Critical Review: Do Expressive Language Characteristics Distinguish Children with Language Impairment from Children with English as a Second Language?

Andrea Chance
M.Cl.Sc. (SLP) Candidate
University of Western Ontario: School of Communication Sciences and Disorders

This critical review examines whether expressive language characteristics are able to distinguish children with Specific Language Impairment from children who are learning English as a Second Language. Overall, the research suggests that there is significant overlap between the two groups in the amount and types of errors made on expressive language tasks such as use of tense and morphology. However, tasks less dependent on previous language such as non-word repetition can provide informative results in assessment, but is still limited in its utility as a diagnostic tool.

Introduction

Children identified as having Specific Language Impairment (SLI) show significant limitations in language skills that cannot be attributed to problems of hearing, neurological status, nonverbal intelligence or other known factors (Leonard, 1997). SLI identification is based on both inclusion and exclusion criteria including language test scores of at least 1.25 standard deviations below the mean language ability of a child's peers, and age-appropriate nonverbal intelligence. A list of criteria can become complex when having to distinguish SLI from other conditions involving similar language deficits. Children learning English as a Second Language (ESL) may show below-average linguistic abilities that mimic those of children with SLI (Paradis, 2005). Problems may manifest regarding over-identification and under-identification of SLI in children who are culturally and linguistically diverse.

The Canadian population is exponentially growing in multicultural diversity with 21.4% of children aged 0-14 being of linguistic minority, that is, have neither English nor French as their first language (Statistics Canada, 2006). This statistic implies that potentially one fifth of speech-language pathologists' caseloads will be ESL children. The amount of research available on ESL language abilities is lacking when compared to SLI. Up until 2008 there have been many thorough studies of children with SLI and their linguistic competencies (Leonard et al., 1992; Rice & Wexler, 1996; Rice, Wexler & Cleave, 1995). However, the few studies that have made comparisons between monolingual children with SLI and their typically developing bilingual peers did not focus on populations of English speakers (Paradis & Crago, 2000; Hakansson, 2001). These studies have shown that there may be some overlap between the monolingual SLI and bilingual groups with respect to their expressive language skills. Therein lies the need for further research to directly compare expressive language profiles of ESL children and children with SLI.

There are important clinical implications with regards to the possibility of erroneous assessment and inaccurate identification of typically developing ESL children as having SLI. Misidentification may lead to the inappropriate use of speech-language pathology resources. Conversely, there also lies the possibility of under-identifying ESL children who have SLI by not proceeding with speech-language pathology intervention. Linguistic errors may be attributed to the child being an English language learner, and not to the language impairment. The most accurate linguistic criterion must be provided to speech-language pathologists, to allow for more informed expectations of language skills during assessment, and therefore a more accurate identification of SLI.

Objectives

The primary objective of this paper is to critically evaluate the existing literature regarding the comparison of expressive language skills of typically developing ESL children and monolingual English-speaking children with SLI. Results are organized into categories of morphology and other notable expressive language measures. Recommendations regarding the use of expressive language measures in speech-language pathology assessments will also be discussed.

Methods

Search Strategy

Research articles were found using a computerized database search including ProQuest, PsycINFO, and Scopus. The following key terms and search strategies were used: ((ESL) OR (ELL) OR (bilingual)) AND (specific language impairment). The search was limited to articles written in English between 1995 and 2008.

Selection Criteria

The studies selected for this critical review paper examined the similarities and differences of expressive language characteristics of monolingual children with SLI and ESL children. Selection was limited to English as the only language in children with SLI, and English as a Second Language in multilingual children. Perceptual-motor task results will not be considered in this critical review, as the research findings focus on expressive language characteristics.

Data Collection

Results of the literature search yielded the following types of articles congruent with the aforementioned criteria: one single group design, two cohort studies, and one summary of studies. The cohort study by Paradis, Rice, Crago and Marquis (2008) was not directly related to the research question, but provided comparisons between ESL children and children with SLI through statistical analysis of expressive language assessment results.

Results

For a brief outline of relevant results from each article, refer to Appendix A. The results summarized relate specifically to the comparison of ESL children to children with SLI through expressive language measures. The main feature that arose from the literature search was morphology. Other notable features such as non-word repetition will also be featured in this critical review.

Morphology

Morphology involves the structure and form of words that give them meaning. It is a major component of expressive language. Morphology has a prominent role in forming proper grammar through bound (eg. –ed for regular past tense) and unbound markings (eg. he is playing) to give tense, plurals, possessives and other word forms.

Johanne Paradis (2005) conducted a single group study to examine whether expressive language characteristics of typically developing ESL children had similarities to that of monolingual children with SLI. Grammatical morphology accuracy and error patterns were obtained from spontaneous and elicited speech. She hypothesized that ESL children's data would show the same three error patterns in grammatical morpheme production as those found for monolingual English-speaking children with SLI. This particular pattern of error and accuracy rates was obtained from previous prominent studies conducted by Rice, Wexler, & Cleave (1995) and Rice & Wexler (1996). Paradis also hypothesized that most, if not all, of ESL children's grammatical performance on the Test of Early Grammatical Impairment (TEGI; Rice & Wexler, 2001) would fall below the range of a typically developing population.

Twenty-four typically developing ESL children aged 4 years 4 months to 7 years 10 months participated in the study. All children had nonverbal IQ in the normal range as determined by the Columbia Mental maturity Scale (CMMS; Burgemeister, Hollander Blum & Lorge, 1972), and all were administered a phonological probe from the TEGI to rule out phonological influence of native language constraining a second language learner's ability to pronounce target morphemes.

An analysis of covariance (ANCOVA) was performed to compare the means of various morphological tense composite scores as a within-subjects factor with the covariate as months of exposure to English (MOE). No significant interaction was found between morpheme scores and MOE, but a significant effect was found for the composite scores. A series of post-hoc paired two-tailed t test comparisons, and paired one-tail t tests were used to further investigate the relationship between the means of the composite scores. The composite score means of the ESL children resembled that of children with SLI on two of three patterns originally hypothesized. The apparent overlap shows the weak ability of expressive language characteristics to provide a clear distinction between the two groups. Elicited grammar composite scores of the ESL children were compared to criterion scores of the TEGI to see if they fell within SLI range individually and as a group. Nineteen of the 24 scores fell within SLI range of performance, indicating the group was below the criterion cutoff for a typically developing population. This supports the portion of Paradis' hypothesis about overall performance of ESL children being similar to SLI. Expressive language, elicited through a grammar probe, was again unable to provide a distinction between ESL children and children with SLI.

One of the merits of the study was careful control for extraneous variables. MOE, phonological constraints, and standardized tests were used to establish normal nonverbal intelligence baselines so that the scores obtained were due solely to the participant's expressive language profile, and not due to a confounding variable. Appropriate sample size was achieved to give sufficient power for a medium effect size, and the data collected was measured against SLI criterion of the TEGI, a valid, standardized formal assessment tool. The use of ANCOVA, post-hoc tests and t tests were fit to establish patterns and relationships between the composite score means of ESL children, to compare them against the patterns hypothesized for SLI.

This study provided level II experimental evidence, one level below the most ideal, 'gold standard' of experimental design. The researcher had more than one hypothesis, and not all were supported. A comparison group of SLI children tested concurrently with the ESL children would have yielded stronger, more clinically relevant results to provide more compelling evidence. A design strategy that included a comparison group may rule out possible extraneous effects of the testing procedure or environment. Despite this weakness, a clinician may consider this study's results during the language assessment process. There was a significant overlap in morphological patterns and scores obtained in a standardized assessment setting, and this should be taken into consideration during language assessments of ESL children in clinical practice.

An extensive cohort study was conducted by Paradis, Rice, Crago and Marquis (2008) that examined the pattern of acquisition of tense morphology in 24 typically developing monolingual children, 24 typically developing ESL children with various first language backgrounds and 20 children with SLI. This highly theoretical study sought to document the extent of similarities and dissimilarities between typically developing monolingual (TDL1), ESL, and SLI tense morpheme acquisition in English. The researchers wanted to determine whether English as a second language acquisition was similar to one of two hypothesized patterns. Although the study does not directly answer the research question posed by this critical review, the data yielded from the study provides important information that may be used to support this review's findings. Each group was tested at a single point in time and morphology was compared thorough statistical analysis. Observations of ESL children and children with SLI and are the focus of this review, and the number of similarities between the two populations was close to the number of differences.

Equivalence among all groups of children was determined by selecting children within the same mean length of utterance in morphemes (MLUM) range of the ESL group. Elicitation probes from the TEGI were given to all groups, followed by the TEGI grammatical judgement task administered to the ESL group and SLI group only, as the L1 group was too young to participate. Spontaneous speech samples were also taken.

Several measures were taken in effort to control extraneous variables from affecting the results. A within-groups analysis for the ESL data was conducted to see if months exposure to English (MOE) and if background in native language skewed the results. Pearson correlations were performed between each of the outcome variables for elicitation and grammatical judgement. MOE was not revealed to be influential enough to skew the results for this group. The ESL children were then divided by the presence of tense versus nontense in their native language. The Mann-Whitney U nonparametric unpaired group test was used to compare tense and nontense groups for each of the outcome variables. The researchers sought to find whether there was a pre-existing relationship between each of the outcome variables and the presence or absence of tense in a native language. None were significant at $p < .05$ level, and therefore presence or absence of tense was not deemed influential.

Level II evidence was obtained from this study. It was well designed in its efforts to control for many extraneous factors that could have potentially influenced the results by correlating MOE, and presence of tense in the native language to the outcome measures. Selection of subjects within a specific MLUM range also strengthened this study to ensure MLUM was not a factor skewing the quality of speech samples obtained. A potential limitation is that there was no mention of whether the researchers were blinded to the subjects when administering the tests or when obtaining spontaneous language samples. Through complex statistical analysis, the authors appropriately sought to fit the morphological profiles of each group into a hypothesized pattern of acquisition. In turn, the comparisons made between ESL children and children with SLI provided information for the research question presented in this review. Overlap is seen among the data yielded from expressive language testing. This overlap further blurs the line of distinction between expressive language characteristics of children with SLI and that of ESL children. Expressive language characteristics such as morphology may not have the ability to provide a robust distinction between ESL children and children with SLI, and assessment test results spawned from these measures should be interpreted with caution.

Other Notable Language Measures

Several other expressive language measures were noted in the available literature. Kohnert, Windsor, and Danahy Ebert (2008, in press) summarized their laboratory's research findings comparing typically bilingual children (BI), monolingual children with Primary Language Impairment (PLI) and typically developing monolingual children who spoke English only (EO). The goal of this summary of studies was to examine points of commonality and divergence across 17 tasks that were perceptual-motor, nonlinguistic, or linguistic in their demands. The results of the linguistic and nonlinguistic tasks were taken from several different studies mostly within the one laboratory of researchers, on the same sample of children. All participants met criterion performance for typical development in language, nonverbal intelligence, hearing, and educational measures, with the exception

of the PLI group in which their language impairment was documented. With regards to PLI and BI performance specifically, the two groups performed equivocally during the word recognition language tasks. During the confrontation naming language task, the PLI group performed similarly to the BI group until the most difficult task level where the PLI group performed better. The authors did not hypothesize this result.

This summary amalgamated the results of a group of studies belonging to the same laboratory, using the same population. No information on the statistical methods used was included to support the findings. The reader is left to seek out the previous studies referenced by the authors in order to judge its validity and reliability. This summary of studies is not the most useful in providing statistically sound evidence to support a change in clinical practice, however it is a basis to be expanded upon in future research. The comparisons made between many expressive language tasks raises questions about the uncertain potential of expressive language measures to distinguish between bilingual children and children with language impairments.

One of the studies featured in the summary of studies was a cohort study by Kohnert, Windsor and Yim (2006). It examined whether the Competing language Processing Task (CLPT) and Non-Word Repetition (NWR) task would be able to provide a clearer distinction between children with SLI and ESL children. These measures are the most common, potentially nonbiased alternatives to the traditional experience-dependent measures for the purpose of identifying a language impairment in culturally or linguistically diverse learners (Kohnert, Windsor & Yim, 2006). It was expected that the bilingual children and monolingual English-speaking children would perform significantly better than the group of children with language impairment.

A sample size of 100 participants aged 7 years 10 months to 13 years 11 months were recruited. The sample size was sufficient to detect a large effect size, as reported in this study. Each child was assigned to a group resulting in 28 monolingual English children who were language impaired (LI), 50 monolingual typically developing children who spoke English only (EO), and 22 typically developing bilingual children (BI). All passed hearing, vision and oral-motor screening and had no other history of neurological or social-behavioural concerns. Children were administered the Clinical Evaluation of Language Fundamantals-3 (CELF-3; Semel, Wiig, & Secord, 1995), the Test of Nonverbal Intelligence (TONI; Brown, Sherbenou, & Johnsen, 1997), and the Goldman-Fristoe Test of Articulation-2 (GFTA-2; Goldman & Fristoe, 2000) to establish a common baseline for each of the groups. Children were excluded from the BI group if they were receiving speech-language services or if parents expressed concern about their first language development.

For both the CLPT and NWR tasks, two sets of analyses were conducted: ANOVA was used to compare group means for each task, and likelihood ratios were calculated to see if CLPT and NWR were accurately able to identify a language impairment. First, a between-group comparison using one-way ANOVA was completed to show differences between the groups. Results from the CLPT task showed that the LI group's comprehension and recall scores were similar to both EO and BI groups. However, ANOVA revealed differences among groups in the NWR task with the EO group performing better than the LI and BI groups. The performance of the BI group was significantly greater than the LI group, but there was overlap between one, two and three-syllable nonwords. It wasn't until the four-syllable nonword level that the groups appeared to clearly separate. Second, likelihood ratios were calculated. The likelihood ratios for the CLPT task indicated limited diagnostic power for ruling in and for ruling out a language impairment. As for the NWR task, it appeared to be more informative in ruling out the presence of LI in the sample, but not for identifying LI.

It was not noted whether blinding procedures were implemented, and it is assumed that the researchers would have strengthened the validity of their experiment if they were blinded to which participant group each child belonged to when the CLPT and NWR tasks were administered. This study was characteristic of a level II study, providing the second strongest type of experimental evidence. The researchers had a broad list of inclusion and exclusion criteria, such as hearing and vision tests, and tests of articulation to rule out possible covariates. ANOVA had revealed similarities and differences between the groups, but the use of likelihood ratios gave practical meaning to the accuracy of the tests themselves.

Despite using less experience-dependent measures to separate bilingual children from monolingual children with SLI, overlap in data is still evident. The CLPT and NWR tasks were unable to show significant differences among groups in their test results. The study's findings suggest there exists a possibility that assessment of expressive language output may not acknowledge nor explain the underlying mechanisms that separate children with SLI from ESL children. For these reasons, a clinician may take this study's results into consideration in order to provide the most accurate and comprehensive assessment of expressive language in a clinical setting.

Discussion

The expressive language measures addressed in the studies reviewed by this critical appraisal were unable to provide a clear and definite distinction between children with SLI and ESL children. Overlap in data between the two groups was seen in each of the studies. However, the strength of evidence from each of these studies was limited due to level of evidence provided and methodology used to yield results. The lack of compelling evidence from these studies does not support a change in clinical practices, but the information addressed in these studies should be taken into consideration by clinicians during the assessment process.

Morphology

Neither of the studies by Paradis (2005) nor Paradis, Rice, Crago and Marquis (2008) stated whether they had employed blinding procedures, possibly contributing to experimental bias when gathering spontaneous and elicited expressive language samples from the different groups of children. However, both of the aforementioned studies used many statistical methods to control for the confounding effect of extraneous variables, so that the resulting data was not attributed to factors such as exposure to English and phonological constraints. The Paradis (2005) study would have provided more clinically relevant evidence had it employed the use of a comparison group of monolingual children with SLI. A valid comparison of expressive language within an assessment setting could have been drawn between ESL children and children with SLI. The cohort study by Paradis, Rice, Crago and Marquis (2008) was highly theoretical in nature, but served its purpose to address theories of tense acquisition among groups of ESL children and children with SLI. Due to its experimental design, this study was unable to provide the most clinically relevant evidence to suggest changes in practice in a clinical setting.

Other Notable Language Measures

The summary of studies by Kohnert, Windsor, and Danahy Ebert (2008, in press) would have provided more compelling evidence relevant to clinical practice had it used a more rigorous method of analysis and data collection. One major flaw of the summary was the absence of statistical evidence supporting each of the findings, leaving the reader to question whether the results were valid and reliable. The evidence yielded from this summary of studies would have been stronger had the researchers included data relevant to their research question, objectively collected from other laboratories in addition to their own. Experimental bias may be reduced if such methods are used, and more clinically useful evidence would have resulted.

The NWR and CLPT study by Kohnert, Windsor and Yim (2006) would have strengthened its results with the use of blinding procedures when administering the tests to each of the children. The examiners' potential biases during testing would have been minimized, limiting the possibility of skewing the results. Despite this weakness, the study provided valuable information regarding the usefulness of NWR and CLPT as tools of identification of a language impairment.

Recommendations

Future research may answer the question posed in this paper through stronger experimental design and employment of improved methodology. In order to provide more compelling evidence to affect change in clinical practice, future research should consider the following suggestions:

a) Future experimental studies should concurrently assess the expressive language profiles of both ESL children and children with SLI. Direct statistical comparison between the two groups receiving the same assessment protocol will provide the most valid and clinically relevant results to answer the research question.

b) Blinding procedures should be implemented within the experimental methodology. Expressive language tests and spontaneous speech samples often include interaction between the examiner and the child, and blinding will help reduce the incidence of bias that may skew the results should the examiner know which participant group child belongs to.

c) Systematic reviews should be included in future research. An analysis of data from a collection of empirical studies through objective search criteria may provide a more valid, less biased view of all available evidence concerning the research question.

d) Further studies examining expressive language characteristics should expand to include semantics, syntax, and vocabulary and other aspects of expressive language that were not prevalent in the literature to date. Availability of evidence in these areas was limited.

Clinical Implications

The evidence provided by the studies included in this critical review was not strong enough to advocate a change in clinical practice. However, the overlap seen in expressive language characteristics implies that there is a potential for misidentification. ESL children may be over-identified as having SLI should the clinician see their expressive language errors as part of second language learning, not the language impairment. Conversely, ESL children who have SLI may be under-identified should their errors due to language impairment be attributed to the process of second language learning. Since a definite distinction between the two groups was not seen, it is recommended that clinicians strive to complete the most thorough and comprehensive assessment possible to give the most accurate assessment findings.

References

Brown, L., Sherbenou, R. & Johnsen, S. (1997). *Test of Nonverbal Intelligence-3*. Austin, TX: PRO-ED.

Burgemeister, B., Hollander Blum, B. & Lorge, I. (1972). *Columbia mental maturity Scale.* New York: The Psychological Corporation.

Dunn, L. & Dunn, L. (1981). *Peabody Picture Vocabulary test-revised (PPVT-R)*. Circle Pines, MN: American Guidance Service.

Goldman, R. & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation-2*. Circle Pines, MN: American Guidance Service.

Hakansson, G. (2001). Tense Morphology and Verb-Second in Swedish L1 Children, L2 Children and Children with SLI. *Bilingualism: Language and Cognition, 4*(1), 85-99.

Kohnert, K., Windsor, J., & Danahy Ebert, K. (in press). Primary or "specific" language impairment and children learning a second language. *Brain and Language*, doi:10.1016/j.bandl.2008.01.009.

Kohnert, K., Windsor, J., & Yim, D. (2006). Do Language-Based Processing Tasks Separate Children with Language Impairment from Typical Bilinguals? *Learning Disabilities Research & Practice, 21*(1), 19-29.

Leonard, L. B. (1997). *Children With Specific Language Impairment: Language, Speech, and Communication.* Cambridge, MA: MIT Press.

Leonard, L.B., Bortolini, U., Caselli, M.C. & McGregor, K.K. (1992). Morphological Deficits in Children with Specific Language Impairment: The Status of Features in the Underlying Grammar. *Language Acquisition, 2*(2), 151-179.

Paradis, J. (2005). Grammatical Morphology in Children Learning English as a Second Language: Implications of Similarities With Specific Language Impairment. *Language, Speech, and Hearing Services in Schools,36*, 172-187.

Paradis, J. & Crago, M. (2000). Tense and Temporality: A Comparison between Children Learning a Second Language and Children with SLI. *Journal of Speech, Language, and Hearing Research, 43*(4), 834-847.

Paradis, J., Rice, M., Crago, M., & Marquis, J. (2008). The acquisition of tense in English: Distinguishing child second language from first language and specific language impairment. *Applied Psycholinguistics,* 29(4), 689-722.

Rice, M., & Wexler, K. (2001). *Test of Early Grammatical Impairment,* New York: The Psychological Corporation.

Rice, M.L. & Wexler, K. (1996). Toward tense as a clinical marker of specific language impairment in English-speaking children. *Journal of Speech, Language, and Hearing Research, 39*(6), 1239-1257.

Rice, M.L., Wexler, K. & Cleave, P.L. (1995). Specific language impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research,38*(4), 850-863.

Semel, E., Wiig, E., & Secord, W. (1995). *Clinical Evaluation of language fundamentals-3*. San Antonio, TX: The Psychological Corporation.

Statistics Canada. (2006). Retrieved November 10, 2008, from http://www12.statcan.ca/english/census06/data/topics/RetrieveProductTable.cfm?Temporal=2006&PID=97298&GID=771240&METH=1&APATH=3&PTYPE=88971&THEME=70&AID=&FREE=0&FOCUS=&VID=&GC=99&GK=NA&RL=0&TPL=NA&SUB=&d1=1&d2=0

Appendix A- A Summary of Relevant Results

Expressive Language Characteristic, Profile or Task	**Similar Performance between ESL and SLI**	**Different Performance between ESL and SLI**
Paradis (2005)		
Three-part hypothesized pattern of SLI		✓
Below criterion cutoff for typically developing	✓ (majority)	
Paradis, Rice, Crago and Marquis (2008)		
Auxiliary/unbound morpheme production of BE	✓	
Auxiliary/unbound morpheme production of DO		✓
Inflectional/bound production of PLURAL [-s] and PAST TENSE [-ed]		✓
Grammatical Judgement of auxiliary and inflectional morpheme	✓	
More Accurate in use of BE compared to DO	✓	
Grammatical Detection of ungrammatical BE vs lexical targets		✓
Kohnert, Windsor, and Danahy Ebert (2008, in press)		
Confrontation naming (at most difficult level)		✓
Word Recognition	✓	
Kohnert, Windsor and Yim (2006)		
CLPT comprehension	✓	
CLPT Recall	✓	
NWR (at 4-syllable level)		✓