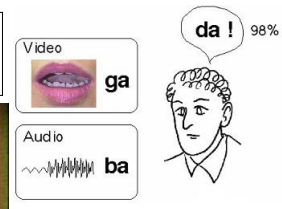
Life-span development of audiovisual speech perception: Examination by behavioral and ERP data

Kaoru Sekiyama
Cognitive Psychology Lab
Kumamoto University

Audiovisual speech perception

The McGurk effect
Perceptual fusion between incongruent visual and auditory speech



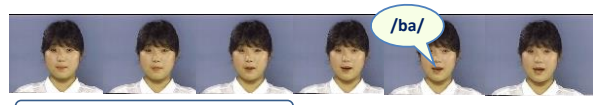
McGurk & MacDonald (1976)

Overview

The way how visual speech affects auditory speech perception is not uniform
(Kind of experience and developmental status matter)

- Interlanguage differences
- Developmental changes over life span
- Examinations of these groups:
  - Does visual speech affect the percept ? (the McGurk effect)
  - Does congruent visual speech speed up auditory processing? (RT & ERP)
  - Do perceivers focus on the mouth of the talker? (eye tracking)
- Language or culture?

Background

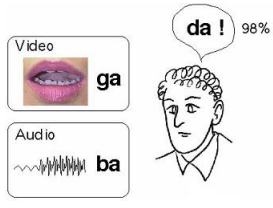


Visual lip-read information

- starts slightly before the audio onset
  - may predict when voice will start (Stekelenburg & Vroomen, 2007)
  - may predict what sound will come (van Wassenhove et al., 2006)
- generally improves perception of degraded speech
  - e.g., in noisy environment: *Sumby & Pollack, (1954)*
- also affects perception of undegraded speech
  - the McGurk effect: *McGurk & MacDonald, (1976)*

Is the McGurk effect universal?

The McGurk ~~effect~~
Perceptual fusion between discrepant visual and auditory speech



McGurk & MacDonald (1976, Nature)

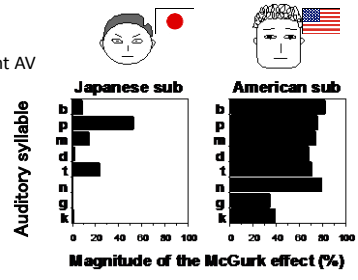
Sekiyama & Tohkura (1991, JASA)
Weaker visual influence in native speakers of Japanese

Interlanguage Differences

Weaker visual influence in Japanese participants when auditory speech is highly intelligible

(Sekiyama, 1994; Sekiyama & Tohkura, 1991, 1993)

For incongruent AV



Possible factors for the weak McGurk effect in the Japanese

- Linguistic factor

  Fewer phonemes in Japanese than English
  (e.g., 5 vs. 14 vowels)
  - Japanese may have less need for visual support

- Cultural factor

  Asian tend to look at eyes, Western at mouth
  (in emotion recognition: *Jack, 2009, Current Biology*)

Developmental cross-linguistic study

— McGurk effect & Reaction times —

(Sekiyama & Burnham, 2008, Developmental Science)

At what age do the interlanguage differences appear?

Background of developmental study

- Young children are less influenced by visual speech (*McGurk & MacDonald, 1976; Massaro et al, 1986; Hockley & Polka, 1994*)
- It is perhaps due to poorer lipreading in children (*Massaro et al, 1986*)
- Is the developmental tendency also the case for native speakers of Japanese?

Method

participants

	6 yos	8yos	11yos	Adults
Japanese	16	16	16	24
Aus. English	16	16	16	24

stimuli

V-only (V): ba, da, ga

A-only (A): Ba, Da, Ga

AV congruent (AV+): Bb, Dd, Gg

AV discrepant (AV-): Bg, Db, Gb

band noise (300-1.2kHz)

SNR:
-4 dB, +4 dB, +12 dB, High

task To report what they perceived, by pressing one of 3 buttons

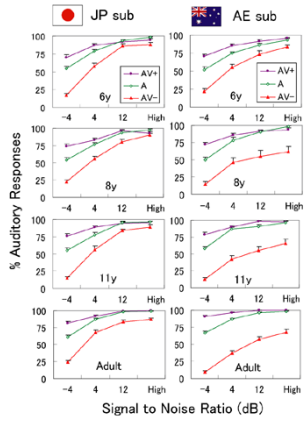
AV vs. A-only

% Correct

Avc (congruent)

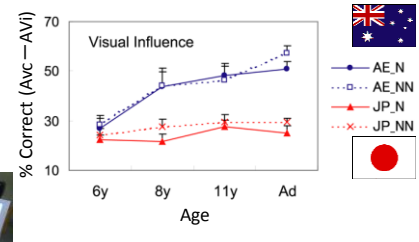
AO

Avi (incongruent)



Language differences emerge in development

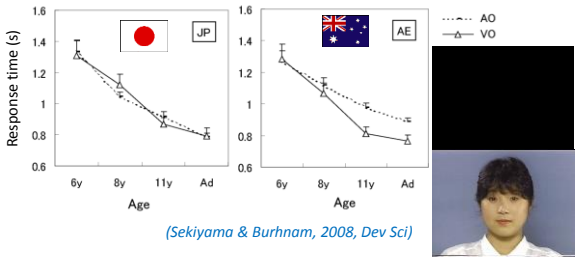
- No language differences at 6years of age
- Language differences appear between 6 and 8 years of age
- Developmental increase is specific to English



(Sekiyama & Burhnam, 2008, Dev Sci)

Speed of processing unimodal speech (AO vs. VO)

English natives are faster for VO speech (Visual precedence)



(Sekiyama & Burhnam, 2008, Dev Sci)

→ "Visual precedence" in ENG from around 8y

〜ENG adults are faster in lipreading than hearing

Interim Summary 1-1

The weaker visual influence in Japanese speakers compared with English speakers become evident over time after 6 years of age

- As a result of developmental increase of the visual influence only in English speakers
  - Encountering many talkers in school may promote the use of visual cues in English, perhaps from linguistic necessity due to many phonemes (?)

Interim Summary 1-2

Response times revealed that English speakers become faster in lipreading than in listening speech as they grow, whereas Japanese speakers did not develop such a visual precedence

- The stronger visual influence in English speakers may be associated with the visual precedence:
  - Is the McGurk effect like a visual priming effect?

Aging affects AV speech perception

(Sekiyama, Soshi, & Sakamoto, 2014, Frontiers in Psychology)

Does hearing decline in older adults affect AV speech perception?

Aging and AV speech perception

Hearing declines (e.g., thresholds increase) with age (CHABA, 1988; Pichora-Fuller & MacDonald, 2009)

- Do older adults use visual speech more than younger adults to compensate for deteriorated hearing?

- An aging-related increase of the McGurk effect has been found in experiments under the physically same stimuli across ages (*Thompson, 1995; Setti et al., 2013*)
- No aging-related increase of the use of visual speech has been reported when signal-to-noise ratios (SNRs) were calibrated to cancel hearing level differences (*Cienkowski and Carney, 2002; Sommers et al., 2005*)

Aging-related changes

- Not only hearing but also lipreading ability decline with age, especially after 70 years of age (*Shoop & Binnie, 1979*)
- However, the previous research has often included older adults over 70 years

Research question

- Is the McGurk effect stronger for older adults when age is limited under 65 years and hearing level differences are canceled out by calibrating SNRs ?

Method

participants Older group 60 - 65 years
Younger group 19 – 21 years

stimuli  Talkers: 3 native speakers of Japanese

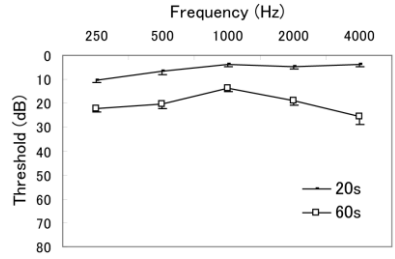
V-only (VO): ba, da, ga
A-only (AO): Ba, Da, Ga
AV compatible (AV+): Bb, Dd, Gg
AV incompatible (AV-): Bg, Db, Gb

band noise (300-1.2kHz)
SNR
Exp 1： 0, 6, 12 , 18 dB
Exp 2： 4 dB diffs btwn groups

task To report what they perceive from 3 alternatives
（also a 2 alternative task for VO）

Hearing thresholds in older and younger adults

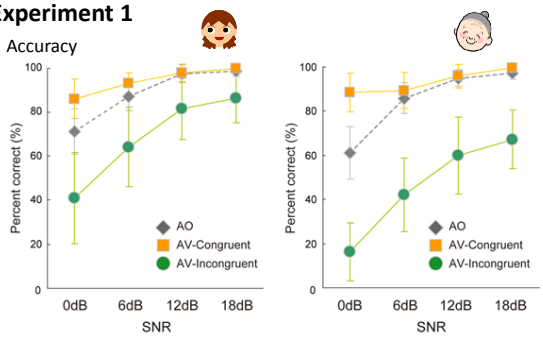
Our participants *



* Limited to normal hearing participants (average < 25 dB)

Experiment 1

Accuracy



Visual influence（AV c － AV i）was larger for the older
AO: The older were inferior to the younger
VO: No group differences

Interim Summary 3-1

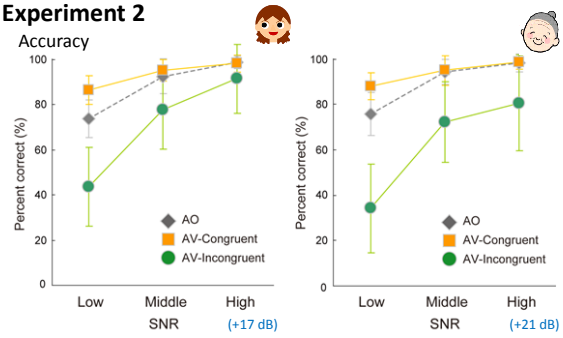
The increased McGurk effect was observed for older adults under the physically same stimuli (as previous research)

- Is it because the accuracy in the AO condition was poorer for older adults?

- Are the results same when SNRs are calibrated (by 4 dB) to cancel out hearing level differences?

Experiment 2

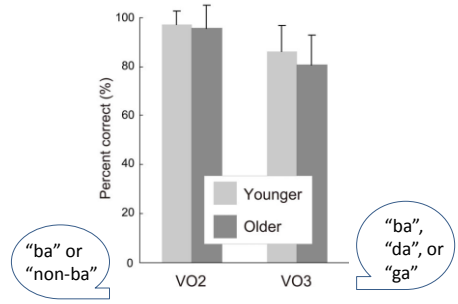
Accuracy



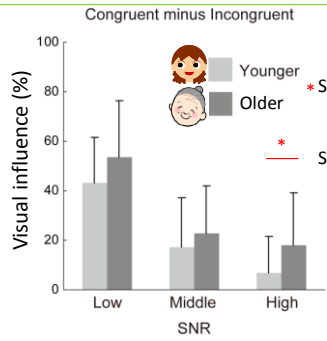
The visual influence was still larger for the older (esp in High SNR)
No group differences for AO accuracy (successful calibration)

Exp 2: VO (lipreading) accuracy

No significant differences were found between groups



Age differences in visual influence (AVc – AV i)



* Significant main effect of age (ANOVA)

Significant age diff in High SNR (Planned comparison)

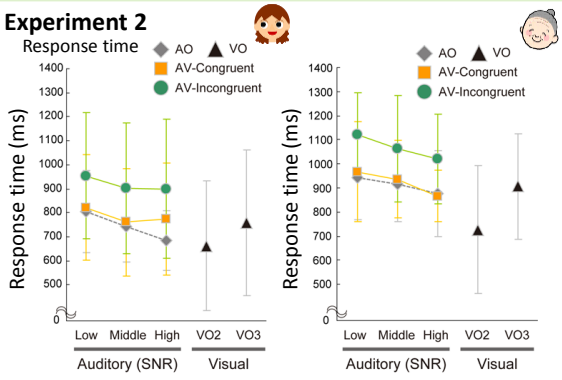
Interim Summary 3-2

The aging-related increase of the McGurk effect was still observed when hearing level differences were cancelled by calibrated SNRs (Novel finding)

- Older adults use visual speech more than younger adults

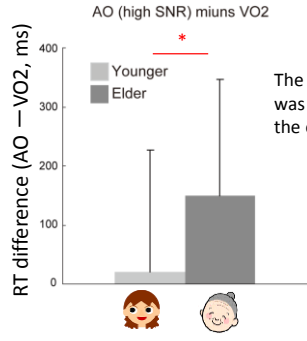
- Is it associated with a larger visual priming effect on the older?

Experiment 2



The older were slower than the younger for AO and AV.
No group differences for VO2.

Concerning the visual priming hypothesis



The visual priming measure was significantly larger for the older at the High SNR.

- consistent with the age difference in the McGurk effect at the High SNR.

Interim Summary 3-3 & Discussion

The increased McGurk effect for the older (with AO and VO accuracy equivalent to that of the young) may be related to their delayed auditory processing, which will cause a larger visual priming effect.

- Is it manifested in brain potentials?

General discussion

- - Linguistic pressure
  To distinguish many phonemes
  (English and Korean speakers)
  - Developmental status
  Delayed auditory processing (Japanese older adults)

- Both may foster attitude to rely on visual speech for dealing with difficult speech perception

- The group differences in the size of the visual influence and temporal measures (RT, ERP) suggest a "visual priming hypothesis" as a basis for the visual influence in speech perception

Conclusion

- The way how visual speech affects speech perception is modulated by kind of (linguistic) experience and developmental status