The train goes 'choo choo'
The train goes ‘choo-choo’: A corpus analysis of onomatopoeic words in child-directed speech and early production

Barbora Skarabela, Emma Pool and Mits Ota
University of Edinburgh

ONOMATOPOEIC WORDS (OWs)

- Words like woof and choo-choo with phonological forms that resemble sounds and objects in the real world.
- A common feature of child-directed speech across different languages.

What role do OWs play in early language development?

Sound-symbolism bootstrapping hypothesis: The non-arbitrary nature of OWs provides infants with referential insight into sound-meaning mappings in words (Imai & Kita, 2014).

Easier articulation: The phonological characteristics of OWs, including the limited inventory of sounds, articulatorily easier segments, and less complex syllables, may facilitate the initial stages of children’s word production (Laing, 2014; Massaro & Perinany, 2017).

- Do OWs provide referential insight in real-life language learning? Are they used in contexts where word-referent mappings can be made transparent?
- Do infants preferentially produce OWs above and beyond their frequencies in the input?

METHOD

Data. Five typically developing children and their mothers from the Providence Corpus (Demuth, Culbertson & Alter, 2006).

<table>
<thead>
<tr>
<th>CHILD NAME</th>
<th>AGE RANGE</th>
<th>Total No. of Words (MOT)</th>
<th>Total No. of Words (CHI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alex</td>
<td>01:04:27-01:11:32</td>
<td>36,387</td>
<td>7,769</td>
</tr>
<tr>
<td>Lily</td>
<td>01:01:02-01:11:26</td>
<td>157,659</td>
<td>6,709</td>
</tr>
<tr>
<td>Naima</td>
<td>01:02:14-01:11:26</td>
<td>152,082</td>
<td>36,434</td>
</tr>
<tr>
<td>Violet</td>
<td>01:02:24-01:11:26</td>
<td>41,904</td>
<td>3,325</td>
</tr>
<tr>
<td>William</td>
<td>01:04:27-01:11:19</td>
<td>47,200</td>
<td>4,860</td>
</tr>
</tbody>
</table>

Data Extraction and Coding. All lexical variants of 22 conventional OWs and corresponding Adult words (AWs) from the MacArthur-Bates CDI:

- Each item was coded as either OW (e.g., woof) or AW (e.g., dog).
- Each OW was examined in the context of the utterance and further coded as referential expression or sound effect (a la Laing, 2014).

RESULTS

- OWs represented less than 1% of the adult input and less than 2% of the children’s production (see also Ota & Skarabela, 2016; Laing et al., 2017; and Ota et al., 2018).
- All children except Alex produced more AWs than OWs.
- Three of the five children used OWs at higher rates than their mothers.

CONCLUSIONS

- However, both mothers and children used OWs as sound effects in the vast majority of cases (88% and 93% on average for children and mothers respectively).
- In these contexts, there are no alternative adult forms. Children, therefore, had no choice but select the OW.

- In referential contexts where either OWs or AWs were available (e.g., ‘There is a choo-choo/train’), the children overwhelmingly chose to produce the AWs.
- While some children had a higher proportion of OWs than their mothers, both children and mothers overwhelmingly used AWs (96% and 98% respectively).

- When they have a choice to represent referents with either OW or AW, children, like adults, use AWs most of the time.
- No direct evidence that OWs provide an articulatory easier alternative to learners.
- In real-life language learning, OWs may not provide referential insights to canonical lexical sound-meaning mappings. They are primarily used as sound effects, whose role in word learning is still poorly understood.