Pitch first, stress next? Prosodic effects on word learning in an intonation language
1. Introduction

- Acquisition of phonology requires learning to interpret phonetic variation.
- Across languages, prosodic properties may vary in their acoustic correlates and the phonological domains they signal:
  - Stress signals word contrasts in Spanish or English but not in French (phrasal level); Pitch signals word contrasts in Chinese or Japanese but not in English (phrasal level).
  - Pitch is an important cue to stress in Spanish (Hualde 2005), but not in European Portuguese.
- The task for the young learner: which variation is meaningless?; which variation conveys meaning, at what prosodic level (word level/phrasal level)?
1. Introduction

- Discrimination studies have shown

- Studies on word recognition have shown that both stress and pitch are relevant to English learning infants by 7.5 months; pitch register is discarded by 9 months; pitch interacts with stress (Jusczyk et al. 1999, Singh et al. 2008; Fikkert & Chen, in press)

- Recent research on word learning has shown that English infants at 1;0 use stress, and at 2;5 disregard pitch contours as lexically relevant (Curtin 2009, Quam & Swingley 2010)

- Language-specific vowel contrasts are discriminated before 6 mos, and used in word learning before 2;0 (Kuhl et al. 2006; Curtin et al. 2009)
1. Introduction

- European Portuguese (EP) is an intonation language with lexical stress

- **Stress** is a **word level** property:
  - Penult, final, antepenult stress; morphology plays a role in stress location; no sensitivity to weight (Mateus & Andrade 2000)
  - In the absence of vowel reduction, duration is the main cue to word stress (Andrade & Viana 1989, Delgado Martins 2002)
  - Stress can be contrastive: bambo ‘unsteady’ bambu ‘bamboo’

- **Pitch** is a property of **phrase level** phonology
  - Pitch contrasts signal phrase level meanings (pragmatic/discoursive meanings, sentence types – Frota 2002)
  - Sparse pitch accent distribution: Besides the nuclear word, only c. 17% of PWs are pitch-accented (Vigário & Frota 2003)
1. Introduction

- The acquisition of stress and pitch contrasts in EP
  - No prior perception studies
  - Meaningful pitch contrasts are produced quite early (1;5)
- We tested whether EP learners were sensitive to stress and pitch contrasts in a word learning task, using an eyegaze-based procedure (and compared this with a vowel contrast).
  - Will young learners notice stress differences and/or intonation differences in ‘new words’?
  - When do young learners interpret phonetic variation at the appropriate levels according to the native language?
2. Method: Procedure

- We used an eyegaze-based procedure (similar to Quam & Swingley 2010) where visual fixation to the labelled picture is the response variable.
  - Children sat on the parent’s lap in front of the screen of an SMI eye-tracker, on which they viewed pictures. Concealed speakers played the recorded utterances that referred to the pictures.
  - Experiment consisted of 3 phases (lasted 2 minutes):
    - 1. Animation phase: a doll introduces two toys, but only one of them is labelled (the ‘A’ toy) [Hi, I have a new toy. This is ‘A’. Let’s play with ‘A’. I have another toy. We can also play with it.]
    - 2. Ostensive-labelling: the ‘A’ toy is repeatedly labelled.
    - Phases 1 and 2 teach a novel word (the trained word).
    - 3. Test phase: pictures of the two toys appear side by side while children listen to the trained word and to stress/pitch deviant versions.
2. Method: Procedure

Animation + Ostensive-labelling with Dolphin and Turtle as ‘A’ counterbalanced;
Test: Order of picture presentation (left/right) counterbalanced

ISI (attention getter) 1500ms; total: 2 min.

Animation

3500ms x 5

Ostensive-labelling

3000ms x 10

Test

4000ms x 12
2. Method: Materials

- **Auditory stimuli:** Recorded by a native EP speaker in CDS.

  The disyllabic word form \[ milu \] preceded by the article ‘the’ or as the nuclear word in the short utterances ‘Look at…’ ‘Where’s…’
  
  - The most frequent word shape (in ADS 44% and CDS 45% - *FrePOP*); high vowels were chosen to avoid vowel reduction

**Stress contrast:** penult / final \[ 'milu \] / \[ mi\'lu \]

  - The most frequent stress patterns (in ADS 76/22% and CDS 67/32%)

**Pitch contrast:** declarative / interrogative \[ 'milu \] / \[ mi\'lu \]

  - Falling contour / Fall-Rise: H+L* L% / H+L* LH%

Trained pronunciation: penult-decl; final-decl; penult-int

Deviant pronunciations: stress change (SC), intonation change (IC), both

Test phase included 4 trained trials + 8 change trials (4+2+2)

Target word started at about 360 (NPs) 670 ms (utts) from trial onset
2. Method: Materials

- Auditory stimuli: Phonetic properties
  Stress contrast: penult / final [ˈmilu] / [miˈlu]
    - Stress location cued by relative duration (stressed > 58-132 ms) and by the alignment of the pitch fall (through the stressed syllable)
  Pitch contrast: declarative / interrogative (H+L* L% / H+L* LH%)
    - Intonation contrast cued by the low versus rising boundary and by the longer duration of the final syllable in interrogatives (117-165ms)
2. Method: Participants

- 93 children between 1;0 and 4;9 were tested, all from monolingual EP homes. 49 children have successfully performed the task (i.e. learned the trained word: >50% fixation to the labelled picture) Table 1.

<table>
<thead>
<tr>
<th>Group</th>
<th>Participants</th>
<th>Mean age (months)</th>
<th>Children included</th>
<th>Mean age (months)</th>
<th>CDI mean score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>26</td>
<td>16.9</td>
<td>11</td>
<td>16.3</td>
<td>19.64</td>
</tr>
<tr>
<td>Group 2</td>
<td>25</td>
<td>30.5</td>
<td>11</td>
<td>29.9</td>
<td>65.16</td>
</tr>
<tr>
<td>Group 3</td>
<td>25</td>
<td>42.4</td>
<td>14</td>
<td>43.5</td>
<td>84.77</td>
</tr>
<tr>
<td>Group 4</td>
<td>17</td>
<td>53.5</td>
<td>13</td>
<td>52.9</td>
<td>89.13</td>
</tr>
</tbody>
</table>

- Parents filled in an EP version of the CDI form. A correlation analysis (Pearson) showed that age is a better predictor of the results than the CDI scores.
Coding for analysis: For each subject in each test trial, proportion looking time at the labelled object picture ‘A’ (looking at ‘A’ divided by the total looking time for both pictures)

- We used two time windows after the onset of the target word: 367+2000 ms for 1&2 year-olds, 367+1500 ms for 3&4 (Fernald et al. 1998, Swingley & Aslin 2002, Fikkert & Chen in press, Gredebäck et al. 2010)
3. Results

- The analysis of looking behaviour before the target word was heard showed no bias towards one of the object pictures (mean = .47, t(92) = -1.51, p = .13).

- We determined whether children had learned the trained word by comparing children’s fixation to the labelled picture to chance fixation. Only children with a fixation > 50% were included in the analysis - Table 1 (mean = .67, t(48) = 8.92, p < .001).

- Next we asked whether there was any difference between children’s responses to the NP (‘the milo’) and the short utt (‘Look at the milo’) trials: No difference both for trained (t(48) = -1.46, p = .15) and deviant (t(48) = -.91, p = .37).
3. Results

- The main analysis examined whether children responded to the deviant pronunciations (SC, IC, both).

- An ANOVA on the % looking time to the labelled picture, with condition (trained, SC, IC, both) as the within-subject factor and training phase (penult-decl, final-decl, penul-int), object familiarized (dolphin, turtle) and age (younger=groups 1-2, older=groups 3-4) as the between-subjects factors, revealed

A significant main effect of condition \( (F(3,108) = 8.29, p < .001) \)

and a significant interaction between condition and age \( (F(3,108) = 2.86, p < .05) \)

Post-hoc analysis revealed significant differences between trained pronunciation and all the other conditions, and IC and Both.

- Further analysis within the younger group showed no effects; within the older group, effect of condition only.
3. Results

Detailed analysis of children’s response to the deviant pronunciations by age. If children were sensitive to any of the changes, (A) they should NOT fixate the labelled picture above chance, within these conditions; (B) the deviant conditions should be significantly different from the trained condition.

Younger: 1-year olds and 2-year olds

Older: 3-year olds and 4-year olds

Figure 1. Proportion looking time at the labelled object picture across the 4 conditions, by age group. Significant t-test results (two-tailed) against chance and significant paired t-tests with trained signalled by * (.05), ** (.01) and *** (.001)

Development from sensitivity to pitch & stress to sensitivity to stress only
4. Testing a Vowel contrast

- Same procedure, but the stress contrast was replaced by a vowel contrast, with two goals:
  - To allow a comparison between the interpretations of the stress change, the pitch change and the segmental change
  - To replicate the pitch change results of the previous experiment

**Vowel contrast:** change in height and backness ['milu] / ['malu]
- These vowels are the 4th and 5th most frequent ones (in 14 vowels)

**Pitch contrast:** declarative / interrogative ['milu] / ['malu]
- Falling contour / Fall-Rise: H+L* L% / H+L* LH%.

Trained pronunciation: ['milu]-decl
Deviant pronunciations: vowel change (VC), intonation change (IC), both
4. Testing a Vowel contrast

- **Participants (so far):** 26 children between 1;0 and 4;9 were tested, all from monolingual EP homes. 17 children have successfully performed the task (i.e. learned the trained word: >50% fixation to the labelled picture)

Table 2.

<table>
<thead>
<tr>
<th>Group</th>
<th>Participants</th>
<th>Mean age (months)</th>
<th>Children included</th>
<th>Mean age (months)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>12</td>
<td>16.9</td>
<td>7</td>
<td>16.6</td>
</tr>
<tr>
<td>Group 2</td>
<td>5</td>
<td>29.0</td>
<td>3</td>
<td>30.0</td>
</tr>
<tr>
<td>Group 3</td>
<td>5</td>
<td>44.0</td>
<td>4</td>
<td>44.0</td>
</tr>
<tr>
<td>Group 4</td>
<td>4</td>
<td>54.7</td>
<td>3</td>
<td>54.7</td>
</tr>
</tbody>
</table>
4. Testing a Vowel contrast

- Preliminary Results: The same pattern as in the stress/pitch experiment.

- The younger group shows sensitivity to both the vowel contrast and the pitch contrast, whereas the older group only interprets the vowel contrast as relevant.

- The pitch change results seem to be replicated.

- VC seems to be a stronger contrast than SC, for the older group.

Fig. 5. Proportion looking time at the labelled object picture by age group.
4. Discussion

- Our results demonstrate that
  
  - Pitch contour variation is regarded as relevant in new words by 1-year olds and 2-year olds, at odds with native language phonology.
  
  - 3-year olds already regard pitch variation as NOT relevant: Only at 3;0 do young learners interpret phonetic variation at the appropriate levels according to the native language.
  
  - Stress variation is regarded as meaningful, both by the younger and older age groups (similar to segmental variation).
4. Discussion

- The pitch change results deviate from the results reported for English (Quam & Swingley 2010), as pitch contour differences are already discarded by English learning 2-year olds (earlier than in EP). However,
  - It is not known whether English-learning 1-year olds include pitch contours among the dimensions of variation relevant to lexical meaning.
  - Is the early sensitivity to intonation contrasts in word learning a more general property (akin to early sensitivity to pitch in discrimination) or a language-specific feature? (more studies!)
  - Production studies for EP showed that pitch contrasts are produced early in the 2nd year and that words form single word prosodic phrases until the end of the 2nd year (Vigário et al. 2011) >> Further research
4. Discussion

- Early sensitivity to stress variation in EP, as in English (Curtin 2009)
  - It is known that 1-year olds have difficulties in detecting subtle phonetic contrasts when learning novel words (e.g., Werker & Yeung 2005, Swingley & Aslin 2007), however they are sensitive to stress variation. It seems that the acoustic cues to stress are salient enough in EP (even in the absence of vowel reduction).
  - However, production studies for EP have shown difficulties with stress contrasts in early word production (e.g. Correia 2010). Is early sensitivity to stress just an instance of a general early sensitivity to (salient) phonetic contrasts?
  - Implications for word segmentation: Does word stress in EP play a role in segmentation as it does in English (e.g. Mersad et al. 2010)? >> Further research
4. Discussion

The persistence of the pitch contrast in word-object associative learning can hardly be accounted for by a general sensitivity to any (salient) phonetic variation (as previous findings suggest fundamental changes ~20 months in children’s interpretation of phonetic properties).

Phonological development in word representations proceeds from pitch & stress to stress only, by fine tuning to the dimensions relevant in the native phonology.
Obrigada

Thanks to the audience of ICPC at York for comments on a previous version of this study.
Thanks to all the children and parents, and the nurseries that collaborated in the study. Marisa Cruz helped with data collection.

DEPE: Grant PTDC/CLE-LIN/108722/2008 , FCT

http://www.fl.ul.pt/laboratoriofonetica/babylab/

sonia.frota@mail.telepac.pt jbutler@fl.ul.pt
References


References


References


