L2 proficiency modulates attention to the mouth during speech processing: An extended replication of Birulés et al. (2020)

Theres Grüter, Jieun Kim, Hitoshi Nishizawa, Jue Wang, Raed Alzahrani, Yu-Tzu Chang, Hoan Nguyen, Michaela Nuesser, Akari Ohba, Sachiko Roos & Mayuko Yusa
University of Hawai'i at Mānoa

Fall 2021, SLS 750: Seminar in Second Language Acquisition: Eye-tracking methods in language research
Collective class project

Visual speech cues in L1 and L2 listening

Under **adverse listening conditions** (e.g., noise, hearing impairment),

- visual speech cues improve speech recognition (Grant & Bernstein, 2019; Sumby & Pollack, 1954)
- listeners look at the talker's mouth more (adults: Vatikiotis-Bateson et al., 1998; Drijvers et al., 2019; children: Król, 2018)

*Do language **experience** and **proficiency** modulate attention to the mouth?*

- attention to the mouth changes over the first year(s) of life (Lewkowicz & Hansen-Tift, 2012; Morin-Lessard et al., 2019)
- toddlers with higher L1 vocabulary skills look at the mouth more (Król, 2018; Morin-Lessard et al., 2019)
- no evidence that bilingual children look more at the mouth when listening to their non-dominant language (Morin-Lessard et al., 2019)

➢ *Evidence from children (monolingual and bilingual) is mixed/unclear.*

- adults listening to an unfamiliar vs a native language look at the mouth more (Barenholtz et al., 2016)

➢ ***What about adults listening to a familiar L2?***

RQ1: Do L2 listeners look at a talker's mouth more than L1 listeners?
RQ2: Does L2 proficiency modulate attention to the talker's mouth?

Birulés, Bosch, Pons & Lewkowicz (2020, Exp2)

Participants

- 4 groups (*N*=19 each)

- L1 English
- L1 Catalan/Spanish L2 learners of English at three proficiency levels: low (A1-lowA2), intermediate (highA2-B1), high (B2-C2)

Materials

- Cambridge English Test (www.cambridgeenglish.org/test-your-english/general-english)
- Three 20-second videos; native English-speaking talker; 9 multiple-choice comprehension questions
  - ➢ Proportion of Total Looking Time (PTLT) scores
    - $PTLT_{Mouth}$ = Looks to Mouth / Looks to Face
    - $PTLT_{Eyes}$ = Looks to Eyes / Looks to Face

Results

- 2 (AOI: eyes, mouth) x 4 (Group) ANOVA on PTLT scores
  - ➢ less fixation on mouth in L1 vs all L2 groups
  - ➢ no differences between L2 groups
  - ➢ no correlation between PTLT difference scores ($PTLT_{Eyes}$ - $PTLT_{Mouth}$) and Cambridge Test scores (nor comprehension test scores)

re RQ1: YES
re RQ2: NO *unexpected*

*Do these findings **generalize***

- *to a different group of L2 listeners,*
- *when using different listening passages,*
- *and additional measures of proficiency?*

Participants

- recruited through participant pool + short-term English programs at UH
- minimum sample sizes determined through a-priori power analysis on Birulés et al.'s data

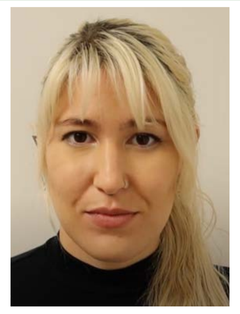
Table 1. Participant information (means and ranges)

	(self-identified) native English speakers L1 group		non-native English speakers L2 group	
N	38	of which 25 early monolinguals	45	of which 36 L1 Japanese
Age	23 (18-39)	23 (18-39)	29 (20-64)	21 (20-35)
Cambridge English Test (/25)	22.7 (15-25)	22.8 (15-25)	12.4 (5-25)	11.3 (5-18)
LexTALE (/100)	91.7 (77.5-98.8)	91.4 (77.5-98.8)	58.6 (42.5-81.3)	58.1 (42.5-70.0)
Self-rated Proficiency (/10)	9.2 (7-10)	9.3 (7-10)	4.8 (1-9)	4.4 (1-7)

- All participants included in analyses; analyses of more homogeneous subgroups only (L1 = early monolinguals, L2 = L1 Japanese) yielded the same pattern of results

Materials

Video 1



- Caucasian female
- identifies as native speaker of English
- nativeness rating (/10) *M* = 8.6 (*SD* = 2.1)

- ~1-minute monologue; 249 wpm (Shoe store employee speaking at staff meeting)
- 9 multiple-choice comprehension questions
- Materials adapted from listening comprehension test (Papageorgiou et al., 2012)

Video 2

more typical native speaker of English in Hawai'i

- Chinese-American male
- identifies as native speaker of English (age of onset: 6 yrs; self-id. profic.: 8/10)
- early multilingual
  - Cantonese (age of onset: 0; self-id. prof.: 9/10)
  - Mandarin (age of onset: 3; self-id. prof.: 7/10)
- nativeness rating (/10) *M* = 6.3 (*SD* = 2.7)

- ~1 minute monologue; 190 wpm (Student talking about daily life)
- 9 multiple-choice comprehension questions
- Materials created for this study

watch Video 1 *watch Video 2*

Procedure

- Language Background Questionnaire (online, before lab visit)

In lab (SMI RED250 eye-tracker, 60 Hz)

- Video 1 (~1 minute) + comprehension questions (*k*=9)
- Video 2 (~1 minute) + comprehension questions (*k*=9)
- Nativeness ratings of both talkers
- Cambridge English Test
- LexTALE (www.lextale.com)

Open Questions & Future Directions

- Why did the difference between L1 and L2 listeners replicate only in the video that was LESS similar to the original materials?
- Why did L2 proficiency measured through the Cambridge Test modulate looks to the mouth in this study but not in Birulés et al.?

➢ Under what (experimental and environmental) conditions do effects of nativeness and proficiency on attention to visual speech cues emerge?
Some suspects to explore: speech rate, passage difficulty, and properties of the talker, such as gaze direction and (assumed) nativeness

Results

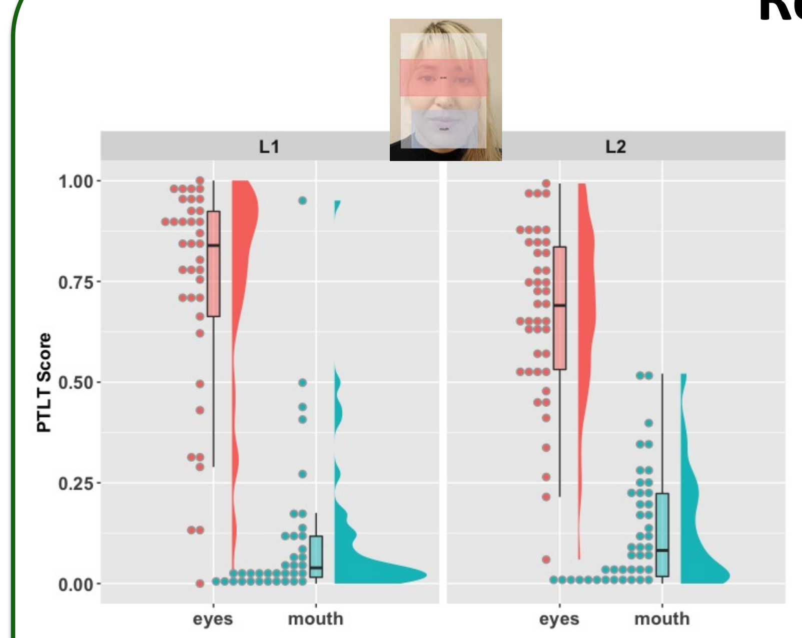


Fig.1. Video 1: Mean proportion Total Looking Time (PTLT) to the talker's eyes and mouth by group

2 (AOI) x 2 (Group) ANOVA

- Interaction [$F(1, 78) = 0.97$, $p = .33$, $\eta^2 = .01$] n.s.

Post-hoc independent-sample *t*-tests

- Looks to Eyes: $t = 1.10$, $p = .27$, $d = .25$
- Looks to Mouth: $t = 0.57$, $p = .57$, $d = .13$

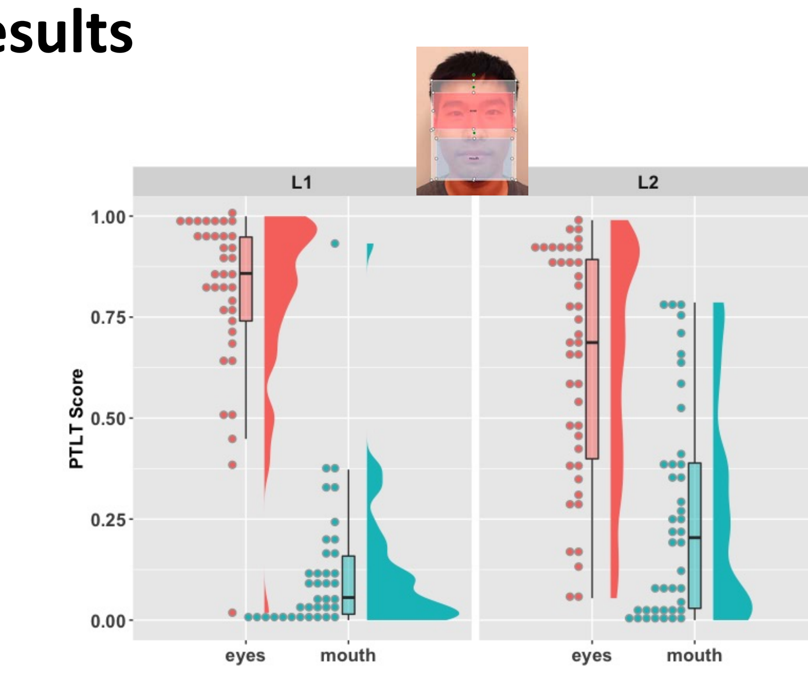


Fig.2. Video 2: Mean proportion Total Looking Time (PTLT) to the talker's eyes and mouth by group

2 (AOI) x 2 (Group) ANOVA

- **Interaction [$F(1, 77) = 8.64$, $p = .004$, $\eta^2 = .10$]**

Post-hoc independent-sample *t*-tests

- **Looks to Eyes: $t = 3.11$, $p = .003$, $d = .69$**
- **Looks to Mouth: $t = 2.82$, $p = .006$, $d = .62$**

➢ **re RQ1: L2 listeners were more likely to look at the mouth, but only in Video 2.**

Modulation by proficiency?

Table 2. Correlations between PTLT difference scores ($PTLT_{Eyes}$ - $PTLT_{Mouth}$) and proficiency measures in the L2 group (N=45; Spearman correlations)

	Cambridge Test	LexTALE	Self-rating	Mean proficiency z-score	Comprehension accuracy
Video 1 PTLT difference	**$\rho = 0.31$** **$p = 0.046$**	$\rho = 0.04$ $p = 0.78$	$\rho = 0.20$ $p = 0.207$	$\rho = 0.26$ $p = 0.091$	$\rho = 0.28$ $p = 0.071$
Video 2 PTLT difference	**$\rho = 0.34$** **$p = 0.025$**	$\rho = -0.01$ $p = 0.954$	$\rho = 0.30$ $p = 0.056$	$\rho = 0.24$ $p = 0.124$	$\rho = 0.29$ $p = 0.059$

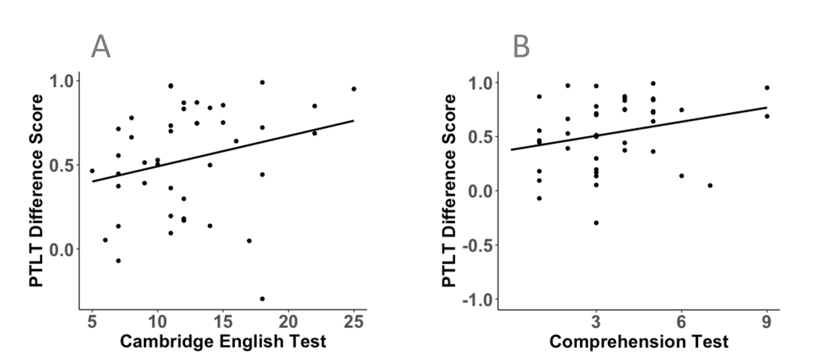


Fig.3. Video 1: Correlation between PTLT difference scores and scores on the Cambridge Test (panel A) and on the post-viewing comprehension test (panel B) in the L2 group (*N*=43).

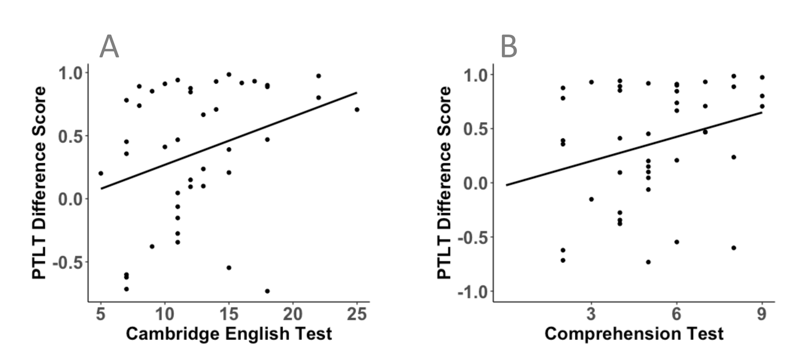


Fig.4. Video 2: Correlation between PTLT difference scores and scores on the Cambridge Test and on the post-viewing comprehension test in the L2 group (*N*=42).

➢ **re RQ2: In both videos, less proficient L2 users were more likely to look at the mouth, but only with proficiency measured through the Cambridge English Test.**

Comparing Original and Replication Studies

RQ1: Do L2 listeners look at a talker's mouth more than L1 listeners?

Birulés et al: YES   Video 1: NO   Video 2: YES

➢ *partial replication*

RQ2: Does L2 proficiency modulate attention to the talker's mouth?

Birulés et al: NO   Video 1: YES   Video 2: YES

➢ *null effect not replicated; observed effect Birulés et al. had predicted*

Barenholtz., Mavica & Lewkowicz (2016). Language familiarity modulates relative attention to the eyes and mouth of a talker. *Cognition*.
Birulés, Bosch, Pons & Lewkowicz (2020). Highly proficient L2 speakers still need to attend to a talker's mouth when processing L2 speech. *LCN.*
Drijvers, Vaitonytė & Özyürek (2019). Degree of language experience modulates visual attention to visible speech and iconic gestures during clear and degraded speech comprehension. *Cognitive Science*.
Grant & Bernstein (2019). Toward a model of auditory-visual speech intelligibility. In Lee et al. (Eds.), *Multisensory Processes.* Springer.
Król. (2018). Auditory noise increases the allocation of attention to the mouth, and the eyes pay the price: an eye-tracking study. *PLoS ONE.*
Lewkowicz & Hansen-Tift (2012). Infants deploy selective attention to the mouth of a talking face when learning speech. *PNAS*.
Morin-Lessard, Poulin-Dubois, Segalowitz & Byers-Heinlein (2019). Selective attention to the mouth of talking faces in monolinguals and bilinguals aged 5 months to 5 years. *Developmental Psychology*.
Papageorgiou, Stevens & Goodwin (2012). The relative difficulty of dialogic and monologic input in a second-language listening comprehension test. *Language Assessment Quarterly*.
Sumby & Pollack (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America.*
Vatikiotis-Bateson, Eigsti, Yano & Munhall (1998). Eye movement of perceivers during audiovisual speech perception. *Perception & Psychophysics.*

Acknowledgments: Many thanks to Joan Birulés for generously sharing materials and data, and to the two talkers for their patience and permission to use their images and video recordings