

Audiovisual benefit for stream segregation in elderly listeners

Esther Janse^{1,2} & Alexandra Jesse²

¹Utrecht institute of Linguistics OTS

²Max Planck Institute for Psycholinguistics Nijmegen

About us

Esther Janse

NWO VENI grant on speech perception and aging

Overall aim:

Establish relative contribution of age-related auditory and cognitive factors to problems in speech perception

Alexandra Jesse

NWO VENI grant on audiovisual speech processing

Complaint #1 for elderly listeners

Speech perception in noisy environments
Competing speech being most difficult
distracting sound for elderly



Advice

Try to get best view of target speaker's face

→ But does that actually help in this situation?

Audiovisual speech processing

Seeing a speaker helps infants (7.5mo) segregating speech streams (Hollich, Newman, Jusczyk, 2005)

Visual info and elderly:

- o Older adults are worse at lipreading than young adults but not at integrating visual with auditory information (Sommers, Tye-Murray, Spehar, 2005; Cienkowski & Carney, 2002)
- o Older adults' ability to lipread does not vary with their hearing sensitivity (Tye-Murray, Sommers, & Spehar, 2007)
- o Rather: Variation in lip-reading ability is explained by cognitive factors (processing speed and spatial working memory: Feld & Sommers, 2007)

AV benefit for stream segregation

Does seeing the face help elderly in stream segregation? If so, how?

Two possibilities

1. The *global synchrony* between speech and face movements helps listeners attending to the speaker.
2. *Local segmental cues* to segment identity help comprehension.

Aim

- What is relative contribution of background measures (auditory, lip-reading and cognitive skills) to elderly listeners' performance in competing speech conditions?
- What is relative contribution of these same measures to audiovisual (AV) benefit?

Method

Phoneme monitoring

Press the button once you perceive the target speaker saying the sound **p** (or **k**)

1. Familiarisation phase: see and hear target speaker speak
2. Phoneme monitoring.
Audio: mix of two female speakers: SNR +2
Crucial manipulation: whether or not one saw the target speaker's face

Per target phoneme:

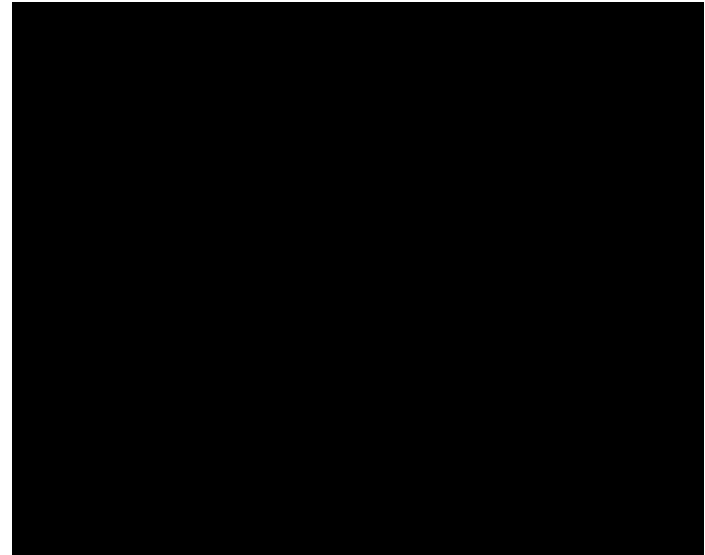
64 test sentences (with target phoneme)

64 filler sentences (without target phoneme)

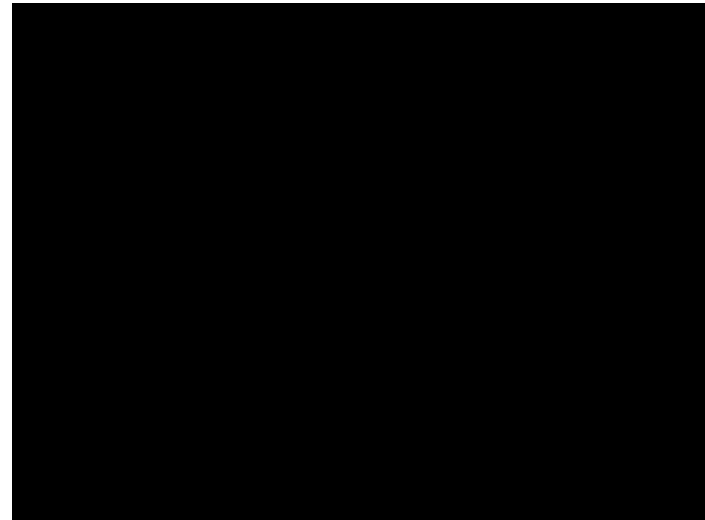
Method

Examples

/p/



/k/



Background information

40 older listeners (17 M, 23 F)

Mean (SD)

Age	72 yrs (5)
Hearing loss (better ear: loss over 1, 2, 4 kHz)	32 dB (12)
Education (1-5)	3.3 (1.1)
Information processing speed	1.1 sec (0.4)
Executive functioning (Trail making)	54 sec (33)
Selective attention (Stroop)	??

Lipreading task

Presented with silent videos of target speaker saying

{p}: peu / meu

{f}: feu / veu

{s}: seu / zeu

{k}: keu / geu

{t}: teu / neu

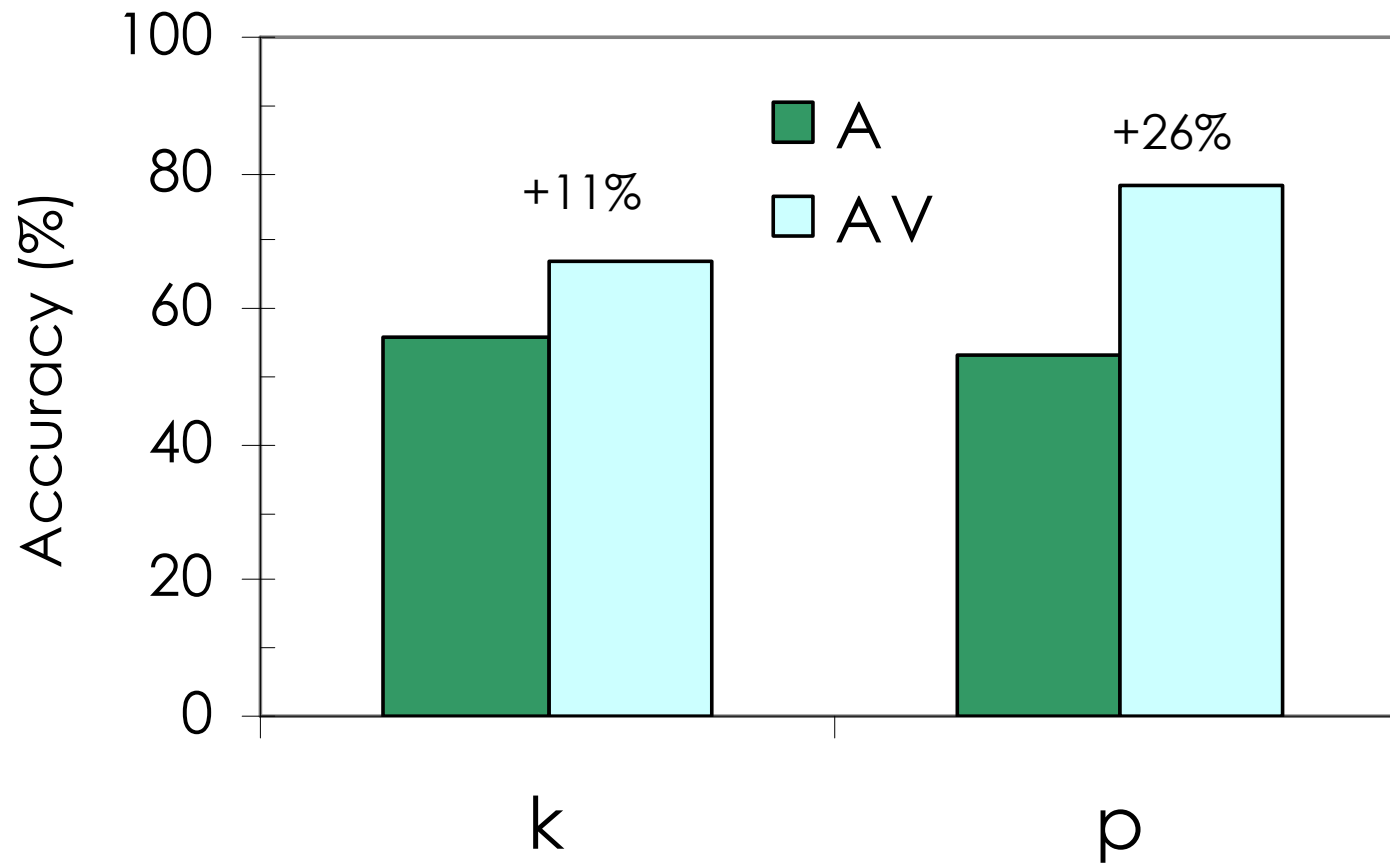
8 repetitions of each phoneme (or 8 x 2 per viseme)

Average identification accuracy (SD)

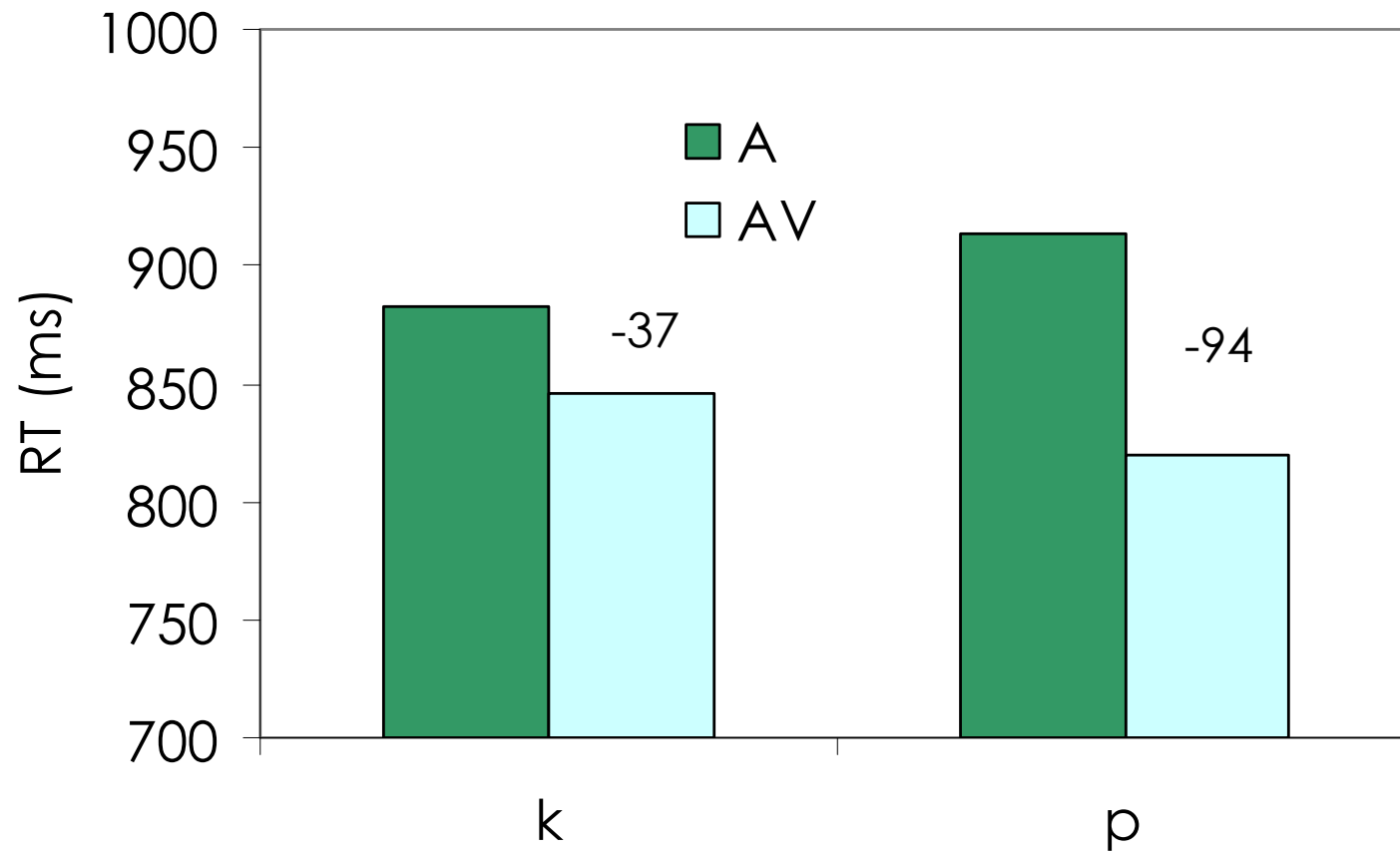
Phoneme (chance at 10%) 41% (7)

Viseme (chance at 20%) 73% (10)

Phoneme detection: Accuracy



Phoneme detection: RT



Conclusion Phoneme detection

Strong modality effect on performance (accuracy & RT)
And stronger effect for /p/ than for /k/

⇒ Both synchrony and local cues help elderly listeners to segregate speech streams

Individual differences

Background measures predicting overall performance

Accuracy

- Hearing loss (-)
- Age (-)
- Lipreading ability (+)

RT

- Hearing loss (+)

Background measures predicting AV benefit

Accuracy

- Hearing loss (+)

RT

- Lipreading ability (+)
- Information proc. speed
(the slower, the greater the benefit)

Conclusion

AV benefit for stream segregation both via

- ✓ overall synchrony between face and speech
- ✓ local segmental cues

Direct link between offline lip-reading ability and on-line use of visual cues

The AV benefit for stream segregation is predicted by:
the participant's hearing loss (+) and lip-reading ability (+) as well as by slower information processing speed