولی ضروری نیست یک و ضروری نیست صفر بررسی شد. میزان CVR به دست آمده برای داستان فیل بازیگوش ۸۵ درصد است که روایی بسیار بالایی را نشان می‌دهد. سپس شاخص روایی محتوایی (Content Validity Index: CVI) که در بررسی روایی ضروری است نیز محاسبه شد. CVI داستان فیل بازیگوش نیز برای آزمون ۸۹ در صد است که روایی بسیار مطلوبی است و مناسب تشخیص داده می‌شود.

با بررسی آزمون‌های گوناگون، یک بخش تصویری نیز برای آزمون در نظر گرفته شد، که بر اساس بخش تصویری آزمون

ENNI طراحی شد. شش تصویر برای داستان آزمایشی و ۱۰ تصویر برای داستان اصلی توسط یک کارتون‌ساز حرفه‌ای تهیه شد و سپس با استفاده از نرم افزار Adobe photoshop 7.0ME توسط یک گرافیست رنگ‌آمیزی و ویرایش شد. تصاویر تا حد امکان ساده بوده و مفهوم اصلی داستان را می‌رساند. تصاویر از هرگونه عامل حواس‌پرت‌کن خالی بوده و تعداد آنها متناسب با تعداد وقایع داستان است. تمام تصاویر از نظر نوع طراحی و اندازه یکسان هستند. به منظور تدوین نهایی داستان، با یک مدیر دوبلاژ مشورت شد و دو داستان‌گو (زن) معرفی شدند. برای ضبط صدا، از آنها خواسته شد که داستان را مطابق با معیارهای آزمون بیان کنند (اغراق‌آمیز نباشد، روی کلمه و جمله خاصی تأکید نشود، و سرعت گفتار مناسب باشد). همچنین، در صورت عدم توافق بین گوینده و محقق، داستان تکرار می‌شد.

برای ضبط اصلی داستان، صدای گویندگان در یک اتاق اکوستیک و با نرم افزار Jet Audio 2007 با فرمت ۴۴۱۰۰ کیلوهرتز، ۱۶ bit mono، و با استفاده از لپ‌تاپ Vaio, SR390 ساخت آمریکا، ضبط شد. سپس با استفاده از همین نرم‌افزار صداها ویرایش شد و نویز اضافی آنها حذف و شدت بلندی تنظیم شد. پس از آن صدای هر دو گوینده در اختیار پنج نفر از اساتید آسیب‌شناس گفتار و زبان و زبان‌شناس قرار گرفت و موردی که روایی بالاتری داشت انتخاب شد. در ادامه، دو داستان همراه با تصاویر آنها با استفاده از نرم‌افزار Movie maker تنظیم شدند. پس از تنظیم آزمون، داستان همراه با تصاویر توسط پنج متخصص آسیب‌شناس گفتار و زبان و زبان‌شناس نظرسنجی شد. روایی داستان همراه با تصاویر ۱۰۰ درصد بدست آمد که با توجه به مقیاس Lawshe (۱۹۷۵) بسیار مطلوب است. در نهایت، اشکالات موجود در تصاویر و داستان تصحیح و آزمون برای اجرا آماده شد.

پایایی آزمون با استفاده از دو روش تکرارپذیری (نسبی و مطلق) و آلفای کرونباخ (ثبات درونی) مورد بررسی قرار گرفت. برای بررسی پایایی آزمون به روش تکرارپذیری، ۷۲ کودک، ۳۶ دختر و ۳۶ پسر، در گروه سنی شش تا ۷هفت سال که به روش تصادفی از مهدکودک‌های مناطق متوسط شهر تهران، پنج، هفت

و نه انتخاب شدند در این آزمون شرکت کردند. این آزمون شامل دو زیر آزمون ساختار بزرگ و ساختار کوچک است که در ادامه به چگونگی اجرای آنها، پاسخ کودک، آماده‌سازی کودک و امتیازدهی آزمون می‌پردازیم.

در مرحله اجرای آزمون، ابتدا با استفاده از اطلاعات موجود در پرونده بهداشتی کودکان و دریافت اطلاعات از معلمان، اطلاعات مربوط به انتخاب نمونه به دست آمد. سپس، فرم رضایت‌نامه برای امضا در اختیار آنها قرار گرفت. پس از کسب اجازه، آزمون مورد نظر برای هر کودک به طور جداگانه و در اتاقی ساکت انجام شد. در ابتدا، کودک برای اجرای آزمون آماده شد که در آن، آزمونگر با کودک رابطه دوستانه برقرار می‌کرد و نحوه اجرای آزمون را به طور کامل به وی شرح می‌داد. برای آشنایی وی با روش کار، در ابتدا داستان آزمایشی که از لحاظ ساخت کاملاً مطابق با داستان اصلی است به کودک ارائه می‌شد.

دستورالعمل اجرای آزمون بدین‌گونه است که آزمونگر در کنار کودک می‌نشیند و به کودک می‌گوید «یک داستان همرا با تصاویر آن برای تو پخش می‌شود، تو خوب گوش می‌دهی، وقتی که تمام شد تو باید داستان را برای من تعریف کنی.» سپس، تصاویر به کودک نشان داده می‌شود و به وی اجازه داده می‌شود که برای چند لحظه خوب به تصاویر نگاه کند. بعد از اظهار آمادگی کودک، داستان ضبط شده از طریق گوشی و با استفاده از لپ‌تاپ برای کودک پخش شده و در حین تعریف، تصاویر مربوط به هر واقعه نشان داده می‌شود. داستان تنها یک‌بار پخش می‌شود. بعد از پخش داستان، آزمونگر به کودک می‌گوید: «حالا نوبت توست که داستان را بگویی.» بعد از اعلام آمادگی کودک، فرصت کافی برای بازگویی داستان به وی داده می‌شود (در این مرحله، زمان خاصی در نظر گرفته نمی‌شود. سپس یک میانگین از آنها به عنوان مدت زمان گرفته شد). به جز در لحظه شروع، آزمونگر مجاز به پرسیدن یا راهنمایی در حین تعریف داستان نیست. تشویق نیز تنها در پایان تکلیف انجام می‌شود. بعد از اطمینان از اینکه کودک روش کار را فهمیده است، همین مراحل برای داستان اصلی انجام می‌شد. صدای کودک با استفاده از میکروفن و با نرم‌افزار Jet

در این مطالعه روش تجزیه و تحلیل داستان ملاک-مرجع (criterion-referenced test) است. بنابراین، در ابتدا با توجه به متغیرهای آزمون و بر اساس آزمون BST برای هر بخش یک ملاک مشخص در نظر گرفته شد و در اختیار دو ارزیاب قرار گرفت. برای بررسی پایایی آن، داستان‌های بازگو شده توسط ۲۵ کودک که به‌طور تصادفی انتخاب شدند، در اختیار دو ارزیاب قرار گرفت و روش‌های بررسی پایایی بر آنها اعمال شد.

امتیازدهی این آزمون به شرح زیر بود:
امتیاز زیرآزمون ساختار بزرگ، شامل حفظ موضوع (۵ امتیاز)، اطلاعات اصلی (۱۸ امتیاز) و توالی (۱۰ امتیاز)، در مجموع ۳۳ بود.

امتیاز زیرآزمون ساختار کوچک، شامل ارجاعات (۱۰ امتیاز)، حروف ربط (۸ امتیاز)، جملات پیچیده (۱۰ امتیاز) و میانگین طول گفته (۱۵ امتیاز)، در مجموع ۴۳ بود.

در پایان تمام امتیازهای به‌دست آمده به رتبه‌های صفر تا چهار تبدیل و به‌صورت درصد محاسبه می‌شوند.

پایایی آزمون با بررسی ضریب همبستگی بین میانگین‌های آزمون-بازآزمون (Interclass Correlation Coefficient: ICC، خطای معیار اندازه‌گیری Standard Error Measurement: SEM)، ثبات درونی (آلفای کرونباخ) و ضریب همبستگی پیرسون مورد بررسی قرار گرفت. پایایی مقیاس امتیازدهی با بررسی ضریب همبستگی بین میانگین بین ارزیاب‌ها، SEM و ضریب همبستگی مورد بررسی قرار گرفت. در پژوهش حاضر برای بیان درجه تکرارپذیری نسبی از تعریف Munro (۲۰۰۵) در تفسیر ضریب همبستگی استفاده شد. ضریب همبستگی صفر تا ۲۵ درصد به عنوان ارتباط اندک، ۲۶ درصد تا ۴۹ درصد ارتباط ضعیف، ۵۰ درصد تا ۶۹ درصد ارتباط متوسط،

جدول ۱- شاخص‌های آماری امتیاز بخش‌ها و زیرآزمون‌ها در آزمون-بازآزمون (n=۷۲)

بخش	آزمون	
	میانگین (انحراف معیار)	بازآزمون
حفظ موضوع	۴/۹۱ (۰/۲۷)	۴/۹۱ (۰/۲۷)
توالی	۷/۹۳ (۱/۰۵)	۸/۱۳ (۱/۱۰)
اطلاعات اصلی	۱۲/۷۹ (۱/۷۵)	۱۳/۴۴ (۱/۷۶)
ساختار بزرگ	۲۵/۶۳ (۲/۵۶)	۲۶/۵۰ (۲/۵۵)
حروف ربط	۳/۹۸ (۱/۲۵)	۴/۱۶ (۱/۵۰)
ارجاعات	۵/۴۱ (۱/۶۵)	۶/۰۴ (۱/۶۰)
پیچیدگی نحوی	۳/۸۶ (۱/۸۸)	۴/۱۹ (۲/۱۴)
طول جمله	۱۰/۰۱ (۱/۷۸)	۹/۸۹ (۱/۶۷)
ساختار کوچک	۲۳/۲۸ (۵/۰۳)	۲۴/۳۰ (۵/۳۱)
کل آزمون	۴۸/۹۲ (۶/۷۶)	۵۰/۸۱ (۶/۹۷)

بازگویی داستان ۰/۷۵ و آلفای کرونباخ نیز ۰/۷۷ به دست آمد. جدول ۲ نتایج پایایی را نشان می‌دهد. همچنین تفاوت میانگین امتیازهای آزمون-بازآزمون با روش آماری t زوجی مقایسه شد. میانگین پاسخ‌های کل آزمون در دو بار انجام ۰/۱۱ بود که این تفاوت معنی‌دار نبود ($p > ۰/۰۵$). در تحقیق حاضر، پایایی امتیازهای حاصل از هر یک از دو ارزیاب برای هر یک از بخش‌ها برای ۲۵ نمونه، از جنبه تکرارپذیری در دفعات آزمون و تکرارپذیری مطلق مورد بررسی قرار گرفت. براساس نتایج بررسی، مقدار ICC بین ارزیاب‌ها در هر دو بخش اصلی ۰/۹۷ و مقدار کل آن برای کل آزمون در دو بار اندازه‌گیری ۰/۹۳ بود که نشان‌دهنده همبستگی و پایایی بسیار بالا بود (جدول ۳).

از طرفی مدت زمان یک آزمون در طی مراحل ساخت حتماً باید در نظر گرفته شود. نتایج این بررسی نشان داد که اجرای این آزمون با احتساب دستورالعمل، مواد آزمون، فرصت پاسخ‌دهی آزمودنی‌ها و امتیازدهی حداکثر ۲۵ دقیقه طول می‌کشد که از

هفتاد درصد تا ۸۹ درصد ارتباط بالا و ۹۰ درصد تا ۱۰۰ درصد ارتباط بسیار بالا در نظر گرفته شد (۱۴). در بررسی روایی نیز از نسبت روایی محتوایی و شاخص روایی محتوایی و روایی صورتی استفاده شد.

یافته‌ها

با توجه به آنکه آزمون از دو بخش اصلی ساختار بزرگ و کوچک تشکیل شده است، امتیاز کل آزمون برای هر فرد می‌تواند از مجموع امتیازهای بخش‌های اصلی به دست آید. جدول ۱ میانگین و انحراف معیار امتیاز بخش‌ها و زیرآزمون‌های حاصل از آزمون و بازآزمون را نشان می‌دهد. نتایج نشان می‌دهد که میانگین امتیازها در دو بار انجام آزمون، ضریب همبستگی بسیار بالایی دارد که از لحاظ آماری معنی‌دار بود ($p = ۰/۰۰۰$). مقدار ضریب همبستگی حاصل از میانگین‌های آزمون و بازآزمون ۰/۸۳ و SEM، برای آزمون حاضر ۲/۷۶ و ضریب همبستگی کل آزمون

جدول ۲- نتایج بررسی پایایی آزمون بازگویی داستان (n=۷۲)

بخش‌ها	ضریب همبستگی پیرسون	همبستگی بین میانگین‌های آزمون بازآزمون	خطای معیار اندازه‌گیری	آلفای کرونباخ	p
اطلاعات اصلی	۴۲/۴	۰/۵۹	۱/۱۷	-	۰/۰۰۰
حفظ موضوع	۱	۱	۰/۲۷	-	۰/۰۰۰
توالی وقایع	۰/۷۹	۰/۸۷	۰/۳۷	-	۰/۰۰۰
ساختار بزرگ	۶۳/۷	۰/۷۴	۱/۳۰	۰/۶۸	۰/۰۰۰
حروف ربط	۰/۴۹	۰/۶۵	۰/۷۳	-	۰/۰۰۰
ارجاعات	۰/۴۹	۰/۶۴	۰/۹۹	-	۰/۰۰۰
پیچیدگی نحوی	۰/۵۲	۰/۸۰	۰/۸۴	-	۰/۰۰۰
طول جمله	۰/۵۹	۰/۷۴	۰/۹۰	-	۰/۰۰۰
ساختار کوچک	۷۲/۲	۰/۸۲	۲/۱۳	۰/۹۵	۰/۰۱
کل آزمون	۷۴/۵	۰/۸۳	۲/۷۶	۰/۷۷	۰/۰۰۰

میانگین مدت زمان اجرای ۷۲ نفر آزمودنی به دست آمد (جدول ۴).

بحث

اکثر آزمون‌های موجود در زبان فارسی توانایی زبانی کودک را در سطح کلمه و جمله مورد بررسی قرار می‌دهند. این روش‌های اندازه‌گیری توانش زبانی واقعی کودک را نشان نمی‌دهند، زیرا توانایی زبانی کودکان در بافت زبانی، بسیار متنوع و متفاوت است. استفاده از آزمون‌هایی همانند داستان‌گویی می‌تواند ویژگی‌های زبانی را دقیق‌تر ارزیابی کرده و اطلاعات جامع‌تری در اختیار متخصصان بالینی قرار دهد.

آزمون فیل بازیگوش با هدف ارزیابی زبان کودک در یک بافت زبانی ساخته شده است و هدف از ساخت این آزمون فراهم کردن ابزاری کاربردی برای کسب اطلاعات بیشتر درباره‌ی زبان کودکان است. یکی از آزمون‌های غربالگری که از مهارت بازگویی داستان استفاده کرده است، آزمون BST (۱۹۹۴) Renfrew است

که از رایج‌ترین و پرکاربردترین ابزارهای غربالگری برای بررسی ساختارهای زبان است که از دیر باز در اکثر مطالعات استفاده شده و برای کودکان سه تا هشت سال هنجار شده است. ساختار داستان طوری تنظیم شده است که قادر به ارزیابی سریع برخی از عناصر ساختار بزرگ (اطلاعات) و کوچک زبان (پیچیدگی و طول جمله) باشد. این آزمون با استفاده از روش آزمون-بازآزمون امتیاز متوسط و در بررسی پایایی مقیاس امتیازدهی امتیاز بسیار بالایی کسب کرده است (۸).

این آزمون و آزمون حاضر هر دو با هدف استفاده از تکلیف بازگویی داستان برای ارزیابی، غربالگری و حتی پیش‌بینی برخی از اختلالات در کودکان پیش‌دبستان طراحی شده‌اند. BST تنها دو جنبه از عناصر ساختار کوچک زبان (پیچیدگی و طول جمله) و یک جنبه از عناصر ساختار بزرگ (اطلاعات) را مورد بررسی قرار می‌دهد، در حالی که آزمون فیل بازیگوش نیم‌رخ از تمام ابعاد ساختار بزرگ و کوچک زبان را به طور همزمان در اختیار درمانگران قرار می‌دهد. در BST نیز همانند آزمون فیل بازیگوش

جدول ۳- نتایج بررسی پایایی بین ارزیاب آزمون بازگویی داستان (تعداد ۲۵ نمونه)

بخش‌ها	ضریب همبستگی	همبستگی بین میانگین‌های ارزیاب اول و دوم	خطای معیار اندازه‌گیری	p
حفظ موضوع	۰/۹۵	۱	۰	۰/۰۰۰
توالی وقایع	۰/۹۵	۰/۹۷	۰/۲۴	۰/۰۰۰
اطلاعات اصلی	۰/۹۴	۰/۹۶	۰/۴	۰/۰۰۰
ساختار بزرگ	۰/۹۶	۰/۹۷	۰/۵۴	۰/۰۰۰
حروف ربط	۰/۵۷	۰/۷۴	۰/۶۵	۰/۰۰۰
ارجاعات	۰/۸۹	۰/۹۳	۰/۳۵	۰/۰۰۰
پیچیدگی نحوی	۰/۸۹	۰/۸۰	۰/۷۰	۰/۰۰۰
طول جمله	۰/۷۳	۰/۸۴	۰/۸۹	۰/۰۰۰
ساختار کوچک	۰/۸۳	۰/۸۹	۰/۴۷	۰/۰۰۰
کل آزمون	۰/۹۰	۰/۹۳	۱/۷۰	۰/۰۰۰

به‌دست آمده از پایایی بین ارزیاب ۸۹ درصد و پایایی به‌دست آمده از بخش‌های مشابه BST برای بخش اطلاعات ۷۸ درصد، برای بخش پیچیدگی نحوی ۹۴ درصد و برای بخش طول جمله ۸۴ درصد است (۸). آزمون فیل بازیگوش در مقایسه با BST با وجود یکسان بودن روش‌های پایایی و اجرا، آزمونی جامع‌تر و کامل‌تر برای بررسی ابعاد زبان است، زیرا عناصر ساختار کوچک بیشتری را بررسی کرده، تکرارپذیری نسبی بالاتری دارد و در خلال این تکلیف به بررسی عناصر ساختار بزرگ زبان نیز می‌پردازد.

از دیگر آزمون‌هایی که قبلاً مطرح شد NAP است که توسط Bliss و همکاران (۱۹۹۸) برای تهیه نیمرخ زبانی کودکان و ارائه راهکارهای درمانی طراحی شده است. این آزمون مواردی از قبیل حفظ موضوع، توالی وقایع، اطلاعات، ارجاع، حروف ربط و روانی کلام را مورد بررسی قرار می‌دهد. امتیازدهی در مقیاس صفر تا چهار رتبه‌بندی شده و در بررسی پایایی آن از روش آزمون-بازآزمون استفاده شده است که نتایج بیانگر این است که از پایایی بسیار بالایی برخوردار است (۱۰). این آزمون مانند آزمون

از آزمون-بازآزمون برای بررسی پایایی استفاده شده است که ۷۹ درصد برای بخش اطلاعات، ۷۲ درصد برای بخش طول گفته و ۵۸ درصد برای بخش پیچیدگی جمله به دست آمده است که پایایی بالایی است (۸). در پژوهش حاضر نتیجه تکرارپذیری به‌دست آمده از آزمون-بازآزمون ۸۳ درصد است که تکرارپذیری بالایی را نشان می‌دهد و عدد حاصل از بخش‌های مشابه BST، ۵۹/۶ درصد برای بخش اطلاعات، ۷۴ درصد برای بخش طول جمله، ۸۰ درصد برای بخش پیچیدگی نحوی به‌دست آمد. همچنین، مدت زمان کل BST ۲۵ دقیقه است که با مدت زمان آزمون حاضر مطابقت دارد.

یکی دیگر از روش‌هایی که برای بررسی پایایی این آزمون استفاده شده پایایی بین ارزیاب است که برای مقیاس امتیازدهی در نظر گرفته شده است. این آزمون با این روش پایایی بالایی را کسب کرده است، یعنی ۹۲ درصد برای بخش اطلاعات، ۸۳ درصد برای بخش طول جمله، و ۶۰ درصد برای بخش پیچیدگی جمله (۸). در پژوهش حاضر نتیجه تکرارپذیری نسبی

جدول ۴- شاخص‌های مرکزی و پراکندگی مربوط به کل زمان آزمون

میانگین (انحراف معیار)	حداکثر	حداقل	نما
زمان کل آزمون	۲۳/۴۸ (۱/۲۱)	۲۵/۶۳	۲۵

ممکن است دانش زبانی اصلی کودک را منعکس نکند، ولی برای یک غربالگری سریع از زبان حائز اهمیت است (۱۵). در بررسی پایایی آزمون NAP نیز از آزمون-بازآزمون استفاده شده است و میزان آن ۸۵ درصد است. در پژوهش حاضر نیز با استفاده از این روش پایایی کل آزمون ۸۳ درصد بود (۱۰).

آزمون دیگر TNL است که Pearson و Gilman (۲۰۰۴) در ارزیابی درک و بیان مهارت داستان‌گویی استفاده کرده‌اند و برای گروه سنی ۵ تا ۱۱ سال هنجار شده است. این آزمون با استفاده از روش آزمون-بازآزمون دارای پایایی بسیار بالایی است. در این آزمون ابزارهای ارزیابی به سه شیوه بدون تصویر، با تصاویر متوالی و یا با یک تصویر ارائه می‌شوند. مدت زمان آزمون ۱۵ تا ۲۰ دقیقه است که با مدت‌زمان آزمون حاضر تقریباً مطابقت دارد. زمان آزمون حاضر کمی بیشتر است که به نظر می‌رسد علت آن دارا بودن داستان آزمایشی باشد. مقیاس امتیازدهی نیز از پایایی بسیار بالایی برخوردار است (۱۲). یکی از تکالیف آن آزمون، بازگویی داستان است. در بررسی پایایی این آزمون، از روش آزمون-بازآزمون و بین ارزیاب استفاده شده است، که پایایی حاصل از کل آزمون ۸۱ درصد و برای بین ارزیاب ۹۶ درصد است که پایایی بالایی است (۱۲). در پژوهش حاضر نیز نتیجه تکرارپذیری نسبی و ضریب همبستگی تکرارپذیری بسیار بالایی معادل ۸۳ درصد را نشان می‌دهد. به رغم پایایی کمتر، آزمون فیل بازیگوش از لحاظ ساختارهای مورد بررسی و قابلیت تشخیص بر آزمون TNL ارجح است (۱۳).

یکی دیگر از آزمون‌هایی که توانایی زبانی را در یک تکلیف داستانی فرمول‌بندی شده ارزیابی می‌کند ENNI است که توسط Schnider و همکاران (۲۰۰۳) مطرح شده است. این آزمون توانایی داستان‌گویی را در یک تکلیف زبانی ساختاریافته ارزیابی می‌کند. پایایی آزمون با استفاده از روش آزمون-بازآزمون امتیاز بالایی را کسب کرده است و مقیاس امتیازدهی آن نیز از پایایی بسیار بالایی برخوردار است (۱۴). این ابزار در اکثر مطالعات استفاده شده و سیستم امتیازدهی بسیار پایایی دارد. در بررسی پایایی آن، از پایایی بین ارزیاب استفاده شده است و در آن پایایی

فیل بازیگوش با استفاده از برخی عناصر ساختار بزرگ و کوچک، زبان را ارزیابی می‌کند، ولی برخلاف آزمون حاضر از تکلیف داستان‌گویی شخصی استفاده می‌کند. Bliss و MacCabe (۲۰۰۳) معتقدند که تکلیف داستان‌گویی شخصی الگوی کاملی از نحوه طرح‌ریزی، توالی و سازماندهی یک پیام برای ارزیابی فراهم می‌کند و قادر به سنجش دانش زبانی اصلی کودک است، زیرا در آن برخلاف بازگویی داستان، الگویی به کودک ارائه نمی‌شود، ولی با این حال نسبت به بازگویی داستان دارای معایی است. تکلیف داستان‌گویی شخصی بار شناختی زیادی را درگیر می‌کند و برای سنین پیش‌دبستان پیچیده است و برای برخی از کودکان با نقایص زبانی مناسب نیست و بسیار وقت‌گیر است. در مقابل، بازگویی داستان بار شناختی کمی را درگیر می‌کند و برای ارزیابی ارجاعات و حروف ربط، به دلیل داشتن چهارچوب مشخص، مناسب‌تر است. از طرفی، به دلیل ثبات تکلیف از امتیازدهی آسان‌تر و با ثبات‌تری برخوردار است و زمان زیادی برای اجرا و تجزیه و تحلیل نیاز ندارد. با این وجود، بازگویی داستان نیز معایی دارد و برای بررسی برخی عناصر ساختار بزرگ از قبیل حفظ موضوع و توالی وقایع تکلیف مناسبی نیست، زیرا در واقع این دو مورد در قالب داستان از قبل به کودک ارائه می‌شود (۱۵).

در بررسی پایایی عناصر ساختار بزرگ آزمون فیل بازیگوش، برای بخش حفظ موضوع ۹۰ درصد، برای بخش اطلاعات ۵۹ درصد و برای بخش توالی ۸۷ درصد به دست آمده است که بیانگر این است که این آزمون برای بررسی عناصر ساختار بزرگ نیز از پایایی بالایی برخوردار است، ولی مطالعات موجود نشان می‌دهند که با وجود این مقدار پایایی، این تکلیف

لوازم اصلی هر آزمون جدیدی است. از روش‌های مقبول اخیر در این زمینه، محاسبه CVI است. عدد به‌دست آمده برای روایی معادل ۸۹ درصد است که نشان می‌دهد آزمون از روایی قابل قبولی برخوردار است. با این همه، برای تعیین اینکه آیا آزمون حاضر قادر به تمایزگذاری بین گروه هنجار و گروه مبتلا به اختلال است، می‌توان در پژوهش‌های بعدی این آزمون را در گروه‌های هنجار و دچار انواع اختلالات مورد بررسی قرار داده و نتایج به‌دست آمده را با گروه هنجار مقایسه کرد (۱۴).

نتیجه‌گیری

نتایج نشان می‌دهد که این ابزار از نظر محتوا و تناسب با زبان فارسی، روایی لازم برای بررسی کودکان شش تا هفت ساله را دارد و در آن، سطح عملکرد کودکان در دو مرتبه به میزان بالایی یکسان بوده یا به عبارت دیگر آزمون از قابلیت تکرارپذیری بالایی برخوردار است. همچنین ضریب همبستگی آزمون در دو بار تکرار از لحاظ آماری معنی‌دار است. این آزمون دارای یک سیستم امتیازدهی ملاک-مرجع است که از پایایی بسیار مطلوبی برخوردار است و همچنین با این سیستم امتیازدهی آسان، روا و پایا، قادر به ارزیابی جامع، کامل و سریع از زبان است. تمامی کودکان، این آزمون را در حداکثر ۲۵ دقیقه انجام دادند.

سپاسگزاری

نویسندگان این مقاله، مراتب تشکر و سپاس خود را از مسئولان سازمان بهزیستی تهران، به ویژه جناب آقای خلیل‌زاده، مدیران مهدکودک‌هایی که در این پژوهش همکاری کردند و اساتید ارجمند Dr. Lynns Bliss از دانشگاه هاستون، Dr. Phyliss و Alissa Langel از دانشگاه بوفالو آمریکا و Shnider از دانشگاه آلبرتا که در انجام این تحقیق از هیچ کمکی دریغ نکردند و همواره راهنمایی‌های مفید ارائه کردند و کلیه اساتید و عزیزانی که در روند انجام این آزمون شرکت داشتند، اعلام می‌دارند.

بالایی معادل ۸۵ درصد کسب کرده است. در پژوهش حاضر نیز نتیجه تکرارپذیری نسبی به‌دست آمده ۸۹ درصد است که پایایی بسیار بالایی را نشان می‌دهد (۱۱).

یکی دیگر از روش‌های پایایی که در آزمون فیل بازیگوش استفاده شده است، بررسی تکرارپذیری مطلق با روش SEM است. SEM به‌دست آمده از آزمون میزان تفاوت امتیاز کودک از امتیاز واقعی را نشان می‌دهد. هنگامی که بخواهیم آزمون را برای یک فرد دو بار تکرار کنیم، SEM یک محدوده یا درجه اطمینان فراهم می‌کند. بنابراین هرچه میزان SEM کمتر باشد، نشان دهنده آن است که آزمون از تکرارپذیری مطلق بیشتری برخوردار است. میزان SEM در دو بار تکرار آزمون فیل بازیگوش ۲/۷۶ بوده است. با توجه به این که انحراف معیار کل آزمون ۶/۷۶ بوده و در میزان SEM مؤثر است، می‌توان نتیجه گرفت که SEM نسبت به میزان انحراف معیار در محدوده مطلوبی است. از این روش فقط در آزمون TNL استفاده شده است و متوسط SEM آن یک است که نسبت به پژوهش حاضر از تکرارپذیری مطلق بالاتری برخوردار است، و همراه شدن با دو نوع تکلیف دیگر ممکن است توجهی برای افزایش دقت این آزمون باشد (۱۲).

در روند استفاده از یک آزمون برای تشخیص اختلال، بررسی تکرارپذیری آزمون برای تعیین اعتبار، به‌ویژه زمانی که از آن برای تعیین راهکارهای درمان بیماران استفاده می‌شود، بسیار مهم است. هرچه مقدار تکرارپذیری نسبی به عدد یک نزدیک‌تر باشد، نشان دهنده تکرارپذیری بالاتر آزمون است و بیانگر آن است که میانگین دو بار انجام آزمون تفاوت معنی‌داری با یکدیگر ندارد. همچنین، هر چه مقدار SEM کمتر باشد، نشان‌دهنده آن است که متغیر از تکرارپذیری مطلق بیشتری برخوردار است. به‌طور کلی، با توجه به نتایج به‌دست آمده از انواع روش‌های پایایی می‌توان بیان کرد که آزمون بازگویی داستان از تکرارپذیری نسبی ۸۳ درصد و تکرارپذیری مطلق ۲/۷۶ و ثبات درونی ۸۵ درصد و ضریب همبستگی ۷۸ درصد برخوردار است. این امر بیانگر این است که آزمون از تکرارپذیری نسبی، مطلق، ثبات درونی و همبستگی بالایی برخوردار است. تعیین روایی محتوایی یکی از

REFERENCES

1. Shipley KG, McAfee JG. Assessment in speech-language pathology. 4th ed. USA: Delmar Cengage Learning; 2009.
2. Clark EV. First language acquisition. 1st ed. Cambridge: Cambridge University Press; 2003.
3. Peña ED, Gillam RB, Malek M, Ruiz-Felter R, Resendiz M, Fiestas C, et al. Dynamic assessment of school-age children's narrative ability: an experimental investigation of classification accuracy. *J Speech Lang Hear Res.* 2006;49(5):1037-57.
4. McCabe A, Bliss L, Barra G, Bennett M. Comparison of personal versus fictional narratives of children with language impairment. *Am J Speech Lang Pathol.* 2008;17(2):194-206.
5. Justice LM, Bowles RP, Kaderavek JN, Ukrainetz TA, Eisenberg SL, Gillam RB. The index of narrative microstructure: a clinical tool for analyzing school-age children's narrative performances. *Am J Speech Lang Pathol.* 2006;15(2):177-91.
6. Scott CM, Windsor J. General language performance measures in spoken and written narrative and expository discourse of school-age children with language learning disabilities. *J Speech Lang Hear Res.* 2000;43(2):324-39.
7. Merritt DD, Liles BZ. Narrative analysis: clinical applications of story generation and story retelling. *J Speech Hear Disord.* 1989;54(3):438-47.
8. Pankratz ME, Plante E, Vance R, Insalaco DM. The diagnostic and predictive validity of the Renfrew Bus Story. *Lang Speech Hear Serv Sch.* 2007;38(4):390-9.
9. Capps L, Losh M, Thurber C. "The frog ate the bug and made his mouth sad": narrative competence in children with autism. *J Abnorm Child Psychol.* 2000;28(2):193-204.
10. Bliss LS, McCabe A, Miranda AE. Narrative Assessment Profile: discourse analysis for school-age children. *J Commun Disord.* 1998;31(4):347-62.
11. Schneider P, Hayward D. Who does what to whom: introduction of referents in children's storytelling from pictures. *Lang Speech Hear Serv Sch.* 2010;41(4):459-73.
12. Damico JS, Müller N, Ball M J. The handbook of language and speech disorders. 1st ed. West Sussex: Wiley-Blackwell; 2010.
13. Shultz KS, Whitney DJ. Measurement theory in action: case studies and exercises. 1st ed. London: Sage publications, Ltd.; 2005.
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. McCabe A, Bliss LS. Patterns of narrative discourse: a multicultural, life span approach. 1st ed. Boston: Allyn and Bacon; 2003.