INTONATION AND BELIEF STATES IN ADULT AND CHILD SPEECH

Meghan Armstrong-Abrami
UMass Amherst

Brown University – 4 April 2018
HOW IS INTONATION USED TO CONVEY SPEAKER BELIEF STATES?
HOW IS INTONATION USED TO CONVEY SPEAKER BELIEF STATES?

Pierrehumbert & Hirschberg 1990:

Intonation contours are chosen by speakers:

(1) to create a relationship between propositional content and previous/upcoming utterances

(2) to create a relationship between the propositional content of an utterance and the beliefs of the speaker and the hearer.
HOW IS INTONATION USED TO CONVEY SPEAKER BELIEF STATES?

Pierrehumbert & Hirschberg 1990:

Intonation contours are chosen by speakers:

1. to create a relationship between propositional content and previous/upcoming utterances

2. to create a relationship between the propositional content of an utterance and the beliefs of the speaker and the hearer.
Earlier work argued that specific tunes were used to convey speaker attitude.
P&H argue that though speaker attitude may be inferred from tune choice, attitude may be better understood as derived from tune meaning as interpreted in a context rather than representing tune meaning itself.

E.g. Ward & Hirschberg (1985) showed that speaker uncertainty, incredulity, politeness & irony can all be derived from L*H LH%.
INTERPRETATION OF TUNES

- Interestingly, as we have advanced in our knowledge of intonational phonology, descriptions tend to make lists of possible pragmatic meanings of intonation contours, rather than focusing on the core meaning of a given tune.

- Especially in Romance (since there has a great “boom” in intonational descriptions of Romance varieties).

- Specifically, we for questions, we use divisions such as “information-seeking vs. confirmation-seeking” or “neutral vs. biased” to describe the pragmatic division of labor found for tunes.
### Yes-no questions

<table>
<thead>
<tr>
<th>Information-seeking yes-no questions</th>
<th>H+L* H%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H+L* L%</td>
</tr>
</tbody>
</table>

#### Biased yes-no questions

<table>
<thead>
<tr>
<th>Echo yes-no questions</th>
<th>H+L* H%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H+L* L%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Counterexpectational echo yes-no questions</th>
<th>H+L* H%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H+L* L%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Invitation yes-no questions</th>
<th>L+H* M%</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Confirmation yes-no questions</th>
<th>H+L* H%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>H+L* L%</td>
</tr>
</tbody>
</table>

Dominican Spanish – Willis 2010
But what is driving contour choice???
### Puerto Rican Spanish Polar Questions

Contour distribution for types of questions using Prieto & Roseano’s Discourse Completion Task.

<table>
<thead>
<tr>
<th>Type</th>
<th>iH* L%</th>
<th>L* HL%</th>
<th>H+L* L%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral (6)</td>
<td>93% (28/30)</td>
<td>0% (0/30)</td>
<td>7% (2/30)</td>
</tr>
<tr>
<td>Focus (3)</td>
<td>100% (15/15)</td>
<td>0% (0/15)</td>
<td>0% (0/15)</td>
</tr>
<tr>
<td>Echo (4)</td>
<td>100% (20/20)</td>
<td>0% (0/20)</td>
<td>0% (0/20)</td>
</tr>
<tr>
<td>Counter-expectation echo (2)</td>
<td>50% (5/10)</td>
<td>50% (5/10)</td>
<td>0% (0/10)</td>
</tr>
<tr>
<td>Tag (4)</td>
<td>100% (20/20)</td>
<td>0% (0/20)</td>
<td>0% (0/20)</td>
</tr>
<tr>
<td>Imperative (2)</td>
<td>100% (10/10)</td>
<td>0% (0/10)</td>
<td>0% (0/10)</td>
</tr>
<tr>
<td>Offers/invitations (3)</td>
<td>100% (15/15)</td>
<td>0% (0/15)</td>
<td>0% (0/15)</td>
</tr>
<tr>
<td>Rhetorical (1)</td>
<td>100% (5/5)</td>
<td>0% (0/5)</td>
<td>0% (0/5)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>94% (118/125)</td>
<td>4% (5/125)</td>
<td>2% (2/125)</td>
</tr>
</tbody>
</table>

¿Hay reunión mañana?  
‘Is there/there’s a meeting tomorrow?’

**PUERTO RICAN SPANISH POLAR QUESTIONS**

1A (unbiased)

Tu amigo: Te quiero llevar a comer comida típica de Puerto Rico. ¿Qué quieres comer?

Tú: ¿Hay por aquí algún lugar que venda piononos? Es que me encantan los piononos!

Your friend: I want to take you out to eat traditional Puerto Rican food. What do you want to eat?

You: *Is there a place that sells piononos? It’s just that I love piononos!*

1B (Inner negation polar question) – confirm \( \neg p \)

Tu amigo: Vamos a tener que ir a Piñones para comprar piononos. En este barrio va a ser difícil.

Tú: ¿Por aquí no hay ningún lugar que venda piononos?

Your friend: We’re going to have to go to Piñones to buy piononos. It’ll be difficult in this neighborhood.

You: *There’s no place that sells piononos around here?*

1C (Outer negation polar question) – confirm \( p \)

Tu amigo: ¿Dónde quieres comer esta noche?

Tú: (Tú crees que hay cerca un lugar que vende piononos) ¿No hay por aquí un lugar que vende piononos?

Your friend: Where do you want to eat tonight?

You: (You think that there’s a place nearby that sells piononos) Isn’t there a place that sells piononos around here?

---

**Discourse Completion Task taking into consideration evidence and belief states**

<table>
<thead>
<tr>
<th></th>
<th>( \hat{H} ) L%</th>
<th>L * H L%</th>
<th>H + L * L%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unbiased</td>
<td>100% (28/28)</td>
<td>0% (0/28)</td>
<td>0% (0/28)</td>
</tr>
<tr>
<td>INPQ</td>
<td>79% (22/28)</td>
<td>21% (6/28)</td>
<td>0% (0/28)</td>
</tr>
<tr>
<td>ONPQ</td>
<td>57% (16/28)</td>
<td>0% (0/28)</td>
<td>43% (12/28)</td>
</tr>
<tr>
<td>Total</td>
<td>79% (66/84)</td>
<td>7% (6/84)</td>
<td>14% (12/84)</td>
</tr>
</tbody>
</table>

*Both INPQs and ONPQs are confirmation-seeking – what drives contour choice?*

Armstrong, M.E. (2017)). Accounting for intonational form and function in Puerto Rican Spanish polar questions. *Probus,* 29(1), 1-40
Both INPQs and ONPQs are confirmation-seeking – why is H+L*L% only preferred for one of these?

- **Epistemic bias** – speaker chooses contour to mark her belief that p
- This explains why H+L*L% never occurs in echo questions
- **Pragmatic uses** – desire

(2) Organizaste una comida y decides cambiar la fecha para que todos los invitados puedan ir. Pregúntales si van a poder venir si la comida es el primer domingo de mayo.

¿Pueden venir a la comida, si la hacemos el primer domingo de mayo?

You organized a dinner and you want to change the date so that everyone can come. Ask them if they can come to the dinner if you do it the first Sunday in May.

Can you guys come to the dinner if we have it on the first Sunday in May?
American English appears to have a similar intonational morpheme that has barely received attention in the literature, save Sadock (2013)
“EARLY RISE FALL-FLAT” CONTOUR

Did they lose your bags?
DEFINING THE ERFF - FORM

“Are you insane?” – speaker thinks Mindy is acting like she is insane
“Didn’t Tyshawm move his spin class to Bushwick?” – speaker has the idea that Tyshawn moved his class to Bushwick
DEFINING THE ERFF - MEANING

Sadock (2013) “an unusual falling but flat intonation pattern” that he proposes conveys bias on the part of the speaker

Contour suggests a conclusion of the speaker based on 1st-hand evidence, “a sort of epistemic in English” – Sadock goes as far to suggest that this contour is a case of grammaticalized evidentiality

Examples of both positive and negative bias – proposes that these are defined situationally
OTHER (INTUITED) OBSERVATIONS

ERFF may occur with plain inverted PQs, but doesn’t work with declarative questions.

(3) #John’s from Iowa? ERFF

Can appear with, but not obligatory for PQs that have verum markers (Romero, 2002) that trigger epistemic implicatures – e.g. preposed negation, really

(4) Isn’t John from Iowa? (rise or ERFF)

(5) Is John really from Iowa? (rise or ERFF)

N.B. it seems that in AE preposed negation is much more common to confirm p rather than ~p –

(6) ?Isn’t Jane coming either?

✔ Isn’t Jane coming too?
ERFF appears with morphemes that are known to convey positive (preposed negation of the form Aux n’t) or negative (really) epistemic implicatures (Romero & Han 2004)

What exactly is the contribution of ERFF and what is it doing when there are other verum morphemes present?
PERCEPTION EXPERIMENT 1

Objective: Compare listeners’ perceived belief scores for decontextualized utterances produced with L*H-H% vs. ERFF with different morphemes known to generate epistemic implicatures (i.e.

Polar questions of the form *Are you from [U.S. city or state]*?

3 structural conditions:

Are you from Texas? **plain inverted**

Aren’t you from Texas? **preposed** *(Auxn’t)*

Are you really from Texas? **Really**

------------------------------------------------------------------------

2 intonational conditions:

Rise (L*H-H%)

ERFF
PERCEPTION EXPERIMENT 1

- 102 native speakers of American English
- 60 females, 42 males from varying regions

Stimuli produced by 1 male and 1 female
Inverted PQ: 4 per speaker x 2 speakers = 8
Preposed neg: 4 per speaker x 2 speakers = 8
Really: 4 per speaker x 2 speakers = 8
Fillers = 16

Experiment administered on SurveyGizmo
24 trials/speaker analyzed – total of 2448 trials analyzed
EXPERIMENT 1 - PROCEDURE

For each trial, participants heard a recording of a decontextualized question

Click on the file to hear what this person says.

5. Choose the number you think best represents what this person thinks.

-5 They are not from Springfield
-4
-3
-2
-1
0 He doesn't know either way
1
2
3
4
5 They are from Springfield

Negative polarity
Positive polarity
RESULTS — EXPERIMENT 1

Are you from Texas? [rise] vs. Are you from Texas [ERFF]
RESULTS — EXPERIMENT 1

Aren’t you from Texas? [rise] vs. Are you from Texas [rise]
RESULTS — EXPERIMENT 1

Aren’t you from Texas? [rise] vs. Are you from Texas [ERFF]
RESULTS — EXPERIMENT 1

Aren’t you from Texas? [rise] vs. Aren’t you from Texas [ERFF]
RESULTS — EXPERIMENT 1

Are you from Texas? [rise] vs. Are you really from Texas [rise]
RESULTS — EXPERIMENT 1

Are you really from Texas? [rise] vs. Are you really from Texas [ERFF]
EXPERIMENT 1 - SUMMARY

In the absence of verum, ERFF conveys more positive bias than rise

If verum is present → ERFF may reinforce/enhance:

Auxn’t ERFF sig. reinforces positive ep. implicature

May also disambiguate: negative epistemic implicature for really + ERFF

Auxn’t PQs decontextualized tend to be interpreted as positively biased (contra prior claims that these are ambiguous in polarity – Ladd 1981; Romero & Hahn 2004)
EXPERIMENT 2

ERFF triggered positive epistemic implicature, but according to Sadock negative epistemic bias can also be conveyed with this contour

Sadock claims this is modulated by “situational” context

Can both positive and negative implicatures be triggered by ERFF?
EXPERIMENT 2

25 speakers of American English
21 females, 4 males

Bias types: baseline (16 contexts) vs. +belief (16 contexts)
   Of the 16 biased contexts half positive (p) half negative (~p)
2 contours: rise vs. ERFF
16 fillers (incredulity and contrastive focus PQs pitted against rises)
Stimuli produced by 1 male and 1 female speaker (male was different from Exp. 1, female same)
EXPERIMENT 2 - MATERIALS

Sara told Ann she would help find a DJ for Ann's husband's birthday party. Sara has never met Ann's husband but needs to know about his music preferences. She says:

Does he like country music

No bias

Sara's mom and sister were supposed to wait for her outside the restaurant. Sara comes back and her sister isn't there. Her sister has no patience, and Sara is sure she already went inside. She says to her mother:

Did she go inside

Positive bias

Sara's roommate was supposed to buy a new phone today, but Sara gets home and sees she hasn't even gotten out of bed. Sara suspects she did *not* buy a new phone yet. Sara says:

Did you buy a new phone

Negative bias
EXPERIMENT 2 - PROCEDURE

Sara told Ann she would help find a DJ for Ann’s husband’s birthday party. Sara has never met Ann’s husband but needs to know about his music preferences. She says

Does he like country music

Option 1

Option 2

39. Which question sounds best given this situation? *

- Option 1
- Option 2

32 targets * 20 participants = 640 trials
EXPERIMENT 2 - RESULTS

Clear preference for rise in contexts where speaker has no specific bias, ERFF when speaker has some sort of belief.
EXPERIMENT 2 - RESULTS

Same relationship for both types of conditions, though slightly less so for negative contexts.

No effect of bias type (pos vs. neg), i.e. ERFF just as likely to be chosen for pos. contexts as for neg.
EXPERIMENT 2 - SUMMARY

• Results confirm that ERFF is preferred over rise in context where speaker has some sort of belief about p, dispreferred in baseline contexts

• The rise is also acceptable for listeners in for the belief contexts

• Whether the ERFF gives rise to a positive or a negative implicature is context dependent (confirms Sadock’s proposal)

• Slight preference for ERFF in cases where a positive epistemic implicature would be appropriate, though not sig. (this is reflected in Experiment 1’s results as well) – in absence of context when listeners are forced they choose positive epistemic implicature
Like other verum forms, ERFF triggers epistemic implicatures.

In the presence of other verum morphemes (preposed negation, really) the ERFF may have a reinforcing/enhancing effect (preposed negation) or disambiguate (really).

In the case of AE PQs, ERFF is an epistemic intonational morpheme that must be hosted by true interrogatives (i.e. syntactically inverted).

In the absence of other verum morphemes ERFF does all the work.

In the presence of other verum morphemes (preposed negation, really) the ERFF may reinforce (Auxn’t) or disambiguate (really).

ERFF can be seen as an epistemic operators used to assert that the speaker is certain that $p$ (or $\neg p$) should be added to the Common Ground.
PRAGMATIC USES OF ERFF

Epistemic meaning can be extended for ERFF:

Child in a car on a long journey:
Are we there yet? (ERFF)
Here ERFF indicates desire on the part of the child

Parent: Do your homework!
Child: Do I have to? (ERFF)
Belief meaning still there (child is aware that he does have to) but contour also indicates his lack of desire
WHAT ABOUT CHILDREN?
MENTAL STATE LANGUAGE

- Mental state language is known to be acquired late and often does not occur in children’s speech until late in the 3rd year.

- Following this, we would predict that contours that convey mental states would also be acquired late.

- In PRS jH*L% is a more general contour that does not convey any specific belief information – this is more frequent in child directed speech and is the contour that children produce from the onset of the corpora.
Longitudinal study of two female PRS-acquiring toddlers (19-43 mo.) and their caretakers shows “epistemic” contours are much lower frequency in child-directed speech (CDS)

Table 1. % production of yes-no question contours in child-directed speech (CDS)
Toddlers barely used epistemic yes-no question contours (99% iH* L% > 1% H+L* L% > 0% L* HL%) (N=661)

2 examples of H+L* L% per child, all occurred after 2;6

In comprehension, awareness of response eliciting nature:

<table>
<thead>
<tr>
<th></th>
<th>Ana</th>
<th>Cristina</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>iH* L%</td>
<td>988/1480 (67%)</td>
<td>1735/2789 (62%)</td>
<td>2723/4269 (64%)</td>
</tr>
<tr>
<td>H+L* L%</td>
<td>8/11 (72%)</td>
<td>33/72 (49%)</td>
<td>41/83 (49%)</td>
</tr>
<tr>
<td>L* HL%</td>
<td>45/61 (74%)</td>
<td>95/161 (59%)</td>
<td>140/222 (63%)</td>
</tr>
<tr>
<td>Total</td>
<td>1041/1552 (67%)</td>
<td>1863/3022 (62%)</td>
<td>2904/4574 (63%)</td>
</tr>
</tbody>
</table>
PERCEPTION OF BELIEF STATE INTONATION

When are children able to perceive these meanings?
L*HL% - DISBELIEF INTONATION IN PRS

36 PRS-speaking: 8 4-year-olds, 16 5-year-olds, 12 6-year-olds (M=54 mo.)

All from schools in San Juan Metro Area
Figure 1. ¿Un mono? ‘A monkey?’ produced with jH* L%
Figure 2. ¿Un mono? ‘A monkey?’ produced with L* HL%
PROCEDURE

Verónica

Marisol

Jeni

Yo vi...
Verónica
Marisol
Jeni
METHODS

12 test trials, 6 fillers/participant
Participants were familiarized (2 trials) with feedback
432 child trials
QUANTITATIVE RESULTS

- All groups above chance
- No sig. differences between 4- and 5-year-olds
- 6-year-olds outperformed 4-year olds (*), and 5-year-olds (*)

Figure 3. Mean trials correct for 4-, 5- and 6-year olds
OTHER OBSERVATIONS

6-year-olds produced facial gestures known to be associated with question marking and incredulity (Crespo-Sendra et al. 2013)
Hübscher, Esteve-Gibert, Igualada and Prieto (2017) investigated Catalan-speaking (ages 3-5) children’s comprehension of certainty vs. uncertainty through lexical, intonational and gestural cues. Younger and older children performed better in detecting uncertainty when visual cues (e.g. facial gestures related to uncertainty) were present, and suggest that visual information may help bootstrap children into linguistic meaning, as has been proposed in other work (Kelly, 2001; Butcher & Goldin-Meadow, 2000; McNeill, Cassel & McCullough, 1994).
Moore, Harris and Patriquin (1993) compared the ability of children aged 3 to 6 to comprehend degrees of certainty conveyed through prosody vs. mental state verbs like think, guess and know.

Youngest children could not use either type of cue, while older children showed an advantage for lexical information over prosody.

Authors stress the fact that children in this age group are developing the ability to make inferences about mental states as conveyed through prosody and the lexicon.

Suggest a child’s ‘representational Theory of Mind’ must be developed to a certain degree in order to comprehend mental state language, regardless of whether it is expressed prosodically or lexically.
DISBELIEF IN CENTRAL CATALAN

Are there differences in how children perceive gesture through intonation vs. facial cues (or the combination of the two)?

Does ToM reasoning help children to perceive disbelief through these different modalities?

Armstrong, Esteve-Gibert, Hübscher, Igualada & Prieto, under revision
187 Central Catalan-speaking children (89 female and 98 male), mean age 4;5 (ranging from 2;10 to 6;3)

30 Central-Catalan speaking adult controls
MATERIALS

Powerpoint with either audio-only (AO), audiovisual (AV) or video-only (VO) stimuli

Between subjects design (children got just one of the three conditions)

Child had to decide which of a set of twins did not believe his/her friend about an animal that the friend claimed to have seen while on vacation

Theory of Mind task (will not be reported on in this talk but see Armstrong, Esteve-Gibert & Prieto 2014)
Pitch track, spectrogram and waveform for the neutral echo question **Una balena?** ‘A whale?’

Pitch track, spectrogram and waveform for the disbelief echo question **Una balena?!** ‘A whale?!’
MATERIALS

Two blocks

Before Block 1 – 4 familiarization trials

Block 1 – 6 targets/2 fillers

Before 2 – 2 more familiarization trials

Block 2 – 6 targets/2 fillers

12 targets x 79 participants = 948 trials

Counterbalanced blocks
…% no creer, gos"
MATERIALS

Sally Ann type false belief task

Child asked where Princess will look for ball when she arrives.

Scored 0 or 1
ANALYSIS/RESULTS

DV: Correct (0 or 1)

Predictors: Age (months), Condition (AO, VO, AV), ToM

Best-fit model included Age, Condition and ToM (no interactions)

- Sig. effect of Age — children improved on all three tasks with age
RESULTS

Sig. effect of Condition when AV compared to AO, also when AV compared to VO (not too different from age range where we see other types of mental state language being developed, e.g. epistemic modals – Ozturk & Papafragou, 2014)

However no sig. difference when AO compared to VO

Overall effect of ToM - in general children who passed the ToM task outperformed children who failed it
DISCUSSION

Recent work shows clearly that the age window of 3-5 is important for intonational development, this seems to be true for perceiving belief through facial gesture as well.

Children improved with age on all three conditions.

Children with more sophisticated belief reasoning skills fare better.

Difference in performance on unimodal tasks (AO vs. VO not statistically significant) – no evidence for bootstrapping through gesture.

Hübscher et al. showed advantage for VO condition – why different?
Hübscher et al. explored uncertainty meaning as expressed by the L* H% contour in Central Catalan, this study focus on the L* LH% contour for expressing disbelief.

Specific tune-meaning relationship may be important in terms of how much children might rely on gestures – role of biological codes?

Perhaps rely more on gesture for less reliable form-meaning pairings.

Cross-linguistic differences have been found for the degree to which adults rely on facial gesture – Valencian Catalan speakers relied on facial gesture more heavily than Dutch speakers in interpreting incredulity meaning.

In any case results suggest that intonation and facial gesture are developing at about the same time, and it is helpful to children when the two cues reinforce each other.
SUMMARY

Contour choice is often driven by a speaker’s epistemic state.

As with any pragmatic analysis, important to understand what unifies intonational meaning vs. what sorts of interpretations or contextually driven.

Important to sort out meaning in order to test children’s understanding of intonational meaning.

Even though we know infants are very sensitive to prosody, ages 3-5 important for development of intonational meaning, perhaps not surprising given trajectory of ToM development (need for studies that include other cognitive factors).

The ability to understand intonational meaning may develop in tandem with the ability to understand facial gestures, possible that these influence each other.
THANK YOU!!