Psych229: Language Acquisition

Lecture 12
Word Forms

Werker & Yeung 2005

Word learning: mapping among concept, word, and word's variable acoustic signal

Perceptual system plays a significant role: perceptual units change throughout word learning (facilitate)

Important ability: "bootstrapping"

= using existing knowledge to facilitate acquisition

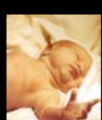
(use existing perceptual knowledge to learn words)

Werker & Yeung 2005

Bootstrapping:

- initial perceptual biases enable infants to initially extract early word forms (yields more precise knowledge of acoustic and phonetic properties of the native language)

- once linked with concepts, infants realize different sounds yield different word meanings and phonological representations is "bootstrapped" from existing perceptual system (phonological system online ~ 18-20 months)



Neonate (and fetus) perceptual biases:

prefer mother's voice, stories & songs in native language

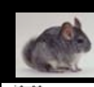
Werker & Yeung 2005

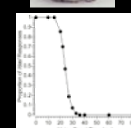
Perceptual biases shared with other animals:

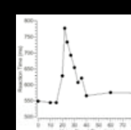
- discriminate native language rhythm only when played forward, not backward

- categorical discrimination of some contrasts

(ex: voice onset time "d" vs. "t")







Werker & Yeung 2005

Perceptual biases possibly shared with other animals:

- preference for speech over acoustically matched non-speech

- sensitivity to phonetic cues that indicate word boundaries

- (from cognitive neuroscience studies): unique cortical activation to forward speech vs. backward speech, phonetic vs. non-phonetic attributes

Werker & Yeung 2005

The effect of early exposure: links with later language proficiency

- vowel discrimination at 6 months predicts vocabulary size at 13-24 months

- reading proficiency correlated with phonetic discrimination as neonate

- word-object association ability delayed if ear infections/initially deaf

- bilingual evidence: don't have true bilingual phonetic discrimination if exposed to sound system at 3-4 years of age

Werker & Yeung 2005

Time course of speech perception: functional reorganization

For example, infants begin life discriminating both native and non-native phonetic contrasts, but by 6–12 months of age show a decline in discrimination of many non-native distinctions and an enhancement of sensitivity to native ones [13,14]. The timing and extent of this reorganization is influenced by several factors, including the acoustic/articulatory characteristics of the phonetic contrast [15] and the similarity of the contrast to those used in the native language [16]. Within the same time period, infants learn many other phonological properties of the native language; as reviewed by Jusczyk [17], infants by 9–10 months prefer well-formed words that correspond to frequent patterns in the input.

…due to statistical learning

Statistical learning: a mechanism for reorganization? One mechanism that underlies functional reorganization might be statistical – as infants are exposed to language-specific input, emergent properties of the input may shape the perceptual system. By at least 9 months, infants are sensitive to the frequency, distribution, and other statistical properties of perceptual input in speech [13]. Highly frequent phonetic contrasts and phonotactic patterns (i.e. permissible combinations of sounds) are categorized in a language-specific manner at younger ages than less frequent ones [13,21]. After repeated exposure to lists of nonsense words with recurring sound patterns, infants can make generalizations about syllable structure [22], stress [23] and phonotactic patterns [24].

Life's easier when the distribution is bimodal, though

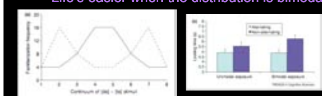


Figure 1. (a) English-learning infants were exposed to an 8-step continuum of stimuli modeled on a phonetic difference that is not used contrastively in English (voiced and unaspirated voiceless alveolar obstruents [da] and [ta]). For 2.3 min, these stimuli were presented in either a Bimodal (dotted line) or a Unimodal (solid line) frequency distribution [25]. (b) Following exposure, infants were presented with alternating (both stimuli 1 and 8) and non-alternating (either stimulus 3 or 6) test trials. Infants in the Bimodal condition discriminated between non-alternating and alternating trials, whereas infants in the Unimodal condition listened equally to both. Data from [25].

Werker & Yeung 2005

Word segmentation

- first segmentation at 7-8 months
- recognize familiar words, words with trochaic bias (if English)
- 10 months: recognize iambic words (guiTAR), using phonetic & phonotactic cues

Strategies:

- language-specific: trochaic/iambic bias, phonotactic cues, familiar words
- language-independent: transitional probabilities

Question: which cues take precedence when both are available?

Johnson & Jusczyk (2001): 8-month olds prefer prosodic information over statistical information

Thiessen & Saffran (2003): 6-month olds prefer statistics to prosody

Werker & Yeung 2005

Word forms

- 9-10 months: encode phonetic detail ("tup" vs. "cup") and indexical detail (speaker identity & emotional affect); don't generalize if indexical details are changed (no recognition)

Task	Conclusions	Stimuli examples	Phonetic transcription[a]	Age (months) 7–9	10–12	13–15	17–24
Word forms							
Discrimination of novel word forms	Phonetic sensitivity at 6 months.	bih / dih [47]	[bɪ] / [dɪ]	✓			
	At 9-months, indexical/ stress information is weighed as heavily as phonetic information, but by 10–12 months, phonetic information is weighed more heavily	cup / tup [29]	[illegible]	✓			
		cup / cup[*] [37]	[illegible]	✓	×		
		cup₁ / cup₂ [*] [36]	[illegible]	✓	×		
		diNNER / Dinner [*] [38]	[illegible]		×		

✓ = stimuli pairs treated as distinct word forms. × = not treated as distinct word forms in task.

Werker & Yeung 2005

Word forms

- 10-11 months: recognition of word form even if indexical details are altered…but still treat mispronounced words as real words when differences are "perceptually confusable"

Task	Conclusions	Stimuli examples	Phonetic transcription[a]	Age (months) 7–9	10–12	13–15	17–24
Recognizing familiar word forms		Didder / Dinner [*] [38]	[illegible]		×		
	At 10–12 months, use of phonetic information still depends on perceptual salience	bonJOUR / ponJOUR [*] [39]	[illegible]		×		
		bonJOUR / bonGOUR [*] [39]	[illegible]		?[†]		
	Task is easier when critical phonetic detail is in stressed, word-initial, and/or syllable-initial positions	paart / paarp [*] [40]	[illegible]		?[†]		
		paart / daart [*] [40]	[illegible]		✓		
		Dinner / Ninner [*] [38]	[illegible]		✓		

✓ = stimuli pairs treated as distinct word forms. × = not treated as distinct word forms in task.

Werker & Yeung 2005

Word form-object pairings

- 14 months: can learn novel word-object pairings, but phonetic detail is fuzzy till about 17 months
- younger infants can learn (8-12 months), but only if synchrony between word presentation and movement of object; also require perceptual/social cues like visual salience and eye gaze

Task	Conclusions		Stimuli examples	Phonetic transcription[a]	Age (months) 7–9	10–12	13–15	17–24
Word-object pairings								
Learning novel word-object pairings[b]	Learning of novel word-object pairings by 14 months, but phonetic detail not accessed until 17 months		lif / neem [44,45]	[illegible]	×	×	✓	
			pin / din [46]	[illegible]			×	
			bih / dih [47]	[bɪ] / [dɪ]			×	✓
	Phonetic detail accessed in easier tasks	Audiovisual synchrony	tah / gah [42]	[illegible]	✓			
		Word forms familiar	ball / doll [50]	[illegible]			✓	
		Preferential Looking Task	tuk / duk [49]	[illegible]			✓	
Recognizing familiar word-object pairings[c]	From 14–24 months, phonetic detail accessed in word recognition		ball / doll [49,72]	[illegible]			✓	✓
			baby / vaby [51,71]	[illegible]			✓	✓

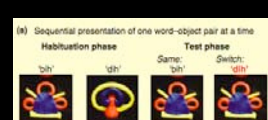
✓ = stimuli pairs treated as distinct word forms. × = not treated as distinct word forms in task.

Werker & Yeung 2005

Word form-object pairings

The "Switch" Task

(presentation of word 7-10 times while object moves)



Preferential Looking Task

(word recognition)

Swingley & Aslin 2002: Word Representation

Word form representation:

Children must encode sound forms…but different instances may have different acoustic properties (speaker voice, rate of speaker, word context)

Word recognition: attend to *linguistically relevant* distinctions in the language (Dietrich, Swingley, & Werker 2007: English *tam* = *taam*; Dutch *tam* ≠ *taam*)

By 12 months, infants seem to know which sounds are linguistically relevant - this is right at the beginning of the word learning phase

…but several studies show that children of this age don't seem to encode a lot of phonetic detail

(novel word *dak* = *gak* for 2 year olds)

Swingley & Aslin 2002: Word Representation

Why?

Maybe the task was too demanding for young children. (Learning a new name for an object, and then being asked to give that object to the experimenter.)

Question: Are children's representations of words adult-like in their level of phonetic detail, or not?

One idea: Encode detail only if necessary

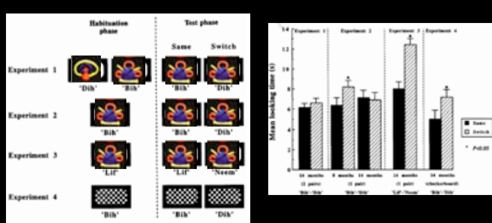
If children have small vocabularies, it may not take so much detail to distinguish one word from another. (*baby, cookie, mommy, daddy*…)

Neighborhood structure idea: When a child knows two words that are acoustically similar, more attention to phonetic detail is required to distinguish them.

Swingley & Aslin 2002: Word Representation

Going with the neighborhood idea, look at Stager & Werker (1997)

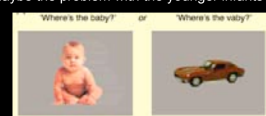
"bih" and "dih" are too close, and kids don't know any words close enough to motivate attention to the "b"/"d" difference when word-learning



Follow-up study (Werker et al. 2002): 17-month olds and 20-month olds do notice the "bih"/"dih" switch

Swingley & Aslin 2002: Word Representation

Maybe the problem with the younger infants was that these were *novel* words



18-23-month olds did better on this task

Maybe younger kids will, too…

Procedure

The procedure was identical to that used in our previous study (Swingley & Aslin, 2000), except that the displayed pictures slowly moved up and down in synchrony, a manipulation intended to help maintain infants' interest in the display. Infants were seated on their parent's lap facing an 80-cm video monitor. Before the experiment began, the infants' attention was attracted to the monitor by a pair of short animated movies. The experiment proper consisted of 28 trials: 24 test trials and 4 fillers. Each test trial began with the presentation of two horizontally aligned pictures of objects. After 3 s, a prerecorded speech stimulus naming one of the pictured objects was played from a central loudspeaker. Sentences were of the form *Where's the [target]?* followed by *Can you find it, Do you like it,* or *Do you see it?* Filler trials with four pictures of other familiar objects were interspersed with test trials to help maintain the infants' interest.

Swingley & Aslin 2002: Word Representation

Table 1. *Correctly pronounced (CP) target words and their mispronounced (MP) versions*

CP	MP-close	MP-distant
apple (/æpl/)	opple (/ɑpl/)	opal (/opl/)
baby (/be'bi/)	vaby (/ve'bi/)	raby (/ɹe'bi/)
ball (/bɔl/)	gall (/gɔl/)	shawl (/ʃɔl/)
car (/kaɹ/)	cur (/kɝ/)	kier (/kiɹ/)
dog (/dɔg/)	tog (/tɔg/)	mog (/mɔg/)
kitty (/kɪti/)	pity (/pɪti/)	yitty (/jɪti/)

Note. International Phonetic Alphabet transcriptions are provided in parentheses. Bisyllabic words were stressed on the first syllable.

What kids heard

Prediction if neighborhood idea is right: close mispronunciations should be harder to distinguish than distant mispronunciations

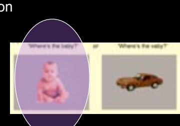
Subjects: 50 14-15-month old infants

24: correct vs. distant mispronunciation

26: correct vs. close mispronunciation

Tracking time spent fixated on target picture

(367-2000ms after word pronounced)



Swingley & Aslin 2002: Word Representation

Results

Correctly pronounced words easier to recognize than all mispronounced words (so task is reasonable & infants notice the difference in pronunciation) [p <.001]

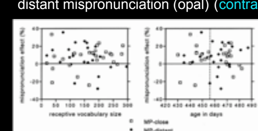
…but both were significantly different from chance (50%)

Correct word fixation on target: ~60%

Mispronounced word fixation on target: ~54%

So infants can recognize mispronounced words, but they have a harder time

But there was no effect for whether it was a close mispronunciation (opple) or a distant mispronunciation (opal) (contrary to prediction of neighborhood)



No relationship between mispronunciation effect and age or receptive vocabulary size (possible neighbors)

Swingley & Aslin 2002: Word Representation

Results

Conclusion: Since ability to notice mispronunciations wasn't dependent on age or vocabulary size, infants (14 month olds) must initially store words with a lot of phonetic detail - it doesn't matter how many neighbors there are

Idea: Any mispronunciation may be noticeable (so not necessarily a difference between close mispronunciations & distant mispronunciations)

Good thing: If infants initially store words with phonetic detail, they don't need minimal pairs to force them into noticing more phonetic detail (minimal pair = *ball* vs. *doll*, which require some semantic knowledge to know they're different)

Swingley & Aslin 2002: Word Representation

Questions

But what about studies with older kids (17-month olds) that had trouble distinguishing mispronunciations of novel words?

Key: *novel* words

Idea: Maybe phonetic detail involves hearing the word a number of times - get a little more phonetic detail each time

{p/b/d/g}{a/o/u}{l/r}

…

(p/b}{a}{l/r}

…

{b}{a}{l}

If it's a novel word, kids haven't heard it enough yet. (Stager & Werker, 1997 = only 7 times of repetition}

Swingley & Aslin 2002: Word Representation

Another idea

Maybe Stager & Werker (1997) kids had the correct pronunciation "bih" activated by the mispronunciation "dih" (remember that kids here did recognize the mispronunciations - they were just harder to get)…and this results from the word not having been heard enough times

This account predicts that "knowledge of neighbors should actually *inhibit* word learning", since word learning requires the recognition of a separate word form for a separate concept in the world.

"Look, it's a *ball!*"

"Look, it's a *doll!*"

Swingley 2005: Word Representation (Yes, really, they have phonetic detail)

11-month olds & familiar words

Using familiar words so there's no issue of generalization of novel items, and training/testing conditions that overestimate children's abilities

Previous work: Hallé & de Boysson-Bardies 1994 (less phonetic detail)

- infants prefer familiar words even if initial consonant mispronounced
  (*bonjour ~ ponjour*)
- infants prefer familiar words even if onset of second syllable mispronounced
  (*bonjour ~ bongour*)
- only if initial onset missing did infants not like mispronounced words
  (*bonjour ~ onjour*)

Swingley 2005: Word Representation (Yes, really, they have phonetic detail)

Previous work: Vihman et al. 2004 (more phonetic detail), 11 month olds

- infants prefer correct pronunciations over unfamiliar words
  dirty more than *budget*
- no preference for mispronunciations over unfamiliar words
  nirty or *dirny* not more than *budget*

Reconciliation:

Hallé & de Boysson-Bardies 1994: French children don't notice change if to onset of initial cluster

Vihman et al. 2004: English children don't notice change if to onset of medial cluster

Combined: children have trouble noticing difference if change is to beginning of *unstressed* syllable

(ponJOUR vs. bonJOUR, bonGOUR vs. bonJOUR for French)

Swingley 2005: Word Representation (Yes, really, they have phonetic detail)

The Plan: Dutch 11-month olds

Use least salient phonetic contrast in language - if infants can track this, they can detect more salient changes, too.

Expt 1: Correct pronunciations of familiar words vs. correct pronunciations of unfamiliar or nonce words (baseline)

Expt 2a: mispronounced familiar words vs. unfamiliar words

Expt 2b: familiar words vs. mispronounced familiar words

Expt 3a: mispronounced familiar words (offset) vs. unfamiliar words

Expt 3b: familiar words vs. mispronounced familiar words (offset)

Swingley 2005: Word Representation (Yes, really, they have phonetic detail)

Experimental Stimuli

Table 1 *International Phonetic Alphabet transcription of the Dutch test stimuli used in each of the experiments, with an English gloss. Columns labeled change give a featural description of the difference between the correct pronunciation and the mispronunciation*

Familiar	English	Nonword	Onset-MP	Change	Offset-MP	Change
[illegible]	leg	[illegible]	[illegible]	place	[illegible]	place
[illegible]	bear	[illegible]	[illegible]	place	[illegible]	place, voice
[illegible]	tummy	[illegible]	[illegible]	place, voice	[illegible]	place
[illegible]	duck	[illegible]	[illegible]	[no change]	[illegible]	place
[illegible]	hand	[illegible]	[illegible]	place	[illegible]	place
[illegible]	hair	[illegible]	[illegible]	place	[illegible]	place, manner
[illegible]	dog	[illegible]	[illegible]	place	[illegible]	place
[illegible]	cow	[illegible]	[illegible]	manner	[illegible]	[no change]
[illegible]	mouth	[illegible]	[illegible]	place	[illegible]	place
[illegible]	nose	[illegible]	[illegible]	place	[illegible]	place
[illegible]	horse	[illegible]	[illegible]	place, voice	[illegible]	place
[illegible]	cat	[illegible]	[illegible]	place	[illegible]	place
[illegible]	sheep	[illegible]	[illegible]	place	[illegible]	place, manner
[illegible]	toe	[illegible]	[illegible]	place	[illegible]	place
[illegible]	fish	[illegible]	[illegible]	manner	[illegible]	place
[illegible]	foot	[illegible]	[illegible]	place, manner	[illegible]	place

Table 2 *Acoustic measurements for stimuli used in all experiments. MP is an abbreviation for 'mispronunciation'. Standard deviations are shown in parentheses. Durations are given in milliseconds, and frequency measures in Hertz*

Condition	Expt.	Duration	Mean F_0	Min F_0	Max F_0	Range F_0
Familiar	1, 2b, 3b	681 (143)	227 (38)	144 (39)	322 (55)	178 (53)
Nonword	1, 2a, 3a	644 (112)	222 (31)	159 (35)	282 (34)	146 (54)
Onset MP	2a, 2b	705 (124)	221 (37)	139 (36)	313 (50)	174 (65)
Offset MP	3a, 3b	684 (118)	220 (26)	142 (41)	304 (42)	162 (61)

Swingley 2005: Word Representation (Yes, really, they have phonetic detail)

Procedure

-16 test trials per infant

-24 words per sound file, 1 second pause in between words

- Experimenter gauged infant's looking time

- Infant fixates on flashing green light, which is expt start

- Then, red light flashes. When infant looks, experimenter starts playing sound file. Experimenter keeps playing sound file until infants looks away for 2 consecutive seconds. (Does not play when infant not looking, though, even if less than 2 seconds.)

Swingley 2005: Word Representation (Yes, really, they have phonetic detail)

Results 1: Familiar vs. Non-Word

familiar: average 9.20 seconds

non-word: average 7.39 seconds

Infants prefer familiar over non-words

Prediction 2a: Mispronounced familiar vs. Non-word

If have phonetic detail, should not have preference

If don't have phonetic detail, treat mispronounced as familiar and have preference.

Prediction 2b: Mispronounced familiar vs. familiar

If have phonetic detail, should have preference for familiar

If don't have phonetic detail, treat as same and have no preference

Swingley 2005: Word Representation (Yes, really, they have phonetic detail)

Results 2a: Mispronounced Familiar vs. Non-Word

mis-pronounced familiar: average 9.01 seconds

non-word: average 9.65 seconds

No preference = phonetic detail

Results 2b: Mispronounced Familiar vs. Familiar

mis-pronounced familiar: average 8.94 seconds

familiar: average 10.07 seconds

Preference for familiar = phonetic detail

Experiments 3a & 3b: But what if mispronunciations are in the offset?

√ onset: *ben ~ den*   offset: *ben ~ bem*

Swingley 2005: Word Representation (Yes, really, they have phonetic detail)

Results 3a: Mispronounced Familiar (offset) vs. Non-Word

mis-pronounced familiar (offset): average 8.85 seconds

non-word: average 8.99 seconds

No preference = phonetic detail

Results 3b: Mispronounced Familiar (offset) vs. Familiar

mis-pronounced familiar (offset): average 9.07 seconds

familiar: average 9.02 seconds

No preference for familiar = no phonetic detail???

What's going on?

Swingley 2005: Word Representation (Yes, really, they have phonetic detail)

Results 3b: Why infants don't prefer "bem" over "ben"

One idea: Maybe they don't notice the offsets period. (less phonetic detail)

…but then why didn't they prefer the mispronunciations over the non-words (expt 3a)?

Another idea: Infant recognition process is over-eager

They start activating "ben" as soon as they hear "be", and can't de-activate it in time to notice the difference between "ben" and "bem". Support for this from ERP studies.

…but then why didn't infants treat mispronunciations as familiar words in expt 3a?

Swingley 2005: Word Representation (Yes, really, they have phonetic detail)

Results 3b: Why infants don't prefer "bem" over "ben"

Offsets forms are more fragile in the infant mind.

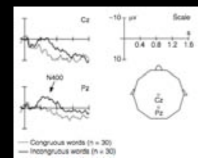
- less informative, less salient, etc.

Prediction: Different trajectories for onset vs. offset?

Question: How does this differ from having less phonetic detail for offsets?

Werker & Yeung 2005

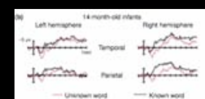
Neurological Data: Brain Activity at 14 months



High negative deflection at 400ms after presentation for incongruous vs. known congruous words (similar to N400 effect in adults)

Effects for known vs. unknown words at 200-400 ms after presentation (word recognition)

- 14 months: even if words are mispronounced ("tup")
- 20 months: only if words are correctly pronounced ("cup")

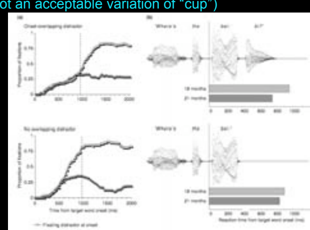


Shift from bilateral activation at 13 months to left-hemisphere dominant at 20 months

Werker & Yeung 2005

What triggers the ability to recognize the important phonetic detail at 17 months? ("tup" is a mispronunciation and not an acceptable variation of "cup")

Suggestion (PRIMIR): critical threshold of word-object mappings has been reached. Infants recognize which phonetic differences signal meaning differences (contrastive phonological categories). Phonological categories are easy cues to differences in meaning - make child's job easier for subsequent recognition and acquisition of new word-object mappings.



Incremental processing at 24 months and 18-21 months (shift to correct picture even after partial information).