Intonational and gestural structures are temporally coordinated

Núria Esteve-Gibert & Pilar Prieto

Group of Prosodic Studies, Universitat Pompeu Fabra
Previous findings in the literature

• Integration between gesture and speech

• McNeill’s (1992) five main reasons:

  1. Gestures occur with speech in 90% of the cases
  2. Gesture and speech develop together in children

  3. Gesture and speech are phonologically synchronous

  4. Gesture and speech are semantically and pragmatically co-expressive

  5. Gesture and speech break together in aphasia

The stroke of the gesture precedes or ends at, but does not follow, the phonological peak syllable of speech (Kendon, 1980)
The temporal coordination is between the most prominent part of the gesture (stroke or apex) and the phonological peak syllable of speech (the most prominent part of speech).

The most prominent part of speech has been interpreted differently.
Speech prominence that coordinates with gesture

- **A noun, a verb, an adverb, or an adjective**, so that gestures tend to start in a pause just before them. *(Butterworth and Beattie, 1978)*

- **Prosodic focus**, so that the most prominent part of the pointing gesture overlaps with it *(Roustan & Dohen, 2010)*

- **Pitch-accented syllable**, so that gestural apices align with them more often than with non-accented syllables *(Loehr, 2004, 2007)*.

- **Stressed syllable**, so that the gestural stroke co-occurred with the jaw cycle for the stressed syllable *(Rochet-Capellan, Laboissière, Galván & Schwartz, 2008)*.

- **Unstressed syllables**, so that gesture strokes align with them rather than with prosodically prominent syllables *(Rusiewicz, 2010)*

- Prominent syllables and **syllables with f0 peak** coordinate with gestural strokes more often than syllables with intensity peak *(Nobe, 1996)*

- **Stressed syllable only in a contrastive focus condition**, so that gestural apices coordinate with them. *(De Ruiter, 1998)*
Results in the previous literature

- Gesture prominence coordinates with the **focused word** (Butterworth and Beattie, 1978; Roustan & Dohen, 2010)

- Gesture prominence coordinates with the **stressed syllable** (Loehr, 2004, 2007; Rochet-Capellan, Laboissière, Galván & Schwartz, 2008)

- Gesture prominence coordinates with syllables with **intonation peaks** (Nobe, 1996; De Ruiter, 1998)
General assumption: **gestural prominence** and **speech prominence** co-occur in time.

Yet, a more detailed analysis of the prosodic prominence is required.

**Main goal**: to investigate which is the anchoring region in speech that aligns with the gesture prominence.

**Related goal**: to test the hypothesis that the gestural peak temporally aligns with the peak of the focused pitch accent.
Methodology

**Subjects:** 15 adult Catalan-speakers

**Task:** pointing-naming task

**RECORDA...**

Imagina’t que m’has d’indicar on és la cara perquè m’hi pugui fixar i que, alhora, has d’anomenar-la correctament, no pas com ho he fet jo.

**mama, es diu,**  
i no **mamà**
Materials

• **Target words** with different **metrical patterns** (ma, mama, mama) in a contrastive focus condition to trigger different temporal position of the intonation peaks with the stressed syllable.

• The prosodic prominence is related to the **stress**.

• In a contrastive focus condition, the **intonation peak** is realized differently when the stress is in phrase-final (ma’ma) or in non-phrase-final ('mama) position. *(Prieto & Ortega-Llebaria, 2007; Ortega-Llebaria & Prieto, 2010)*
Hypotheses

• **H1**: if the apex occurs in the same place when the stress is in the penultimate or in word-final position, the apex coordinates with the **stressed syllable**.

![Diagram of H1](image1)

• **H2**: if the apex occurs in a different place when the stress is in penultimate position or in word-final position, the apex coordinates with the **intonation peak**.

![Diagram of H2](image2)

Materials consist of words with different **metrical patterns** to trigger different temporal positions of the intonation peaks with the stressed syllable.
Gestural and intonational labeling

<table>
<thead>
<tr>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>gaze</td>
<td>pointing</td>
</tr>
<tr>
<td>gesture</td>
<td>preparation</td>
</tr>
<tr>
<td></td>
<td>stroke</td>
</tr>
<tr>
<td></td>
<td>traction</td>
</tr>
<tr>
<td>pitch acc</td>
<td>pitch accent</td>
</tr>
</tbody>
</table>

[Image of audio software interface with waveforms and labeled events]
Does the position of intonation peaks within the accented syllable change when the metrical structure of the word is different?

**Distance > 0** The $f(0)$ peak occurs after the end of the stressed syllable.

**Distance < 0** The $f(0)$ peak occurs before the end of the stressed syllable.

**LMM**
- **Dependent:** distance between...
- **Fixed factor:** metrical structure (3 levels: monosyllable, trochee, iamb)
- **Random factors:** subject, item (i.e. consonant diffs)
Does the position of gestural peaks with respect of the accented syllable change when the metrical structure of the word is different?

**Distance > 0**  The apex occurs after the end of the stressed syllable.

**Distance < 0**  The apex occurs before the end of the stressed syllable.

**LMM**

**Dependent:** distance between...

**Fixed factor:** metrical structure (3 levels: monosyllable, trochee, iamb)

**Random factors:** subject, item (i.e. consonant diffs)
Results

Do the intonation peaks and the gestural peaks align?

Distance > 0  The f(0) peak occurs before the apex.

Distance < 0  The f(0) peak occurs after the apex.

LMM
Dependent: distance between…
Fixed factor: metrical structure (3 levels: monosyllable, trochee, iamb)
Random factors: subject, item (i.e. consonant diffs)
Results

Do the stroke and the stressed syllable start together?

Distance > 0  The stroke starts after the stressed syllable

Distance < 0  The stroke starts before the stressed syllable

LMM
Dependent: distance between...
Fixed factor: metrical structure (3 levels: monosyllable, trochee, iamb)
Random factors: subject, item (i.e. consonant diffs)
ma'ma

ma'ma
Conclusion

- **H1**: the apex is coordinated with the stressed syllable because it occurred in the same place when we changed the metrical structure of the focused word.

- **H2**: the apex is coordinated with the $f_0$ peak because it occurred in a different place when we changed the metrical structure of the focused word.

Gesture peaks are temporally coordinated with intonation peaks.
Conclusion

• Both intonation peaks and gesture peaks are **retracted** in final position of intonation domain (monosyllables and iambs).

• Gesture peaks appear after the intonation peaks in monosyllables with respect to iambs: there is a ‘**lagging effect**’.

This may be due to the fact that the pointing gesture starts in utterance-initial position in monosyllables and the whole trajectory has to be realized during the accented syllable. By contrast, in the iambic condition, the pretonic syllable already contains part of the forward-pointing gesture.
Moltès gràcies!