

# Diagnostic Accuracy of Working Memory Tasks Depending on Scoring Unit and Condition in Preschool Children

Dongsun Yim, Shinyoung Kim, Jeewon Yoo, Yunjung Lee, Sangeon Lee, Haeun Chung

Department of Communication Disorders, Ewha Womans University, Seoul, Korea

**Correspondence:** Dongsun Yim, PhD  
Department of Communication Disorders,  
Ewha Womans University, 52 Ewhayeodae-gil,  
Seodamun-gu, Seoul 03760, Korea  
Tel: +82-2-3277-6720  
Fax: +82-2-3277-2122  
E-mail: sunyim@ewha.ac.kr

Received: April 10, 2017  
Revised: May 16, 2017  
Accepted: June 5, 2017

**Objectives:** The purpose of the current study was to examine the diagnostic accuracy of working memory (WM) tasks, including nonword repetition (NWR), sentence repetition (SR), and matrix tasks, as clinically efficient diagnostic tools depending on the scoring systems and task condition. **Methods:** Forty children with vocabulary delay (VD group) and 40 typically developing children (TD group) were administered a series of WM tasks (NWR, SR, and matrix tasks). Different scoring units were applied to each task (syllable vs. word for NWR task and word vs. sentence for SR task) while different conditions (forward matrix vs. backward matrix) were used for the matrix task. Separate discriminant function analyses were conducted with the scores calculated by the different scoring units or the different condition of the tasks as predictors. **Results:** Multiple discriminant function analyses for the NWR task yielded word units with higher overall classification values than syllable units, whereas for the SR task sentence units had a higher overall classification value than word units. For the matrix task, none of the conditions reached the level of significance in the discriminant function analyses. **Conclusion:** The results in this study support the diagnostic accuracy of the NWR task with word units and the SR task with sentence units in discriminating among the VD group and the TD group. However, caution should be used when regarding these matrix tasks as an accurate diagnostic tool, especially in young children.

**Keywords:** Nonword repetition, Sentence repetition, Matrix, Discriminant function analyses, Diagnostic accuracy

양육자 및 교사의 의뢰로 언어장애의 가능성이 있는 아동들이 임상 현장에 오게 되면 임상가는 선별검사와 진단검사를 통해 아동의 언어장애 여부를 판별한다. 이때 임상가는 아동의 증상과 연령, 보호자의 보고를 종합적으로 고려하여 아동에게 나타나는 언어적 결함의 세부적 영역에 대해 구체적으로 평가한다. 언어평가는 공식 검사(formal test; standardized test)와 비공식 검사(informal test), 직접검사와 간접검사 등으로 나눌 수 있는데(Kim, 2014), 일반적으로 타당도와 신뢰도가 검증되었으며 기준을 제공할 수 있는 공식 검사를 우선적으로 실시하되 대상 아동이 가진 구체적인 언어문제에 대한 심화평가의 수단으로 다양한 비공식 검사를 추가적으로 실시한다(Tyler & Tolbert, 2002). 그런데 현재 임상에서 사용되고 있는 검사도구들은 매우 제한적인 실정이며(Kim, 2014; Kim,

Park, & Lee, 2005; Lee, 2006), 많은 임상가들은 본인에게 의뢰된 아동의 언어능력을 정밀하게 평가하기 위하여 외국의 공식 검사를 번역하여 사용하거나, 혹은 다양한 연구들에서 개발한 언어평가 과제들을 활용하고 있다(Oh & Yim, 2013; Vandierendonck & Szmalec, 2004; Yim, 2017; Yim, Jo, Han, & Sung, 2016).

그러나, 공식 검사와 달리 비공식 검사는 객관적인 기준이 없고, 대부분의 경우 연구에서 사용된 검사를 차용함으로써 각 연구의 목적에 따라 검사의 진행 방식이나 채점 방식 등이 일치되지 않아 임상에서 활용할 때 범용이 어렵다는 제한점이 있다. 특히 작업기억(working memory)이 언어능력의 주요 기저요인으로 주목받아 온 이후 작업기억을 평가하는 다양한 과제들이 임상에서 사용되고 있으나(Kane et al., 2004; McNamara & Wong, 2015; Menghini,

Finzi, Carlesimo, & Vicari 2011; Park et al., 2002), 각 과제의 수행 결과를 채점하는 방식이 검사도구마다, 혹은 과제를 사용한 연구들마다 일관적이지 않다. 특히 비공식 검사에서 이러한 과제들을 사용하는 경우 어떤 방식을 적용하느냐에 따라 과잉진단(over diagnosis) 또는 오진단(under diagnosis)의 가능성으로 인해 선별 및 진단의 정확도가 저해될 수 있음을 주의해야 한다. 따라서 아동의 언어능력을 세밀하게 평가하기 위한 비공식 검사들의 진단도구로서의 기능을 검증하기 위하여, 언어장애를 정확하고 효과적으로 판별할 수 있는 채점 체계에 대한 연구가 필요하다.

국내외에서 자주 사용되는 비공식 검사들 중, 비단어따라말하기 과제, 문장따라말하기 과제는 Baddeley의 다중요소 작업기억 모델을 기반으로 자주 논의되어 왔다. Baddeley가 제안한 다중요소 작업기억 모델에 의하면, 작업기억은 언어적 정보 또는 비언어적 정보 등 영역-특정적(modality-specific) 정보를 저장하는 동시에 해당 정보를 조작하고 처리하는 능력을 말한다(Baddeley & Hitch, 1974; Baddeley, 1986, 2000). Baddeley의 작업기억 모델은 핵심 구성요소인 중앙집행기(central executive)와 3개의 하위체계로 구성되어 있다. 중앙집행기는 작업기억 모델에서 저장된 정보와 처리하는 정보 사이에 가용할 수 있는 자원을 할당하며 정보를 조작 및 처리하는 데 중심적 역할을 하는 구성요소로 알려져 있으며, 3개의 하위체계는 언어적 정보를 저장하는 음운루프(phonological loop), 비언어적 정보를 저장하는 시공간 스케치패드(visuospatial sketch pad), 그리고 장기기억에 저장된 정보들과 중앙집행기 및 언어적, 비언어적 정보들을 통합하며 처리하는 다차원적 저장소인 임시완충기(episodic buffer)가 있다(Baddeley, 2007; Baddeley, Allen, & Hitch, 2010). 임상에서는 특히 비단어따라말하기, 문장따라말하기 검사가 많이 사용되고 있는데(Martin, & Schwartz, 2003), 이들 과제는 각각 작업기억의 구성요소 중 음운루프와 임시완충기를 측정하는 과제로 알려져 있다(Petrucelli, Bavin, & Bretherton, 2012).

비단어따라말하기 과제(nonword repetition task)는 피실험자가 다음절로 구성된 비단어를 듣고 음운 정보를 저장하여 그대로 다시 따라 말하는 과제로, Baddeley의 작업기억 모델의 구성요소 중 음운루프가 관여한다. 비단어따라말하기 과제는 많은 선행연구들에서 다양한 언어능력, 예를 들어 어휘력, 어휘 습득능력(Botting & Conti-Ramsden, 2001; MacRoy-Higgins & Dalton, 2015; McKean, Letts, & Howard, 2013; Yim et al., 2016), 언어 이해능력(Yim, Kim, & Yang, 2015; Yang, Yim, Kim, & Han, 2013), 읽기능력(Melby-Lervag & Lervag, 2012; Rispens, Baker, & Duinmeijer, 2015; Rispens & Baker, 2012) 등과 정적 상관관계를 가지고 있으며, 나이가 단순 언어장애(Archibald & Joanisse, 2009; Bishop, North, & Donlan,

1996; Conti-Ramsden, Botting, & Faragher, 2001), 자폐범주성장애(Botting & Conti-Ramsden, 2003), ADHD (Redmond, Thompson, & Goldstein, 2011) 등 다양한 장애군을 선별하는 유용한 도구가 될 수 있음이 검증되어 왔다. 현재 국내의 공식적·비공식적 언어처리능력 평가 과제로 비단어따라말하기 과제가 빈번하게 사용되고 있으나(Dollaghan & Campbell, 1998; Hong & Yim, 2014; Windsor, Kohnert, Lobitz, & Pham, 2010; Yim, Kim, et al., 2016), 수행결과에 대한 채점단위는 통일되어 있지 않다.

비단어따라말하기 과제는 제시된 비단어의 음절 모두를 정확하게 따라 말했을 경우 1점을 부여하는 낱말 단위 채점방식(Tattersall, Nelson, & Tyler, 2015; Yim et al., 2016)과, 음절 단위로 각 1점씩 부여하는 음절 단위 채점방식이 주로 행해지고 있으며(Burke & Coady, 2015; Oh & Yim, 2013; Yim, Kim, & Yang, 2016), 음소 단위 채점방식을 적용한 선행연구도 찾아볼 수 있다(Brandeker & Thorardottir, 2015; Dollaghan & Campbell, 1998; Hong & Yim, 2014; McKean et al., 2013). 국내에서도 비단어따라말하기의 임상적 선별 도구로서의 가능성을 검토한 많은 선행연구들이 진행되어 왔으며(Hong & Yim, 2014; Yim, Kim, et al., 2016), 채점단위로는 낱말 단위(Yim et al., 2015)와 음절 단위(Oh & Yim, 2013; Yang et al., 2013)로 점수를 산출한 연구가 주를 이룬다. 국내의 경우 4세 이상 아동들의 초기 문해 수준을 평가하는 공식검사(Basic Academic Skills Assessment, BASA; Early Literacy; Kim, 2008)에서는 낱말 단위 채점방식을 적용하고 있으며, 읽기 검사를 평가하는 한국어 읽기검사(Korean Language-based Reading Assessment, KOLAR; Pae, Kim, Yoon, & Jahng, 2015) 및 읽기 성취 및 읽기 인지 처리능력 검사(Test of Reading Achievement & Reading Cognitive Processes Ability, RA-RCP; Kim, Kim, Hwang, & Yoo, 2014)에서도 낱말 단위 채점방식을 적용하고 있다. 국외의 경우 음운처리능력을 평가하는 Comprehensive Test of Phonological Processing-second edition (CTOPP-2; Wagner, Torgesen, Rashotte, & Pearson, 2013)에서는 낱말 단위 채점방식을 적용하고 있다. 다수의 선행연구에 의하면 채점 방식과 무관하게 거의 대부분의 연구들에서 언어장애 아동 집단과 일반아동 집단 간 과제 수행력의 차이가 유의하였다(Dollaghan & Campbell, 1998; Yang et al., 2013). 또한 비단어따라말하기 과제의 채점단위 간 차이를 비교 검토한 선행연구(Guiberson & Rodriguez, 2013)에서는, 낱말 단위 채점방식과 음소 단위 채점방식이 모두 유의하게 언어장애 집단과 일반아동 집단을 유의하게 판별하는 것으로 났으나, 민감도(sensitivity)와 특이도(specificity)로 판단한 언어장애 집단 판별력은 낱말단위 채점방식이 더 높은 것으로 나타났다.

문장따라말하기 과제는 피실험자가 문장을 듣고 그대로 따라 말하는 과제로, 과제를 수행하면서 어휘 및 문장구조에 관한 정보를 기억하고 재구성해야 한다. 따라서 Baddely의 작업기억 모델 구성 요소 중 장기기억 및 집행기능이 통합적으로 개입하는 임시완충기가 관여하여, 이를 측정하는 과제로 사용되어 왔다(Alloway, Gathercole, Willis, & Adams, 2004; Petruccelli et al., 2012). 문장따라말하기 과제는 구문능력 등 일반적 언어능력을 측정할 뿐 아니라(Klem et al., 2015) 단순언어장애 및 언어능력 결함을 진단하는 검사로 폭넓게 사용되고 있다(Archibald & Joanisse, 2009; Eadie, Fey, Douglas, & Parsons, 2002; Hansson, Nettbelbladt, & Leonard, 2003; Hwang, 2003; Poll, Betz, & Miller, 2010; Riches, 2012). 국외의 경우 문장따라말하기 과제는 Clinical Evaluation of Language Fundamentals-fourth edition (CELF-4; Semel, Wiig, & Secord, 1989), Test of Language Development Primary-third edition (TOLD-P3; Newcomer, & Hammill, 1997)과 Wechsler Preschool and Primary Scale of Intelligence-Revised (WIPPSI-R; Wechsler, 1989) 등과 같은 공식검사의 하위검사에서 사용되고 있는데, 각 검사들에서 사용하는 과제의 채점단위가 서로 통일되어 있지 않다. CELF-4 (Semel et al., 1989)에서는 오류의 개수에 따라 채점하는데, 문장을 완전하게 따라 하면 3점, 오류가 1개일 때 2점, 오류가 2-3개일 때 1점, 오류가 4개 이상일 때 0점으로 처리한다. 이에 반해, TOLD-P3 (Redmond, 2005)에서는 문장 단위의 채점방식이 적용되고 있고, WIPPSI-R의 하위검사인 문장따라말하기 과제를 사용한 연구에서는(Eadie et al., 2002) 문법형태소 단위의 채점방식이 적용하였다. 국내의 경우 공식검사인 RA-RCP (Kim et al., 2014)의 하위검사 중 하나인 문장따라말하기 과제에서는 문장 단위 채점방식을 사용하고 있으며, 취약전 아동의 수용언어 및 표현언어 발달척도(Preschool Receptive-Expressive Language Scale, PRES; Kim, Sung, & Lee, 2003)에서는 일부 문항에서 문장 단위 채점방식의 문장따라말하기 검사를 실시하고 있다. 공식검사 외에 말 늦은 아동과 일반 아동(Oh & Yim, 2013), 단순언어장애 아동과 일반 아동(Kang & Kang, 2016; Lee, Choi, & Hwang, 2014), 읽기부진 아동과 일반 아동(Hwang, 2014) 간 문장따라말하기 과제의 수행력 차이를 비교한 국내의 다양한 선행 연구에서 사용된 채점방식은 일관적이지 않으며, 음절 단위로 채점하여 점수를 산출한 연구(Kang & Kang, 2016; Lee et al., 2014)와 문장 단위로 채점하여 점수를 산출한 연구(Oh & Yim, 2013; Park, Yoon, Han, & Yim, 2014) 등이 혼재되어 있다. 한편, 문장따라말하기 과제의 채점 체계를 네 가지(문장 단위, 문장 성분, 오류 횟수, 음절 단위)로 분류하여 아동의 수용어휘 능력을 가장 잘 예측하는 방식이 무엇인지 검토한 선행연구(Park et al., 2014)에 의하면, 대상

아동의 수용어휘 능력을 가장 잘 예측하는 것은 문장 단위의 평가 방식인 것으로 나타났다. 다만 이 연구에서는 문장 단위 채점 방식은 점수 획득이 어려워 아동의 반응을 과소평가할 가능성이 높으므로 결과 해석에 주의해야 함을 밝혔다. 5-6세 아동들을 대상으로 한 선행연구(Hwang, 2012)에서는 6세 아동의 경우 문장단위 점수나 낱말단위 점수의 판별력에 차이가 없었으나, 5세 아동의 경우 문장단위 점수가 낱말단위 점수보다 판별력이 높게 나타났으며, 이에 따라 문장단위 점수로 진단정확도를 검증한 결과 5세와 6세 집단 모두에서 민감도(sensitivity), 특이도(specificity), 예측도(predictive value)가 모두 85% 이상으로 높게 나타나 문장따라말하기 검사의 단순언어장애 진단도구로서의 유용성을 확인하였다.

한편, 언어적 작업기억에 대한 연구에 비해 시공간 작업기억과 언어장애의 관계를 밝힌 연구는 상대적으로 드물다. 하지만, 난독증 아동 집단과 일반 아동 집단 간 비언어적 작업기억의 유의한 차이를 밝힌 선행연구(Menghini et al., 2011) 및 학습장애 아동들이 일반 아동들에 비해 비언어적 작업기억 과제에서 유의하게 낮은 수행결과를 보였다는 선행연구(McNamara & Wong, 2015)에 의하면, 작업기억이 영역-일반적인(domain-general) 인지능력으로써 언어적 작업기억과 비언어적 작업기억이 모두 언어능력과 관련이 있다고 볼 수 있다. 다양한 언어적 작업기억과 비언어적 작업기억 과제를 대상으로 요인분석을 실시하여 두 영역이 상당한 양의 분산을 공유함을 밝힌 선행연구(Kane et al., 2004; Park et al., 2002)와 언어장애 아동들이 비언어적 기억 능력에도 결함이 있음을 보인 선행 연구(Im-Bolter, Johnson, & Pascual-Leone, 2006; McNamara & Wong, 2015; Menghini et al., 2011) 등은 이러한 입장을 지지하나, 반면 단순언어장애 아동들의 경우 비언어적 작업기억의 능력은 언어적 작업기억에 비해 상대적으로 보존된다는 선행연구(Archibald & Gathercole, 2006) 등은 반대의 입장을 견지한다. 즉, 작업기억이 영역-일반적인(domain-general) 인지능력인지 아니면 영역-특정적인(domain-specific) 인지능력인지에 대한 합일된 의견이 있다고 할 수 없지만, 언어의 사용이 제한적인 장애 아동들에게 있어 비언어적 과제는 언어적 관여를 최소화할 수 있어 아동의 능력을 평가하는 데 더욱 적합할 수 있다는 점을 고려한다면, 비언어적 기억과제의 수행과 언어적 작업기억의 결함이 뚜렷한 언어장애군과의 연관성에 대한 연구는 필요하다.

언어장애의 기저 원인으로 제시되는 음운 작업기억 등 언어적 작업기억에 대한 연구는 다양한 과제들을 통해 광범위하게 진행되어 온 반면, 시공간적 작업기억(visuospatial sketchpad)은 시각적 패턴이나 시각적 자극제를 이용하여 일련의 움직임을 순서대로 혹은 역순으로 회상하는 과제를 이용하여 측정해왔는데, 대표적으로

코지블록 과제(Corsi block task; Cowan, Donlan, Newron, & Lloyd, 2005), 시각패턴 과제(visual pattern task; Della Sala, Gray, Baddeley, Allamano, & Wilson, 1999), 점 매트릭스 과제(dot-matrix task; Vugs et al., 2014, 2016) 등이 있다. 일반적으로 시각적 자극을 기억했다가 순서대로 회상하는 과제는 피실험자의 시공간적 단기 기억을 측정하는 과제로 주로 사용되는 반면, 역순으로 회상하는 과제는 시공간적 작업기억을 측정하는 과제로 사용되고 있다(Vandierendonck & Szmalec, 2004). 그러나 8-11세 아동을 대상으로 세 가지 시각자극을 순서대로 회상하는 과제와 역순으로 회상하는 과제를 비교한 선행연구(Hutton & Towse, 2001)에서는 성인과 달리 아동들의 경우 두 가지 과제가 동일한 인지적 기억(memory) 영역을 측정함을 밝힌 바 있다. 국외의 선행연구에서는 아동의 시공간적 작업기억을 측정하는 과제로 Automated Working Memory Assessment (AWMA; Alloway, Gathercole, & Pickering, 2004)와 Working Memory Test Battery for Children (WMTB-C; Pickering & Gathercole, 2001)의 하위검사들이 주로 사용되고 있는데, 해당 검사들의 경우 모두 점이나 도형, 블럭들의 위치를 순서대로 회상하게 한다.

본 연구는 작업기억의 각 하위 체계들을 측정하는 과제들의 언어장애 진단정확도를 분석하고자 한다. 이를 위하여 각 케이스가 속한 집단을 예측하여 집단을 분류하는 방식인 판별분석을 통하여 실제 언어장애 아동 중 해당 과제를 통하여 언어장애 집단으로 진단되는 비율인 민감도(sensitivity) 및 실제 일반아동 중 해당 과제를 통하여 언어장애 집단에서 배제되는 비율인 특이도(specificity)를 검토하고자 한다. 또한 본 연구에서는 장애의 유병률에 영향 받지 않는 진단 정확도로서 우도비를 측정하였는데, 비장애임에도 해당 검사에서 장애로 진단되는 경우를 반영한 양성우도비와 실제 장애가 있음에도 비장애로 진단되는 경우를 반영한 음성우도비는, 각각 해당 검사에서 장애로 진단된 경우 실제로 장애가 있을 진단 정확도(rule-in)와 해당 검사에서 비장애로 진단된 경우 실제로 장애가 없을 진단정확도(rule-out)를 의미한다고 할 수 있다(Guiberson & Rodriguez, 2013). Guiberson과 Rodriguez (2013)는 양성우도비가 10 이상일 때 해당 장애의 진단정확도(rule-in)가 높다고 할 수 있으며, 3 이상일 때에는 제한적이지만 용인할 수 있는 수준, 양성우도비가 1일 때에는 진단적 도구로서 사실상 의미가 없다고 보았다. 또한 음성우도비가 0.1 이하인 경우 진단의 정확도(rule-out)가 높다고 할 수 있으며, 0.3 이하인 경우 제한적이지만 용인할 수 있는 수준, 음성우도비가 1인 경우 진단적 도구로서 사실상 의미가 없다고 보았다. 다른 연구(Dollaghan & Horner, 2011; Fischer, Bachmann, & Jaeschke, 2003)에서는 양성우도비가 10 이상 또는 음성

우도비가 0.1 이하인 경우 임상적 진단의 잠재력이 있는 것으로 보았으며, 양성우도비가 5에서 10 사이, 음성우도비가 0.1에서 0.2 사이일 때에는 부가적으로 유용한 정보를 제공해줄 수 있을 정도의 진단적 가치를 갖고 있는 것으로 보았다. 또한 양성우도비가 약 3, 음성우도비가 약 0.30의 범위에 있는 경우 다소 제한적으로만 임상적 결정에 영향을 미칠 수 있다고 하였으며, 양성우도비와 음성우도비가 1에 가까울수록 임상적인 유용성이 낮은 것으로 보았다.

요약하자면, 비단어따라말하기 과제와 문장따라말하기 과제는 많은 선행연구들에서 다양한 언어장애 아동 집단과 일반아동 집단 간 수행력 차이가 유의하게 나타나는 등 그 임상적 유용성이 검증되어 왔다. 하지만, 동일한 내용의 과제를 사용하면서도 실시 방법이나 채점 방식은 여전히 연구들마다 다르며, 이에 대해 특히 진단정확도만을 목적으로 한 판별연구는 많지 않다. 또한 명확히 구분되는 임상적 집단을 대상으로 작업기억의 세 하위 체계를 다각적으로 측정하면서 각 과제가 해당 집단을 얼마나 정확하고 효과적으로 변별할 수 있는지를 검토하는 연구는 찾아보기 어렵다. 한편 학령전기 아동의 어휘력은 이후의 언어능력(Lee, 2011; Storch & Whitehurst, 2002) 및 학업 성취도(Morgan, Farkas, Hillemeier, & Hammer, 2015; Walker, Greenwood, Hart, & Carta, 1994)를 예측하는 주요 요인인 바, 어휘발달이 또래에 비해 지체된 아동에 대한 정확한 진단과 이에 따른 신속하고 효과적인 개입이 필요하다. 따라서, 본 연구에서는 어휘발달지체 아동과 일반 아동을 대상으로 Baddeley (1986, 2000)가 제시한 작업기억의 세 가지 하위 체계 중 언어적 요소가 관여하는 두 체계인 음운루프(phonological loop), 임시완충기(episodic buffer)를 측정하는 비단어따라말하기 과제와 문장따라말하기 과제의 채점 방식에 따른 진단적 정확도를 검토해보고자 한다. 또한, 비언어적 요소가 관여하는 시공간 스케치패드(visuospatial sketchpad)를 측정하는 매트릭스 과제의 수행조건에 따라 어휘발달지체 아동과 일반 아동이 얼마나 정확하게 구분되는지를 검토해 보고자 한다.

## 연구방법

### 연구대상

본 연구는 서울 및 경기 지역에 거주하는 생활연령 5-6세의 어휘발달지체(VD) 아동 40명(남아 23, 여아 17), 그리고 이 아동들과 생활연령 및 성별을 일치시킨 일반아동(NL) 40명(남아 23, 여아 17)을 대상으로 하였다. 모든 아동들은 (1) 카우프만 아동용 지능검사(Korean Kaufman Assessment Battery for Children, K-ABC; Moon & Byun, 2003)의 비언어성 지능지수가 85 (-1 SD) 이상이고, (2) 주

**Table 1.** Participants' characteristics by subgroup

	VD (N=40)	NL (N=40)
Age (mo)	70.9 (3.92)	71.03 (3.97)
Nonverbal IQ <sup>a</sup>	104 (10.42)	107.4 (7.48)
Receptive vocabulary <sup>b</sup>	50.7 (8.13)	68.2 (9.98)
Expressive vocabulary <sup>b</sup>	60.3 (10.41)	75.48 (9.2)

Values are presented as mean (SD).

VD=children with vocabulary delay; NL=children with normal language.

<sup>a</sup>Korean Kaufman Assessment Battery for Children (Moon & Byun, 2003), <sup>b</sup>Receptive & Expressive Vocabulary Test (Kim, Hong, Kim, Jang, & Lee, 2010).

양육자 또는 어린이집 및 유치원 교사에 의해 정서 및 감각(시각, 청각), 기타 신경학적 결함이 없으며 인지 및 신체기능이 정상 범주인 것으로 보고된 아동으로 선정하였다. 대상 아동들의 평균 생활연령은 70.98개월(SD=3.92), 비언어성 지능지수의 평균은 105.71(SD=9.18)로 나타났다.

연구에 참여한 아동들은 공식 언어검사인 수용 및 표현어휘력 검사(Receptive & Expressive Vocabulary Test; Kim, Hong, Kim, Jang, & Lee, 2009) 수행력에 따라 수용어휘력 또는 표현어휘력 점수가 또래 대비 10%ile 미만인 경우 어휘발달지체 아동 집단으로, 수용어휘력 및 표현어휘력 점수가 모두 또래 대비 10%ile보다 높은 경우 일반아동 집단으로 선정되었다. 각 집단의 생활연령, 비언어성 지능지수, 수용 및 표현어휘력 점수의 평균 및 표준편차는 Table 1에 제시하였다.

독립표본 *t*-검정으로 집단 간 차이를 검증한 결과, 수용어휘력( $t(78)=8.601, p<.001$ ) 및 표현어휘력( $t(78)=6.909, p<.001$ )이 모두 유의한 차이가 있는 것으로 나타났으며, 생활연령과 비언어성 지능지수는 통계적으로 유의한 차이가 없었다.

## 과제실시 방법

### 비단어따라말하기

아동의 음운작업기억 능력을 측정하기 위하여 비단어따라말하기 과제(Lee, Yim, & Sim, 2012; Oh & Yim, 2013)를 사용하였다. 과제의 문항은 2음절, 3음절, 4음절, 5음절, 6음절이 각 4개씩, 총 20개의 비단어 목록으로 구성되었으며, 과제의 타당도 평가를 위해 언어병리학 석사학위 소지자 2명에게 해당 20개 비단어 목록의 단어 유사성(wordlikeness)을 평가하게 하였다. 5점 척도(1점 '전혀 유사하지 않다', 2점 '유사하지 않다', 3점 '보통이다', 4점 '유사하다', 5점 '매우 유사하다')로 내용타당도를 평가한 결과, 평균 1.02점으로 과제 문항의 단어 유사성이 낮은 것으로 나타났다. 검사자는 아동에게 비단어 문항이 녹음된 파일을 헤드폰 없이 컴퓨터로 들려주고, 앵무새처럼 똑같이 따라 말하게 지시하여 과제를 수행하도록

하였다. 문항을 들려주는 동안 컴퓨터 화면에는 음성 파일의 재생 화면만이 제시되었으며, 조용한 방에서 아동과 검사자가 일대일로 검사를 진행하였다. 검사자는 연습문제를 통하여 아동이 검사의 수행 방법을 숙지했다고 판단이 되면 검사를 진행하였으며, 아동의 반응을 듣고 즉각적으로 정오를 판단하면서 전사하였다. 아동의 반응은 녹음기로 녹취하여 검사 종료 후 재확인하였다.

비단어따라말하기 과제는 두 가지 방식으로 채점되었는데, 낱말 단위의 채점방식은 제시 문항에서 음절길이에 관계없이 비단어를 정확하게 따라 말할 경우 1점으로 처리하고, 음소를 하나라도 부정확하게 따라 말할 경우 0점으로 처리하였다(Tattersall, Nelson, & Tyler, 2015). 이 채점방식에서는 총 20문항의 낱말수의 총합이 총 점으로, 최고점수는 20점이 된다. 음절 단위의 채점방식은 제시 문항에서 아동이 산출한 구어반응이 정확하고 위치도 일치하는 경우 정반응으로 처리하였으며, 정확하게 따라 말한 음절의 총합을 측정하였다(Oh & Yim, 2013; Yim, Kim, et al., 2016). 선행연구를 참고하여 정확한 발음과 경미한 왜곡은 정반응으로, 음절내의 음소의 생략과 대치, 첨가는 오류로 처리하였다(Dollaghan & Campbell, 1998; Oh & Yim, 2013). 총 20문항의 총 음절수는 80음절로, 이 채점방식에서 최고점수는 80점이 된다. 원점수로 총점을 계산한 후, 과제 간 수행력 비교의 편의를 위하여 정반응률도 함께 계산하여 기록하였다.

### 문장따라말하기

아동의 임시완충기(episodic buffer) 능력을 측정하기 위하여 문장따라말하기 과제(Ahn & Kim, 2000; Oh & Yim, 2013; Park et al., 2014)를 일부 수정하여 사용하였다. 본 과제는 문장길이와 문장구조에 따라 난이도가 구분되어 3어절 단문, 5어절 단문, 5어절 접속문 각 9개씩, 5어절 내포문 8개, 총 35개의 문장으로 구성되었다. 검사자는 아동에게 앵무새처럼 똑같이 따라 말하기를 지시하며 과제 수행 방법을 설명하면서 연습 문항을 실시한 후, 아동이 검사 내용을 숙지했다고 판단이 되면 검사를 진행하였다. 검사 문항은 사전에 음성파일로 녹음되어 헤드폰 없이 컴퓨터로 제시되었으며, 검사자는 아동의 반응을 즉각적으로 정오 판단하면서 전사하였다. 문항이 제시되는 동안 컴퓨터 화면에는 음성 파일의 재생 화면만이 보여졌으며, 조용한 방에서 아동과 검사자가 일대일로 검사를 진행하였다. 아동의 반응은 녹음기로 녹취하여 검사 종료 후 재확인하였다.

문장따라말하기 과제는 비단어따라말하기 과제와 마찬가지로 두 가지 방식으로 채점되었는데, 문장 단위 채점방식은 문장 전체를 오류 없이 정확하게 따라 말한 경우 1점, 문장의 어느 한 부분이

라도 정확하게 따라 말하지 못한 경우 0점으로 처리하였다(Kim & Chung, 2011; Hwang, 2012; Lee et al., 2014). 제시 문항은 총 35문항으로 이 채점방식에서 최고 점수는 35점이 된다. 검사자는 아동이 정확하게 따라 말할 문장수의 총점을 기록하고 정반응률을 계산하였다. 어절 단위 채점방식은 문항별로 각각 3어절 문장은 3점 만점에서 오류가 나타난 어절의 횟수만큼 제하였고, 5어절 문장은 5점 만점에서 오류 어절의 횟수를 제하는 방식으로 처리하였다(Stokes, Wong, Fletcher, & Leonard, 2006; Park et al., 2014). 본 과제에서 제시 문항은 총 35문항으로 총 어절수는 157어절, 최고 점수는 157점이 된다. 아동이 정확하게 따라 말할 어절수의 총점을 기록하고 정반응률도 함께 계산하여 기록하였다.

### 매트릭스

비언어적 작업기억인 시공간 스케치패드(visuospatial sketchpad)를 측정하는 과제로 매트릭스 과제를 사용하였다. 본 연구에서 사용된 과제(Yim et al., 2015)는 선행연구(Gathercole & Pickering, 2000)에서 사용한 역동적 매트릭스 과제(dynamic matrices test)를 수정·보완하여 사용한 과제로, 3×3 총 9개의 정사각형으로 구성된 매트릭스에 불이 깜빡이는 순서를 기억하여 회상하는 과제이다. 먼저 컴퓨터 화면 중앙에 동일한 크기로 분할된 3×3 매트릭스가 제시된 이후, 0.5초의 간격으로 9개 매트릭스 중 한 칸에 주황색 불이 1초 동안 점등된다. 정지 화면이 나타난 후, 처음에 제시되었던 9개의 빈 매트릭스가 나타나면, 아동은 불이 들어왔던 칸의 순서를 기억하여 손가락으로 가리키면서 반응한다.

본 과제는 매트릭스에 점등되는 순서를 그대로(forward) 기억하는 방식과, 역순으로(backward) 기억하는 두 가지 방식으로 진행되었다. 총 검사 문항 수는 연습 문항 5개와 본 문항 18개이며, 난이도 1은 점등하는 칸이 2개(span-2), 난이도 2는 점등하는 칸이 3개(span-3), 난이도 3은 4개(span-4), 난이도 4는 5개(span-5)로 각 단계별로 회상해야 하는 칸의 개수가 하나씩 늘어난다. 각 난이도별 검사 문항 수는 난이도 1과 2는 4문항, 난이도 3과 4는 5문항으로 구성되었다. 연습 문항을 실시하여 아동이 검사 방법을 충분히 숙지했다고 판단한 이후에 검사 문항을 실시하였다. 과제의 점수는 정반응은 1점, 오반응은 0점으로 기록하였으며, 원점수로 총점을 계산한 후 정반응률도 함께 기록하였다.

### 자료분석 방법

본 연구의 모든 통계분석은 SPSS ver. 23.0을 사용하였다. 본 연구에서는 각 과제별로 두 가지 분석을 실시하였는데, 우선 세 가지 작업기억 과제의 집단 간 차이를 검증하기 위하여 비단어따라말하

기 과제 및 문장따라말하기 과제는 채점단위별로, 그리고 매트릭스 과제는 수행조건별로 독립표본 *t*-검정을 실시하였다. 그리고 각 과제별 채점단위 또는 수행조건에 따른 언어발달지체에 대한 진단 정확도를 측정하기 위하여 판별분석(discriminant analysis)을 실시하였으며, 이에 따라 각 과제의 민감도(sensitivity), 특이도(specificity), 통합정확도(overall classification), 양성우도비(positive likelihood ratio), 음성우도비(negative likelihood ratio)를 산출하였다. 민감도는 실제 장애가 있는 사람들 중 장애로 진단되는 정확도, 특이도는 실제 장애가 없는 사람들 중 비장애로 진단되는 정확도이며, 통합정확도는 전체 대상자 중 장애와 비장애로 바르게 진단되는 해당 검사의 정확도를 가리킨다(Hwang, 2012; Lugo-Neris, Peña, Bedore, & Gillam, 2015). 민감도와 특이도는 해당 장애의 유병률(prevalence)에 영향을 받는다는 한계가 있는데, 우도비(likelihood ratio)는 유병률에 영향받지 않는 값으로서 해당 검사의 진단정확도를 나타낸다. 우도비는 민감도와 특이도로부터 산출할 수 있는데, 양성우도비는 민감도/(1-특이도), 음성우도비는 (1-민감도)/특이도의 값으로 구해진다(Archibald & Joanisse, 2009).

또한 평가자 간 신뢰도를 측정하기 위하여 전체 자료 중 10%를 임의로 선정하여 언어병리학과 박사과정에 재학 중인 연구원 1명이 재측정하여 상관관계를 분석하였다. 그 결과 피어슨 상관계수가 비단어따라말하기 과제는 .921 ( $p < .05$ ), 문장따라말하기 과제는 .892 ( $p < .05$ ), 매트릭스 과제는 .997 ( $p < .01$ )로 유의하게 높은 상관관계를 보였다.

## 연구결과

### 언어적 작업기억 과제의 집단 간 차이 검증 및 판별분석 결과 언어적 작업기억 과제의 집단 간 차이 검증

어휘발달지체 아동 집단 및 일반 아동 집단의 비단어따라말하기, 문장따라말하기 과제의 수행력에 대한 기술통계 결과는 Table 2와 같다. 각 과제는 채점단위에 따라 수행력을 제시하였다.

독립표본 *t*-검정으로 집단 간 과제 수행력 차이를 검증한 결과, 비단어따라말하기 과제의 경우 낱말단위 채점방식( $t(78) = 2.429$ ,  $p < .05$ ) 및 음절단위 채점방식( $t(78) = 2.228$ ,  $p < .05$ ) 모두 집단 간 차이가 통계적으로 유의하였으며, 문장따라말하기 과제도 마찬가지로 문장단위 채점방식( $t(61.417) = 4.400$ ,  $p < .001$ ) 및 어절단위 채점방식( $t(74.690) = 2.625$ ,  $p < .05$ ) 모두 집단 간 차이가 통계적으로 유의하였다. 집단 간 평균차이를 통합표준편차(pooled standard deviation)로 나눈 값인 Cohen's *d*로 효과크기를 추정하였으며, Cohen's *d*값은 비단어따라말하기(낱말단위)는 .543, 비단어따라

말하기(음절단위)는 .498, 문장따라말하기(문장단위)는 .984, 문장 따라말하기(어절단위)는 .587로 나타났다.

**언어적 작업기억 과제의 판별분석 결과**

비단어따라말하기 및 문장따라말하기 과제의 채점방식에 따른 어휘발달지체 진단정확도를 분석하기 위하여 판별분석을 실시하였으며, 그 결과는 Table 3에 제시하였다.

판별분석 결과, 비단어따라말하기 과제의 경우 특이도는 두 가지 채점방식에서 동일하게 나타났으나 민감도가 낱말단위 채점방식(65%)이 음절단위 채점방식(52%)에 비해 높았으며, 따라서 통합정확도는 낱말단위 채점방식이 67.5%로 음절단위 채점방식(61.3%) 보다 높게 나타났다. 문장따라말하기 과제의 경우는 특이도는 어절단위 채점방식(87.5%)이 문장단위 채점방식(77.5%)에 비해 높았으나, 민감도는 문장단위 채점방식(60%)이 어절단위 채점방식(45%)에 비해 높았다. 통합정확도는 문장단위 채점방식(68.8%)이 어절단위 채점방식(66.3)에 비해 높은 것으로 나타났다. 우도비는 비단어 따라말하기 과제의 낱말단위 채점방식의 LR+는 2.16, LR-는 0.5로, 음절단위 채점방식의 LR+는 1.73, LR-는 0.69로 나타났다. 문장따

라말하기 과제의 경우 낱말단위 채점방식의 LR+는 3.6, LR-는 0.63으로, 문장단위 채점방식의 LR+는 2.67, LR-는 0.51로 나타났다. 또한 과제 조합의 판별력을 검토하기 위하여 비단어따라말하기 과제와 문장따라말하기 과제의 채점방식을 조합하여 판별분석을 실시하였다. 그 결과, 비단어따라말하기(음절단위)+문장따라말하기(어절단위)를 제외한 나머지 세 조합의 민감도(77.5%)와 특이도(60%), 그에 따른 통합정확도(68.8%)가 모두 동일한 것으로 나타났다.

**비언어적 작업기억 과제의 집단 간 차이 검증 및 판별분석 결과**

**비언어적 작업기억 과제의 집단 간 차이 검증**

어휘발달지체 아동 집단 및 일반 아동 집단의 매트릭스 과제의 검사 수행조건에 따른 수행력에 대한 기술통계 결과는 Table 4와 같다.

독립표본 t-검정으로 집단 간 과제 수행력 차이를 검증한 결과, 순서대로 회상하는 방법( $t(78) = .334, p = .148$ ) 및 역순으로 회상하는 방법( $t(78) = 1.074, p = .286$ ) 두 가지 모두 집단 간 차이가 통계적으로 유의하지 않았다.

**비언어적 작업기억 과제의 판별분석 결과**

매트릭스 과제의 수행조건에 따른 어휘발달지체 진단정확도를

**Table 2.** Descriptive statistics on verbal working memory tasks by groups

Task	Accuracy (%)	
	VD (N=40)	NL (N=40)
NWR		
Word	62.88 (14.41)	71.00 (15.49)
Syllable	84.78 (7.86)	89.03 (9.15)
SR		
Sentence	67.08 (22.12)	84.73 (12.43)
Word	86.23 (14.01)	93.70 (11.31)

Values are presented as mean (SD).  
VD=children with vocabulary delay; NL=children with normal language; NWR=nonword repetition; SR=sentence repetition.

**Table 4.** Descriptive statistics on nonverbal working memory tasks by groups

Task	Accuracy (%)	
	VD (N=40)	NL (N=40)
MATRIX		
Forward	55.90 (20.50)	57.30 (16.84)
Backward	39.93 (16.88)	44.38 (20.06)

Values are presented as mean (SD).  
VD=children with vocabulary delay; NL=children with normal language.

**Table 3.** Discriminant functions in verbal working memory

Discriminant function	Overall classification	Sensitivity	Specificity	LR+	LR-	Wilks's λ	χ <sup>2</sup>	Canonical correlation	p-value
NWR_word	67.5	65	70	2.16	0.5	.930	5.651	.067	.017
NWR_syllable	61.3	52	70	1.73	0.69	.940	4.782	.117	.029
SR_sentence	68.8	60	77.5	2.67	0.51	.801	17.181	.056	<.001
SR_word	66.3	45	87.5	3.6	0.63	.919	6.562	.079	.010
NWR_word + SR_sentence	68.8	77.5	60	1.94	0.38	.801	17.075	.446	<.001
NWR_word + SR_word	68.8	77.5	60	1.94	0.38	.901	8.604	.315	.018
NWR_syllable + SR_sentence	68.8	77.5	60	1.94	0.38	.801	17.121	.447	<.001
NWR_syllable + SR_word	67.5	80	55	1.78	0.36	.909	7.329	.301	.026

LR+=positive likelihood ratio; LR-=negative likelihood ratio; NWR=nonword repetition; SR=sentence repetition; NWR\_word=NWR task with a word unit in scoring; NWR\_syllable=NWR task with a syllable unit in scoring; SR\_sentence=SR task with a sentence unit in scoring; SR\_word=SR task with a word unit in scoring.

**Table 5.** Discriminant functions in nonverbal working memory

Discriminant function	Overall classification	Sensitivity	Specificity	LR+	LR-	Wilks's $\lambda$	$\chi^2$	Canonical correlation	$p$ -value
MATRIX_F	52.5	55	50	1.1	0.9	.999	.111	.053	> .05
MATRIX_B	57.5	62.5	52.5	1.32	0.71	.985	1.137	.054	> .05

LR+= positive likelihood ratio; LR-= negative likelihood ratio; MATRIX\_F= matrix forward task; Matrix\_B = matrix backward task.

분석하기 위하여 판별분석을 실시하였으며, 그 결과는 Table 5에 제시하였다.

판별분석 결과, 매트릭스 과제의 경우 두 가지 수행조건 모두 판별함수가 통계적으로 유의하지 않은 것으로 나타났다.

## 논의 및 결론

본 연구는 언어장애 아동을 진단하는 데 있어 임상 및 연구에서 다양한 방식으로 사용되고 있는 작업기억 과제들의 진단정확도를 비교 검증하기 위하여 언어적 작업기억을 측정하는 과제인 비단어 따라말하기 과제와 문장따라말하기 과제의 채점방식에 따른 진단정확도를 비교하고, 비언어적 작업기억을 측정하는 매트릭스 과제의 순행 및 역행방식의 진단정확도를 검토하고자 하였다. 이를 위하여 어휘발달지체 아동 집단 및 일반아동 집단을 대상으로 작업기억의 하위영역인 음운루프, 임시완충기, 시공간 스케치패드를 측정하는 각 과제(비단어따라말하기, 문장따라말하기, 매트릭스)의 수행력을 비교한 후 판별분석을 실시하였다.

본 연구에서는 비단어따라말하기 과제의 경우 낱말 단위 및 음절 단위의 두 가지 채점방식으로 어휘발달지체 아동과 일반아동을 대상으로 그 수행력을 분석하였는데, 두 가지 채점방식에서 모두 집단 간 차이가 유의한 것으로 나타났다. 두 가지 채점방식의 특이도에는 차이가 없었으나, 낱말 단위 채점방식의 민감도가 음절단위 채점방식보다 높아 민감도와 특이도를 모두 고려한 진단정확도는 낱말 단위 채점방식이 더 높은 것으로 나타났다. 또한 양성우도비와 음성우도비의 의미를 고려하였을 때, 양성우도비가 높을수록, 음성우도비가 낮을수록 진단정확도가 높음을 알 수 있는데, 본 연구에서는 낱말 단위 채점방식이 음절 단위 채점방식에 비해 양성우도비가 더 높고 음성우도비가 더 낮은 것으로 나타났다. 즉, 낱말 단위 채점방식이 장애군을 장애로 진단하고 비장애군을 비장애로 배제시킬 수 있는 가능성이 높은 것으로 나타났다. 또한 Cohen's  $d$  값으로 측정된 효과크기도 낱말 단위 채점방식(Cohen's  $d = .543$ )이 음절 단위 채점방식(Cohen's  $d = .493$ )보다 높은 것으로 나타나 판별분석 결과를 지지해주었다. 5세 아동을 대상으로 낱말 단위 채점방식을 사용한 비단어따라말하기 과제에 대한 판별분석을 실시

한 선행연구(Conti-Ramsden & Hesketh, 2003)에서는 민감도가 59%-66%, 특이도가 85%-89%로 나타났다. 해당 선행연구에서는 특히 비단어따라말하기 과제수행력이 백분위수 25 미만인 아동이 단순언어장애로 진단될 가능성이 높음을 밝힘으로써, 비단어따라말하기 과제의 선별도구로서의 유용성을 다시 한 번 확인하였다. Conti-Ramsden과 Hesketh (2003)의 연구와 비슷한 연령의 아동을 대상으로 한 본 연구에서는 낱말 단위 채점방식의 경우 민감도가 65%, 특이도가 70%로, 해당 선행연구에 비해 특이도는 낮았으나 민감도는 비슷한 수준인 것으로 나타났다. 따라서 본 연구에서 사용한 비단어따라말하기 과제의 임상적 유용성을 집단 간 차이 검증을 통해 확인함과 동시에, 선행연구와 비교하였을 때에도 본 연구에서 사용된 과제의 경우 시간적 효율성이 높은 낱말 단위 채점방식이 언어발달지체를 선별해내는 진단정확도도 높음을 확인하였다.

문장따라말하기 과제의 경우 문장 단위 채점방식과 어절 단위 채점방식의 수행력을 분석하였는데, 민감도는 문장단위 채점방식이 높았으며 특이도는 어절단위 채점방식이 높았다. 민감도와 특이도를 함께 고려한 통합정확도는 문장단위 채점방식이 높은 것으로 나타났다. 우도비를 고려하였을 때에는 양성우도비의 경우 어절단위 채점방식이 더 높았으며 음성우도비는 문장단위 채점방식에서 더 낮은 것으로 나타나 민감도 및 특이도 분석 결과와 상반되는 결과를 보였다. 즉, 실제 장애가 있는 사람들 중 장애로 진단되는 비율인 민감도는 문장단위 채점방식이 높았으나, 검사 결과가 해당 장애가 있다고 말할 수 있는 임상적 정확도는 어절단위 채점방식이 더 높은 것으로 나타났으며, 실제 장애가 없는 사람들 중 비장애로 진단되는 비율인 특이도와 검사 결과가 해당 장애가 없다고 배제할 수 있는 임상적 정확도 간의 관계도 상반되게 나타났음을 의미한다. 이는 우도비의 계산방식을 고려해보았을 때, 어절단위 채점방식의 민감도가 50% 미만의 낮은 값을 가졌으며 특이도는 87.5%로 가장 높게 나타나, 어절단위 채점방식에서 민감도와 특이도의 격차가 커진 것에서 비롯되었음을 알 수 있다. 즉, 양성우도비를 산출할 때 (1-특이도)의 값으로 계산되는 분모값이 작아지는 폭이 커짐으로써 민감도가 작아지는 폭을 상쇄하였으며, 음성우도비를 산출할 때에는 (1-민감도)의 값으로 계산되는 분자값의 증가폭이 분모값



인 특이도가 작아지는 폭을 상쇄하여 민감도와 특이도, 우도비 간 상충되는 결과를 나타낸 것으로 보인다. 민감도와 특이도의 경우 Plante와 Vance (1994) 등은 80%를 최소한의 기준으로 제시하고는 있으나, 일반적으로 통용되는 기준은 없다고 할 수 있다. 그러나 어절 단위 채점방식은 민감도가 50% 미만으로 낮게 나타났으며, 문장 단위 채점방식의 경우 특이도는 어절 단위 채점방식보다 낮은 수준이나 80%에 가까운 정도이고 민감도 및 통합정확도가 높게 나타나 본 연구에서 진단의 정확도가 더욱 우수한 수준인 것으로 판단할 수 있다. 이는 문장단위 채점방식이 언어장애를 진단하는 데 더욱 정확한 정보를 제공했음을 밝힌 선행연구(Leclercq, Quemart, Magis, & Maillart, 2014; Hwang, 2012; Park et al., 2014)와도 일치하는 결과라 하겠다. Cohen's  $d$ 값으로 측정된 효과크기는 문장단위 채점방식이 0.984, 어절단위 채점방식이 0.587로, 문장단위 채점방식의 경우 Cohen (1992)의 기준에 의할 때 매우 높은 수준인 것으로 나타나, 선행연구 결과를 지지하였다. 즉, 본 연구에서 사용된 문장따라말하기 과제가 어휘발달지체 집단을 유의하게 구별해주며, 특히 문장단위 채점방식이 집단을 구분하는 임상적 정확도가 높은 것으로 나타났다.

매트릭스 과제의 경우 두 가지 과제 수행조건에서 모두 집단 간 차이가 통계적으로 유의하지 않았으며, 판별분석 결과 또한 유의하지 않았다. 이는 두 가지 가능성으로 해석할 수 있을 것으로 보인다. 첫째, 연구 대상 아동들의 비언어적 작업기억을 측정하는 과제인 매트릭스는 9개 정사각형 매트릭스에 점멸하는 칸을 기억했다가 순서대로 또는 역순으로 회상하여 반응하는 과제인데, 본 연구에서 사용된 매트릭스 과제는 2개 칸이 점멸하는 span-2가 4개 항목, span-3이 4개 항목, span-4가 5개, span-5가 5개 항목으로 구성되었다. 두 집단 모두 순서대로 회상하는 방식에서 정반응률이 60%에 미치지 못하였으며, 이는 대부분의 아동들이 span-4에서 정반응률이 급격히 낮아졌음을 의미한다. 역순으로 회상하는 방식에서는 두 집단 모두 정반응률이 50%에 미치지 못하였다. 본 연구에서 사용된 비단어따라말하기 과제의 정반응률이 어휘발달지체 아동 집단은 62.88%, 일반아동 집단은 71%임을 고려해볼 때 언어적 작업기억 과제의 경우 두 집단 모두 평균적으로 span-4의 수행률을 보였으며, 5-6세 아동을 대상으로 한 선행연구(Jones, Tamburelli, Watson, Gobet, & Pine, 2010)에서 비단어따라말하기 과제 수행력의 최대치를 span-5로 제시한 것과는 대조적이다. 본 연구의 결과로 비언어적 작업기억은 언어적 작업기억과 다른 부하(load)와 기억폭(span)을 가짐을 추측해볼 수 있다. 이는 다양한 작업기억 과제들을 대상으로 요인분석을 실시한 국내의 선행연구(Yim et al., 2015)에서 매트릭스 과제가 복합처리 또는 다중처리 작업기억 과제

로 분류된 것에서도 확인할 수 있다. 즉, 비언어적 작업기억 과제인 매트릭스 과제는 학령전기의 두 집단 아동 모두에게 과제를 수행하는 데 있어 보다 높은 수준의 처리 부하(load)를 요구함으로써 집단 간 변별력이 낮게 나타났을 가능성이 있다. 또 다른 가능성으로는 어휘발달지체 아동의 비언어적 작업기억이 언어적 작업기억에 비해 비교적 손상되지 않았을 가능성을 검토해볼 수 있다. 단순언어장애 아동을 대상으로 한 선행연구(Alt, 2013; Ricco, Cash, & Cohen, 2007)에 의하면 비언어적 작업기억 과제의 수행력이 또래 일반 아동과 통계적으로 유의한 차이가 나타나지 않았으며, 이러한 연구결과는 단순언어장애 아동이 비언어적 정보를 순간적으로 유지하고 조작 및 처리하는 능력은 언어적 정보를 조작하는 것과 달리 또래와 크게 다르지 않을 것임을 의미한다. 비언어적 작업기억 과제의 어휘발달지체 아동에 대한 판별력이 낮게 나타난 본 연구의 결과는 작업기억의 하위체계들 간의 관계에 있어서 영역-일반적인(domain-general) 작업기억을 강조하는 입장(Engle, Tuholski, Laughlin, & Conway, 1999; Kane et al., 2004)보다는 영역-특정적인(domain-specific) 작업기억을 강조하는 선행연구들(Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001; Shah & Miyake, 1996)을 지지하는 결과라고 할 수 있다.

본 연구는 5-6세 어휘발달지체 아동 및 일반아동을 대상으로 Baddeley가 제안한 작업기억의 하위체계인 음운루프, 임시완충기, 시공간 스케치패드를 측정하는 세 가지 과제들의 채점단위 및 과제 수행조건에 대한 진단 정확도를 검토하였다. 음운루프를 측정하는 비단어따라말하기 과제 및 임시완충기를 측정하는 문장따라말하기 과제에서는 집단차이 검증 및 판별분석에서 모두 통계적으로 유의한 결과가 도출되었으나, 시공간 스케치패드를 측정하는 매트릭스 과제에서는 그렇지 않았다. 그 원인으로 영역-특정적인 작업기억 이론에 기반하여 비언어적 작업기억 능력이 언어적 작업기억 능력에 비해 비교적 손상되지 않았을 가능성을 제시하였으나, 과제 난이도 조절의 실패를 또 다른 원인으로 지적할 수 있을 것이다. 다시 말해, 본 과제에서 사용한 매트릭스 과제가 5-6세 아동들에게 지나치게 어려워 과제의 변별력이 저하되었을 가능성이 있다. 따라서 후속연구에서는 해당 과제를 고연령층의 아동들에게 확대 실시하여 매트릭스 과제의 두 가지 수행조건 중 어느 것이 더 집단을 명확히 구분해줄 수 있을지 검증할 필요가 있다. 또한 본 연구는 어휘발달지체 아동으로 대상군을 제한한 점과, 기억폭 과제인 세 가지 과제 모두 조건을 폭의 범위별로 세분화하여 정밀하게 측정하지 못한 한계점이 있다. 그러나 많은 선행연구들에서 이미 임상적으로 유용성이 인정된 비단어따라말하기, 문장따라말하기 과제의 채점단위 간 직접적인 비교를 통하여 연구 및 임상에서의 효용성을 제

고하고자 했다는 데에 연구의 의의가 있을 것이다.

## REFERENCES

- Ahn, J., & Kim, Y. (2000). The effect of syntactic complexity on sentence repetition performance and intelligibility between specific language impairment and normal children. *Speech Sciences, 7*, 262-275.
- Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2004). *Automated working memory assessment*. London: Pearson Assessment.
- Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology, 87*, 85-106.
- Alt, M. (2013). Visual fast mapping in school-aged children with specific language impairment. *Topics in Language Disorders, 33*, 328-346.
- Archibald, L. M., & Gathercole, S. E. (2006). Visuospatial immediate memory in language impairment. *Journal of Speech, Language, and Hearing Research, 49*, 265-277.
- Archibald, L. M., & Joanisse, M. F. (2009). On the sensitivity and specificity of nonword repetition and sentence recall to language and memory impairments in children. *Journal of Speech, Language, and Hearing Research, 52*, 899-914.
- Baddeley, A. D. (1986). *Working memory*. New York, NY: Oxford University Press.
- Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences, 4*, 417-423.
- Baddeley, A. D. (2007). *Working memory, thought, and action*. New York, NY: Oxford University Press.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The Psychology of Learning and Motivation, 8*, 47-90.
- Baddeley, A. D., Allen, R., & Hitch, G. (2010). Investigating the episodic buffer. *Psychologica Belgica, 50*, 223-243.
- Bishop, D. V., North, T., & Donlan, C. (1996). Nonword repetition as a behavioural marker for inherited language impairment: evidence from a twin study. *Journal of Child Psychology and Psychiatry, 37*, 391-403.
- Botting, N., & Conti-Ramsden, G. (2003). Autism, primary pragmatic difficulties, and specific language impairment: Can we distinguish them using psycholinguistic markers? *Developmental Medicine & Child Neurology, 45*, 515-524.
- Botting, N., & Conti-Ramsden, G. (2001). Non-word repetition and language development in children with specific language impairment (SLI). *International Journal of Language & Communication Disorders, 36*, 421-432.
- Brandeker, M., & Thordardottir, E. (2015). Language exposure in bilingual toddlers: performance on nonword repetition and lexical tasks. *American Journal of Speech-Language Pathology, 24*, 126-138.
- Burke, H. L., & Coady, J. A. (2015). Nonword repetition errors of children with and without specific language impairments (SLI). *International Journal of Language & Communication Disorders, 50*, 337-346.
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155-159.
- Conti-Ramsden, G., & Hesketh, A. (2003). Risk markers for SLI: a study of young language-learning children. *International Journal of Language & Communication Disorders, 38*, 251-263.
- Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry, 42*, 741-748.
- Cowan, R., Donlan, C., Newron, E. J., & Lloyd, D. (2005). Number skills and knowledge in children with specific language impairment. *Journal of Educational Psychology, 97*, 732-744.
- Della Sala, S., Gray, C., Baddeley, A.D., Allamano, N., & Wilson, L. (1999). Pattern span: a tool for unwelding visuo-spatial memory. *Neuropsychologia, 37*, 1189-1199.
- Dollaghan, C. A., & Horner, E. A. (2011). Bilingual language assessment: a meta-analysis of diagnostic accuracy. *Journal of Speech, Language, and Hearing Research, 54*, 1077-1088.
- Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1136-1146.
- Eadie, P. A., Fey, M. E., Douglas, J. M., & Parsons, C. L. (2002). Profiles of grammatical morphology and sentence imitation in children with specific language impairment and Down syndrome. *Journal of Speech, Language, and Hearing Research, 45*, 720-732.
- Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental Psychology: General, 128*, 309-331.
- Fischer, J. E., Bachmann, L. M., & Jaeschke, R. (2003). A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Medicine, 29*, 1043-1051.
- Gathercole, S. E., & Pickering, S. J. (2000). Working memory deficits in children with low achievements in the national curriculum at 7 years of age. *British Journal of Educational Psychology, 70*, 177-194.

- Guiberson, M., & Rodriguez, B. L. (2013). Classification accuracy of non-word repetition when used with preschool-age Spanish-speaking children. *Language, Speech, and Hearing Services in Schools, 44*, 121-132.
- Hansson, K., Nettelbladt, U., & Leonard, L. (2003). Indefinite articles and definite forms in Swedish children with specific language impairment. *First Language, 23*, 434-362.
- Hong, S., & Yim, D. (2014). The assessment of language impairment in bilingual children through learning and memory tasks. *Communication Sciences & Disorders, 19*, 31-44.
- Hutton, U. M., & Towse, J. N. (2001). Short-term memory and working memory as indices of children's cognitive skills. *Memory, 9*, 383-394.
- Hwang, M. (2003). Sentence comprehension of Korean children with specific language impairments. *Korean Journal of Communication Disorders, 8*, 1-21.
- Hwang, M. (2012). Sentence repetition as a clinical marker of specific language impairment in Korean-speaking preschool children. *Korean Journal of Communication Disorders, 17*, 1-14.
- Hwang, M. (2014). Working memory of children with reading comprehension difficulty: sentence repetition and nonword repetition. *The Korea Journal of Learning Disabilities, 11*, 53-72.
- Im-Bolter, N., Johnson, J., & Pascual-Leone, J. (2006). Processing limitations in children with specific language impairment: the role of executive function. *Child development, 77*, 1822-1841.
- Jones, G., Tamburelli, M., Watson, S. E., Gobet, F., & Pine, J. M. (2010). Lexicality and frequency in specific language impairment: accuracy and error data from two nonword repetition tests. *Journal of Speech, Language, and Hearing Research, 53*, 1642-1655.
- Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: a latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*, 189-217.
- Kang, E., & Kang, D. (2016). Characteristics of non-word repetition and sentence repetition performance of preschool children with specific language impairment. *Journal of Special Education & Rehabilitation Science, 55*, 559-573.
- Kim, A. H., Kim, U. J., Hwang, M., & Yoo, H. S. (2014). *Test of Reading Achievement and Reading Cognitive Processes (RA-RCP)*. Seoul: Hakjisa.
- Kim, D. (2011). *Basic academic skills assessment: early literacy*. Seoul: Korea Psychological Services.
- Kim, J. S., & Chung, S. M. (2011). Sentence repetition performance according to length and structure of sentences in 3 to 5 year-old children. *Journal of Speech-Language & Hearing Disorders, 20*, 19-36.
- Kim, Y. T. (2014). *Assessment and treatment of language disorders in children*. Seoul: Hakjisa.
- Kim, Y. T., Hong, G. H., Kim, K. H., Jang, H. S., & Lee, J. Y. (2009). *Receptive & expressive vocabulary test (REVT)*. Seoul: Seoul Community Rehabilitation Center.
- Kim, Y. T., Park, S. H., & Lee, H. R. (2005). A qualitative study for development of a Korean child language assessment model for preschool children with language disorders. *Korean Journal of Communication Disorders, 10*, 24-40.
- Kim, Y. T., Sung, T. J., & Lee, Y. K. (2003). *Preschool receptive & expressive scale (PRES)*. Seoul: Seoul Community Rehabilitation Center.
- Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A. H., Gustafsson, J. E., & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental Science, 18*, 146-154.
- Leclercq, A., Quemart, P., Magis, D., & Maillart, C. (2014). The sentence repetition task: a powerful diagnostic tool for French children with specific language impairment. *Research in Developmental Disabilities, 35*, 3423-3430.
- Lee, J. (2011). Size matters: early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics, 32*, 69-92.
- Lee, J., Choi, S., & Hwang, M. (2014). Production of case-markers during sentence repetition in Korean children with specific language impairment. *Communication Science & Disorders, 19*, 477-485.
- Lee, Y. (2006). A study for the development of language assessment model for Korean school-age children with language disorders: a qualitative inquiry. *Korean Journal of Communication Disorders, 11*, 30-50.
- Lee, Y., Yim, D., & Sim, H. (2012). Phonological processing skills and its relevance to receptive vocabulary development in children with early cochlear implantation. *International Journal of Pediatric Otorhinolaryngology, 76*, 1755-1760.
- Lugo-Neris, M. J., Peña, E. D., Bedore, L. M., & Gillam, R. B. (2015). Utility of a language screening measure for predicting risk for language impairment in bilinguals. *American Journal of Speech-Language Pathology, 24*, 426-437.
- MacRoy-Higgins, M., & Dalton, K. P. (2015). The influence of phonotactic probability on nonword repetition and fast mapping in 3-year-olds with a history of expressive language delay. *Journal of Speech, Language, and Hearing Research, 58*, 1773-1779.
- Martin, K., & Schwartz, R. G. (2003). Working memory capacity and language processes in children with specific language impairment. *Journal of Speech,*

- Language, and Hearing Research*, 46, 1138-1153.
- McKean, C., Letts, C., & Howard, D. (2013). Developmental change is key to understanding primary language impairment: the case of phonotactic probability and nonword repetition. *Journal of Speech, Language, and Hearing Research*, 56, 1579-1594.
- McNamara, J. K., & Wong, B. (2015). Memory for everyday information in students with learning disabilities. *Journal of Learning Disabilities*, 36, 394-406.
- Melby-Lervag, M., & Lervag, A. (2012). Oral language skills moderate nonword repetition skills in children with dyslexia: a meta-analysis of the role of nonword repetition skills in dyslexia. *Scientific Studies of Reading*, 16, 1-34.
- Menghini, D., Finzi, A., Carlesimo, G., & Vicari, S. (2011). Working memory impairment in children with developmental dyslexia: is it just a phonological deficit? *Developmental Neuropsychology*, 36, 199-213.
- Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visual spatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, 130, 621-640.
- Moon, S. B., & Byun, C. J. (2003). *Korean Kaufman assessment battery for children (K-ABC)*. Seoul: Hakjisa.
- Morgan, P. L., Farkas, G., Hillemeier, M. M., & Hammer, C. S. (2015). 24-month-old children with larger oral vocabularies display greater academic and behavioral functioning at kindergarten entry. *Child Development*, 86, 1351-1370.
- Newcomer, P. L., Hammill, D. D. (1997). *Test of Language Development (TOLD-P3)*. Austin, TX: Pro-Ed.
- Oh, D., & Yim, D. (2013). Non-word repetition and sentence repetition performance in 2-3 years old late talkers and normal children. *Communication Science & Disorders*, 18, 277-287.
- Pae, S., Kim, M., Yoon, H. J., & Jahng, S. (2015). *Korean language based reading assessment (KOLRA)*. Seoul: Hakjisa.
- Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N., Smith, A. D., & Smith, P. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging*, 17, 299-320.
- Park, W. J., Yoon, S. R., Han, B. Y., & Yim, D. (2014). A comparison of scoring methods on the sentence repetition test in Korean children with delayed language development. *Journal of Speech-Language & Hearing Disorders*, 23, 17-29.
- Petrucelli, N., Bavin, E. L., & Bretherton, L. (2012). Children with specific language impairment and resolved late talkers: working memory profiles at 5 years. *Journal of Speech, Language, and Hearing Research*, 55, 1690-1703.
- Pickering, S. J., & Gathercole, S. E. (2001). *Working memory test battery for children*. London: Psychological Corporation Europe.
- Plante, E., & Vance, R. (1994). Selection of preschool language tests. *Language, Speech, and Hearing Services in Schools*, 25, 15-24.
- Poll, G. H., Betz, S. K., & Miller, C. A. (2010). Identification of clinical markers of specific language impairment in adults. *Journal of Speech, Language, and Hearing Research*, 53, 414-429.
- Redmond, S. M. (2005). Differentiating SLI from ADHD using children's sentence recall and production of past tense morphology. *Clinical Linguistics and Phonetics*, 19, 109-127.
- Redmond, S. M., Thompson, H. L., & Goldstein, S. (2011). Psycholinguistic profiling differentiates specific language impairment from typical development and from attention-deficit/hyperactivity disorder. *Journal of Speech, Language, and Hearing Research*, 54, 99-117.
- Riccio, C. A., Cash, D. L., & Cohen, M. J. (2007). Learning and memory performance of children with specific language impairment (SLI). *Applied Neuropsychology*, 14, 255-261.
- Riches, N. G. (2012). Sentence repetition in children with specific language impairment: an investigation of underlying mechanisms. *International Journal of Language & Communication Disorders*, 47, 499-510.
- Rispens, J., & Baker, A. (2012). Nonword repetition: the relative contributions of phonological short-term memory and phonological representations in children with language and reading impairment. *Journal of Speech, Language, and Hearing Research*, 55, 683-694.
- Rispens, J., Baker, A., & Duinmeijer, I. (2015). Word recognition and nonword repetition in children with language disorders: the effects of neighborhood density, lexical frequency, and phonotactic probability. *Journal of Speech, Language, and Hearing Research*, 58, 78-92.
- Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals, fourth edition (CELF-4)*. San Antonio, TX: The Psychological Corporation.
- Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: an individual difference approach. *Journal of Experimental Psychology: General*, 125, 4-27.
- Stokes, S. F., Wong, A. M. Y., Fletcher, P., & Leonard, L. B. (2006). Nonword repetition and sentence repetition as clinical markers of specific language impairment: the case of Cantonese. *Journal of Speech, Language, and Hearing Research*, 49, 1031-1041.

- ing Research*, 49, 219-236.
- Storch, S. A., & Whitehurst, G. J. (2002). Oral language and code-related precursors to reading: evidence from a longitudinal structural model. *Developmental Psychology*, 38, 934-947.
- Tattersall, P. J., Nelson, N. W., & Tyler, A. A. (2015a). A pilot study comparing two nonword repetition tasks for use in a formal test battery. *Communication Disorders Quarterly*, 36, 172-176.
- Tattersall, P. J., Nelson, N. W., & Tyler, A. A. (2015b). Associations among nonword repetition and phonemic and vocabulary awareness: implications for intervention. *Child Language Teaching and Therapy*, 31, 159-171.
- Tyler, A. A., & Tolbert, L. C. (2002). Speech-language assessment in the clinical setting. *American Journal of Speech-Language Pathology*, 11, 215-220.
- Vandierendonck, A., & Szmalec, A. (2004). An asymmetry in the visuo-spatial demands of forward and backward recall in the Corsi blocks task. *Imagination, Cognition and Personality*, 23, 225-231.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *CTOPP-2: Comprehensive test of phonological processing*. Austin, TX: Pro-Ed.
- Walker, D., Greenwood, C., Hart, B., & Carta, J. (1994). Prediction of school outcomes based on early language production and socioeconomic factors. *Child Development*, 65, 161-172.
- Wechsler, D. (1989). *Wechsler Preschool and Primary Intelligence Scale-Revised*. San Antonio, TX: Psychological Corporation.
- Windsor, J., Kohnert, K., Lobitz, K. F., & Pham, G. T. (2010). Cross-language nonword repetition by bilingual and monolingual children. *American Journal of Speech-Language Pathology*, 19, 298-310.
- Yang, Y., Yim, D., Kim, S., & Han, J. (2013). The relationship among receptive vocabulary, non-word repetition, and quick incidental learning in preschoolers with and without delay in vocabulary development. *Communication Sciences & Disorders*, 18, 379-391.
- Yim, D. (2017). Common versus unique findings on processing-based task performance in Korean speaking children with cochlear implants. *Otology & Neurotology*, 38, 339-344.
- Yim, D., Jo, Y., Han, J., & Sung, J. (2016). Executive function in Korean-English bilingual children with and without vocabulary delay. *Communication Sciences & Disorders*, 21, 472-487.
- Yim, D., Kim, S., & Yang, Y. (2015). Factor analysis of working memory tasks based on information processing characteristics: Predictive factors of receptive vocabulary and quick incidental learning in children with typically developing and receptive vocabulary delay. *Communication Sciences & Disorders*, 20, 304-318.
- Yim, D., Kim, Y., & Yang, Y. (2016). Exploring the utility of verbal and visuo-spatial working memory for identifying children with language impairment. *Communication Sciences & Disorders*, 21, 193-205.

## 국문초록

### 채점단위 및 수행조건에 따른 작업기억 과제의 언어장애 진단정확도 비교 연구

임동선 · 김신영 · 유지원 · 이윤정 · 이상연 · 정하은

이화여자대학교 대학원 언어병리학과

**배경 및 목적:** 비단어따라말하기, 문장따라말하기, 매트릭스는 언어능력의 기저요인으로 주목 받아온 작업기억의 각 하위체계를 평가하기 위한 과제로, 많은 선행연구들과 임상에서 사용되어 왔다. 그러나 각 과제의 채점단위 또는 수행조건이 일치되지 않아왔던 바, 본 연구에서는 각 과제별로 언어장애를 보다 정확하고 효과적으로 판별할 수 있는 방식을 재검토하고 정확한 시행방법을 제시하고자 한다. **방법:** 본 연구는 5-6세의 어휘발달지체 아동(N=40)과 일반 아동(N=40)을 대상으로 비단어따라말하기 및 문장따라말하기의 채점단위 및 매트릭스 과제의 수행조건에 따른 진단정확도를 민감도, 특이도, 양성우도비 및 음성우도비를 통하여 분석하였다. **결과:** 판별분석 결과, 비단어따라말하기 과제는 음절 단위 채점방식에 비해 낱말 단위 채점방식이, 문장따라말하기 과제는 어절 단위 채점방식에 비해 문장 단위 채점방식의 진단정확도가 더욱 높은 것으로 나타났다. 매트릭스 과제는 집단 간 수행력 차이와 판별분석 결과가 순행 및 역행 조건 모두에서 유의하지 않았다. **논의 및 결론:** 비단어따라말하기 과제 및 문장따라말하기 과제는 시간적 효율성이 높은 채점 방식이 진단정확도도 높은 것으로 나타나 그 임상적 효용성을 입증하였다.

**핵심어:** 비단어따라말하기, 문장따라말하기, 매트릭스, 진단정확도, 판별분석, 작업기억

## 참고문헌

- 강은희, 강대옥(2016). 학령 전 단순언어장애 아동의 비단어따라말하기와 문장따라말하기 수행능력 특성. *특수교육재활과학연구*, 55, 559-573.
- 김동일 (2011). *기초학습기능 수행평가체제: 초기문해*. 서울: 학지사 심리검사연구소.
- 김애화, 김의정, 황민아, 유현실(2014). 읽기 성취 및 읽기 인지처리능력 검사(RA-RCP). 서울: 학지사.
- 김영태(2014). *아동언어장애의 진단 및 치료*. 서울: 학지사.
- 김영태, 박소현, 이희란(2005). 학령전 언어장애아동 진단모델 정립을 위한 질적 연구. *언어청각장애연구*, 10, 24-40.
- 김영태, 성태제, 이윤경(2003). *취학전 아동의 수용언어 및 표현언어발달척도(PRES)*. 서울: 서울장애인종합복지관.
- 김영태, 홍경훈, 김경희, 장혜성, 이주연(2009). *수용·표현어휘력검사(REVT)*. 서울: 서울장애인종합복지관.
- 김정숙, 정승문(2011). 문장의 길이와 구조에 따른 3-5 세 아동의 문장 따라말하기 수행력. *언어치료연구*, 20, 19-36.
- 문수백, 변창진(2003). *한국 카우프만 아동지능검사(K-ABC)*. 서울: 학지사.
- 박원정, 윤사라, 한보연, 임동선(2014). 한국어 문장따라말하기 검사의 점수체계 비교 연구. *언어치료연구*, 23, 17-29.
- 배소영, 김미배, 윤효진, 장승민(2015). *한국어 읽기 검사(KOLRA)*. 서울: 학지사.
- 안지숙, 김영태(2000). 단순언어장애 아동과 정상 아동의 구문적 난이도에 따른 문장따라말하기: 수행력 및 명료도 비교. *음성과학*, 7, 249-262.
- 양윤희, 임동선, 김신영, 한지운(2013). 학령 전 어휘발달지체 및 일반 아동의 비단어 따라 말하기, 빠른 우연학습(Quick Incidental Learning)과 수용 어휘와의 관계. *언어청각장애연구*, 18, 379-391.
- 오다연, 임동선(2013). 2-3세 말 늦은 아동과 정상 아동의 비단어따라말하기와 문장 따라말하기 수행 능력. *언어청각장애연구*, 18, 277-287.
- 이윤경. (2006). 학령기 아동언어장애 진단 및 평가에 관한 질적연구: 진단 및 평가 모형 정립을 위한 기초연구. *언어청각장애연구*, 11, 30-40.
- 이은주. (2010). 국외화용언어평가도구분석. *언어치료연구*, 19, 109-132.
- 이정미, 최소영, 황민아(2014). 문장따라말하기에서 나타난 단순언어장애아동의 조사 처리 능력. *언어청각장애연구*, 19, 477-485
- 임동선, 김신영, 양윤희(2015). 정보처리 특성에 따른 작업기억 과제의 탐색적 요인분석. *언어청각장애연구*, 20, 304-318.
- 임동선, 김영태, 양윤희(2016). 언어장애아동 판별을 위한 구어 및 시공간적 작업기억의 효용성 탐색. *언어청각장애연구*, 21, 193-205.

- 임동선, 조연주, 한지윤, 성지민(2016). 한국어-영어 이중언어사용 아동의 어휘발달지연 유무에 따른 집행기능 비교. *언어청각장애연구*, 21, 472-487.
- 홍성미, 임동선(2014). 학습 및 기억과제를 통한 이중언어아동의 언어발달지체 평가. *언어청각장애연구*, 19, 31-44.
- 황민아(2003). 단순언어장애아동의 문장 이해: 단서이용 양상을 중심으로. *언어청각장애연구*, 8, 1-21.
- 황민아(2012). 문장따라말하기 검사의 학령 전 단순언어장애 진단 정확도. *언어청각장애연구*, 17, 1-14.
- 황민아(2014). 읽기이해부진 아동의 작업기억 특성: 문장 따라말하기 및 비단어 따라말하기 검사 중심으로. *학습장애연구*, 11, 53-72.