Phonological development: The acquisition of a (really) complex system

Jan Edwards¹, Mary E. Beckman², Benjamin Munson³, Eunjong Kong¹, and Fangfang Li⁴

¹University of WI – Madison, ²Ohio State University, ³University of MN, ⁴Lethbridge University
A complex web of collaborators
A complex web of collaborators

Edwards

Todd
Arbisi-Kelm
Chung
Syrika

Beckman

Schellinger

Munson

Urberg-Carlson

Li

Johnson

Holliday

Yoneyama
A complex web of collaborators
Is phonological learning trivial?

• Speech sounds are often thought of as the mere 'front end' of higher-level language.

• Pinker: “[Learning of] words and grammar are different. The sequence of sounds making up a word is not capturable by rules (monkey cannot be understood as a combination of mon and key), but must be memorized. And because there are a finite number of words, they all can be recorded.” (Science, 1997)
Assumptions underlying a traditional "phonology as mere front end" view

1. Children are acquiring abstract phonological categories when they are learning to produce sounds correctly.

2. There are more-or-less universal patterns of development.

3. Acquisition can be studied through alphabetic phonemic transcriptions alone.

4. Acquisition after about age 5-6 is primarily related to fine-tuning of motor skills.
Organization of talk

1. Model of phonological knowledge.

2. Do children acquire abstract phonological categories such as the phonemes /s/ vs. /ʃ/ (“sh”) directly?

3. Do children acquire the same sounds, such as /s/ and /ʃ/, in the same way across languages?

4. Is alphabetic transcription of sounds like /s/ and /ʃ/ adequate to capture phonological development?

5. Is acquiring categories like the phonemes /s/ and /ʃ/ all there is to phonological development?

6. Clinical implications.
Levels of knowledge about speech sounds

categorical phonological knowledge

[\textit{t}]onset

\textit{stressed}[\textit{i}] \textbf{back}[\textit{k}]\textit{onset}

\textit{front}[\textit{k}]\textit{onset}

\textit{stressed}[\textit{u}]

\textit{trochee}

words

\{key\} \quad \{tuna\} \quad \{cougar\}

language-specific phonetic detail

language-specific phonetic detail

EF:

man's <key> \quad man's <tuna> \quad girls' <cougar>

voices

socio-indexical knowledge

Buffalo

gay

adult

NYC

male

child

female

people

Aunt Jan

Marie

Ben
1. Children learn sounds in words

• Claim: Abstract phonological categories develop gradually.
  – Relationships between phoneme frequency and production accuracy across languages.
  – Relationships among phoneme frequency, production accuracy, and vocabulary size in English.
Phoneme frequency and accuracy: English (Vodopivec, 2004)

- Picture naming task
- 3- to 5-year-olds with phonological disorders and typically developing age controls.
Consonant accuracy, frequency, and vocabulary size: English (Edwards et al, 2004)

- Nonword repetition task
- 3- to 7-year olds and adults

High frequency sequences
Low frequency sequences

EVT Raw Score (LN)
Lexical generalization hypothesis

• Edwards et al., 2004 interpretation: Children make phonological generalizations over their lexicon.
  – The larger the lexicon, the more robust and context-independent these phonological generalizations are.
  – This is why children with larger lexicons are more accurate at producing familiar sounds in novel contexts.

• Alternative interpretation:
  – Perhaps some sounds and sound sequences are low in frequency because they are more difficult to produce or perceive.
Why we need to examine phonological development cross-linguistically

• The frequency of sounds and sound sequences differs across languages.
  • /tʃ/ (“ch”)
    – high-frequency in Japanese
    – low-frequency in English
    – non-existent in Greek
  • /sɪ/ (“see”)
    – high-frequency in Greek
    – non-existent in Japanese
  • /kʲo/ (“kyo”)
    – non-existent in English
    – mid-frequency in Japanese
  • /tʊ/ (“too”)
    – high-frequency in English
    – non-existent in Cantonese
The παιδολογός database

• Targets: word-initial lingual obstruents in 5 vowel contexts in:
  – Cantonese, English, Greek, Japanese
  – Also, Korean, Mandarin, Taiwanese, French, ….

• Participants:
  – About 20 2-, 3-, 4-, 5-year-olds and adults / language

• Procedure:
  – Elicit single word repetitions of target CVs in familiar words and nonwords.
  – Data collected in Hong Kong, Tokyo, etc.

• Measures:
  – Native-speaker transcriptions of target word-initial consonants
  – Acoustic measures
  – Naïve listeners’ perceptions
Example stimuli for /k/ in English

- kaytush
- key
- coffee
- cube
- cutting
- cougar
- coat
- quick
Example stimuli for /k/ in Japanese

[kimono] [keki] [kuri] [kuma] [kuruma] [kibi] [kaba] [koara]
Relating accuracy to frequency: Cantonese, English, and Japanese
Consonant accuracy, frequency, and vocabulary size (again): English

- Significant relationship between accuracy and frequency.
- The slope of this function differs across participants.
- Slope = frequency effect.
1. Conclusion: Children learn sounds in words

- CV frequency is correlated with consonant accuracy across languages.
- This influence of CV frequency on accuracy decreases as vocabulary size increases.
  - It’s not simply the case that low-frequency sounds and sequences are more difficult to produce/perceive.
  - Low-frequency sounds and sequences are less accurate because children have less practice hearing and producing them.
2. Phonological learning is language-specific

- Cross-linguistic differences in phoneme frequency do not explain all language-specific patterns.
  - For example: Sibilant fricative contrast in Japanese acquired later than similar contrast in English, although phoneme frequencies are similar.

- At least some of these cross-linguistic differences seem to be related to differences in fine phonetic detail across languages.
Cross-linguistic differences in fine phonetic detail: Fricatives and fricative development
Questions to address

• Why is /s/ produced with such low accuracy by Japanese-speaking 2- and 3- year olds?

• Why is /s/ produced with such high accuracy by English-speaking 2- and 3- year olds?
Cross-linguistic differences in fine phonetic detail: Fricatives and fricative development (from Li et al., in press)

- Both English and Japanese have a contrast between /s/ and /ʃ/.  
- Large-scale studies show opposite patterns of acquisition and different error patterns.  
  - English:  
    - /s/ is mastered earlier than /ʃ/ and [s] is commonly substituted for /ʃ/ (Smit et al. 1991).  
      
      | shoe | safe |
      |------|------|
      | ∈ | ∈ |
      
  - Japanese:  
    - /ʃ/ is mastered earlier than /s/ and [ʃ] is often substituted for /s/ (Nakanishi et al., 1972).  
      
      | Shukurimu “cream puff” | semi “cicada” |
      | ∈ | ∈ |
      
      English: Fronting error  
      Japanese: Backing error
Acoustic analysis of adults’ productions: English vs. Japanese

English (female)

Japanese (female)

Onset F2 Frequency (Hz)

Centroid (Hz)

S /s/
S /ʃ/
Acoustic analysis of children’s productions:
English vs. Japanese

[Graph showing onset F2 frequency vs. centroid frequency for English and Japanese children's productions]
Cross-linguistic perception experiment: Rationale and methods (from Li et al., in press)

• Research questions:
  – To what extent is the apparent cross-linguistic asymmetry due to differences in perceptual norms?

• Prediction:
  – Given the production differences, we might expect that adult native speakers of English and Japanese would parse the multidimensional acoustic space differently.

• Participants:
  – 20 naïve adult native English speakers (Minneapolis, MN)
  – 20 naïve adult native Japanese speakers (Tokyo, Japan)
Cross-linguistic perception experiment: Methods

• **Stimuli:**
  – Correct productions of /s/ and /ʃ/ by children and adults, prototypical substitutions of children in each language.

• **Task:**
  – Each listener heard two blocks of the same 400 tokens.
  – In one block: “Is it an “s”? In other block: “Is it an “sh”?”
  – Responded by pressing “Yes” or “No” button as quickly as possible.
  – Naïve listeners didn’t know they were listening to multiple languages.

• **Analysis:**
  – Determine the ‘community consensus’ for each token by examining whether it received a ‘yes’ response by 70% or more listeners.
Cross-linguistic perception experiment: Results

**English listeners:**
- acceptable range for /s/ is larger than acceptable range for /ʃ/

**Japanese listeners:**
- acceptable range for /ʃ/ is larger than acceptable range for /s/
Phonological learning is language-specific: Conclusion

Questions:
• Why is accuracy of /s/ so high in English?
• Why is accuracy of /s/ so low in Japanese?

• Production: The contrast between /s/ and /ʃ/ is more robust in English than in Japanese.

• Perception: English listeners accept a wider range of productions as correct for /s/; Japanese listeners accept a narrower range of productions as correct for /s/.

Conclusion: At least some cross-linguistic differences in acquisition related to language-specific fine phonetic detail in perception and production.
3. How informative is alphabetic transcription?

- Assumption: acquisition can be studied through alphabetic phonemic transcriptions.
- However, accuracy judgments depend on listeners’ experience.
- Also, children do not always progress directly and categorically from incorrect to correct productions.
  - Covert contrast: systematic acoustic difference that is not perceptible (Macken, 1980)
  - Other intermediate productions
    - English: [k] or [g]
    - [f] or [θ]
    - Greek: [k] or [t]
    - [s] or [θ]
- Clinical importance of intermediate productions (Tyler, 1995)
Perception experiment: Intermediate productions and visual analog scaling (Schellinger et al., 2008)

• Research questions:
  – Can naïve listeners reliably categorize productions as intermediate between /s/ and /θ/ (“th”)?

• Prediction:
  – Naïve listeners would be able to do so, given the right task.

• Participants:
  – 20 naïve adult listeners in Minneapolis, MN

• Method: Visual analog scaling
  – Allows people to scale where a token falls relative to fixed endpoints.
  – The visual space is made essentially analogous to the perceptual space.
The “s” sound

The “th” sound
Perception experiment: Intermediate productions and visual analog scaling (Schellinger et al., 2008)

- Stimuli:
  - 200 CV sequences from single-word productions of English-speaking children, aged 2 through 5 years.
    - correct /s/
    - [s] for /θ/
    - intermediate: closer to [s] than [θ]
    - Intermediate: closer to [θ] than [s]
    - [θ] for /s/
    - correct /θ/
Perception experiment: Intermediate productions and visual analog scaling (Schellinger et al., 2008)

Mean response from more “th” like to more “s” like

Transcription category

“s” for target [s]  "s" for target [T]  "s:T" for either  "T:s" for either  "T" for [s]  "T" for [T]
Effect of experience (Munson, Johnson, & Edwards, 2010)

Speech-Language pathologists outperform inexperienced listeners on this task.

Their responses better differentiate among transcription categories.

They don't have as strong a bias to label sounds as 's'.

They have superior intra-rater reliability.
3. How informative is alphabetic transcription: Conclusion

- Not informative enough.
- Influenced by listeners’ experience.
- Children don’t always proceed directly and categorically from incorrect to correct productions.
- Children produce intermediate productions that can be reliably categorized even by naïve listeners, given an appropriate task.
- Acoustic analysis and/or perception tasks are needed to describe these productions.
4. There’s more to phonological development than phonemes

• Speech sounds encode at least two kinds of information:
  – Lexical information
  – Socio-indexical information
    • Information about social identity such as gender, age, geographic origin, ethnicity, formality, etc. (e.g., Labov 1990; Purnell et al., 1999; Clopper & Pisoni, 2004).
    • For example, what do you know about these speakers?
Speech sounds and socio-indexical categories

• Gender-marking: /s/
  – Glaswegian English (Stuart-Smith, 2004)
  – Systematic differences in fricative spectrum for males and females.
  – Interacts with social class and age.

• Marking of sexual orientation: /s/ and vowels
  – American English (Munson et al., 2006; Munson, 2007)
  – Listeners use different acoustic parameters to judge male sexual orientation and masculinity.

• Very little research on acquisition of socio-indexical categories
Mandarin sibilant fricatives: Lexical phonetic contrast
(from Li & Kong, 2008)

• Mandarin has two post-alveolar fricatives:
  • /ç/ (“she”)  
  • /ʂ/ (“shr”)

• Onset F2 frequency (y-axis values) differentiates /ç/ (“she”) and /ʂ/ (“shr”).

• Socio-indexical coding for /ç/ (“she”) uses centroid frequency (x-axis).
Speech sounds and socio-indexical categories: Gender-marking by adults in Mandarin (Li & Kong, 2008)

- The difference between /ç/ (“she”) and /ʂ/ (“shr”) is greater for women than for men.
- The women seem to be fronting /ç/ (“she”).
- This results in a higher centroid value and mimics the effect of having a smaller vocal tract.
  - Without fem. accent:
  - With fem. accent:
Speech sounds and socio-indexical categories:
Acquisition of gender-marking in Mandarin (Li & Kong, 2008)

Boy:                          Girl w/o F.A.                 Girl w/F.A.

Development of Gender-marked Phonetic Variant in Mandarin
4. There’s more to phonological development than phonemes: Conclusion

- Mandarin-speaking children can correctly produce both /ç/ ("she") and /ʃ/ ("shr") by about 3 years of age.

- However, gender-marking of /ç/ ("she")-male and /ç/ ("she")-female isn’t seen until 4 or 5 years of age.
Phonological acquisition is complex

1. Children gradually learn sounds and sound sequences in words of their language.

2. Phonological learning is highly language-dependent.

3. Transcription must be supplemented with other methods.

4. Children continue learning after they can produce speech sounds correctly.
Levels of knowledge about speech sounds

categorical phonological knowledge

[\texttt{t}] onsets

\textit{stressed}[\texttt{i}] \textit{back}[\texttt{k}] onsets

\textit{stressed}[\texttt{u}]\textit{ trochee}

words

\{\texttt{key}\} \{\texttt{tuna}\} \{\texttt{cougar}\}

language-specific phonetic detail

man's <\texttt{key}> man's <\texttt{tuna}> man's <\texttt{cougar}>

socio-indexical knowledge

Buffalo

gay

NYC

male adult

child female

people

Ben Marie Aunt Jan

girl's <\texttt{cougar}> voices

\{\texttt{key}\} \{\texttt{tuna}\} \{\texttt{cougar}\}
Clinical implications:
Vocabulary size and phonological acquisition

• Children with phonological disorders typically have slightly smaller vocabularies than their typically developing peers.

• What is the direction of this relationship?
  – Do children with smaller vocabularies have difficulty learning sounds because they have a smaller set of words to generalize over?
  – Do children with difficulty learning sounds have smaller vocabularies because they have difficulty parsing and remembering the sounds in words they hear?

• Clinical implications: Need to consider the words a child knows as well as the sounds he/she knows.
Clinical implications:
Socio-indexical knowledge and language disorders

• Pragmatic disorders characterized by difficulties understanding social cues.
  – Autism, Asperger Syndrome, Specific Language Impairment
• Many social cues are signaled by sociophonetic features.
  – Formal vs. informal speech
  – Literal vs. figurative language
  – Use of a particular dialect
  – Sexual orientation
• Perhaps one aspect of the pragmatic disorder is difficulty in perceiving socio-phonetic cues?
Acknowledgments

• Lab colleagues: Tim Arbisi-Kelm, Hyunju Chung, Eden Kaiser, Sarah Schellinger, Asimina Syrika, Kari Urberg-Carlson

• Help with local arrangements by Catherine McBride-Chang, Katerina Nicolaidis, Areti Okalidou, Kiyoko Yoneyama

• Support from NIDCD Grant 02932 and NSF Grant BCS-0729140 to Jan Edwards

• Participation of the children and cooperation from their parents

     For all of which, a heartfelt:

         謝謝     thank you     ευχαριστώ πολύ     ありがとう
Clinical implications: Phonological knowledge and reading

• What about children who are learning non-standard dialects of English with different phonological systems?
  – For example, African-American English (AAE).

• The phonological system of the spoken language has a tremendous impact on decoding and spelling.
  – Example from AAE:
    • “Ms. Four”

• We know very little about interactions between phonological knowledge and learning to read in non-standard dialects.
Clinical implications: Phonological knowledge of children with cochlear implants

• Children with cochlear implants have much better speech production skills relative to children with hearing aids.
• However, their speech intelligibility is reduced relative to peers with normal hearing.
• Sibilant fricative production of children with cochlear implants (Todd et al., 2010; Todd et al., in preparation).
  – Centroid frequencies for /s/ are lower relative to normal hearing peers, even for correct productions.
  – The relationship between consonant accuracy and CV frequency is weaker for children with CI’s relative to either chronological-age or vocabulary-age peers.
Clinical implications

• With a more complex understanding of phonological knowledge,
  – Potentially, we can have a much finer-grained understanding of how to assess and treat phonological/language breakdowns
  – We can assess and treat breakdowns at different levels of representation.