Attentional Shifts and Cognitive Load in Simultaneous Interpreting: An Eye-Tracking Study

Kilian G. Seeber (kilian.seeber@unige.ch)[1], Dongpeng Pan (dongpeng.pan@unige.ch)[1]

[1]Faculty of Translation and Interpreting, University of Geneva, Switzerland

Background In spoken-language comprehension, listeners actively direct their attention to relevant visual representations even before they are explicitly mentioned [2]. Similarly, in spoken language production, speakers' eye movements attend to visual representations prior to articulation, reflecting conceptual and linguistic planning [5]. Simultaneous interpreting (SI) requires concurrent comprehension and production under severe time pressure [6]. During such tasks, cognitive load is often indexed by pupil dilation [3]. Whether the predictive and planning mechanisms persist under the dual demand of SI is unknown, as no prior study seems to have compared comprehension, production, and SI within a single paradigm. Our study combines the visual world paradigm with pupillometry to analyse the attentional-shift and cognitive-load patterns associated with these tasks.

Method Forty-two professional conference interpreters (M_{age} = 40.83; $M_{\mathrm{experience}}$ = 13.38 years), all with English as a working language and French (21), Spanish (8), Russian (5), German (4), or Italian (4) as L1, completed three tasks: comprehension (listening while viewing the display), production (describing a scene in English), and SI (simultaneously interpreting into their L1). Stimuli were auditory descriptions of spatial relations following the structure "The X (smallest object) is next to the Y (largest object)", presented in a blank-screen Visual World Paradigm [1] across 30 critical trials (10 per task). Fixation proportions and pupil dilation were recorded via an SR EyeLink Portable Duo. Generalised additive mixed models (GAMMs) with AR1 correction captured non-linear dynamics in fixation [4] and pupillometric data [8], and fixation models compared target and competitor gaze patterns within each task [7]. Analyses were time-locked to four windows aligned with the unfolding sentence: *Pre* (preview), *X*, *Next* ("is next to"), and *Y*.

Results During comprehension, gaze diverged progressively toward each named referent (Figure 1, left), with GAMM difference curves confirming significant TargetX–TargetY differences throughout the four windows (Figure 2, top). During SI, target fixation was attenuated (Figure 1, right), and divergence was absent except for the last window (Figure 2, bottom), suggesting a shift from anticipatory to verification processing. During production, significant target-directed fixation emerged early: on TargetX during the *Pre* window and on TargetY as early as the *X* window (Figure 3, top), reflecting the forward-looking nature of speech formulation. In SI, the early planning-driven pattern was absent in TargetX, although there was a mid-window divergence toward TargetY fixation during the *Next* window (Figure 3, bottom). Pupillometric analysis indicated that SI imposed significantly greater cognitive load than comprehension across all windows (Figure 4). These findings suggest that SI suppresses predictive attention allocation during comprehension and the planning-driven fixation during production, reflecting delayed verificational strategies at the cost of elevated cognitive load.

Conclusion Although SI represents an extreme case of language use, it provides a uniquely controlled window into how our language system manages the concurrent demands of comprehension and production. The selective suppression of anticipatory mechanisms under dual-task load reveals the resource-dependent nature of predictive processing.

References

[1] Gerry T. M. Altmann. Language-mediated eye movements in the absence of a visual world: the “blank screen paradigm”. *Cognition*, 93(2):B79–87, 2004.

[2] Gerry T. M. Altmann and Yuki Kamide. Incremental interpretation at verbs: restricting the domain of subsequent reference. *Cognition*, 73(3):247–264, 1999.

[3] Jackson Beatty. Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91(2):276–292, 1982.

[4] Sun-Joo Cho, Sarah Brown-Schmidt, Paul de Boeck, and Matthew Naveiras. Space-time modeling of intensive binary time series eye-tracking data using a generalized additive logistic regression model. *Psychological Methods*, 27(3): 307–346, 2022.

[5] Zenzi M. Griffin and Kathryn Bock. What the eyes say about speaking. *Psychological Science*, 11(4):274–279, 2000.

[6] Alexis G. Hervais-Adelman and Laura Babcock. The neurobiology of simultaneous interpreting: Where extreme language control and cognitive control intersect. *Bilingualism: Language and Cognition*, 23(4):740–751, 2020.

[7] Kate Stone, Sol Lago, and Daniel J. Schad. Divergence point analyses of visual world data: applications to bilingual research. *Bilingualism: Language and Cognition*, 24(5):833–841, 2021.

[8] Jacolien van Rij, Petra Hendriks, Hedderik van Rijn, R. H. Baayen, and Simon N. Wood. Analyzing the time course of pupillometric data. *Trends in Hearing*, 23:2331216519832483, 2019.

Figures

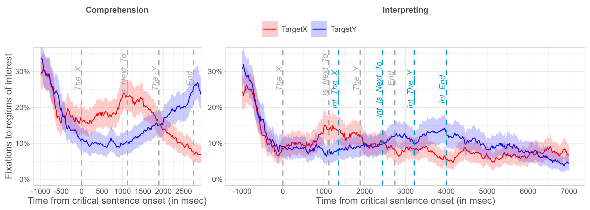

Fig. 1. Fixation proportions to TargetX (red) and TargetY (blue) during comprehension (left) and SI (right). Grey dashed: audio onsets; cyan: speech onsets in SI.

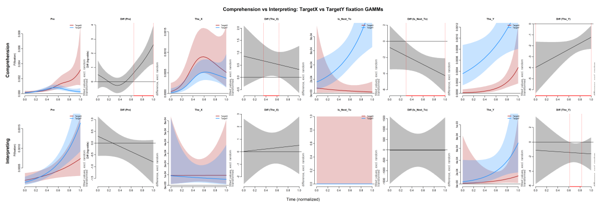

Fig. 2. Comprehension (top) vs. interpreting (bottom): GAMM-estimated fixation. Left: TargetX (red) vs. TargetY (blue) with 95% CI; right: log-odds difference (red = significant).

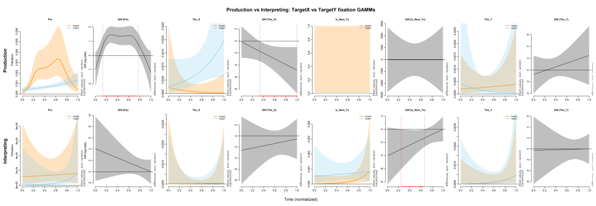

Fig. 3. Production (top) vs. interpreting (bottom): GAMM-estimated fixation. Left: TargetX (orange) vs. TargetY (blue) with 95% CI; right: log-odds difference (red = significant).

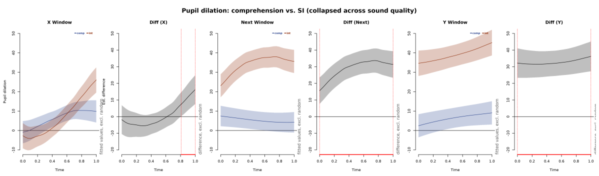

Fig. 4. GAMM pupil dilation: SI vs. comprehension for three windows (*X*, *Next*, *Y*). Left: smooth curves (blue = comprehension, red = SI) with 95% CI; right: difference (red = significant).