The Use of Nonlinguistic Statistical Learning as a Clinical Marker in Bilingual Children

Dongsun Yim§
Department of Communication Disorders, Ewha Womans University

§Correspondence to
Prof. Dongsun Yim, PhD,
Department of Communication Disorders, Ewha Womans University, 11-1, Dachyun-dong, Seodaemun-gu, Seoul, Korea
e-mail: sunyim@ewha.ac.kr
tel.: +82 2 3277 2120

Background & Objectives: The goal of this study was to explore the value of nonlinguistic statistical learning tasks that can be used to more accurately identify Language Impairment (LI) in children with diverse language learning experiences. The theoretical rationale for investigating statistical learning is that it has been found to be a basic cognitive learning mechanism that is critical for language learning. First, performances in visual and auditory statistical learning were compared between typically developing bilingual and monolingual children. Second, based on Spanish and English standardized tests, bilingual children were divided into two groups: children who scored within the normal range and those who did not achieve the normal range. The performances of the two groups on visual and auditory statistical learning tasks were compared. Methods: Participants were 37 Spanish-English bilingual children and 33 native English speaking monolingual children. They all met the criteria for normal development based on nonverbal IQ and a parental questionnaire. All children completed visual and auditory statistical learning tasks in addition to standard language tests. Results: The bilingual children performances were comparable to those of the monolingual children on visual and auditory nonlinguistic statistical learning tasks. These findings illustrate that, if there is no internal error in processing, performance involving statistical learning should be normal even when the subjects were systematically exposed to two different languages. When the bilingual children were divided into two groups based on standard English language scores (below average vs. above average), there were no significant differences regarding visual or auditory statistical learning. These findings held constant when the group was divided based on the ability to speak Spanish. These results supported the current diagnostic issues of bilinguals in which certain normally developing bilingual children perform poorly on standardized language tests. In addition, our results suggest that statistical learning is a more accurate assessment of underlying cognitive-linguistic processing. Discussion & Conclusion: These study results indicate that nonlinguistic statistical learning reduces the bias toward bilingual children. These research findings foster our ability to answer practical questions and identified a new way to accurately identify LI in bilingual children. This study performed the initial step of providing information on fundamental learning aspects of nonlinguistic cognitive areas in children and explored a stronger theoretical basis for accurate assessment practices in bilingual children. (Korean Journal of Communication Disorders 2011;16;13-22)

Key Words: nonlinguistic statistical learning, bilingual children, accurate (non-biased) assessment

I. Introduction

In order to identify Language Impairment (LI), standardized assessment tools are generally used. However, these standardized assessment tools are heavily influenced by previous knowledge which may be confounded by exposure to a particular language (Kohnert, Windsor & Yim, 2006). Additionally, these measures rely on static abilities such as vocabulary size that may reflect a child's
language capability but are confounded by exposure to a particular language, educational and social experiences. Thus, performance on these standardized tests is found to be biased for children with diverse cultural and language backgrounds (Campbell et al., 1997; Dollaghan & Campbell, 1998; Paradis & Crago, 2000; Windsor & Kohnert, 2004). As a result, these bilingual children from culturally and linguistically diverse backgrounds may be misidentified as learning difficulties or underidentified for LI (Silliman, Wilkinson & Brea-Spahn, 2004). Thus, bilingual children’s performance on standardized assessment tools may be informative but are not always accurate.

A lack of accurate assessment tool can be partly solved by using a parental report. It has been found that a parental report is a strong and reliable measure to identify LI in bilingual children (Paradis, Emmerzael & Duycan, 2010; Restrepo, 1998). Restrepo (1998) examined parent questionnaire data along with direct observation of child’s language abilities data for testing the best discriminators of LI among bilingual children. Findings supported that the questions given to parents including child’s current home language abilities, family history of speech, language best predicted the children affected with LI. Most recent study by Paradis, Emmerzael & Duycan (2010) developed and investigated whether using parent report on first language development for bilingual children can differentiate between language differences and language disorders. Findings indicated that the scores obtained from this parental report could be useful for assessment in bilingual children in which speech-language pathologist cannot directly examine children’s first language.

An increasing body of empirical literature has emphasized the underlying cognitive constructs in monolingual children with LI. Along with these findings, there is a growing attention on language assessment procedures that also consider the role of cognitive-linguistic underpinnings as opposed to linguistic knowledge. One of the outcome of considering cognitive-linguistic components in assessment is ‘linguistic processing-dependent tasks’ which attempts to measure the integrity of the underlying language learning or cognitive processing mechanism rather than language knowledge. These linguistic processing-dependent measurements have been proposed as potentially nonbiased alternatives to traditional standardized tests for bilingual children. Linguistic processing-dependent measurements minimize the roles of prior experience or knowledge by using basic linguistic units that are equally familiar to participants (high frequency vocabulary) or equally unfamiliar to participants (nonsense words). The basic idea is to minimize the role of prior language experience that may have on performance. It was found that these linguistic processing-dependent measurements were reliable at identifying LI in culturally diverse children such as children who use dialect such as African-American or culturally diverse children (Dollaghan & Campbell, 1998; Ellis Weismer et al., 2000). However, these linguistic processing-dependent tasks were still found to be biased for children who use two languages who are linguistically diverse (Kohnert, Windsor & Yim, 2006).

Recent findings support that children with LI show subclinical limitations in processing tasks that require little or no language ability (Leonard, 1998). For example, relative slowness and/or lower accuracy by children with LI on a range of nonlinguistic tasks, including perceptual-motor tasks such as auditory and visual detection and more complex cognitive tasks such as mental rotation and visual search have been documented by several researchers (Johnston & Ellis Wesimer, 1983; Kohnert & Windsor, 2004; Leonard et al., 2007; Miller et al., 2001; Windsor et al., 2008). Certain cognitive ability is more closely related to language skills and as a result, responsible for their language outcomes later on. Thus, based on the interactive processing information theory, nonlinguistic processing-dependent task was developed which was another outcome of emphasizing the integrity of the underlying language learning (Windsor et al., 2008; Yim, Kohnert & Windsor, 2005).

It was found that some nonlinguistic processing-dependent tasks were able to differentiate children with LI from typically developing bilingual children at a group level but not at an individual level. It is clear that not all nonlinguistic processing tasks are equally
sensitive and specific to the presence of LI (Kohnert & Windsor, 2004). Additionally, there are still unanswered questions concerning whether nonlinguistic processing tasks will be less biased than traditional measures (Yim, Kohnert & Windsor, 2005). Thus, it is crucial to find out which nonlinguistic processing task can rule in and to rule out LI in both monolingual and bilingual children. It is also important to test not any random nonlinguistic tasks but theoretically driven nonlinguistic tasks that predict language skills in children.

The theoretical rationale for using statistical learning is described as follows. As opposed to a theoretical view in which language is considered autonomous, general interactive information-processing approaches support the idea that basic cognitive mechanisms need to be integrated in order to efficiently learn and use language (Elman et al., 1996). One basic cognitive learning mechanism that has been proposed to underlie both language and nonlinguistic performance is statistical learning, that is the ability to learn new information in which rules are embedded incidentally or without explicit instruction (Gomez & Gerken, 2000; Marcus et al., 1999; Saffran, Aslin & Newport, 1996).

Statistical learning appears to be an important feature in the development of complex abilities such as language, social, and motor skills. It has been suggested that statistical learning is a fundamental cognitive mechanism that allows us to learn not only language but also nonlinguistic information, such as patterns of tones or visual shapes (Fiser & Aslin, 2002; Saffran et al., 1999). Ullman and colleagues propose that LI is the result of difficulty in general procedural learning, specifically, the learning of rule-based grammar (Ullman & Gopnik, 1999; Ullman & Pierpont, 2005). In this declarative/procedural model of language acquisition, syntactic ability is gained through implicit learning, in this case procedural learning. One of the methods of testing procedural learning is statistical learning and this learning system allows for learning of rule-like relations required for grammar learning. Recently, a direct link between statistical learning and language skills especially overall language performance including grammar has been found (Conway et al., 2010; Yim & Windsor, 2010).

Many studies have tested whether language difficulties for children with LI (Connell & Stone, 1992; Kiernan & Snow, 1999; Swisher et al., 1995) and reading difficulties (Catts & Kamhi, 1986; Scarborough & Dobrich, 1990; Vicari et al., 2003) are due to inefficient rule extraction, and found that children with LI and/or reading difficulties were due to statistical learning difficulties.

More recent findings from Evans, Saffran & Robe-Torres (2009) showed that children with LI performed statistically significantly poorly than typically developing children on both linguistic and nonlinguistic statistical learning. Based on the procedural model of language acquisition (Ullman & Gopnik, 1999; Ullman & Pierpont, 2005), Evans et al. (2009) concluded that children with LI have limited general procedural learning which is the reason why we observe their difficulties in both linguistic and nonlinguistic information processing. Thus, it is evident that statistical learning is a fundamental learning mechanism that is responsible for efficient language learning and is worthwhile pursuing for its value as an assessment measurement.

The purpose of this study is first to examine whether typical bilingual children can perform comparable to monolingual children on nonlinguistic statistical learning and second to investigate whether bilingual children who perform poorly on standardized language tests would perform similarly on the experimental tasks to those who scored within normal limits on these language tests.

Thus, as a ground work, we will explore typically developing bilingual children compared to typically developing monolingual children on visual and auditory statistical learning. First, we will examine whether bilingual children perform comparable to monolingual children. Our hypothesis is that if statistical learning truly taps the underlying cognitive construct of language learning mechanism, typically developing children should perform comparable whether they are monolinguals or bilinguals. Second, we will then investigate how standardized measurement scores can be biased via examining bilingual
children who scored less than 1.5 SD below the mean and look like LI. We will compare bilingual children who scored within normal range on language tests and those who scored less than normal on statistical learning. We hypothesize that bilingual children with the face of LI will perform as well as children who are typical on standardized measurement on statistical learning since they are normally developing children.

II. Methods

1. Participants

A total of 70 typically developing children who were aged between 8;1 and 13;11 completed the experimental tasks. Children were recruited from Chicago metropolitan school district area. Thirty-three children were English native speaking monolingual children, hereafter MO, with no medical concerns or any other cognitive and/or language related difficulties (Mean age = 8;10, SD = 2;3). The other thirty-seven children were bilingual children, hereafter BI, who spoke Spanish from birth as their mother tongue and were exposed to English from 3 at school settings (Mean age = 8;2, SD = 2;6). Based on parental reports (Restrepo, 1998), all children were typically developing with no academic issues and passed hearing screening (pure tones presented at 25 dB at 1, 2, and 4 kHz). All participants showed a nonverbal intelligence test score within the normal range on the Leiter International Performance Scale-Revised (Roid & Miller, 2002). BI had a mean Leiter standard score of 101 (SD = 15) and MO had a mean of 111 (SD = 11).

Previous studies (Conway et al., 2010; Yim & Windsor, 2010) found that language skills especially knowledge learned for a long time was correlated with statistical learning. In this study, a standardized test, the Clinical Evaluation of Language Fundamentals-IV (CELF)(Semel, Wiig & Secord, 2003) in Spanish (CELF-S) and English (CELF-E) were used to assess spoken semantic and syntactic knowledge, with the total standard score reflecting both receptive and expressive language skills. Trained research assistants who were fluent in both Spanish and English ran subjects for the whole experiments. Table 1 shows the group characteristics on age, LEITER and CELF-E and CELF-S.

<table>
<thead>
<tr>
<th>Age (in months)</th>
<th>LEITER</th>
<th>CELF-E</th>
<th>CELF-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI (N=37)</td>
<td>98 (30)</td>
<td>101 (15)</td>
<td>96 (19)</td>
</tr>
<tr>
<td>MO (N=33)</td>
<td>104 (27)</td>
<td>111 (11)</td>
<td>117 (13)</td>
</tr>
</tbody>
</table>

2. Experimental Tasks

Experimental tasks used in this study were very similar to Yim & Windsor (2010) study (review, Yim & Windsor, 2010) with minor changes. First, for the visual statistical learning, 12 shapes were directly used from Yim & Windsor (2010) but with more randomly reorganized way. This was done because previous study followed a systematic way of grouping each shape within triplets even though no participants recognized the pattern. However, in order to eliminate the pattern within the triplet, we reorganized each shape randomly. For the auditory statistical learning, we eliminated the procedure of coloring picture during listening to the stimuli. This was done because there may be attention shift issues based on the previous findings.

In each of visual statistical learning and auditory statistical learning there was a training session and a test session. In the training session, participants were exposed to sequences of stimuli for several minutes. This was followed by a test session in which participants were tested on whether they had learned the statistical regularities. Visual and Auditory statistical learning were presented on a desktop computer with MATLAB using the Psychophysical Toolbox (Brainard, 1997; Pelli, 1997). All tasks were administered in a counterbalanced order on the same day, but with the language and nonverbal IQ tasks separating the experimental tasks.

There was a predictable sequence of each stimulus presentation. For example, for the A-B-C triplet,
whenever shape A appeared on the screen, shape B immediately followed, followed immediately by shape C. After C, the first shape from one of the three other triplets (chosen at random) appeared next followed by other members of that base pair. The same base triplet never appeared twice consecutively, and the same sequence of two base triplets also never appeared twice consecutively. This paradigm was exactly the same in the auditory statistical learning task. The within-triplet transitional probability was 1.0 and the between-triplet probability was 0.33.

2.1. Visual Statistical Learning

During a training session, a continuous 4.5 minute ‘movie’ of 12 sequenced non-namable shapes was presented on the computer for both groups. Participants were told to pay attention so that they can answer some of the questions related to these shapes afterwards. After the training session, children were shown two shape triplet sets consecutively (presented in random order), and asked to press a response button to indicate the triplet set that looked familiar based on what they had seen during training. Twenty-four test pairs were used, presented in the same format as in the training session. In each, one of the pairs was a triplet shown during the training session and one was an impossible triplet which had never occurred in the training movie. The dependent variable was percentage accuracy in identifying the 24 base triplets.

2.2. Auditory Statistical Learning

During a training session, a 2.4 min sound stream was used. Auditory stimuli were 330ms in duration with different frequencies composed of different types of sounds within the 12 sounds: four different steady-state tones, four glide tones, and four different noises. All sound files were digitized at 22.05 kHz, with 16-bit quantization. Participants heard stimuli under headphones and were asked to listen to sounds. Participants were instructed to pay attention to the sounds, and to be ready to answer questions at the end of the session. After the training session, participants listened to pairs of triplet sounds and were asked to press a response button to indicate the triplet sounds that sounded familiar based on what they had heard in training. Twenty-four test pairs were used, each with one possible and one impossible triplet. The dependent variable was percentage accuracy in identifying the 24 base triplets.

3. Analyses

First, ANCOVA with age as a covariate was used to compare BI and MO on visual and auditory statistical learning. The reason for covarying out the age was because it was found to be correlated with the performance of statistical learning (Evans et al., 2009; Yim & Windsor, 2010). Then we divided the BI group into high CELF-E scores vs. low CELF-E scores based on standardized test scores to find out whether these groups differ on statistical learning. With the same method, we also used CELF-S scores to divide the BI group again into high CELF-S and low CELF-S.

III. Results

1. Group performance on visual and auditory statistical learning

The children’s performance on the visual and auditory experimental tasks is summarized in Table 2. BI children’s mean accuracy for visual statistical learning was 55% and 57% for MO. Auditory statistical learning mean was 56% for BI and 58% for MO. There was no significant difference between BI and MO on both visual and auditory statistical learning ($F_{(1, 60)} = .174, p > .05$ for visual and $F_{(1, 60)} = .41, p > .05$ for auditory).

Table 2. Mean percent accuracy on visual statistical learning and auditory statistical learning in both groups

<table>
<thead>
<tr>
<th></th>
<th>Visual statistical learning (%)</th>
<th>Auditory statistical learning (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BI (N=37)</td>
<td>55 (17)</td>
<td>56 (13)</td>
</tr>
<tr>
<td>MO (N=33)</td>
<td>57 (20)</td>
<td>58 (11)</td>
</tr>
</tbody>
</table>

Note. Standard deviations are given in parentheses.
2. The effect of language scores on statistical learning in the bilingual group

Even though BI were typically developing children based on parental reports, some children performed 1.5 SD below the mean on CELF-E and CELF-S which frequently occur in this population. As noted earlier, BI was divided into two groups. First they were divided based on the CELF-E standard scores and their statistical learning was compared. Second, we used CELF-S standard scores to divide BI group and compared their statistical learning.

There were 7 BI who scored less than 85 which was 1.5SD below the mean, hereafter delay CELF-E group, and the rest 30 children (normal CELF-E group) scored above the mean on CELF-E. Figure 1 shows the overall group performance on visual and auditory statistical learning for these two groups. The mean of the delay CELF-E group on visual statistical learning was 64% (SD = 21) and normal CELF-E group was 53% (SD = 16) which was not significantly different ($F_{(1, 36)} = 2.8, p > .05$). Additionally, the mean of the auditory statistical learning for the delay CELF-E group was 53% (SD = 14) and that of normal CELF-E was 56% (SD = 12) which was not statistically significantly different ($F_{(1, 36)} = .72, p > .05$).

![Figure 1. BI group (Delayed CELF-E vs. Normal CELF-E) performance on visual and auditory statistical learning.](image)

Then for the second analysis, we used CELF-S scores to divide the group. There were 10 children who were below 1.5 SD (delayed CELF-S group) and the rest 27 children (normal CELF-S group) scored above the mean. Figure 2 shows the group performance on statistical learning. The mean of visual statistical learning for the delayed CELF-S group was 52% (SD = 15) and for the normal CELF-S was 56% (SD = 18). As found before with CELF-E score division, there was no statistically significant difference between the groups ($F_{(1, 36)} = .40, p > .05$). For the auditory statistical learning, mean of the delayed CELF-S group was 56% (SD = 13) and normal CELF-S was 55% (SD = 13) which again did not significantly differ ($F_{(1, 36)} = .30, p > .05$).

![Figure 2. BI group (Delayed CELF-S vs. Normal CELF-S) performance on visual and auditory statistical learning.](image)

IV. Discussion and Conclusion

The current study investigated the performance of bilingual and monolingual children on visual and auditory statistical learning tasks. Two primary questions were addressed. The first question was whether typical bilingual children perform similarly to typically developing monolingual children. The second question was whether bilingual children who perform poorly on standardized language tests, CELF-E and CELF-S, would perform similarly to those who scored within normal limits on these language tests. Our findings suggest that bilingual children with no abnormal internal cognitive-linguistic processing mechanism perform comparable to monolingual children even when their performance incorrectly indicate that they might have language difficulties based on standardized tests. Additionally, these results confirm biased assessment errors on standardized tests for bilingual children who are typically developing.

The challenging issue for bilingual population in assessment is that there is an overlap performance...
between monolingual children with LI and typically developing bilingual children. As a result, typical bilingual children sometimes are diagnosed as having LI. Thus, educators and professionals who deal with children with special needs report an additional challenge in effectively serving children whose first language does not match to their native language (Kohnert et al., 2003). As found in our study, even when parents reported their bilingual children were typically developing with no academic problems, some of our bilingual children performed significantly poorly than average which resulted putting them into LI category. However, when we compared those two groups divided by CELF-E and CELF-S, delay and normal on standardized test groups did not inform anything beyond biased results. Both groups performed well and not significantly differ on visual and auditory statistical learning and performance of bilingual children were comparable to that of monolingual children.

Recent studies (Leonard et al., 2007; Mainela-Arnold, Evans & Coady, 2010; Montgomery & Evans, 2009; Windsor et al., 2008) consistently have found that children with LI show sub-clinical deficits on a broad range of cognitive-linguistic tasks that require little or no language ability, such as memory for nonsense figures and visual matching. These findings indicate that the language performance deficits are not isolated but are part of a larger cognitive learning profile. In order to better understand the nature of the relationship between underlying cognitive processing and language and to develop a nonbiased assessment measure, systematic investigation that considers the process of learning linguistic and nonlinguistic information is needed.

The long term objectives of this proposed study was to design and build an understanding whether limitations in statistical learning underlie language difficulties and to further investigate whether statistical learning task may reduce or eliminate assessment bias. Overall, the proposed study has two contributions. From theoretical view, the proposed study enhances our understanding of the relationship between underlying cognitive processing and language. From practical view, this study may provide partial solution of assessment issues that is urgent for bilingual children.

Our study was the very first study to apply nonlinguistic statistical learning to bilingual children as a potential accurate measurement for detecting LI while reducing previous world knowledge or experience that can lead to an assessment bias. Our hypothesis was that if statistical learning taps the underlying learning mechanism that is fundamental for language learning then variation in linguistic experience may not interfere the performance. Our findings supported that nonlinguistic statistical learning was able to tap cognitive-linguistic underpinnings that are above and beyond the superficial language skills. However, in order to confirm the clinical usage of the statistical learning as diagnostic tools, we need to examine children with LI both in bilingual and monolingual groups. Additionally, not only group analysis but also individual analysis should be used to test the sensitivity and the specificity of statistical learning task.

Our results yielded no statistical significance between bilingual children in the language low (scored lower than -1.5 SD) and high groups. This result warrants further examination of characteristics of these participants in terms of their parental report, LEITER scores, CELF receptive and expressive language profiles. Additionally the number of participants between language score high vs. low was unevenly distributed (n=30, n=7 on CELF-E; n=27, n=10 on CELF-S). However, the current results might suggest that the standardized test results were biased against bilingual children. It should be noted that this interpretation requires caution due to the relatively small sample size coupled with the skewed distribution in the language low group.

**REFERENCES**


이중언어 사용 아동을 위한 비언어정보 통계적 학습의 진단 도구로서의 사용

임동선
이화여자대학교 언어병리학과

배경 및 목적: 본 연구는 비언어정보 통계적 학습(nonlinguistic statistical learning)이 언어경험이 다양한 이중언어 사용아동의 언어장애를 판별해 내는데 정확한 검사도구로 사용될 수 있는지를 연구하였다. 통계적 학습은 언어수행능력에 있어서 매우 근본적인 중요한 기제이다. 그리고 이론적으로, 언어경험의 유무를 언어경험에서 다루는 이유로 정확히 구분해 내지 못하는 현재 표준화검사의 단점을 보완할 수 있다는 가능성은 지니고 있다. 본 연구에서는 이중언어 사용 아동과 단일언어 사용 아동의 시각적 및 청각적 통계적 학습을 비교하였다. 두번째, 이중언어 사용 아동들 중, 표준화된 언어검사(영어, 스페인어 각각의 영역)에서 평균 이하의 점수를 받은 아동들과 평균보다 높은 점수를 받은 아동들의 통계적 학습을 비교하였다. 제1, 정상 학령기에 이중언어 사용 아동 37명과 단일언어 사용 아동 33명이 연구에 참여하였으며, 모든 아동들은 정상 인지능력을 지녔고, 부모의 사례 설문지 검사결과, 모두 정상적으로 학교생활을 하고 있으며, 언어능력도 정상적으로 발달하는 아동들이었다. 제2, 영어 표준화 검사에서 언어 수행력이 낮은 이중언어 사용 아동과 높은 이중언어 사용 아동들간에 통계적 학습 능력(시각적 및 청각적)에 유의미한 차이를 보이지 않았다. 이러한 결과는 스페인어 표준화 검사결과로 그룹을 나누었을 때도 동일한 결과를 보였다. 연구결과를 종합적으로 정리해 보면, 표준화 검사에서 이중언어 사용 아동들은 정상적으로 발달하고 있음을 보였고, 평균보다 낮은 언어능력의 수준의 점수를 나타내는 경우가 있었다. 그러나, 이러한 아동들간의 표준 인지능력에 대한 정확한 학습을 측정해 보면, 아동의 언어 경험이 유무를 보다 정확하게 판단 할 수 있었다. 그리고, 이러한 아동들은 단일언어 사용 아동과도 비슷한 수행능력을 보였다. 논의 및 결론: 본 연구 결과, 비언어정보의 통계적 학습은 이중언어 사용 아동들을 위한 진단의 문제점을 해결할 수 있는 도구로써의 정확한 정보를 제공한다는 입증을 하였다. 언어청각장애연구, 2011;16;13-22.

핵심어: 비언어정보 통계적 학습, 이중언어 사용 아동, 정확한 진단

교신저자
임동선
이화여자대학교 언어병리학과 교수
서울시 서대문구 대현동 11-1
e-mail: sunyim@ewha.ac.kr
tel: 02-3277-2120