**Frequency and Universal Principles**

Outi Bat-El

1. **Introduction:** There are two opposing, though partially overlapping approaches, to language acquisition – the generative approach and the usage-based approach – corresponding to the well-known nature vs. nurture controversy. Both approaches grant infants with the innate tools essential for the acquisition of their first language, but they differ with respect to the characteristics of the toolbox from which these tools are drawn. For the usage-based approach (Tomasello 2001), children rely entirely on the general cognitive toolbox, which allows them to draw generalizations from the *frequency* of the input provided by the ambient language. For the generative approach, in addition to the input and the general cognitive capacity, there is a toolbox containing *universal principles* specific to linguistic knowledge.

*Frequency* and *linguistic principles* usually converge, since linguistic principles state the unmarked structure (in relative terms) which often enjoys a high frequency within a language. However, historical change may deal a bad hand, resulting in a frequent marked structure. These are the rare cases that can tease apart frequency from universal principles, and one such case will be introduced in the talk.

Hebrew stress system provides this unique opportunity to tease apart frequency from universal principles.

- **Frequency** – final stress (iambic feet): Stress in Hebrew is predominantly final (71-75% in nouns, depending on the type of corpus), analyzed as a right-aligned iambic foot (e.g. *ma[taná]*, ‘present’).

- **Universal principles (UP)** – penultimate (trochaic feet): Hebrew is expected to be trochaic, i.e. with penultimate stress (e.g. *ba[ná]*, ‘banana’), since its stress system is weight insensitive (Hayes 1995).

The goal of the talk is twofold:

- To provide evidence supporting the role of UP in early infants’ speech and the periphery of the lexicon.

- To rephrase our research question. Given the evidence for UP effects, our question should not be “is it UP or frequency?” but rather “how do frequency and UP interact?”

2. **Infants:** During early stages of speech, infants attempt more trochaic targets than iambic and also produce more trochaic words than iambic (Adam and Bat-El 2009). They may shift stress (e.g. *lekaléf* → *kájef* ‘to peel’), add a syllable at the end of the word (kaxól → óla ‘blue’), or simply produce targets with penultimate stress as disyllabic (*sáfta* → *táta* ‘grandma’) before they produce targets with final stress as disyllabic (*todá* → *da* ‘thanks’). These findings support the “trochaic bias”, according to which “the natural metric form of children’s words is trochaic” (Allen and Hawkins 1978:176). Evidence for the trochaic bias is found mostly in trochaic languages (e.g. English, Dutch, Greek), where we cannot tease apart frequency from universal principles. Hebrew, however, provides a unique opportunity to tease apart these two approaches to language acquisition and language knowledge.

<table>
<thead>
<tr>
<th>Database:</th>
<th>Child</th>
<th>Age</th>
<th># forms</th>
<th>Data type</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR (boy)</td>
<td>1;2-1;7</td>
<td>1368</td>
<td>Production types per session</td>
<td></td>
</tr>
<tr>
<td>YI (boy)</td>
<td>2:00</td>
<td>159</td>
<td>Tokens</td>
<td></td>
</tr>
<tr>
<td>RI (girl)</td>
<td>1;10-1;11</td>
<td>217</td>
<td>Tokens</td>
<td></td>
</tr>
<tr>
<td>ON (boy)</td>
<td>1;3-TBA</td>
<td>TBA</td>
<td>Production types per month</td>
<td></td>
</tr>
</tbody>
</table>
3. Adults: Like many other languages, Hebrew phonology distinguishes between core and peripheral lexica. The core lexicon of Hebrew is phonologically rigid, subject to morphologically-specific processes, while the peripheral lexicon allows the unmarked structures to emerge. This contrast is sound with regard to the Hebrew stress patterns: Final stress (iamb) is dominant in the core lexicon, while penultimate stress (trochee) is dominant in the peripheral lexicon. Which words live in the periphery?

- **Loan words**: Out of the 6811 di- and trisyllabic words in Bolozky and Becker’s (2006) lexicon of nouns, 87% are native and 13% are loan words. 85% (5057/5958) of the native words bear final stress as opposed to 36% (304/853) of the loan words (Fainleib 2013).

- **Hypocoristics**: There are various types of hypocoristics in Hebrew, all bearing non-final stress (Bat-El 2005). Those ending in -i bear penultimate stress (e.g. dán-i, snád-i, mórd-i), and those ending in -le bear antepenultimate stress (e.g. míra-le, anáti-le, akíva-le).

- **Acronym words**: Acronym words, which are structurally identical to native words, seem to prefer penultimate stress more than native words (Bat-El 1994, Zadok 2002). An experimental study supports this impression with only 56% (239/430) final stress (cf. 85% in the native lexicon).

4. Conclusion: The various types of data suggest that the universal preference for trochee emerges in Hebrew despite the high frequency of iamb. Infants’ early speech is characterized with preference for trochee, but soon enough, iamb becomes dominant. Adults’ core lexicon is typically iamb, but in the peripheral lexicon trochee emerges.

Admitting both UP and frequency into language acquisition and language knowledge, our research enterprise should address the interplay between these two resources. In the talk I will present the conditions that allow the effect of UPs to emerge.

---


