Expressing information status through prosody in the spontaneous speech of American English-speaking children

Jill C. Thorson¹, Jill M. Trumbell¹, and Kimberly Nesbitt¹

¹University of New Hampshire
jill.thorson@unh.edu

Abstract

Prosody is used to express information structure and status differences in American English. For this study, our motivation was to analyze these abilities during an ecologically valid interaction where we traded control for more natural spontaneous speech. We ask how children package information when playing with their parents during exhibit exploration in a children’s museum. Specifically, we employed a MAE-ToBI analysis to look at the production of new and given information status differences during these interactions. Parent-child dyads were recorded while playing in a museum exhibit at a children’s museum. Preliminary analyses were conducted on one 4-year-old, one 5-year-old, and one 6-year-old speaker. As predicted, we found a particular set of pitch accents to be commonly found as well as considerable variation in nuclear configuration patterns due to pragmatic effects. While pitch accent types largely stayed the same over the three ages analyzed to date, the H⁺H* pitch accent was only found in the speech of the 4-year-old speaker. These data continue to add to the knowledge of how pitch accent selection relates to both information status and the pragmatics of the discourse.

Index Terms: child language acquisition, information status, intonational phonology

1. Introduction

Speakers of American English package information using different aspects of language including the prosodic system, or the melody and rhythmic patterns of speech. In order to be successful interlocutors, children must acquire the guidelines of this complex prosodic system and learn how to map meaning onto intonational contours. One aspect of how we structure information is called information status. In English, one way to indicate whether a referent is new or not in a discourse context is to employ distinctions in accentuation and acoustic prominence (i.e., referential givenness-newness or the information status of a referent) [1]–[3]. Generally speaking, in American English, new information is typically accented and given information deaccented in adult speech [4], [5].

Children acquire the ability to express information structure and status differences from a young age, with abilities continuing to develop over the first few years of life. Cross-linguistic work in German has shown that children and adults do not have a one-to-one mapping between pitch accent selection and information status, and that children and adults tend to deaccent given information (but not always), and use comparable sets of pitch accents for new and accessible referents [6]. Work in Catalan and Spanish have additionally shown that the number of nuclear configuration patterns (as they reflect semantic and pragmatic meaning) begins more limited at the onset of speech, but that this quickly jumps to a larger set of configurations before the age of three [7].

In American English, particular pitch accents have been associated in the productions of 2.5-year-old toddlers’ productions of new, given, and corrective information categories [8]. This work found that most commonly, H* was associated with new, given, and corrective referents, with secondary preferences for deaccenting given referents and using L+H* in corrective cases. While preferences emerged, this work demonstrated there is not a 1:1 correspondence for mapping meaning to intonation. Importantly, these data were from a semi-controlled spontaneous speech task, where particular words were elicited, but the child spoke in a spontaneous manner. The current work takes this one step further and allows an unstructured interaction to naturally develop between the parent and child.

Children’s museums are places for children to explore new environments and engage in different activities not only with peers but also their families. Specifically, they provide caregivers a unique opportunity to engage and communicate with their children. The child gains more knowledge when the caregivers ask engaging questions and lead discussion about the topics. Recent research has analyzed the relationship between museums and parent-child interactions, showing that the context aids in learning and encouraging conversations [9].

Our goal is to explore how children express information status differences via prosody during a fully naturalistic and spontaneous experience at a children’s museum. We ask:

1) What types of pitch accents and nuclear configuration patterns are found for referents that are either new or given to the discourse context in a naturalistic speech task?

2) Across children aged 4 to 6-years-old, do we observe changes in the intonation patterns used to express information status differences?

To address these questions, we analyzed spontaneous speech collected during parent-child dyad play interactions. Based on previous research, we hypothesized that children would employ a particular set of pitch accents, with a preference for H* across all information, and additional preferences for L+H* in cases of contrast or correcting [8]. Additionally, work in Dutch has shown that sentence position is critical in predicting accentuation, showing that given information is often accented when in sentence-final position [10]. For this reason, we focused our initial analyses solely on referents in final position to help control for potential position effects and predict less deaccentuation. Future work plans to extend analyses to other sentence positions to determine if there are position effects in English as there are in Dutch, or not.
2. Method

2.1. Participants

Data for this study were taken from Project ACME (Advancing Children’s Museum Engagement), a larger study conducted at the Children’s Museum of New Hampshire in Dover, NH. Project ACME attracted a variety of families from the surrounding seacoast communities.

For the overall ACME study, participants were on average 4;11 (n = 30; 22 females, 8 males; range = 3;2 - 6;9). For inclusion, children had to be between 3 and 6 years old, typically developing, native English speakers, have normal or corrected-to-normal vision/hearing, and be able to provide verbal assent. Preliminary data for this study has analyzed three children, one at each year in the range. Table 1 summarizes the individual child data along with their sex, age, and Clinical Evaluation of Language Fundamentals-Preschool Second Edition (CELF-P2) Core Language score [11].

Table 1: Data analyzed per participant to date

<table>
<thead>
<tr>
<th>Child ID</th>
<th>Sex</th>
<th>Age</th>
<th>CELF-P2 Score</th>
<th># of Utterances Analyzed</th>
</tr>
</thead>
<tbody>
<tr>
<td>001</td>
<td>Female</td>
<td>4;01</td>
<td>92</td>
<td>18</td>
</tr>
<tr>
<td>009</td>
<td>Male</td>
<td>5;02</td>
<td>88</td>
<td>11</td>
</tr>
<tr>
<td>014</td>
<td>Female</td>
<td>6;01</td>
<td>108</td>
<td>13</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>****</td>
<td>****</td>
<td>****</td>
<td><strong>42</strong></td>
</tr>
</tbody>
</table>

*Ongoing data analysis

2.2. Procedure

Natural play was observed between caregiver-child dyads during exhibit exploration, with novel materials provided. Two exhibits were utilized. The Castle exhibit included open-ended style toys (e.g., blocks, felt boards, castle figurines). The River exhibit included areas to fish with fishing poles, play and learn from stuffed beavers that explained the history of a beaver and their habitat, and interact with obstacles that challenged the children’s motor skills (e.g., a treehouse, a boat). To collect the language samples, participants wore Samson XPD1 audio headsets which wirelessly uploaded to GarageBand for recording. The first 15 out of the 20-minute spontaneous language samples were used for analysis since all recordings had a minimum of 15 minutes recorded (but not all had 20 minutes).

Upon arrival, the study was explained, and consent was obtained. Caregivers completed a demographic form, and the child was administered the Clinical Evaluation of Language Fundamentals-Preschool, Second Edition (CELF-P2), an assessment designed to analyze a child’s language skills between the ages of 3 and 6;11 [11]. Children completed the sentence structure, word structure, and expressive vocabulary sections to yield a Core Language score (see Table 1).

2.3. Analysis

Language samples were transcribed in Praat, a computer software program designed for the acoustic analysis of speech [12], and then analyzed following the MAE_ToBI transcription system. For each child, all declarative utterances having a sentence-final referent were analyzed. Only utterances of two or more words were analyzed, with utterance length ranging from two to nine words across the three participants. This yielded the number of utterances analyzed as reported in Table 1, with a total of 42 utterances analyzed to date. Data analysis is ongoing. The four-year-old data showed that for referentially new information the pitch accents H*, L+H* and H+!H* were employed (Figure 1). For given information, H* was used less, along with L+H* (typically in cases of emphasis or contrast), H+!H*, and one case of deaccentuation. The five-year-old had only H* and L+H* used for new referents, and had H*, L*, L+H*, and deaccentuation used for those that were given (Figure 2). Finally, for the 6-year-old, H* was the preferred pitch accent for new and given referents, with L+H* and L* also appearing for both categories as well, but to a lesser extent (Figure 3).

![4-Year-Old](image1)

![5-Year-Old](image2)

Figure 1: 4-year-old pitch accent patterns for new and given referents

Figure 2: 5-year-old pitch accent patterns for new and given referents

3. Preliminary Results

The spontaneous speech of each child was analyzed following the MAE_ToBI transcription system. For each child, all declarative utterances having a sentence-final referent were analyzed. Only utterances of two or more words were analyzed, with utterance length ranging from two to nine words across the three participants. This yielded the number of utterances analyzed as reported in Table 1, with a total of 42 utterances analyzed to date. Data analysis is ongoing. The four-year-old data showed that for referentially new information the pitch accents H*, L+H* and H+!H* were employed (Figure 1). For given information, H* was used less, along with L+H* (typically in cases of emphasis or contrast), H+!H*, and one case of deaccentuation. The five-year-old had only H* and L+H* used for new referents, and had H*, L*, L+H*, and deaccentuation used for those that were given (Figure 2). Finally, for the 6-year-old, H* was the preferred pitch accent for new and given referents, with L+H* and L* also appearing for both categories as well, but to a lesser extent (Figure 3).
Example utterances from each child are shown in Figures 4 through 6. Each waveform/spectrogram and corresponding TextGrid shows the phonetic realization and ToBI label for a new referent. In all three cases, the ToBI label was a L+H*.

While the majority of instances of L+H* usage was due to emphasis or contrast, some seem to better reflect another aspect of the child’s speech (perhaps a playful-like manner of interaction).

In terms of nuclear configuration patterns, a relatively diverse set of boundary tones were also annotated. The most common type was L-L%, which occurred in over 2/3 of data. Other types found include H-L%, !H-L%, L-H%, and H-H%.

4. Discussion & Conclusion

Preliminary analyses revealed several interesting patterns that emerged across the ages analyzed. In reference to both of our research questions, we found that all children used H* to express referents that were new or given to the discourse context. Though very preliminary data with only one child per age group thus far analyzed, there appears to be an emerging preference for the H* pitch accent that increases with the age of the child. This general preference for the H* pitch accent is also similar to what has been reported for English-acquiring 2.5-year-olds across information status categories [8].

In addition, we found that the four-year-old was the only speaker to use the H+!H* pitch accent. (across both new and given information status categories). Future analyses will reveal if this is speaker-specific, or a pattern more generally seen at this age. This is in line with work by [15], who has shown that young children employ this pitch accent at age 2- to 3-years-old. Additionally, we observe the use of L* in the five- and six-year-old speakers, but not the four-year-old speaker. Finally, we observed deaccentuation by the four and the five year old, but not the six year old. This may reflect a tendency for younger children to begin accenting given information around this age period. More data is needed to uncover if this is a broader trend.

Data analysis is ongoing for this study. Additional data from other participants are being added for each age group (4, 5, and 6 years old). Analyses will also be extended to referents that occur in sentence-medial position and other types of utterances (e.g., interrogatives). Finally, we would also like to conduct a corollary acoustic analysis alongside the phonological one in order to further tease apart differences across the ages and development. Importantly, we are continuing to delve into the pragmatic variation present in the spontaneous speech dialogues. Further work is needed to understand how pragmatic (and other social-emotional) factors may influence the selection of pitch accents and boundary tones in child speech.

5. Acknowledgements

Thank you to Gabrielle Gagnon, Lydia Sack, Sabrina Salmela, Kaleigh Samowski, Lauren Gabriel, Mikaela Coombs, and all of the members of the CAT Lab at UNH for their help in data collection and data analysis. Special thanks to all the families who participated in this study.
6. References


