How the language we speak guides the way we listen

Anne Cutler

Listening to speech

- We start listening to speech and learning about it even before we are born
- Our native language trains us up to listen to speech in the most efficient way (for that particular language)
- So it can happen that in two different languages, exactly the same speech information occurs, but listeners use it differently

Some speech perception

- It is easier to discriminate between talkers if they are speaking your native language
- The knowledge that underpins this ability is already in place at 7 months of age

Testing early discrimination

1. HABITUATION
   - Repeating stimulus
2. TEST
   - Stimulus changes; is the change noticed? i.e., does the infant again look longer?
Discriminating between talkers at 7 months

Do infants look longer (i.e., do they notice) if a new talker is added? Input: 3 versus 4 voices, saying sentences such as *This rugby season promises to be a very exciting one*.

Discriminating a switch of talkers: what matters is not whether speech is understood, but whether its sound is familiar.

(Johnson, Westrek, Nazzi & Cutler, 2011)

Guided Listening 1: stress

- Listening to speech is efficient and also very quick
- We use speech information continuously as it arrives and consider possible interpretations in parallel
- The speech information is any part of a word’s pronunciation – vowels, consonants, but also tone (in a tone language), stress (in a stress language) – and also whole-utterance features (intonation)

**PERvert (noun) vs. perVERT (verb)**

- In stress languages two different and unrelated words can even have exactly the same vowels and consonants but differ in stress
- e.g. INSight inCITE
  RElay reLay
  FOREbear forBEAR
  TRUSTy trusTEE
- (Actually there are not very many such pairs in English or other stress languages)
stress

• What there is a lot of: two words that begin the same way, except that one has initial stress, the other doesn't
• e.g.: CARton carTOON
• CAMpus camPAIGN
• CASHew cashIER
• MUsic muSEum
• MOtive moTEL etc. etc.!!
• How quickly can we tell what word we're hearing? In the first syllable already?

stress

• Example 1 1. MOtive or moTEL?
• Example 2 2. DIStant or distINCT?
• Example 3 3. TYPHus or tyPHOON?
• Example 4 4. CAMpus or camPAIGN?
• Example 5 5. MUsic or muSEum?
• Example 6 6. CARton or carTOON?

Stress: Results from an experiment

Language-guided listening in stress identification

• Dutch listeners use stress cues to distinguish words
• English listeners don’t (even though cues are clearly present in pitch, loudness and duration):
  mus- from MUSic or muSEum?
• In Dutch, there are a lot of words with unstressed syllables with full vowels; but in English, most unstressed syllables have a weak vowel. So the payoff of using the pitch, loudness, duration cues in Dutch makes using them worth while; the payoff in English is so little that listeners don’t bother

English-speaking listeners are not very good at identifying stress levels of isolated English syllables. Dutch-speakers outdo them. (Cutler & Donselaar, 2001; Cooper, et al., 2002)
Guided Listening 2: speech sounds

• Although every language has its own set of speech sounds, many sounds are pronounced very similarly across languages.
• An example is [s] and [f] – see and fee, soot and foot etc. – many languages have such s/f contrasts, and the pronunciation is always pretty similar.
• This does not mean that listeners identify the sounds in exactly the same way across all languages!!!

s or f?: Results from an experiment

The task: detect a specified sound (/s/, or /f/)
The input: nonsense strings e.g. "dokubapi pekida tikufa"
The crucial factor: Every target s or f came from another recording, e.g. /f/ in tikufa with tiku_a from:
– another utterance of tikufa (vowel has congruent cues)
– or an utterance of tikusa (vowel has mismatched cues)

Background: This experiment was first done (with English-speakers) in 1958. It showed that responses to /f/ were much harder to make in the mismatch case, but responses to /s/ were totally unaffected by mismatch!

(Wagner, Ernestus & Cutler, 2006)

speech sounds

The English results from 1958 were replicated: MUCH bigger effect of mismatch for /f/ than for /s/. In other languages, though....

Polish sounds confusable with [s]

[s] kos – blackbird
[s] koś – mow!
[i] kosz - basket
[ɕ] kość - bone
[ɕ] koci - of the cat (as in koci grzbiet)
[f] koc - blanket

Information about consonants is always available in adjacent vowels, but is only used where it helps listeners to distinguish similar sounds.
In Spanish and English, to distinguish /f/ from the confusable [θ], in Polish, to distinguish [s] from several confusable sounds like it.
Phoneme repertoire size and words within words

- A language with fewer phonemes has longer words
- Longer words allow for more embedded words
- However, just having fewer phonemes should itself produce more spuriously embedded words - independently of word length
- Comparing a very typical vocabulary against a very unusual one:
  - Spanish: 20 consonants, 5 vowels
    LEXESP: 73,000 word forms
  - British English: 24 consonants, 20 vowels
    CELEX: 60,000 word forms

(Cutler, Norris & Sebastian-Galles, 2004)

Phoneme repertoire size and word length

(a) word types in the vocabulary - % words per length in phonemes

Mean word length in English is 6.94 phonemes; Spanish is significantly longer at 8.3 phonemes
Even though actual speech contains mostly short words, the difference was still there: English 3.54 phonemes; Spanish 4.62 phonemes.

\[(\text{b) word tokens in speech (estimated)})\]
- % words per length in phonemes

Even though actual speech contains mostly short words, the difference was still there: English 3.54 phonemes; Spanish 4.62 phonemes.

How the language we speak guides the way we listen

- Every aspect of the way we listen to speech is shaped and guided by our native language
- Speech sounds, prosodic structure, what words are like — they all affect how we listen
- This makes our listening maximally efficient...
- ... FOR THAT LANGUAGE!
- (and incidentally makes it harder to listen to second languages that we acquire later)

(Repertoire size: Results from an experiment)

Stress effects in word recognition

(Soto-Faraco et al., 2001; Donselaar et al. 2005; Cooper et al., 2002)