English-learning infants’ perception of word stress patterns

Katrin Skoruppa  
Department of Speech, Hearing and Phonetic Sciences, University College London, London WC1N 1PF, United Kingdom  
kskoruppa@gmail.com

Alejandrina Cristiá and Sharon Peperkamp  
Laboratoire de Sciences Cognitives et Psycholinguistique, EHESS, DEC-ENS, CNRS, 75005 Paris, France  
alecristia@gmail.com, Sharon.Peperkamp@ens.fr

Amanda Seidl  
Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, Indiana 47907  
aseidl@purdue.edu

Abstract: Adult speakers of different free stress languages (e.g., English, Spanish) differ both in their sensitivity to lexical stress and in their processing of suprasegmental and vowel quality cues to stress. In a head-turn preference experiment with a familiarization phase, both 8-month-old and 12-month-old English-learning infants discriminated between initial stress and final stress among lists of Spanish-spoken disyllabic nonwords that were segmentally varied (e.g., [ni'l'a, 'tui] vs [lu'ta, pu'ki]). This is evidence that English-learning infants are sensitive to lexical stress patterns, instantiated primarily by suprasegmental cues, during the second half of the first year of life.

1. Introduction

Languages differ with respect to the lexical stress patterns they permit. In some languages, such as English and Spanish, word stress can vary freely and convey lexical distinctions (e.g., English: discount ['diskaunt] vs discount [dis'kaunt]; Spanish: sában ['saßana]—‘sheet’ vs sabana [saßana]—‘savannah’). In other languages, stress is fixed and always falls on the same syllable position within words, e.g., Hungarian words are stressed on the first syllable, Swahili words on the penultimate. These cross-linguistic differences impact adult speakers’ stress perception abilities: Speakers of French, Finnish, Hungarian, and Polish, all fixed stress languages, are less proficient than Spanish speakers at distinguishing nonwords that differ only in their stress pattern (e.g., ['ma'pi] vs [ma'pi]) in a sequence recall task (Peperkamp et al., 2010). In contrast, speakers of English, a free stress language, are able to perform a range of phonological and lexical tasks based on stress: They can identify stressed syllables based on suprasegmental cues alone (Fry, 1958) and integrate stress cues in cross-modal lexical priming (Cooper et al., 2002).

However, free stress languages are not a homogenous group. First, there are differences in the distribution of stress patterns, with many free stress languages being less free than it would appear at first glance. For example, roughly 60% of the
disyllabic words have initial stress in the Spanish lexicon, compared to 75% of disyllabic words in the English lexicon (Pons and Bosch, 2010). Second, languages also differ with respect to the acoustic realization of lexical stress. Most, including Spanish, use some combination of suprasegmental cues (most importantly, pitch, duration, and amplitude) (Fry, 1958). Other languages, including English, further recruit segmental cues, by introducing important changes in vowel quality in unstressed syllables (Fry, 1958). Although vowels of shorter duration are centralized due to undershoot in many languages, this phenomenon is greatly amplified for unstressed vowels in English, where vowel reduction is five times larger than in Spanish (Delattre, 1969).

Given that free stress languages are not a homogenous group, it is not surprising that stress perception abilities of speakers of these languages vary as well, and may be modulated by the distributional and acoustic differences sketched previously. For instance, Spanish listeners’ reaction times are always negatively affected by stress mismatches in a cross-modal priming task (Soto-Faraco et al., 2001), whereas English listeners only exhibit slowed responses in short words (Cooper et al., 2002). In fact, suprasegmental differences alone, i.e., the differences between minimal pairs that do not differ in vowel quality, e.g., *forebear* ['fɔːrbiə] vs *forbear* [fɔːrˈbiə]—do not disrupt priming in British English listeners (Cutler, 1986). Further, stress misplacement only affects word recognition if it results in a nonword (Small et al., 1988): for instance, *polite* produced with incorrect initial stress inhibits word recognition, whereas the noun *insert* produced with incorrect final stress does not. Finally, Fear et al. (1995) report that English listeners are more sensitive to changes in vowel quality than to changes in suprasegmental stress cues. To sum up, as stress is more regular and relies on more diverse cues in English than in Spanish, English listeners are overall less sensitive to stress, and when they do attend to it, they seem to give less weight to suprasegmental cues than Spanish listeners.

Cross-linguistic differences in infants’ stress sensitivity in words with varying vowels and consonants have only recently begun to be investigated (see Skoruppa et al., 2009, for a comprehensive review of infant stress perception research with segmentally nonvaried stimuli). The perception of stress patterns in segmentally varied stimuli differs between infants learning fixed and free stress languages already at the age of 9 months. At this age, Spanish-learning infants discriminate stress patterns (such as [ˈlapɪˈnaku] vs [kiˈbuˈluˈta]), whereas French-learning infants do not discriminate stress patterns in variable stimuli, although they can distinguish stress patterns in repeated identical nonwords ([ˈpiˈma] vs [piˈma]) (Skoruppa et al., 2009).

Cross-linguistic differences have also been found among learners of free stress languages. English learners prefer stress-initial over stress-final disyllabic real words (e.g., *pliant*, *falter* vs *comply*, *befall*) at 9 but not at 6 months of age (Jusczyk et al., 1993). Turk et al. (1995) found a similar preference in English-learning 9-month-olds using nonwords (e.g., [ˈæzəlˈʒiːləl] vs [ləˈæzəlbˈʒiːl]). Interestingly, Spanish-learning 9-month-olds show no overall preference for either of these patterns (e.g., [ˈkɪbaˈbuki] vs [niˈka biˈlʊ]) (Pons and Bosch, 2010). This difference could indicate that even children learning different free stress languages begin to show diverse perceptual patterns in the first year of life, similar to those found in adults. Indeed, the preference in English-learning, but not Spanish-learning, infants could be explained on the basis of differences in overall frequency (as stress-initial disyllables occur more frequently in English than in Spanish). Evidence that stress-initial words have a special status for English-learning infants also comes from a study on early word segmentation: Jusczyk et al. (1999) report that 7.5-month-olds can extract unfamiliar stress-initial words such as *hamlet* [ˈhæmlət] from continuous speech, but they missegment stress-final words such as *device* [diˈvɛs] until the age of 10.5 months. A similar preference for initial stress has also been found in an artificial language learning study using longer stimuli (Gerken, 2004): Tested on nonwords with five syllables, 9-month-old American infants showed an overall preference for English-like stress on the first and fourth syllables (e.g., *dotórnemifə*),
despite the fact that half of them had been familiarized with stimuli following the latter pattern. However, infants were able to learn other novel rules involving stress and heavy syllables in this study, suggesting that their stress perception abilities are not entirely rigid. Taken together, these studies suggest that lexical stress perception may be more flexible and less biased toward initial stress in Spanish- than in English-learning infants.

Finally, the weighting of suprasegmental and segmental cues to stress may also develop differently depending on language exposure. In particular, English-learning infants, like English-speaking adults (Fear et al., 1995) may rely more heavily on vowel reduction. As a matter of fact, Jusczyk et al. (1993; Exp. 3) report that 9-month-old English-learning infants’ preference for stress-initial words holds even when stimuli are low-pass filtered. However, this only demonstrates that English-learning infants are sensitive to suprasegmental cues in the absence of segmental content; it does not show whether they can rely mainly on suprasegmental cues when segmental content is available. Evidence that prosodic information may be harder to attend to in the presence of segmentally varied content comes from a study on pitch perception (Lebedeva and Kuhl, 2010), which shows that 11-month-old American infants can detect melody inversions if the pitch changes are presented on four identical syllables (i.e., lalalala), but not if they are presented on four different syllables (i.e., gobiratu).

Thus, the present study set out to investigate whether English-learning infants can encode lexical stress patterns both in stress-initial and in stress-final words, based principally on suprasegmental cues. In order to have an established point of comparison, we used the same (Spanish) stimuli and method (head-turn preference procedure) as in Skoruppa et al. (2009). This also allowed us to address the question of whether infants can discriminate between stress patterns in foreign words, despite the fact that the phonetic realization of both the segments and the stress cues are unfamiliar to them. Considering previous evidence of a shift in stress-related segmentation abilities between 7.5 and 10.5 months in English-learning infants (Juszyk et al. 1999), we tested infants at two age ranges within the second half of the first year of life.

2. Method

2.1 Participants

Fifty-six healthy full-term English-learning infants (27 girls, 29 boys) were tested in West Lafayette, IN. Half of them were around eight months (mean 7;30, range 7;17-8;24) and half of them around 12 months old (mean 11;30, range 11;07-12;20). A further 15 infants participated whose results are not reported for the following reasons: 14 for fussing or crying; and 1 for having a total looking time of less than 1 s for one test trial type.

2.2 Materials

The stimuli, listed in Table 1, were the same as in Skoruppa et al. (2009). They had been produced in infant-directed speech by a female native-speaker of Spanish. There were 16 CVCV nonwords. Eight nonwords with initial stress and eight segmentally

<table>
<thead>
<tr>
<th>Familiarization</th>
<th>Stress-final group</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress-initial group</td>
<td>Stress-final group</td>
<td>All infants</td>
</tr>
<tr>
<td>List 1</td>
<td>List 2</td>
<td>List 1</td>
</tr>
<tr>
<td>'datu'</td>
<td>'latu'</td>
<td>da'tu</td>
</tr>
<tr>
<td>'sapi'</td>
<td>'buki'</td>
<td>sa'pi</td>
</tr>
<tr>
<td>'kifla'</td>
<td>'luma'</td>
<td>ki'fa</td>
</tr>
<tr>
<td>'nuki'</td>
<td>'tiku'</td>
<td>nu'ki</td>
</tr>
</tbody>
</table>
identical nonwords with final stress were used for familiarization. Eight different non-
words, four with initial stress and four with final stress, were used for the test phase. 
Acoustic measurements revealed that stress was instantiated by significant differences in 
duration, intensity, and pitch between stressed and unstressed vowels (all $p$’s < 0.001); for further details can be found in Skoruppa et al. (2009).

2.3 Procedure
A variant of the head-turn preference procedure (Jusczyk and Aslin, 1995) was used. 
Infants were tested in a sound-attenuated, dimly lit room, while seated on a caregiver’s 
lap in the middle of a three-walled enclosure. Behind three openings in the front panel 
of the enclosure, out of sight of the infant, was located an experimenter who observed 
the infant’s head-turns through a camera, and recorded them using a response box. 
Both the caregiver and the experimenter listened to music designed to mask human 
speech over Peltor Aviation headphones. Approximately at eye level for the infant, 
there was a green light on the front panel and a red light on each side panel, and 
behind the latter there were speakers. Each trial started with a green light flashing on 
the front panel. As soon as the infant fixated on it, it was extinguished, and one of the 
red side lights began to flash. Once the infant oriented toward the side light, the stimu-
lus presentation began, and continued until the infant turned away for more than 2 s 
or until the stimulus list had been repeated three times. The time spent oriented toward 
the source of the sound (“looking time”) is the dependent measure, a proxy for infants’ 
attention.

Infants were randomly assigned to the “stress-initial” or “stress-final” group. 
During familiarization, infants in the stress-initial group heard the two stress-initial 
familiarization lists; similarly, the stress-final group heard the two stress-final familiar-
ization lists (see Table 1). The interstimulus interval coherence was fixed at 1.8 s. The 
side of the light and the list being played, alternated until the infant had accumulated 
1 min of total attention time for each list. The subsequent four-trial test phase was 
identical for all infants. There were two trials with a list of new stress-initial nonwords 
and another two with a list of new stress-final nonwords (see Table 1). The order and 
side of presentation of the two lists were randomized.

3. Results
Mean looking times for familiar vs novel stress patterns by age and familiarization 
group are shown in Table 2. A repeated measures analysis of variance with the within-
subject factor Stress Pattern (familiar vs novel) and the between-subject factors Age (8 
months vs 12 months) and Familiarization (stress-initial vs stress-final) revealed a sig-
nificant main effect of Stress Pattern: Looking times were significantly higher for the 
novel than for the familiar stress pattern [$F(1,52) = 11.70$, $p < 0.001$]. There was also a 
marginal main effect of Familiarization: Infants familiarized with stress-initial non-
words had marginally higher looking times than infants familiarized with stress-final 
nonwords [$F(1,52) = 3.80$, $p = 0.057$]. All other effects and interactions were not signifi-
cant ($F < 1$), showing, in particular, that there were no age-related changes. Nonpara-
metric analyses using a Pearson’s $\chi^2$ test confirmed the main result: 40 infants (out of 
56) showed longer looking times to the novel stress pattern [$\chi^2 (1) = 10.26$, $p = 0.001$].

Table 2. Mean looking times (standard deviation) in seconds.

<table>
<thead>
<tr>
<th>Familiarization</th>
<th>All</th>
<th>Stress-initial</th>
<th>Stress-final</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Familiar</td>
<td>Novel</td>
<td>Familiar</td>
</tr>
<tr>
<td>All</td>
<td>6.75 (4.56)</td>
<td>9.94 (6.33)</td>
<td>7.84 (5.23)</td>
</tr>
<tr>
<td>8 months</td>
<td>7.12 (4.63)</td>
<td>10.40 (6.92)</td>
<td>8.49 (4.91)</td>
</tr>
<tr>
<td>12 months</td>
<td>6.38 (4.55)</td>
<td>9.48 (5.77)</td>
<td>7.18 (5.63)</td>
</tr>
</tbody>
</table>
4. Discussion

After 2 min of familiarization with disyllabic nonwords that shared the same stress pattern (stress-initial or stress-final), both 8- and 12-month-old English-learning infants listened longer to novel nonwords with the opposite stress pattern. As test and familiarization stimuli differed in their segmental content, focusing on segmental differences or recalling particular tokens could not help infants in differentiating between the two test lists. Therefore, the novelty preference observed at both ages suggests that, like Spanish learners and unlike French ones (Skoruppa et al. 2009), infants learning English are able to discriminate stress patterns in segmentally varied nonwords. This is all the more remarkable given that the consonants and vowels in the stimuli were produced by a foreign language speaker who used different stress cues as well as language-specific realizations of vowels and consonants.

Skoruppa et al. (2009) showed that the fundamental distinction between speakers of fixed stress languages and speakers of free stress languages (Peperkamp et al., 2010) is present from the first year of life. Given that adult research also documents variation in perception among free stress languages, one might expect English infants’ performance with Spanish stimuli to be intermediate between that of Spanish and French learners. This would be a precursor of the reduced lexical stress sensitivity recorded in adults, particularly when vowel quality is relatively preserved (e.g., Soto-Faraco et al., 2001 vs Cooper et al., 2002). This prediction was not met, lending little support to the hypothesis that differences in the distribution and realization of stress patterns in the Spanish and English input would affect Spanish and English infants’ sensitivity to stress very early on, at least not at the ages we tested and with the task and stimuli we used. Of course, a more sensitive procedure might reveal possible differences between Spanish and English infants; for instance, it would be interesting to use electrophysiological measurements of auditory evoked potentials, which can reveal a more graded response pattern at the neural level.

This is not to say that cross-linguistic differences in lexical stress have no impact in infancy. On the contrary, as noted in the Introduction, previous work documents differences in infants’ preferences (Turk et al., 1995 vs Pons and Bosch, 2010). These were (partly) replicated here, as we found a marginal trend for greater overall looking times in infants familiarized with stress-initial nonwords. However, the lack of an interaction between familiarization group and trial type at test suggests that these preferences did not impede infants’ discrimination of stress patterns. Taken together with previous research, our results suggest that while differences in frequency of occurrence and realization of stress patterns affect prelinguistic infants’ preferences, these differences do not, in their stead, block the representation of stress in segmentally variable material. More generally, the present work suggests that differences in sensitivity to lexical stress across various types of free stress languages that have been documented in adult listeners cannot yet be demonstrated during the first year of life.

Acknowledgments

We thank Laura Bosch and Begona Diaz for providing the stimuli, Ashley Foxworthy, Amanda Schultz, Carrie Wade, and Yuanyuan Wang for help with recruiting and testing, and all babies and parents for their participation. The first and the second author contributed equally to this study.

References and links

