Knowing the Future: Predictive Processing in the ELT Classroom

Edward C. Howland

Hiroshima YMCA

Recent research in neuroscience has confirmed the brain's ability to predict linguistic input before it occurs. Predictive language processing (PLP) is linked to language comprehension and production, yet it is rarely discussed within the ELT academic community. This paper is intended to explore that omission in three stages. First, available research is reviewed. Second, the results of a questionnaire regarding English language teachers' awareness of and attitudes towards PLP are discussed. Finally, a presentation of two techniques designed to train an L2 learner's predictive abilities is offered as a model for future language teaching.

最近の神経科学の研究により、脳が言語的な入力をあらかじめ予想する能力があることが確認されている。予想言語処理（predictive language processing; PLP）は言語の理解力と発話に関わりがあるが、英語教授法（ELT）の分野のなかでめったに論じられていない。本論文では、この分野を深く三段階にわたって研究する。最初に、これまでの研究を総括する。次に、アンケート調査の結果から、PLPに関して英語教師の認識また態度について考察する。最後に、今後に応用可能な言語教授法のモデルとして、第二言語学習者（L2）の予想力を高めるための二つのトレーニング技術が提示されている。

In the past two decades, researchers in the fields of linguistics, psycholinguistics, and neuroscience have conducted extensive research into the mind's ability to predict upcoming linguistic input. It is well established that prediction is a fundamental part of the way the brain processes (Kuperberg & Jaeger, 2016) and acquires (Phillips & Ehrenhofer, 2015) language. Despite its importance to language acquisition, prediction seems to be an overlooked topic in ELT and foreign language teaching (FLT) academic circles. This paper is a summary and expansion of a series of presentations I have given at conferences around Japan. It is intended to help start a discussion on the topic of prediction in our language-teaching community. It is organized into three parts. Part 1 consists of a review of extant research on predictive language processing (PLP), introducing the concepts and establishing their importance to the language educator. Part 2 presents the results of a small questionnaire on teacher awareness of and attitudes towards PLP. This survey should not be considered an authoritative investigation, but rather a first step. In Part 3, a pair of ideas for training and strengthening predictive skills are discussed.

Defining Prediction

It is worth noting at this point that predictive language processing (also referred to as anticipatory processing) as understood in this paper is separate from the technique called "prediction" that is often used in English language classes. The primary distinctions are the *speed* and *automaticity* with which the processing takes place. The classic exercise of having students read the beginning of a sentence or story, then giving them time to write an ending is certainly a form of prediction, but it is not the subject of this paper. Predictive processing takes place in a span of milliseconds, without conscious thought or effort. Researchers theorize that near-simultaneous prediction is active in all aspects of daily life, from driving an automobile (Engström et al., 2018) to playing sports (Ridderinkhof & Brass, 2015), but this paper focuses on the phenomenon of linguistic prediction.

A Thought Experiment

To demonstrate the concept of predictive language processing, imagine a casual conversation between friends. One friend (A) starts a sentence with the phrase "I love . . . ." Their interlocutor (B) relies on their own understanding of English syntax to allow them to make a reasonable guess as to the grammatical construction of the rest of the sentence, while their semantic knowledge helps them predict the meaning. At this point,

the range of possibilities for the next utterance is nearly limitless. "I love you" is quite a common expression as is "I love this weather." As the sentence progresses, however, the options become more limited. For example, if (A) goes on to say, "I love to eat . . . ," it is a near certainty that the next word or phrase will be a noun (unless, of course, they are going to say, "I love to eat late at night") and almost certainly something edible.

Interlocutor (A) continues their sentence, saying, "I love to eat Christmas . . . ." At this point, cultural context becomes a factor in the predictive process. If (A) is British, they are likely to finish the sentence with "pudding," but an American is more likely to talk about "cookies," and a Japanese person may say "cake." Of course, the context of the conversation at hand also plays a major role. If (A) and (B) were discussing Japanese Christmas traditions, the most likely choice would be "cake." If (B) had just expressed their distaste for sponge cake, the stress and intonation choices (A) makes (i.e., "Well, *I* love to eat Christmas cake.") would indicate to an experienced listener the direction the sentence was headed from the second phoneme. Thus semantic, lexical, phonological, and cultural contexts all play a role in the predictive process.

For the purposes of this thought experiment, time was not taken into consideration. In an L1 conversation, the entire sentence would have been uttered in around one to one-and-a-half seconds. During that time, (B)'s mind would have generated, tested, rejected, and reworked numerous predictions about the possible conclusion of the sentence (Pickering & Garrod, 2013). Assuming English is (B)'s first language, they would make these predictions automatically, effortlessly, and without a moment's conscious thought. For those working in their second language, the process is considerably more difficult, requiring conscious effort. This effort is an important consideration for language learners and their teachers.

Literature Review
Prediction in L1
Two main types of experiments have been conducted to provide evidence of predictive processing in language: eye-tracking and event-related potential (ERP). Kuperberg and Jaeger (2016) conducted the most comprehensive survey of the research available to date, and interested parties are encouraged to consult their work to determine the breadth of the field. In this paper I will describe in detail one experiment of each type to provide a concrete introduction to the concepts.

Eye-Tracking
Eye-tracking equipment can pinpoint exactly where on a screen the user is looking at any given moment. If the user looks at an item significantly before pointing at or naming it, it can be inferred that some part of the mind has identified the item before the conscious mind is even aware of it (Cooper, 1974). In 2007, Altmann and Kamide conducted an experiment in which they showed participants a picture of a man standing next to an empty wine glass and a full beer glass. Over headphones, a neutral male voice read one of the following sentences: "(1) The man will drink the beer. (2) The man has drunk the wine." In a significant percentage of the cases, the subject's eye moved to the full beer glass between the words "will drink" and "beer" and to the empty wine glass between the words "has drunk" and "wine." Altmann and Kamide concluded that the listeners (all L1 English users) had used the syntactic clues ("will drink" implying a full glass and "has drunk" implying an empty glass) to anticipate the next word before they heard it.

Event-Related Potential
ERP experiments use electroencephalography equipment to measure changes in brain activity. Semantically inappropriate words have been shown to set off a flurry of electrical activity in the brain, known as the N400 potential (Kutas & Hillyard, 1980). In one experiment, DeLong, Urbach, and Kutas (2005) displayed a sentence on a computer screen one single word at a time and measured the electrical responses in the subject's scalp. The subjects (all L1 English users) saw the sentence "The day was breezy so the boy went outside to fly . . ." followed by either the most likely continuation ("a kite") or a less-contextually-appropriate phrase ("an airplane"). The interesting finding was not that the word "airplane" set off the N400 activity, but rather that the N400 began activating when the article "an" was displayed on the screen. The presence of a reliable electrical signal at the point of divergence from a statistically likely outcome demonstrates that an L1 English user takes input (in this case, the semantic association of the breezy day plus the typical collocations of the verb *to fly*) and uses it to develop an anticipatory model of the rest of the sentence (DeLong et al. (2005) cited a cloze probability of *a* and *kite* at 86% and 89%, respectively) before they receive the input. When the input does not match the model (the article *an* indicates that the following word will begin with a vowel, which *kite* does not), the brain scrambles to rework the model given the newer data.

Prediction in L2

As noted in the previous section, an exhaustive review of the literature regarding differences between the anticipatory abilities of L1 language users and L2 learners already exists, and readers are encouraged to peruse Kaan's (2014) work for a complete overview of the topic. The consensus is that L2 learners do not anticipate upcoming linguistic input with the same rapidity and automaticity as L1 users, identified as the RAGE (reduced ability to generate expectations) hypothesis (Grüter, Rohde, & Schafer, 2014). This reduction of anticipatory ability persists across languages. For example, Lew-Williams and Fernald (2010) conducted an eye-tracking experiment with L1 and L2 users of Spanish. In Spanish, all nouns are assigned a gender; that is, *la pelota* (the ball) is feminine and *el zapato* (the shoe) is masculine. The subjects were shown pictures of various objects (let us assume in this case a ball and a shoe) on the screen at once and listened to a sentence: "*Encuentra* (find) *la pelota* or *el zapato*." Upon hearing the gender-discriminated article *la* or *el*, L1 users' eyes accurately and consistently tracked to the correct picture before they heard the noun to which it referred. L2 learners' eyes only moved to the appropriate picture once the noun itself was read. Although L2 participants were determined to have been familiar with the vocabulary before the experiment, they could not predict the upcoming noun based solely on its gender with anything like the speed and accuracy of the L1 speakers.

The ELT Context

Up to this point all the research referenced in this paper has been drawn from the fields of neuroscience, linguistics, and psycholinguistics for the simple reason that, as far as can be determined, nothing on the topic has been published in ELT/FLT journals. It seems clear that such a significant aspect of the language learner's experience of the L2 should be a topic of discussion among language education professionals.

Questionnaire

As noted previously, this paper is an elaboration on a series of presentations and workshops on PLP that I gave at ELT/FLT conferences in Japan in 2017 and 2018. In discussion, workshop participants tended to be unaware of or unclear on the topic. I designed a questionnaire to quantify English language teachers' awareness of and interest in PLP both before and after the presentations. The results are presented here in the hope that they will serve as a jumping-off point for other researchers to investigate the relationship between PLP and language education.

Format

The questionnaire was used in two sessions, once at the 2018 JALT conference in Shizuoka, Japan and again at an in-service training session at the Hiroshima YMCA School of Languages. On both occasions, the questionnaire was administered immediately following a workshop on PLP that included presentations of the activities outlined in the following section. Responses were on a typical 5-level Likert-type scale with 1 signifying "strongly disagree" and 5 signifying "strongly agree." The statements were as follows:

Before the workshop
1. I was aware of predictive language processing (PLP).
2. I took PLP into account in my planning/classwork.

After the workshop
1. I am confident in my understanding of PLP.
2. I plan to learn more about PLP.
3. I plan to take PLP into account in my planning/classwork.

Results

The mean responses are in Table 2. A more detailed representation of the responses can be found in Figure 1.

Table 1. Mean Responses by Question (*N* = 12)

Question	Mean
1. I was aware of predictive language processing (PLP).	2.58
2. I took PLP into account in my planning/classwork.	2.17
3. I am confident in my understanding of PLP.	3.83
4. I plan to learn more about PLP.	4.17
5. I plan to take PLP into account in my planning/classwork.	4

Note. Responses were from 1 *strongly disagree* to 5 *strongly agree*.

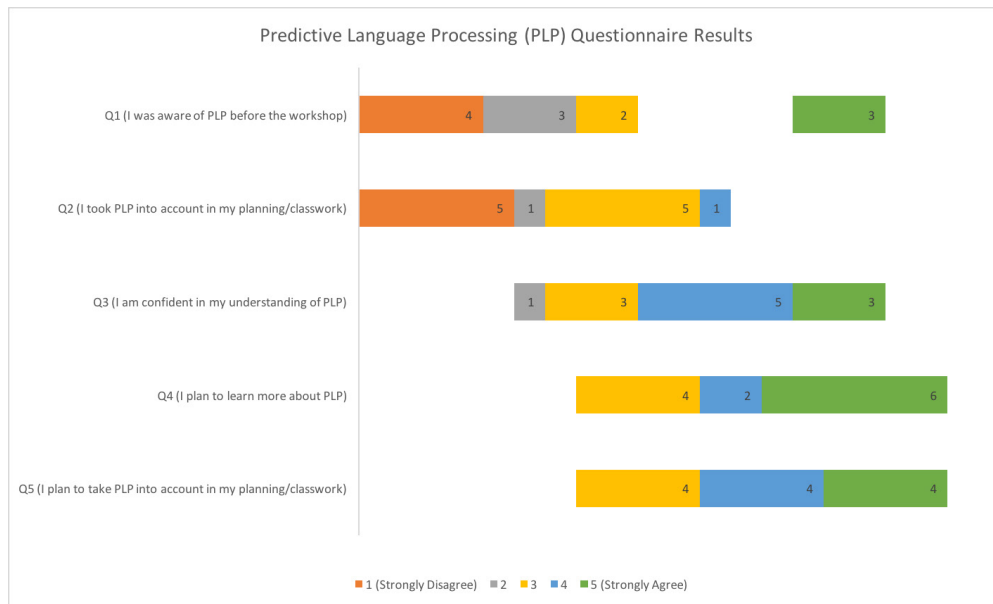**Predictive Language Processing (PLP) Questionnaire Results**

Figure 1. Responses to the questionnaire (*N* = 12).

Discussion

Most of the teachers who participated in the workshops were unfamiliar with the concepts beforehand. The three anomalous results (answers of 5) to Q1 could be explained by the fact that one workshop was presented under the banner of the BRAIN Special Interest Group, the members of which have an interest in the intersection of neuroscience and ELT. Those three aside, the results suggest that this sample of English language teachers were generally unaware of PLP prior to the presentation, answering 1, 2, or 3 to the first two questions. After being introduced to the concepts, they seemed enthusiastic to learn more, with 100% of respondents answering 3, 4, or 5 to questions 4 and 5.

As noted previously, this questionnaire was designed primarily for my own edification and should not be considered an authoritative take on the topics investigated. Its main limitation is the number of respondents. Twelve teachers, several of whom are neuroscience enthusiasts, is not a sufficient sample size to assess the general awareness

of and attitudes towards PLP in the ELT community overall. There is an opportunity here for further research. A more detailed survey of teacher awareness/attitudes towards PLP with a much larger sample is an obvious starting point.

How to Teach Prediction?

If we can assume that predictive processing is a concern for English language teachers, the question then is raised: How exactly can we help our students do it? If we consider English as a skill to be acquired rather than a set of rules to be memorized (DeKeyser & Criado, 2012), we find that automaticity and speed can both be improved by dedicated practice combined with feedback. In this section, two games designed to assist with developing the predictive skills of students will be presented. The first is a video game developed by Swedish researchers; the second is an adaptation of the Japanese game *karuta* that I have developed.

A Video Game

Schremm, Hed, Horne, and Roll (2017) developed a prototype computer game to attempt the task of training predictive L2 processing. In the Swedish language, certain tone inflections of a verb stem indicate the correct suffix that should be applied to demonstrate tense. The study participants would hear the beginning of a sentence, for example *Kungen bygg-* ("The king build-) and were presented with two suffix choices, *-er* (present tense) and *–de* (past tense). If the stem was presented with accent 1, the correct choice would be the present tense suffix, and vice versa. The participants then pushed a button on the screen to indicate their choice of the correct suffix. Crucially, they were scored not only on their accuracy, but also on the speed of their response, as both are vital to the successful prediction of language in a L1-level interaction.

The researchers found that playing the game not only increased the participants' accuracy with the given language feature, but over time reduced the latency of their responses. In other words, they not only got better at predicting, they got faster. Furthermore, their ability to *produce* the tonal cue improved as well. Of particular interest to CALL educators, Schremm et al. (2017) are currently working on developing their simple, text-based prototype into a visually pleasing, fast-paced action game. Such a game design could very easily be adapted to an English-language scenario, training the player's ability to predict any number of syntactical, semantic, or phonological utterances and could be very useful for working with younger, screen-savvy students.

Karuta

競技かるた (*Kyougi karuta*) is a classical Japanese card game. To provide a vast oversimplification of the rules, two players sit across from each other facing a set of *torifuda* [grabbing cards] laid out on the table or *tatami* mat. Each *torifuda* has the conclusion of a poem on it, the beginning of which is displayed on a matching *yomifuda* [reading card], which is held by the judge. To play, the judge reads a *yomifuda* out loud, and as soon as the players feel confident that they know which poem is being read, they reach out and slap the *torifuda* containing the appropriate words. In competitive *karuta*, players often know from the very first syllable which card they should reach for. This focus on both speed and precision is exactly the factor that makes the game appropriate for training students to automatically and unconsciously finish sentences in their head. The head-to-head nature of the game provides a strong motivation to answer quickly and correctly.

Researchers and educators have written extensively about utilizing *karuta* to develop students' pronunciation (Gooch, Saito, & Lyster, 2016), vocabulary (Fachriyani & Syafe'i, 2018), and even motivation (Lukito, 2013), but not their predictive skill or automaticity. As such, with credit to Philip Head for the original idea, I developed an adaptation of *karuta* that is aimed at training the students to predict upcoming language quickly and accurately.

In this version of the game, there are no *yomifuda*, and each *torifuda* has only one word written on it. These words could be random words, as in our example cards, or could be adapted to the needs of the specific class. For instance, a list of vocabulary words that the class has been working on for the past weeks or a set of core vocabulary for the EIKEN exam would make excellent *torifuda*. In the case of the example cards provided in Figure 2, the students were 6 to 7 years old and assessed at pre-A1 on the CEFR scale. As such, they were provided with very simple vocabulary with which they were already familiar.

This version of *karuta* has been played in class sizes ranging from one to eight students, as opposed to the traditional two, with great success. The cards are laid out on the table with the students arranged around them in such a way that all cards are equally accessible to each student. The teacher has prepared a list of sentences, each with the last word missing. To play the game, the teacher reads a sentence from their list, and as soon as a student is certain that they know what the last word will be, they slap the appropriate card. Given the cards in Figure 2, the teacher might read the sentence, "When I grow up, I want to be a _____." The first student to slap the "doctor" card would get a point. As the game progresses, the teacher can use trickier sentences. For example, the sentence, "My father is a teacher, but my mother is a _____,"

Figure 2. Eight torifuda in a karuta deck designed for Japanese first and second graders.

would also require students to slap the "Doctor" card, but in practice, they often slap the "English" card as soon as they hear the word "teacher." This error provides an excellent look at the process of prediction, hypothesis testing, rejection, and reformation. Given that the context of the game is an EFL class, it is not unreasonable to link the word "teacher" with the "English" card and predict that the sentence is headed in that direction. However, as soon as the second clause is introduced, the listener realizes that they should instead be searching for job categories and seeks out a card containing such a word.

The game has many possible adaptations. In traditional *karuta* matches, the *torifuda* is removed as soon as it is matched because there is only one possible match, whereas in this adaptation, multiple sentences may point to the same *torifuda*; as such they remain in play throughout the game. Depending on the age and level of competition between the students, the teacher can keep a running tally of each student's points on the board or potentially not even keep score at all. The lack of a scoreboard can keep the game in-the-moment and the students focused on the task at hand.

Conclusion

While linguists and neuroscientists have been investigating predictive language processing for nearly two decades, ELT/FLT research is lacking. The three parts of this

paper were designed to provide a solid background on the theory of PLP, initiate research in our ELT community, and inspire some classroom activities. Both Schremm et al. (2017)'s video game and this adaptation of *karuta* are designed to help students to know the future without even realizing that they are doing so. It is my hope that the readers of this paper become inspired to conduct their own inquiries into PLP, design their own activities, and share them with the teaching community. A continuing focus on this vital aspect of language acquisition can only benefit our students.

Bio Data

Edward Cooper Howland lives and works in Hiroshima. He enjoys bringing research from outside disciplines to the ELT community's attention. He has recently taken up curling in his free time. <ech@cooperhowland.com>

References

Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language, 57*(4), 502-518. https://doi:10.1016/j.jml.2006.12.004

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6,* 84-107. https://doi.org/10.1016/0010-0285(74)90005-X

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8*(8), 1117. https://doi.org/10.1038/nn1504

DeKeyser, R., & Criado, R. (2012). Automatization, skill acquisition, and practice in second language acquisition. *The Encyclopedia of Applied Linguistics*. https://doi.org/10.1002/9781405198431.wbeal0067

Engström, J., Bärgman, J., Nilsson, D., Seppelt, B., Markkula, G., Piccinini, G. B., & Victor, T. (2018). Great expectations: a predictive processing account of automobile driving. *Theoretical Issues in Ergonomics Science, 19*(2), 156-194. https://doi.org/10.1080/1463922X.2017.1306148

Fachriyani, I., & Syafe'i, A. F. R. (2018). Using *karuta* game in teaching vocabulary to young learners. *Journal of English Language Teaching, 7*(1), 32-39.

Gooch, R., Saito, K., & Lyster, R. (2016). Effects of recasts and prompts on L2 pronunciation development: Teaching English /ɹ/ to Korean adult EFL learners. *System, 60,* 117-127. https://doi.org/10.1016/j.system.2016.06.007

Grüter, T., Rohde, H., & Schafer, A. (2014, May). The role of discourse-level expectations in non-native speakers' referential choices. *Proceedings of the Annual Boston University Conference on Language Development*. Retrieved from http://par.nsf.gov/biblio/10028988

Kaan, E. (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism, 4*(2), 257-282. https://doi.org/10.1075/lab.4.2.05kaa

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience, 31*(1), 32-59. https://doi.org/10.1080/23273798.2015.1102299

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207*(4427), 203-205. https://doi.org/10.1126/science.7350657

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language, 63*(4), 447-464. https://doi.org/10.1016/j.jml.2010.07.003

Lukito, I. E. J. (2013). Improving students' motivation of speaking by using game card tournament: The case of grade 7 students of state junior high school 6 Tegal, in the academic year of 2012/2013. *English Education Journal, 3*(2). https://journal.unnes.ac.id/sju/index.php/eej/article/view/2708

Phillips, C., & Ehrenhofer, L. (2015). The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism, 5*(4), 409-453. https://dx.doi.org/10.1075/lab.5.4.01phi

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329-347. https://doi.org/10.1017/S0140525X12001495

Ridderinkhof, K. R., & Brass, M. (2015). How kinesthetic motor imagery works: a predictive-processing theory of visualization in sports and motor expertise. *Journal of Physiology-Paris, 109*(1-3), 53-63. https://doi.org/10.1016/j.jphysparis.2015.02.003

Schremm, A., Hed, A., Horne, M., & Roll, M. (2017). Training predictive L2 processing with a digital game: Prototype promotes acquisition of anticipatory use of tone-suffix associations. *Computers & Education, 114,* 206-221. https://doi.org/10.1016/j.compedu.2017.07.006