Language Teaching (2026), **59**, 128–133
doi:10.1017/S0261444825100864

CAMBRIDGE UNIVERSITY PRESS

RESEARCH IN PROGRESS

Attending to contextual cues to convey pragmatic meaning: An exploratory eye-tracking study

Naoko Taguchi[1] and Saisai Xu[2]

[1]English, Northern Arizona University, Flagstaff, USA and [2]College of Foreign Languages, Ocean University of China, Qingdao, China
Corresponding author: Naoko Taguchi; Email: naoko.taguchi@nau.edu

(Received 4 June 2025; revised 16 June 2024; accepted 22 June 2025)

This exploratory study was part of a three-year collaboration between the first author and the Pragmatics Interest Group at Ocean University of China (OUC). To advance research in L2 pragmatics, the first phase of collaboration explored the potential of digital games for pragmatics learning, which resulted in several publications (Taguchi, 2023). In the second phase of collaboration, we used the eye-tracking lab at OUC to explore L2 learners' attention to contextual cues when performing communicative functions. This paper presents preliminary findings from the project.

1. Background

Pragmatic competence is defined as L2 learners' ability to perform a variety of communicative acts in diverse social contexts (Taguchi & Roever, 2017). As this definition indicates, contextual sensitivity – or the ability to understand context and adapt linguistic demeanors according to context – is the primary dimension of pragmatic competence. Traditionally, L2 learners' contextual sensitivity has been examined by comparing their linguistics choices across different contexts. Learners are often presented with scenarios encoding varying contextual factors (e.g., speakers' roles and their relationships) and are asked to perform a communicative act assuming the role assigned in the scenarios. Researchers analyze learners' chosen linguistic forms (e.g., directness level of the forms, external and internal modification devices) and the appropriateness of the forms in a given context. They focus on whether learners are able to adjust their linguistic forms in response to different contextual factors. For example, when asking a high-imposition favor from someone with a higher social status and larger social distance, people typically use elaborate indirect forms with mitigation devices (e.g., "If you have time, it'd be great if you can take a quick look at my proposal."). In contrast, short, direct forms are usually appropriate when asking a small favor from someone in a close relationship (e.g., "Pass me the TV remote."). Learners' ability to vary linguistic forms according to different contextual cues has been considered an indication of their contextual sensitivity, and essentially their pragmatic competence.

This study intends to expand the scope of the concept of contextual sensitivity. Moving away from the common analysis of linguistic adjustments across contexts, we explore contextual sensitivity as the ability to attend to visual contextual cues and use them as resources to determine the course of action. Specifically, we present L2 learners with a video clip illustrating a scenario and ask them to perform a communicative act directed to a person in the video. To determine who to look at and direct their speech to, learners need to recognize the critical contextual cues associated with the target person

© The Author(s), 2025. Published by Cambridge University Press.

https://doi.org/10.1017/S0261444825100864 Published online by Cambridge University Press

(see the Method section). As such, learners' eye movement patterns and areas of attention – whether they can fixate their eyes on the target person signaled by key contextual cues – can be a reflection of their contextual sensitivity.

To be clear, in the field of cognitive psychology, it has been established that contextual cues play a critical role in people's cognitive and perceptual processes (Chun, 2000). A number of studies have investigated how visual contextual information guides people's visual attention and subsequently their behavior (Chun, 2000). Contextual information first determines what is salient for a scene currently in view. Then, attention and eye fixation patterns are adopted to the salient area, which in turn facilitates people's reaction to objects in that area. In this study, contextual cues, if noticed, are expected to have such facilitative effects, assisting participants' decisions on who to direct their speech to when performing a communicative act.

In this study, L2 learners' eye fixation patterns are analyzed in relation to their general proficiency. Existing studies have established the impact of proficiency on pragmatic competence (Xiao, 2015). Many studies have used a data-eliciting instrument (e.g., discourse completion task and role-play) involving different situations as prompts to collect samples of learners' communicative acts and analyze linguistic forms in the acts (Nguyen, 2019). Often researchers have analyzed the forms in relation to learners' general proficiency (as assessed with a standardized proficiency measure) to gain developmental insights into their pragmatic competence. Existing findings are largely consistent, revealing a clear impact of proficiency. As learners' general proficiency develops, they typically use longer and more elaborate syntactic forms, along with a wider range of modification devices (e.g., hedging) that can mitigate the tone and force of their act (Xiao, 2015). The close relationship between proficiency and pragmatic competence is plausible because knowledge of grammar and vocabulary – primary indications of general proficiency – can directly support learners' use of linguistic forms in communication. More proficient learners have a larger repertoire of syntactic forms and vocabulary to construct utterances and supply mitigation devices.

The present study investigates whether the established relationship between proficiency and pragmatic competence can extend to learners' ability to attend to visual contextual cues to shape their course of action (i.e., identifying the right person to direct their communicative act to). The study addresses the following research question: Is there a relationship between proficiency and L2 learners' ability to direct their eyes to the appropriate person based on the contextual cues available in a given situation?

2. Method

2.1. Participants

Participants involved 40 speakers of L2 English enrolled in a university in China. They were all volunteer participants recruited from different levels of study (undergraduate and graduate level) to represent wide-ranging English proficiency levels. Of 40, 36 were female and four were male. The mean age was 23.0 ranging from 18 to 28. The undergraduate group (n = 19) included four freshmen, two sophomores, and 14 juniors who were majoring in English. The graduate group (n = 21) included two MA students and 19 Ph.D. students majoring in Languages and Literature. They were all native speakers of Chinese who grew up in China. On average they had 14.6 years of formal English study in China. Only one Ph.D. student had a study-abroad experience in an English-speaking country (1.5 years in Canada).

2.2. Instruments

2.2.1. Video-based speaking task

The video-based speaking task was developed to assess participants' ability to analyze contextual cues and produce communicative acts. The speaking task included 19 items: 8 experimental items, 9 filler

https://doi.org/10.1017/S0261444825100864 Published online by Cambridge University Press

Sample Item 1

Instruction: You will read a scenario written in English on the computer. After reading the scenario, you will watch a short video about the scenario. As soon as the video freezes, speak to the person in the video **in English** following the scenario. **When you speak, please look at the person you are speaking to.**

Scenario: You are in your English class. You forgot to bring a pen. You want to borrow a pen from your classmate. What do you say?

Figure 1. Sample item 1.



Figure 2. Screenshot of the video scene for sample item 1.

items, and 2 practice items. All items took the same format. Participants read a brief scenario in English and then watched a video illustrating the scenario for about 20 seconds. When the video stopped at the end, participants were asked to orally produce the target communicative act directed to a person in the video. Participants' speech and eye movements were recorded on the computer. Figure 1 displays a sample item. The video clips were downloaded from stock videos available on Storyblocks (https://www.storyblocks.com/) and edited so they fit each scenario.

Several people appeared in each video, but participants were asked to direct their eyes to one person when speaking. In the experimental items, some distinct contextual cues signaled the most appropriate person to gaze at while speaking in the given scenario. For example, in Sample Item 1 above (asking to borrow a pen), three people appeared in the video (Fig. 2). The male person on the left was the most appropriate person to talk to because he has an extra pen in front of him. Hence, participants were expected to look at this person and ask him for a pen.

2.2.2. The quick placement test

The Quick Placement Test (QPT) was given to participants via computer to assess their general English proficiency. The QPT is a commercially available standardized English test developed by Oxford University Press and Cambridge ESOL (University of Cambridge Local Examination Syndicate, 2001). The test, consisting of two parts, assesses English proficiency in listening, reading, and grammar through a series of multiple-choice questions (60 items total). It took about 30 minutes to complete the test.

https://doi.org/10.1017/S0261444825100864 Published online by Cambridge University Press

2.3. Data collection procedure

Data collection took place individually in a quiet office on campus. To ensure that participants fully understood the experimental requirements, all instructions were presented in their native language (Chinese). Participants were seated approximately 60 cm from a computer screen and their eye movements were recorded with a screen-based eye tracker (Tobii Pro Spectrum) with a sampling rate of 1200 Hz. Binocular eye data were tracked throughout the session. A standard 9-point calibration procedure was conducted to ensure accurate data recording. Immediately following the calibration procedure, participants completed two practice trials to familiarize themselves with the task.

The main experimental procedure consisted of 17 trials (8 experimental items and 9 filler items), each starting with a centrally presented fixation point displayed for 500 ms as a calibration check. Next, a brief textual description of the scenario was displayed on the screen. Once participants had fully understood the description of the scenario, they clicked the mouse to view a corresponding video. When the video stopped, they were instructed to immediately fixate on one of the characters and perform the video-based speaking task. After completing their response, participants clicked the mouse again to indicate they had finished speaking and entered the next trial.

The entire testing session lasted approximately 30 minutes. Eye-tracking data from the main task were preprocessed using Tobii Pro Lab and then exported to Excel for further analysis.

2.4. Data analysis procedure

Two types of eye tracking data were collected: eye fixation duration and eye fixation counts. To collect these data, we first identified the Area of Interest (AOI) for each of the eight experimental items. The AOI refers to a predetermined specific region within a visual environment where we wanted to analyze participants' attention. To reiterate, the goal of the eye-tracking analysis was to determine whether participants were able to attend to the appropriate person to speak to based on the contextual cues available in the video (see Sample Item 1 in Figs. 1 and 2). Hence, AOIs were drawn around the target person appearing in the last scene when the video stopped and participants started speaking. Figure 3 below shows the sample AOI corresponding to Sample Item 1. The AOI is the red box surrounding the male person on the left. Eye fixation duration refers to the total duration of gaze time at the AOI while speaking. Eye fixation counts refer to the number of times when participants shifted their eye movements to the AOI while speaking. The sum of the fixation duration and counts from eight target items were used for the analysis.

Pearson correlation was used to examine the relationship between eye fixation to AOIs (fixation duration and fixation counts) and general proficiency (QPT score). Because eye fixation data were not normally distributed, they were log-transformed. After the transformation, the data met the assumption of normality. All the statistical analyses were conducted on the log-transformed data.

3. Results

Table 1 displays descriptive statistics of proficiency score (QPT score), eye fixation duration (by seconds), and eye fixation counts. The mean QPT score was 42 points, which roughly corresponds to the B1 level in the Common European Framework of Reference (Council of Europe, 2001). The QPT score range of 33 to 52 indicates that the participants fell into a relatively narrow proficiency range, spreading from A2 to B2 level. Eye fixation duration (sum of duration from all eight experimental items) showed a large range, spanning from 15.50 to over 80 seconds total, indicating a large individual variation in terms of how long they fixated their eyes on the target person while speaking. Similarly, eye fixation counts also showed a large variation with mean counts of over 99 and a standard deviation of almost 70.

https://doi.org/10.1017/S0261444825100864 Published online by Cambridge University Press



Figure 3. AOI for sample item 1.

Pearson correlation revealed no significant relationship between the QPT score and eye fixation duration ($r = .18$, $p = .27$) nor between the QPT score and eye fixation counts ($r = .18$, $p = .27$).

4. Discussion

This study investigated the relationship between general proficiency and pragmatic competence from the point of contextual sensitivity using eye-tracking measures. We examined whether proficiency is related to learners' ability to pay attention to contextual cues to identify the right person to direct their speech to. The learners' eye movements to the target person in a video prompt were measured when they were performing the target communicative act. Then their relationship with general proficiency was assessed using a standardized measure.

Pearson correlation revealed no significant relationship between proficiency and either of the eye fixation data (duration and counts). Hence, the present findings offer counter evidence to the established consensus that general proficiency is related to pragmatic competence from the point of contextual sensitivity (Taguchi & Roever, 2017; Xiao, 2015). Regardless of proficiency levels, learners similarly directed their eyes to the target person when speaking. Hence, proficiency did not influence learners' ability to pay attention to the contextual cues that facilitated their decision about who to look at and direct their speech to.

These findings indicate that the visual contextual sensitivity examined in this study is within the domain of general cognition, rather than the domain of linguistic competence. Among the number of contextual cues competing for visual attention, people are likely to attend to the most salient and relevant cues, which are closely tied to their communicative goal (Chun, 2000). In this study, the prompts given to the participants set the goal of their communicative act. Because of the specific behavioral goals and set intentions, the key contextual cues embedded in the visual environment probably became salient to the participants. It seems that the participants drew on the contextual cues as resources to determine where to look during the target communicative act. In other words, contextual information guided participants' visual processing, directing their attention to the salient object (i.e., target interlocutor) and facilitating their response to the object (Chun, 2000). These series of cognitive processing are largely independent of linguistic processing. With normal cognitive functioning, adult participants, like L2 learners in this study, can execute those processes efficiently during communication. In essence, no relationship between eye tracking data and proficiency makes sense

https://doi.org/10.1017/S0261444825100864 Published online by Cambridge University Press

Table 1. Descriptive statistics

	Mean	*SD*	Min.	Max.	95% CI
Total QPT score	42.13	4.60	33.00	52.00	40.65; 43.60
Total fixation duration (seconds)	28.18	19.97	15.49	80.02	21.79; 34.57
Total fixation counts	99.13	69.23	10.00	316.00	76.99; 121.26

Note: QPT score range: 0–60. Mean refers to the sum of the data from 8 experimental items.

because these cognitive processes involving visual contextual sensitivity are unlikely to be affected by linguistic competence and L2 proficiency.

The present findings offer a new perspective on the concept of contextual sensitivity as the ability to process critical contextual cues in a visual environment and use them to guide own behaviors. Although previous studies focused on the ability to vary linguistic forms across different contexts and revealed proficiency impact on the ability, visual contextual sensitivity examined in this study is unique in that it is free from proficiency influence.

5. Future directions

This study was conducted at Ocean University of China as part of an ongoing project that explores the construct of L2 pragmatic competence and process of pragmatic development. Since almost no studies have used eye tracking to examine visual contextual sensitivity in relation to pragmatic competence, we will continue to assess the generalizability of the present findings. Our next step is to employ data triangulation because the interpretations of the eye fixation data and cognitive processes inferred from the results are somewhat speculative. In the future, we will use introspective interviews or think-aloud protocol to explore where participants pay attention to and why. Such additional data can help us understand how contextual cues actually facilitate participants' visual processing and subsequent actions and essentially inform their contextual sensitivity and pragmatic competence. We will also refine the instrument used in this study. Large variation found in the eye fixation duration and frequency across eight experimental items indicates that the items were not consistent; it seems that some contextual cues were more salient than others in the video clips. Although using real-world scenes has ecological validity, reliability and validity of experimental items are crucial, and thus experimental control needs to be implemented more strategically in our future work.

References

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences*, *4*(5), 170–178. https://doi.org/10.1016/s1364-6613(00)01476-5

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.

Nguyen, T. T. M. (2019). Data collection methods in L2 pragmatics research: An overview. In N. Taguchi (Ed.), *The Routledge handbook of SLA and pragmatics* (pp. 195–211). Routledge.

Taguchi, N. (2023). Technology-enhanced language learning and pragmatics: Insights from digital game-based pragmatics teaching. *Language Teaching*, *57*(1), 57–67. https://doi.org/10.1017/S0261444823000101

Taguchi, N., & Roever, C. (2017). *Second language pragmatics*. Oxford University Press.

University of Cambridge Local Examination Syndicate. (2001). *Quick Placement Test*. Oxford University Press.

Xiao, F. (2015). Proficiency effect on L2 pragmatic competence. *Studies in Second Language Learning and Teaching*, *5*(4), 557–581. https://doi.org/10.14746/sllt.2015.5.4.3

Cite this article: Taguchi, N., & Xu, S. (2026). Attending to contextual cues to convey pragmatic meaning: An exploratory eye-tracking study. *Language Teaching*, 59(1), 128–133. https://doi.org/10.1017/S0261444825100864

https://doi.org/10.1017/S0261444825100864 Published online by Cambridge University Press