Lecture 2

Intonation and emotion, affect, & illocutionary force

CN MSc course ‘Language Acquisition’, 3 March 2011

Instructor: Aoju Chen
The pre-verbal stage (1)

- A variety of vocalisations by two months of age
  - Categories of sounds in English-listening two-month olds (Legerstee 1991)
    - Long sounds with variable pitch contours (melodic)
    - Nasal-like sounds with flat pitch (vocalic)
    - Sounds (cries, fussing, laughing) (emotional)

- Infants modulate sounds depending on contexts
  - from Legerstee’s (1991) study
    - More melodic sounds when conversed with by mother or another woman than in any other context
    - more vocalic sounds when adults were not unresponsive
    - More emotional sounds to people than to objects

- control over vocal productions from early on
The pre-verbal stage (2)

- Changes in pitch range and contour shape as a function of (social) contexts (age range: 4 ~ 9 months)
  - Presence vs. absence of the mother
    - Call cries & request (non-cry) vocalisations vs. discomfort cries
      - Pitch register
        - Call cries > discomfort cries (in level contours and to a less extent rising contours) > request vocalisations
      - Contour types
        - More rises in call cries (mostly rises) & request vocalisations than in discomfort cries (mostly falls and levels)
  - With vs. without eye contact with the mother
    - Higher pitch register and more rises in vocalisations produced with eye contact

(D’Odorico 1984, D’Odorico & Franco 1991)
The pre-verbal stage (3)

- Vary pitch and duration to express communicational intentions at 10 months
  - Expressing infant’s participation in interpersonal exchanges
    - E.g. looks at/touches/seeks/searches mother, shows, offers etc.)
  - Reflecting internal perceptual and cognitive processing (‘thinking aloud’)
    - E.g. holds/inspects objects
- Main findings
  - A higher mean f0, and max f0 in interpersonal exchanges
  - A shorter duration in interpersonal exchanges (why?)

(Papaeliou & Trevarthen 2006)
The late babbling and one-word stage

1. Requests
   - Social gestures central to our everyday life (Garvey 1974)
   - The earliest ‘acts of meaning’ present in children (Halliday 1975)

2. The use of initial pitch and contour shape in requests vs. labelling in French infants (Marcos 1987)
   - Two types of requests
     - Initial requests: first or sole request for an object or for co-operation
     - Repeated requests: only for objects
   - A cross-sectional design involving 2 groups of infants (1;2 ~ 1;10)
   - Elicitation methods for labelling and requests
   - Data: all babbles, proto-words and words
   - Judgments of function based on:
     - The infant’s gestures
     - The mother’s responses
     - The infant’s reactions to the mother’s responses
The late babbling and one-word stage (2)

- **Vocabulary size**
  - At 1;2: only 1 infant produced recognisable words
  - At 1;5: all the infants used one or several words
  - At 1;9: only 2 infants reached the stage of two-word utterances

- **Initial pitch**
  - repeated requests > initial requests (significant at 1;31/2, 1;61/2, 1;8)
  - Initial requests > labeling (significant at 1;2, 1;31/2, 1;5, 1;61/2)
  - No age/vocabulary size related effect on the use of pitch

(Marcos 1987)
The late babbling and one-word stage (3)

Fig. 1. Percentage of falling, rising and level contours for requests and for labelling at each age level. R: requests; L: labelling; ......., rising contour; ———, falling contour; ----, level contour.

(Marcos 1987)
Individual differences in the use of intonation in a range of functions

- 3 English-listening infants
  - age at time of testing: RS - 0;11.09, AB - 1;0, AL- 1;2.24
- Audio- and videotaped three times (1 hour/session) over a period of about 6 months while playing with the mother
  - 1st taping: before the onset of single words
  - 2nd taping: when vocabulary consisted of 10 words
  - 3rd taping: when vocabulary consisted of 50 words
- Data included
  - Babbling
  - Single-word utterances
  - Non-word like vocalisations during the one-word utterance stage

Flax, Lahey, Harris, and Boothroyd (1991)
The late babbling and one-word stage (5)

- Function categories
  - Requests
    - for object or action
    - for attention
    - for response
    - in loud voice and repeated until mother complied
    - Response to other’s speech
  - Comment-(non)Interactive
  - Protest
  - ...

- Acoustic analysis
  - Mean pitch
  - Max pitch
  - Pitch span
  - Final contour (rise vs. non-rise)

(Flax, Lahey, Harris, & Boothroyd 1991)
Differences in the use of rise
- between children (AL-54%, 14%, 12%, for each child respectively)
- between function categories

Similarities in the use of rise
- Consistency over time of the overall % of total rises within each child (AL > RS, AB)
- The % of rises in a particular category was consistent within each child across time
  - The greatest % of rises in requests
  - Frequent use of falls in comment-(non)interactive, response to others’ speech, repeated requests until mother responded

(Flax, Lahey, Harris, & Boothroyd 1991)
The late babbling and one-word stage (7)

- Mean pitch, pitch maximum
  - Higher in requests for attention and protest than in other functions
- Pitch span
  - Requests for attention: greatest span
  - Response to others’ speech: narrowest span

(Flax, Lahey, Harris, & Boothroyd 1991)
The late babbling and one-word stage (8)

- Similarities between French- and English-listening infants and continuity from pre-verbal stage to early verbal stage
  - Similarities in the form-function relations in adult French and English?
  - Intonation universals
    - The biological codes (Gussenhoven 2002, 2004)
  - Teasing apart cross-linguistic similarities in the grammar from (paralinguistic) intonation universals
    - e.g. risings statements in Belfast English (Jarman & Cruttenden 1976) and Chicksaw (Gordon 1999); falling questions in Roermond Dutch (Gussenhoven 2000) and Chicksaw
Sensitivity to intonational cues to emotion and affect (1)

- 2-month-olds
  - Looking longer at visual stimuli upon hearing rising ‘ba’ than hearing falling ‘ba’ in a female voice (Sullivan & Horowitz 1983)

- 3-month-olds
  - showing fear when hearing an angry voice (Popich 2003)

- 4-month-olds
  - Looking longer at visual stimuli upon hearing approving rise-fall than hearing disapproving rise-fall (Papousek et al. 1990)
Sensitivity to intonational cues to emotion and affect (2)

- Intonation, lexical content, & body language packed in a message
  - Words vs. paralanguage (intonation + body language)
    - At 9 months: responding more to paralanguage (Lawrence & Fernald 1993)
    - At 18 months: responding more to lexicon (Lawrence & Fernald 1993)
    - 15-months olds (Friend 2001)
  - Intonation vs. body language (e.g. Quam et al. 2009)
  - Intonation vs. words (e.g. Morton & Trehub 2001)
Infants’ interpretation of messages with approving or disapproving lexical content combined with approving or disapproving intonation and facial expressions

- Don’t touch
- Bad stop
- Nice play
- Good look

63 infants (mean age: 15 months, 14m12d ~ 17m3d)

Friend (2001)
Prior to testing parents were mailed a copy of the CDI with instructions to complete the instrument at home

- A 10-minute warm-up period
- Infant in high chair, parent holding magazine over his/her face (E1)
- One practice trial (familiarising with the testing situation)
- Four test trials
  - E2 placed novel toy on the tray, and moved the toy beyond the curtain
  - E2 played the 1st videotaped stimulus
  - Toy within the infant’s reach
  - Stimulus repeated
  - Toy remained within infant’s reach for 15s
Sensitivity to intonational cues to emotion and affect (5)
Sensitivity to intonational cues to emotion and affect (6)

Don’t touch

Don’t touch
Sensitivity to intonational cues to emotion and affect (7)

- **Between-subject Independent variables**
  - (dis)agreement between lexicon and paralanguage (into, fac.exp)
    - Consistent (lexical content = paralanguage)
    - Discrepant (lexical content ≠ paralanguage)
  - Stimulus order, Gender

- **Co-variants**
  - Lexical item comprehended (total score 8)
    - Girls (5.13, 1 ~ 8) vs. Boys (4.35, 0 ~ 8)
  - Receptive vocabulary
    - Girls (198.16, 51 ~ 370) vs. Boys (162.68, 18 ~ 285)

- **Dependent variables**
  - Difference in ‘delay to approach’ between approving lexicon and disapproving lexicon
  - Difference in ‘manipulation/play time’ between approving lexicon and disapproving lexicon
Sensitivity to intonational cues to emotion and affect (8)

- Differences in ‘delay to approach’

<table>
<thead>
<tr>
<th>Consistent</th>
<th>Approving lexicon</th>
<th>Disapproving lexicon</th>
<th>&lt; 0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Approving paralang</td>
<td>Disapproving paralang</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discrepant</th>
<th>Approving lexicon</th>
<th>Disapproving lexicon</th>
<th>&lt; 0 (m) Lex</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Disapproving paralang</td>
<td>Approving paralang</td>
<td>&gt; 0 (m) para</td>
</tr>
</tbody>
</table>

20
Sensitivity to intonational cues to emotion and affect (9)

- Differences in ‘manipulation time’

**Consistent**
- Approving lexicon - Disapproving lexicon > 0
- Approving paralang - Disapproving paralang

**Discrepant**
- Approving lexicon - Disapproving lexicon > 0 (m) Lex
- Disapproving paralang - Approving paralang < 0 (m) para
Infants’ responsiveness to emotion and affect in the voice (11)

- Lexical comprehension was a significant predictor of infant play.
Sensitivity to intonational cues to emotion and affect (12)

- Why do some 15- and 18-months olds respond more to lexicon?
  - Processing limitations on the expression of meaning (Bloom 1993)
  - Impact of newly emerging lexicon
  - Not knowing what is the most reliable cue when multiple cues are present
Sensitivity to intonational cues to emotion and affect (13)

- Children’s use of intonation to infer a speaker’s emotions
  - Exp. 1: Puppy searches for a toy (e.g., the *Toma*) on each trial. He is excited when he finds the *Toma*, and disappointed when he finds a different toy.
  - Task: Give Puppy the Toma, throw the other toys in the trash

**Condition 1: body-language cue**
- **Toy 1** (not Toma) Puppy shakes head, slumps
- **Toy 2** (Toma) Puppy nods, dances
- **Toy 3** (not Toma) Puppy shakes head, slumps

**Condition 2: pitch cue**
- **Toy 1** (not Toma) Low, flat pitch (on “Mmm”)
- **Toy 2** (Toma) High pitch, wide excursions
- **Toy 3** (not Toma) Low, flat pitch

(Quam et al. 2009)
Children’s use of intonation to infer a speaker’s emotions

- Exp. 1: Puppy searches for a toy (e.g., the Toma) on each trial. He is excited when he finds the Toma, and disappointed when he finds a different toy.
- Task: Give Puppy the Toma, throw the other toys in the trash

(Quam et al. 2009)
Sensitivity to intonational cues to emotion and affect (15)

- Children’s use of intonation to infer a speaker’s emotions
  - Exp. 2: Puppy searches for his lost toy in each trial.
  - Task: If Puppy is happy, point to the happy face.
    If Puppy is sad, point to the sad face.

Condition 1: body-language / facial cues
Toy 1 (not lost toy) Experimenter frowns, she & Puppy slump
Toy 2 (lost toy) Experimenter smiles, she & Puppy dance

Condition 2: pitch cue (on ‘Oh, look at that’)
Toy 1 (not lost toy) Low, flat pitch
Toy 2 (lost toy) High register, wide excursions

(Quam et al. 2009)
Sensitivity to intonational cues to emotion and affect (16)

- Children’s use of intonation to infer a speaker’s emotions
  - Exp. 2: Puppy searches for his lost toy in each trial.
  - Task: If Puppy is happy, point to the happy face. If Puppy is sad, point to the sad face.

(Quam et al. 2009)

<table>
<thead>
<tr>
<th>Age</th>
<th>Body-lang</th>
<th>Pitch</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 yrs</td>
<td>N=4</td>
<td>N=1</td>
</tr>
<tr>
<td>3 yrs</td>
<td>N=6</td>
<td>N=10</td>
</tr>
<tr>
<td>4 yrs</td>
<td></td>
<td>N=11</td>
</tr>
<tr>
<td>5 yrs</td>
<td></td>
<td>N=10</td>
</tr>
</tbody>
</table>
Interpretation of ‘happy’ and ‘sad’ in messages with conflicting or non-conflicting intonation and lexical content

Children’s limited attentional resources
- 3- and 4-year-olds’ stories are rich in prosody but poor in content
- 7- and 8-year-olds’ stories are rich in content but poor in prosody

(Morton and Trehub 2001)
Participants
- 145 English-speaking children aged 4 to 10 (about 20 in each of the 7 age groups)
- 20 adults (native speakers of English)

Stimuli
- 40 spoken utterances spoken by a female speaker
- Duration does not differ for sentences with conflicting and non-conflicting cues
- Task: judge whether the speaker feels happy or sad

(Morton and Trehub 2001)
Sensitivity to intonational cues to emotion and affect (19)

- **Response latency** measured from the stimulus onset to entry of response (‘happy’ or ‘sad’)
  - A longer response time to stimuli with conflicting cues (both children and adults)
- Implication: Children processed both sources of information and seemed to notice abnormality in the sentences

(Morton and Trehub 2001)
Sensitivity to intonational cues to emotion and affect (20)

- Happy/sad judgments
  - Adults: exclusively based on intonation
  - 4-year-olds: primarily based on lexical content
  - Between 5 and 10 years of age: a decrease in their reliance on lexical content
  - Half of the 10-year-olds responded to intonation when the cues were conflicting

- Not an issue of understanding intonation
  - Morton & Trehub’s Exp. 2 and 3
  - But see finding in Quam et al. (2009)

- But an issue of a limited understanding of the role of intonation in communication
  - when the words conflict with the intonation, they “do not treat intonational cues as a basis for qualifying or overriding the literal message” (Morton and Trehub 2001: 841)
Sensitivity to intonational cues to emotion and affect (21)

- **Intonation, lexical content, & body language**
  - Words vs. paralanguage (intonation + body language)
    - 9 month-olds: responding more to paralanguage (Lawrence & Fernald 1993)
    - 15-months olds: responding more to paralanguage but …. (Friend 2001)
    - 18 month-olds: responding more to lexicon (Lawrence & Fernald 1993)
  - Intonation vs. body language (e.g. Quam et al. 2009)
    - Children younger than 5 failed to interpret pitch cues
  - Intonation vs. words (e.g. Morton & Trehub 2001)
    - Children younger than 10 relied on lexical content

- **Concluding remarks**
  - An astonishingly slow process
  - Individual differences?
    - e.g. in development of a sense of humor

Questions vs. statements in natural production at different ages (Patel & Grigos 2006)

- 4 English-speaking children in 3 age groups (4;4, 7;4, 11;3)
- Repetitions of two phrases spoken as statements & questions via a naturalistic elicitation technique
  - 4 puppets: Pop (grandpa), Bob (Sponge Bob), pot (a pot), bot (a robot)
  - Children’s task: instruct the experimenter to perform an action using one character and one object
    - (Bob is lonely. What should I show Bob?) Show Bob a bot.
    - (Bob is hungry and needs something to make soup …). Show a Bob a pot? (not a typical yes-no question)

Four variables (in each syllable): duration, mean F0, F0 slope, mean intensity (slightly higher in Q than in S, only significant in 7-yrs)
Intonation & Illocutionary force (2)

Patel & Grigos (2006)
Intonation & Illocutionary force (3)

Patel & Grigos (2006)
Age-related differences
- 4-year-olds reply mostly on duration, as found in earlier work (Allen & Hawkins 1980)
- 7-year-olds use all three cues in tandem
- 11-year-olds use primarily f0 and their use of duration is less exaggerating than that of younger children, like adults

Why less consistent use of final rise in the 4-year-olds?
- More complex to rise pitch
  - The change in pitch slower in a rising contour than in a falling contour (Xu and Sun 2002)
  - 4-year-olds can imitate rising contours but with a smaller pitch span (Snow 1998)
- Production of voiced plosives
- Or....
Intonation & Illocutionary force (6)

- Question intonation in French children (Gérard & Clément 1998)
  - 3 speakers from each age group: 5 yrs, 7 yrs, 9yrs, adults
  - Picture-description
    - *Maman dit* ‘Mommy said’ … (statements)
    - *Quoi* ‘what?’ (declarative questions)
      - e.g. ‘On emmène Michel en vacances’

- Main findings
  - Overall contour: all groups of speakers used rise in questions
  - Duration (of last syllable of each disyllabic word):
    - Adults: Longer duration in questions (only in 1\textsuperscript{st} disyllabic word)
    - 7- and 9-yrs: opposite pattern (on different words)
    - 5-yrs: overall lengthening in questions
Intonation & Illocutionary force (7)

- English 4-year-olds vs. French 5-year-olds
  - Final lengthening in both groups
  - Final rise in French children
    - Why this difference?
2-year-olds rely on intonation rather than word order to identify questions in English (Leder & Egelston 1982)
- Lack of knowledge of the function of inversion

How accurately can children identify questions using only intonational cues? (a gating experiment by Gérard & Clément 1998)
- At the whole-sentence level (on the 5th word)
  - Adults: perfect identification
  - 9-year-olds: very accurate identification (80%, 85%)
  - 7-year-olds: 65%, 55%
  - 5-year-olds: 20%
- Anticipatory identification
  - Only present in adults (right from the 2nd word)
- The results are surprising ….