How do we investigate speech perception?

Speech perception in adverse listening conditions: Day 3

Overview today

11.15 – 12.30: Lecture

- Last bits of non-native speech perception (yesterday)

- The role of the research question

- 9 often used experimental methods highlighted
  - Task?
  - Measurements?
  - What issues are being investigated with the method?
  - Example

12.45 – 13.15: Lab tour + demos

Speech perception in a non-native language

Why is listening in a **non-native language** harder than in one's **native** language?

Differences in:

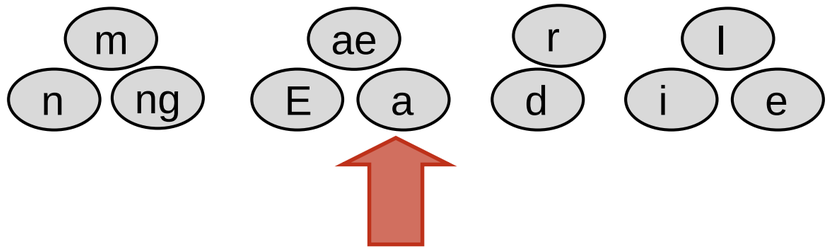
- Lexicon size
- Knowledge of idiomatic expressions
- Differences in sound systems

- Tuning of the perceptual system to the native language

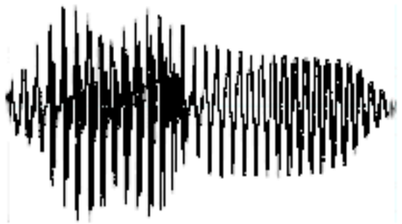
Native listening

marry: /m ae r i/

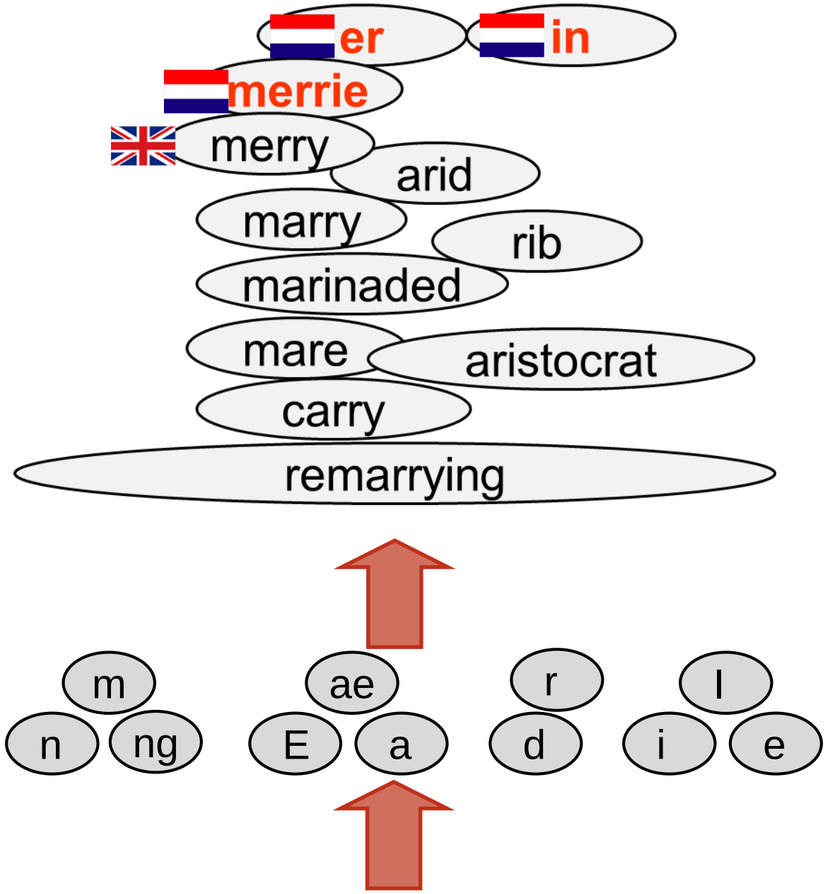
Native listening



marry: /m ae r i/

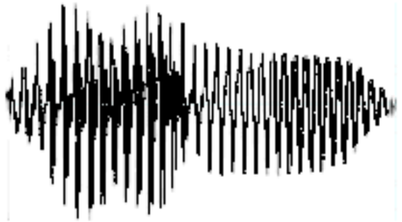


Nonnative listening



Incorrect sound perception during non-native listening leads to the spurious activation of words in the native **and** non-native language

marry: /m ae r i/



[Broersma & Cutler, 2008, 2011; Cutler et al., 2006; Spivey & Marian, 1999; Weber & Cutler, 2004]

Experiment: English word?

chair	word
vabe	non-word
drile	non-word
dask	near-word
church	word
plorn	non-word
thenk	near-word
sneeze	word
lemp	near-word
blonce	non-word
chass	near-word
grish	non-word
guide	word

Courtesy of Mirjam Boersma

Summary

Prelexical level

Incorrect sound perception during non-native listening leads to the spurious activation of words in the native **and** non-native language

Lexical level

Superiously activated words are hard to suppress

More competition during non-native listening than during native listening

Word recognition is slowed down and leads to more errors

The role of the research question

Human spoken-word recognition: What is investigated?

- Goal: uncovering the mechanisms and representations underlying human language

- Research into:
  - Processes involved in word recognition, e.g.,
    - Phoneme perception
    - Word recognition
    - Speech segmentation
  - Structure, organisation, meaning within the mental lexicon
  - Processes involved in comprehension
  - Language production
  - Language development
  - Second language learning
  - Cross-over into other fields, e.g., computer science

Research methods

- Listener experiment methods
  - Behavioural studies
  - Brain studies

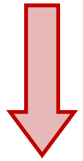
- Computational modelling
  - Computer implementation of a (partial) theory
  - Verification of whether a particular theory can really account for the behavioural data

The research question

= What you try to answer/understand using an experiment

- Usually and ideally theory-driven
- Should be as precise as possible: vague questions → vague interpretations

Theory: a unifying explanation of a set of facts or observations



Predictions and claims

Hypothesis: formally stated expectation

- Must be testable, falsifiable, precise, rational, and parsimonious

Important: a negative piece of evidence can disconfirm a hypothesis, but no piece of evidence can prove a hypothesis true

The research question determines which experimental method you should use

Sometimes multiple experimental methods are used to answer a research question

9 experimental methods

(Auditory) lexical decision

= Speeded classification of words and nonwords

Stimuli: list of (non)words in random order

- In isolation / preceded by priming stimuli

Task: Is stimulus a word or a non-word? Press button as fast as possible

- Issues addressed include:
- Effects of lexical variables on lexical access time
  - E.g., neighbourhood density, word frequency/length, stress
- Effects of semantic and form-based priming
- Effects of L1 on L2
- Effects of ambiguity in the word

- Measurements
- Reaction time (RT)
- Classification accuracy
- Magnitude of priming effect, relative to baseline

Experiment: earlier today

chair	word
vabe	non-word
drile	non-word
dask	near-word
church	word
plorn	non-word
thenk	near-word
sneeze	word
lemp	near-word
blonce	non-word
chass	near-word
grish	non-word
guide	word

Courtesy of Mirjam Boersma

Example

- Comprehension of non-native speech - activation of lexical competitors of non-native listeners
  [Broersma, *ICSLP*, 2002]

- Stimuli: non-words, near-words, words

- Participants: English (native) and Dutch listeners

- Results: Non-natives say ‘yes’ [it’s a word] far more often to near-words than native speakers do

Phonetic categorisation

= Identify (un)ambiguous speech sounds as one of two endpoints

Stimuli: range of speech sounds forming a continuum of ambiguous sounds between two unambiguous endpoints

Context: minimally a CV or VC syllable

Task: two-alternative forced-choice

Issues addressed include:

- (Categorical) perception of speech sounds
- Roles of various sources of information at a number of different stages of spoken-word recognition
- Trading relations (interplay between multiple cues to a phonetic contrast)

Measurements

- RT (slower for more ambiguous sounds)
- Accuracy
- Relative responses proportions (sound A vs. sound B)

Example

Investigate the flexibility of the perceptual system [Drozdova, van Hout, Scharenborg, Bilingualism, 2016]

Task: listen to a story with words containing manipulated sounds 

collect ? correct

Participants: native listeners

Results include (more on Friday!): listeners adapt their phone categories to include the ambiguous sound

Phoneme/Word identification in noise

= Identification of stimuli (partially) embedded in noise

Stimuli: (pseudo)words/syllables/sentences mixed into noise at different SNRs

- Different types of noise
- In isolation or in sentence contexts

Task: recognise/identify phoneme/word

bag *bell* *bat*

Issues addressed include:

- SNR needed for correct identification of phonemes/words in noise
- Modelling early (sublexical) processes in speech perception
- The role of word-initial and word-final information for correct word identification
- The role of hearing loss in speech recognition
- Differences in L1 and L2 speech recognition

Measurements

- Accuracy
- Confidence ratings
- Response confusions

Example

The Consonant Challenge [Cooke & Scharenborg, *Interspeech*, 2008]

Aim: identify effects of different masker types on consonant perception

Stimuli: VCV in seven conditions: clean+ 6 noise types

B Bee	CH CHart	D Dog	F Far	G Guard	H Heart
J Jar	K Key	L Leek	M Moon	N Neat	NG siNG
P Part	R Root	S Sue	SH SHoe	T Tea	TH THought
DH oTHer	V Vase	W Was	Y Yacht	Z Zoo	ZH meaSure

achu: competing talker

asi: modulated SSN

umi: 3-speaker babble

Results:

In clean: /dh, th, f, v, zh, dj/ hardest to recognise correctly

In noise: /dh, th, f, v, b, ng/ hardest to recognise correctly

More difficult

1. Quiet
2. Modulated speech shaped noise
3. Competing talker
4. 8-talker babble
5. Speech shaped noise
6. 3-talker babble
7. Factory noise

Priming

= Subjects listen to a spoken stimulus (prime)

At various times relative to prime presentation, an auditory or visual target - which may be related/associated or unrelated to the prime - is presented

Types of priming:
- Cross-modal semantic
- Form
- Morphological
- Syntactic

Stimuli: primes usually embedded in contexts

Task: combined with lexical decision task and/or recognition test and/or naming task

Examples

Type of priming	Target		Prime
Cross-modal semantic		<*pause*>	doctor
Form	blank	<*pause*>	plank
Morphological	walks	<*pause*>	sleeps
Syntactic	the girl	<*pause*>	the girls

Issues addressed include:

- Access to ambiguous words
- Access to morphologically complex words
- Spoken-word recognition
- Speech segmentation

Monitoring

= Detecting a target in a spoken stimulus

Target types:

- Phoneme (monitoring)
- Sequence or syllable (monitoring)
- Word (monitoring)

Stimuli: dependent on target types

- Embedded in a sentence or list of (non)words

Task: Press button as fast as possible when target is heard

Issues addressed include:

- Processing and representation of various types of linguistic information, e.g., semantics and prosody
- Contribution of the lexicon to word recognition (is there lexical feedback or not, see yesterday)

Measurements

Reaction time (RT)

- The speed with which the listener gives an answer
- How easy is a word to respond to in a listening experiment?
- In: ms
- Slower (=higher) RT: more difficult to respond
- Faster (=lower) RT: easier to respond

Classification accuracy

- How accurate was the listener?
- In: %(in)correct answers

Example

p

Example

Example

The effect of sentence accent on non-native speech perception in noise [Scharenborg et al., *Interspeech*, 2016]

Stimuli: Sentences with manipulated sentence prosody in different degrees of background noise

Participants: English (native), Dutch and Finnish non-native listeners

Results: Tomorrow!

Gating

= A spoken language stimulus is presented in segments of increasing duration

Stimuli: sound, syllable, word, diphone, phrase

- In isolation / with preceding and/or following context

Task:

- Identify the *entire* stimulus
- Give confidence rating

Issues addressed include

- Amount of acoustic-phonetic information needed to identify the stimulus
- Role played by phonetic, lexical, contextual variables during identification
- Nature of lexical representations

Measurements

- Identification point, correct identification with no change at later gates
- Confidence ratings
- Candidates proposed at each segment before isolation point

Example

The Dutch diphone perception experiment [Smits, Warner, McQueen, Cutler, *JASA*, 2003]

Aim: provide temporally detailed perceptual data concerning the unfolding of phonetic information over time

CV 
VC 

Results include:

- *Stops and fricatives:* voiceless consonants recognised better than voiced
- *Liquids and glides:* perceptually relevant information temporally more spread out than for other consonants

Eye-tracking

= Monitoring listener’s eye movements while they listen to speech

Stimuli: sentences containing critical (un)manipulated target words

Tasks:

- Do nothing
- Follow spoken instructions to touch or move (pictures of) objects on a computer screen

Issues addressed include *online* effects of

- Activation of candidate words
- Durational information on word recognition
- Speaking rate on the uptake of durational information
- (Other) perceptual cues on word recognition
- Semantic spreading

Measurements:

- Eye movements
- Percentage eye fixations

Example

Question: Is the native lexicon *routinely* activated when listening in a non-native language? [Hintz & Scharenborg, in preparation]

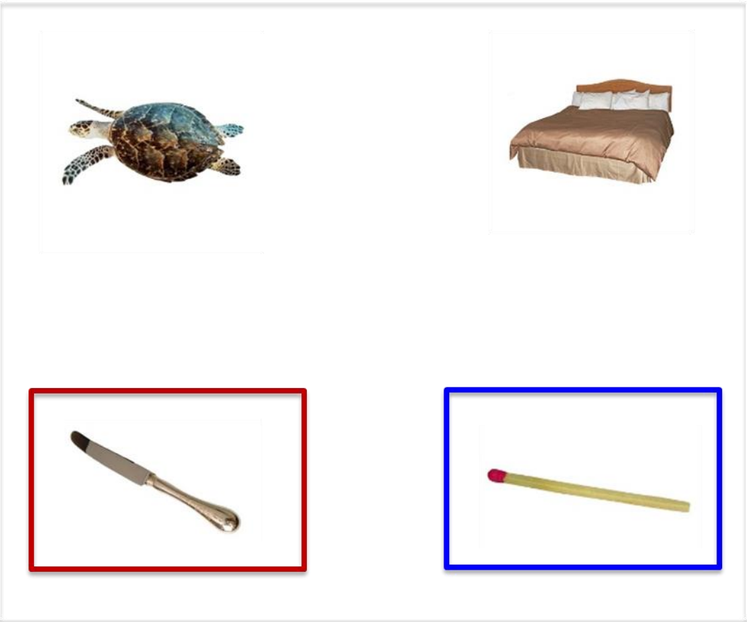
Task: Look and listen

Stimuli: sentences with English target words, pictures of

- Dutch phonological competitor
- English phonological competitor

Results: Thursday!

*"We left the house, but forgot the **map** with all directions."*



Brain studies

Aim: Understanding the neural basis of auditory processing and speech perception

Tasks: see behavioural studies

Event related potentials (ERP)

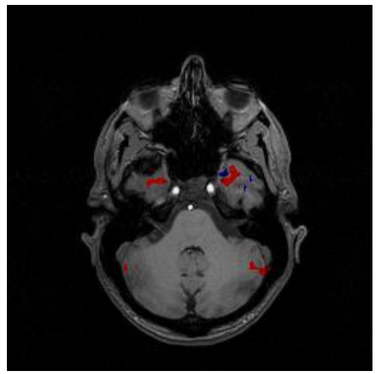
- Measures brain response that is directly the result of a thought or perception
- Can be reliably measured using ectroencephalography (EEG), which measures electrical activity of the brain

fMRI

= functional Magnetic Resonance Imaging

Form of neuroimaging: Measures neural activity in the brain or spinal cord

Berkeley's 4T fMRI scanner.



More information

Language and Cognitive Processes, 1996, 11 (6)

Take-home message

- The research question determines which experimental method you should use
- Different experimental methods can be combined

Next:
To the CLS Lab on the 12th floor!