Using psycholinguistic tasks to predict phonetic accuracy in children with and without speech sound disorder

Benjamin Munson
University of Minnesota, Minneapolis, USA

Supported by NIDCD Grant R03 DC005702, McKnight Foundation
Collaborators: Adriane L. Baylis, Mary E. Beckman, Jan Edwards, Miriam Krause, Dongsun Yim
Progressive Mastery

• Children don’t sound like adults. Sounding like an adult takes a long time
  • “Little Benny is talking” (…at five months? Yes, sort of)
  • Babbling
  • Reduplicated Babbling and Language-Specific Vocalizations
  • First words
  • Pronunciation changes
    – From [wa di] to [rakit sip] for Rocket Ship
    – Acoustic detail changes (from no contrast to inaudible contrast to audible contrast)
Progressive Mastery

• Later Development
  – Gender differentiation
  – Developing sociolinguistic/sociophonetic competence
  – Decreases in trial-to-trial variability
  – Learning about phonotactic constraints, morphophonological alternations
  – Developing metaphonological knowledge
The (Perceived) Big Leaps

• (1) The first word
  – In the U.S., it is this level of analysis that is used to categorize a child’s early language development for clinical and educational services

• (2) Going from making lots of errors in adult forms to making no errors
  – In the U.S., it is this level of analysis that is used to categorize a child’s phonological development for clinical and educational purposes.
    • Kids who continue to say [wə dɪ] Rocket Ship beyond the age when their same-peers do not catch the eyes (and ears) of the people who determine who gets speech therapy and who doesn’t

• Ergo, figuring out why some kids make errors while others do not is an important task.
Specific Errors

• Most studies of phonological development focus on characterizing the trajectory of learning particular contrasts or patterns
  – Properly ranking faithfulness and markedness
• But...the “unmarked” isn’t as transparent as Jakobson would have us believe
Specific Errors

- Edwards, Beckman, and the παιδολογος research collective have found that the apparent error patterns cross-linguistically are an interaction between fine differences in phonetic detail, frequency in the ambient language, and cross-linguistic differences in adults’ perception of children’s errors

  - Beckman & Edwards (2010); Edwards & Beckman (2008a, 2008b); Li, Munson, Edwards, Yoneyama, & Hall (2011)
From Phenotypes to Endophenotypes

• Phenotypes: What the person does that leads us to identify her/him as having a disorder

• Endophenotypes: What causes the person to do whatever it is that causes us to identify him/her as having a disorder

• Maybe it’s time to step back and look at endophenotypes rather than phenotypes: What are the skills that underlie successful phonological acquisition?
Figure 1. Schematized model of the relationships among different levels of phonological knowledge and the lexicon.
One strategy for studying this

a) By measuring children’s phonetic accuracy
b) By measuring skills that we think support accurate speech production
c) By examining how children with an identified phonological disorder differ from children
d) By examining which measures in (b) predict (a)
The project

• Does not have a catchy name like παιδολογος.

• *Acquiring Access and Representation in Phonology (AARP)*
  – Funded by the US National Institutes of Health.
    • PI: Munson; RAs: Adriane Baylis, Miriam Krause, Dongsun Yim

• Contrasted the performance of children with speech sound disorder to that of children with typical speech development

• Age-by-accuracy differences allowed us to examine predictors of phonetic accuracy continuously using regression and related techniques
Measures

• A number of measures were given
  – Two guiding principles in selecting measures
    • Contrasting existing measures with novel experimental measures
    • Contrasting measures of representation-building capacity with measures of on-line access of these representations
      – “Objects” versus “Operations”
      “Representations” versus “Access”
‘Existing’ Measures

• (1) Speech-sound discrimination
  – Minimal-pair identification
• (2) Maximum syllable repetition rate
• (3) Vocabulary Size
  – Receptive
  – Expressive
Two-Stage Models of Production

• Production involves many stages ‘from intention to articulation’ (Levelt, 1989)

• We focused on two stages in single word production: a semantic-conceptual stage, and a phonological stage
Lexical Access

• *Delayed Naming*

• In the delayed naming paradigm, children view a picture and either name it immediately, or wait a specified interval before they must name it.

• Used previously by Lahey and Edwards (1996) to examine lexical access in children with SLI

• A variant of the delayed reading task (Balota & Chumbly, 1985; Munson, 2004)
Lexical Access

• In the immediate-response condition, response characteristics (particularly latency) reflect the combined influence of lexical access and the cognitive processes that occur after lexical access.

• As the delay interval increases, responses reflect progressively less of the influence of lexical access, and progressively more of the influence of post-access processes.
Lexical Access

• Latency of response should decrease as delay interval increases, and accuracy should increase.
• This reflects the benefit associated with allowing time for lexical access to take place.
• If the lower phonetic accuracy in children with SSD is associated with a deficit in lexical access, then we should see the largest group differences between children with better and poorer speech-production accuracy at the shorter delay intervals.
Lexical Access

• Children named 15 pictures each at 3 delay intervals: 1000 ms, 500 ms, and 0 ms
• Pictures were chosen because they had a high likelihood of being named uniformly, and because their names did not include late-acquired phonemes (i.e., the had relatively low motoric demands)
• Accuracy, as well as latencies for correct responses, were tallied
Phonological Encoding

• The second experiment examined phonological encoding (i.e., phonologically ‘instantiating’ the successfully accessed conceptual representation).
• Used a cross-modal picture-word interference paradigm
Phonological Encoding

• A technique developed to study the timing of lexical access and phonological encoding in speech production

• Children see a picture and hear a word. The spoken word (the Interfering Word [IW]) varies in:
  – Its phonological similarity to the picture
  – Its timing relative to presentation of the picture
Phonological Encoding

Extremely Facilitatory: 
*cat* is named quickly
Phonological Encoding

"make"

Extremely Inhibitory:
foot is named slowly
Phonological Encoding

Naming RT for *bed* is intermediate between the other conditions

“red”
“book”
Phonological Encoding

- Responses are facilitated relative to a neutral condition if the word is identical, and inhibited if the word is not identical.
  - The degree of interference depends on the phonological similarity of the word to the picture name
  - Phonological facilitation is greatest when the interfering word occurs concurrent with or after the picture
Phonological Encoding

- The influence of IW similarity on naming latencies can tell us about the efficiency of phonological encoding during speech production.
- If the difference between the phonologically unrelated and phonologically related IWs is relatively small, then we can infer that phonological encoding isn’t efficient.
Phonological Encoding

• Nine pictures were named in the presence of five types of IWs. There were three stimulus onset asynchronies: -150 ms, 0 ms, and +150 ms
• Picture names were trained in advance of the experimental task
• Naming latencies were measured for correctly named pictures
Access versus Representation

- These were our measures of ‘access.’ We needed a measure of ‘representation.’
- Representations in their static form are unmeasurable. They are latent variables.
- Instead, we measured representation-building capacity.
Representation Building

• Fisher et al. (2001): Children listened passively to a string of nonwords presented without a referent

• After a distracter task, they are played a list that includes some of the same nonwords (‘primed’) and some new nonwords (‘unprimed’).
Representation Building

• In Fisher et al. (2001), children repeated primed nonwords more accurately than unprimed ones.

• This was true even when the primed sequences of phonemes were put in a different phonetic frame.

• Children are able to build acoustic-perceptual representations for novel words based on minimal exposures.
Representation Building

• Our priming task was similar to Fisher et al.’s,

• Three types of nonwords were presented in the test phase
  – 1. Unprimed
  – 2. Primed, same talker
  – 3. Primed, new talker

• Stimuli included only early-acquired sounds and simple syllable shapes
Schematic

Study Phase  Distractor Phase  Test Phase  Response

/jʌg/  ?  /fɛɪɡ/  /peɪk/

/jʌɡ/  /jʌɡ/
Representation Building

• Typically developing children should produce primed nonwords more accurately and with shorter latencies than unprimed nonwords, reflecting their ability to create phonological representations based on brief exposures.

• We predict that children with SSD will show a reduced influence of priming, reflecting their decreased ability to learn new perceptual representations.
What did we find?

First, experiment by experiment
Second, predicting phonetic variation
‘Existing’ Measures

• There were small but significant group differences in the measure of perceptual ability.

• There were no group differences on the measure of maximum syllable repetition rate.

• There were small but significant group differences in PPVT-III scores, and a trend toward differences on the EVT.
Lexical Access

Accuracy: TD

Accuracy: SSD
Children with PD were not disproportionally slower than children with TD in the 0 ms delay condition.
Lexical Access

If anything, the children with SSD have smaller differences across conditions than do the children with TD.
Lexical Access

• When LMER (Baayen, Bates, & Davidson, 2008) was used to examine the performance of individual subjects, the effect of condition correlated neither with measures of speech-production accuracy nor with measures of vocabulary size.
Lexical Access

• RTs at each delay interval were strongly correlated with phonetic accuracy for the children with the lowest phonetic accuracy
Lexical Access

• The preliminary conclusion: phonological disorder is not associated with a deficit in lexical access. However, children with SSD do name pictures more slowly than their peers.
Phonological Encoding (TD)

Picture-naming RTs in the presence of onset- and rime-related IWs were faster at the +150 ms SOA than those for unrelated IWs for kids with TD...
Phonological Encoding (TD)

Picture-naming RTs in the presence of onset- and rime-related IWs were faster at the +150 ms SOA than those for unrelated IWs.
Phonological Encoding (SSD)

...And for kids with SSD. SSD does not appear to be associated with reduced phonological encoding abilities.
Phonological Encoding

• Preliminary conclusion: SSD is not associated with a deficit in phonologically encoding lexical items that have been selected from long-term memory
Children with SSD show a reduced influence of priming on the accuracy and latency with which they repeat stimuli in the test phase.
Results I

Typically Developing

SSD
Results: Minimal Pair Identification

- Minimal-Pair Identification mediated the effect of priming. While there was a significant difference between primed and unprimed words for all of the participants, this difference was largest for the children with good speech-sound discrimination.
Results: Word-Specific Production Accuracy

• The magnitude of priming for individual nonwords varied as a function of how accurately their component phones were produced in real words
  – there was more priming for words that contained phonemes in the child’s production repertoire on the GFTA than those not in the repertoire
  – Again, robust when the other significant effects were controlled for
A Mélange of Results

- Supports the notion that perceptual problems are central in SSD
- Provides no clear support for the hypothesis that access processes are implicated in SSD.
What predicts what?

Regressions
What predicts what?

- Dependent measure: Rationalized arcsine transformed percent phonemes correct for all sounds on the GFTA-2
- Age always forced as the first measure
- Regression 1: All of the ‘existing’ measures entered stepwise before the experimental measures
  - Age accounted for 20% of the variance
  - Expressive Vocabulary Size: 23%
  - Minimal-Pair Identification Accuracy: 9%
  - Total: 52%
What predicts what?

- Regression 2: All of the ‘existing’ and experimental measures entered stepwise in a single block
  - Age accounted for 20% of the variance
  - Naming Speed: 30%
  - Receptive Vocabulary: 23%
  - Identity Priming: 6%
  - Total: 72%
Endophenotypes and Phonetic Accuracy

- What is associated with accurate speech production?
- Knowing lots of words
- Naming them quickly
- Hearing them accurately and representing them in detail
A fluke or authentic?

• Was this just a weird fluke that relates more to our subject populations and to our unrestrained use of stepwise regression than to anything else?

• We have access to data from a separate cohort of subjects on related measures.
A fluke or authentic?

• Stoeckel (2007 UMN Dissertation)
• Age, PPVT, and minimal-pair nonword discrimination predict 16%, 13.5%, and 19.5% of variance in phonetic accuracy, respectively
Endophenotypes, Redux

• Yup, accurate speech production really is associated with knowing lots of words and with being able to hear things correctly
• Importantly, these have independent influences on speech-production accuracy…
• …at least statistically
• Three possible path analyses
Phonetic Accuracy

- Vocabulary Size
- Perception
Phonetic Accuracy

Vocabulary Size

Perception
Whew…

- In the AARP data-set, vocabulary size and perception continue to predict a significant proportion of variance in speech-production accuracy even after their shared variance is accounted for.
Why are we surprised?

• Previous work by Edwards and colleagues (i.e., Munson, Edwards, & Beckman, 2005, following the findings of Edwards, Beckman, and Munson, 2004) suggested that SSD was associated with deficits in primary encoding abilities, not in deficits in higher-level phonological knowledge.

• Why would the association between vocabulary size and speech-production accuracy exist in these three corpora?
Learning lots of words leads to the development of robust sublexical units. These…

• Facilitate ongoing word learning
• Provide a lexicon-independent ‘fast map’ that quickly associates perceived acoustic events with the routines needed to reproduce them
Is this really the biggest of the leaps?

• We don’t really know, do we?

• Our overt perception of language and language ability more often than not doesn’t reflect what is truly important.
  – Expectation \(\Rightarrow\) responsivity, subsequent models (Julien & Munson, submitted)

• These might be culturally specific, even within languages
Other Leaps

• Nobody has examined the relative contribution of other factors (covert contrast, acoustic variability, sociophonetic variation) on children’s perceived communicative competence.

• We need to think beyond the traditional divides to best understand the outcomes that we care the most about: educational success, stable social relationships, and vocational success.