Automated Analysis of Natural Language Samples: Comparison of Children with ASD, DLD, and TD

Richard Sproat1,2, Lois Black1, Emily Prud’hommeaux1, Jan van Santen2, Brian Roark1
1Center for Spoken Language Understanding, Oregon Health & Science University; 2BioSpeech, Inc.
{rws, lmblack, emily, vansanten, roark}@cslu.oigi.edu

Background
There are areas of overlap between autism spectrum disorders (ASD) and developmental language disorders (DLD) that pose challenges for differential diagnosis. One such area is language impairment. There have been varying reports on the types of language impairments in ASD, their severity, and their incidence. These studies generally use structured, decontextualized instruments; yet, Language Sampling and Analysis (LSA) methods may provide information that critically complements structured instruments. Since the paucity of LSA-based studies is likely due to the labor-intensiveness of LSA, automated methods are urgently needed.

Objectives
- Demonstrate the feasibility of automating the analysis of natural language samples, focusing on the IPSyn (Scarborough, 1991).
- Apply these methods to document morphology and syntax in high functioning verbal children with ASD, children with DLD, and typically developing children (TD).

Subjects
- Children ages 4-8 given battery of language & neurocognitive measures.
- DLD group (N=20): diagnosis via Tomblin’s EpisLI criteria or a CELF index score at -1 SD plus a spontaneous language measure at -1 SD, and DSM-IV-TR-based diagnosis.
- ASD group (N=36): diagnosis via ADOS, the Social Communication Questionnaire, and DSM-IV-TR-based clinical consensus diagnosis.
- Divided into 2 subgroups: ASD+DLD group meeting DLD criteria (N=25) and ASD-DLD group not meeting DLD criteria (N=11).
- Groups were well-matched on age, but only the DLD and ASD+DLD groups were matched on VIQ and PIQ.
- Stringent exclusion criteria applied to all groups: children with neurodevelopmental disorders, neuropsychiatric disorders, or a sibling with ASD or DLD were excluded from the TD group.

Subjects

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>age</th>
<th>PIQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>TD</td>
<td>24</td>
<td>5.82</td>
<td>117</td>
</tr>
<tr>
<td>DLD</td>
<td>20</td>
<td>6.85</td>
<td>101</td>
</tr>
<tr>
<td>ASD-DLD</td>
<td>11</td>
<td>6.64</td>
<td>109</td>
</tr>
<tr>
<td>ASD+DLD</td>
<td>25</td>
<td>6.77</td>
<td>102</td>
</tr>
</tbody>
</table>

Index of Productive Syntax (IPSyn)
- Requires a corpus of 100 spontaneous sentences.
- Proposes a set of 60 (mostly) English-specific constructions based on noun phrase syntax and morphology, verb phrase syntax and morphology, sentence structure, questions and negations.
- Count the number of occurrences of each structure found in the transcript, up to a count of 2.

Manual IPSyn Annotation
- Trained speech-language pathology graduate students transcribed all child utterances from the following ADOS activities: Play, Conversation, Picture Description, Wordless Picture Book.
- Select the first 100 utterances from each child’s transcript.
- Five labelers (all undergraduate linguistics majors) scored the conversations according to the IPSyn guidelines.
- Two labelers per conversation adjudicated if the difference for any of the major categories (Noun, Verb, Question, Sentence) exceeded 4 points.

Automated IPSyn Annotation
- Probabilistic parsing algorithm applied to each utterance to produce a syntactic parse tree.
- Hand-crafted rules used to extract the 60 IPSyn categories from these parse trees.

Example:
Watch this.

Parse Structure:
\[ (\text{S} (\text{VP} (\text{VBP watch}) (\text{NP} (\text{DT this})))) \]

IPSyn Structures Extracted:
V01: Verb
  \[ (\text{VBP watch}) \]
N02: Pronoun or prolocative excluding modifiers
  \[ (\text{NP} (\text{DT this})) \]
S01: Two word combination
  \[ (\text{S} (\text{VP} (\text{VBP watch}) (\text{NP} (\text{DT this})))) \]
S03: Verb-object sequence
  \[ (\text{VP} (\text{VBP watch}) (\text{NP} (\text{DT this}))) \]

Results
- Automated system replicates human coding quite accurately.
- Automated system differed as little from manual coders as they differed from each other.

Conclusions
- The data show that the method is as accurate as human coding.
- Surprisingly, the methods showed that both ASD groups performed more poorly on IPSyn measures than one would expect based on their IQ characteristics. Specifically, the IQ of the ASD-DLD group is 1.7 SD above that of the DLD group, yet their IPSyn scores are roughly equal.
- These results argue for the importance of LSA methods to complement structured measures.
- Moreover, given the extreme labor intensiveness of LSA, these results argue for developing additional automated LSA methods.

References & Acknowledgements

Sponsors: NICDD 1R10DC007129-01 and an Autism Speaks Innovative Technology for Autism grant awarded to CSLU. NICDD 1R41DC010502-01 awarded to BioSpeech.