Pupillometry reveals reduced effort in children's sentence processing when visual speech cues are available

Rebecca Holt, Laurence Bruggeman & Katherine Demuth (Macquarie University)
rebecca.holt@mq.edu.au

Processing spoken language can be an effortful task, particularly for children and in adverse listening conditions, such as the classroom [1,2]. It is therefore important to understand how children's processing effort can be reduced in these circumstances, to minimise fatigue associated with sustained cognitive effort and to free mental resources for learning activities.

Visual speech cues (e.g., the speaker's facial movements) may assist in reducing the effort associated with children's sentence processing. The integration of auditory and visual cues can confer an "audiovisual (AV) benefit": an improvement in processing *speed* and *accuracy* for AV over auditory-only (AO) presentation. AV benefits are typically found for adults listening in noise [e.g., 3,4], but also occur in quiet [5] and among children [6]. There is mixed evidence for an AV benefit for processing *effort* among adults (i.e., reduced effort for processing AV speech relative to AO) [7-9]. AV benefits for effort have not been examined for children, though they may behave differently to adults as their linguistic development is ongoing. We therefore asked whether children would show an AV benefit for *effort* in either quiet or noise, hypothesising that they would [7]. Alternatively, children might show an AO benefit due to the extra effort of integrating across auditory and visual modalities [7, 8], or might show equivalent processing effort across modalities [9]. We also hypothesised that any AV or AO benefit for effort would be exaggerated in noise compared to quiet, due to the difficulty of speech perception in noise [1].

Thirty-five 7-11-year-old monolingual English-speaking children (M_{age} = 8;11) completed a phoneme monitoring task with concurrent pupillometry. Participants heard sentences in AV or AO modality (Table 1) and made a button-press when they heard pre-specified phonemes. Half the participants completed the task in quiet and half with continuous pink noise overlaid at a -2 dB signal-to-noise ratio (SNR). A Tobii X2-60 eye-tracker recorded participants' pupil dilation during the task: Increased pupil dilation functions as an index of increased cognitive effort [10]. Event-related pupil dilation was calculated per trial using a pre-task baseline period, cf. [11], and two metrics were extracted: peak dilation (maximum pupil dilation relative to baseline) and peak latency (the time at which the peak dilation occurred). We constructed a linear mixed-effects model for each metric with Modality (AV vs. AO) and Listening condition (quiet vs. noise) as fixed factors and a maximal random effect structure. Peak dilation was significantly greater in AO compared to AV modality (β = -0.76, SE = 0.26, p = .005), indicating greater cognitive effort in the AO condition (i.e., an AV benefit). Peak latency was also greater in AO compared to AV (β = -2.94, SE = 0.92, p = .001), showing participants were slower to reach their peak dilation in the AO modality (i.e., processing took longer when visual cues were not available; Figure 1). No significant main effect of Listening condition or significant interaction was found for either metric.

Thus, as hypothesised, 7-11-year-olds showed an AV benefit for effort when processing spoken sentences: When visual speech cues were available, processing was both less effortful and faster. There was no interaction with listening condition, contrasting with the hypothesis that any benefit would be exaggerated in noise. This may be due to the relatively favourable SNR used or extensive variability between participants. Our findings suggest that the effort children expend in sentence processing in classroom situations can be reduced, even in quiet conditions, by ensuring the speaker's face is visible. This in turn may result in improvements in performance on other tasks which rely on efficient sentence processing, such as important learning activities.

References

1. Klatte M, Lachmann T, Meis M. Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting. Noise Health. 2010 Oct 1;12(49):270.
2. Howard CS, Munro KJ, Plack CJ. Listening effort at signal-to-noise ratios that are typical of the school classroom. Int J Audiol. 2010 Dec 1;49(12):928-32.
3. Fort M, Spinelli E, Savariaux C, Kandel S. The word superiority effect in audiovisual speech perception. Speech Commun. 2010 Jun 1;52(6):525-32.
4. Jesse A, Janse E. Audiovisual benefit for recognition of speech presented with single-talker noise in older listeners. Lang Cognitive Proc. 2012 Sep 1;27(7-8):1167-91.
5. Davis C, Kim J. Audio-visual interactions with intact clearly audible speech. Q J Exp Psychol-A. 2004 Aug 1;57(6):1103-21.
6. Fort M, Spinelli E, Savariaux C, Kandel S. Audiovisual vowel monitoring and the word superiority effect in children. Int J Behav Dev. 2012 Nov;36(6):457-67.
7. Fraser S, Gagné JP, Alepins M, Dubois P. Evaluating the effort expended to understand speech in noise using a dual-task paradigm: The effects of providing visual speech cues. J Speech Lang Hear R. 2010 Feb 1;53(1):18-33.
8. Gosselin PA, Gagné JP. Older adults expend more listening effort than young adults recognizing audiovisual speech in noise. Int J Audiol. 2011 Nov 1;50(11):786-92.
9. Picou EM, Ricketts TA, Hornsby BW. Visual cues and listening effort: Individual variability. J Speech Lang Hear R. 2011 Oct 1;54(5):1416-30.
10. Schmidtke J. Pupillometry in linguistic research: An introduction and review for second language researchers. Stud Second Lang Acq. 2018 Sep;40(3):529-49.
11. Wagner AE, Toffanin P, Başkent D. The timing and effort of lexical access in natural and degraded speech. Front Psychol. 2016 Mar 30;7:398.

Table 1 – Sample sentences used in the phoneme monitoring task. Target phoneme shown in bold.

Target phoneme	Sample sentence
/b/	The ladies put their **b**asket in the car.
/p/	The artist left his **p**aint on the floor.
/g/	The auntie bought a **g**ift for the little boy.
/k/	The girl left the **c**abbage on her plate.

Figure 1 – Mean event-related pupil dilation across trials by modality. AO modality (blue) and AV modality (red). Curves smoothed using a generalised additive model.