論文

Implications of the Predictive Processing Model of Human Cognition for Foreign Language Learning and Teaching

Stephen M. Ryan[1)]

Keywords: Predictive Processing, ELT, Language Learning

Predictive Processing Theory is a model of human cognition which offers an explanation of how individuals perceive, learn from, and predict sensory signals from their environment. Since everything we do can be conceptualized in terms of interaction with our environment, this is a very wide-ranging theory which has implications for all aspects of human existence.

In this paper, I shall explore the implications of Predictive Processing Theory for the learning and teaching of a foreign language. The primacy of experience in the learning process, as envisaged by Predictive Processing Theory, means that certain language learning activities, such as extensive reading and discovery learning, are recommended as more effective than others, such as having the teacher explain things or de-contextualised practice activities.

Learning

Thought Experiment

Imagine you are trapped inside a box (Feldman Barrett, 2020). You cannot get out of the box. You are almost completely cut off from the world outside the box, having only a few holes in the sides of the box through which to perceive it. Yet, you are dependent on the world outside the box for the resources you need to live. Over time, you become aware, through the holes in the box, that there are changes in the world outside. Some of those changes represent opportunities to increase your store of resources. Others represent threats to your very existence. In order to take advantage of the opportunities and avoid the threats, you need to reposition yourself and re-distribute the resources that you have access to.

The box is your skull. Your brain sits inside your skull. It cannot leave your skull and can gather information about the world outside your skull only through the very limited “holes” of your five senses: limited physically, because, for example, your eyes can look in only one direction at once, your ears perceive sound over a limited range of frequencies and even that only within a certain distance; and limited in terms of processing capacity, as your brain can process only 40 or so of the estimated eleven million bits of information that are available from your physical environment at any one

1) 山陽学園大学総合人間学部言語文化学科

time (Zimmerman, 1986). Yet, your brain and body are dependent on resources from outside the skull, such as air, food, and water, for their survival. Knowing how to interpret limited sensory information in a way that makes sense of the world outside is the difference between sudden death and surviving to pass on your genes to the next generation.

So, what strategies has the brain, under such evolutionary pressures, developed to get by on the small amount of information available to it? First, it has learned to detect and take note of patterns. In your box, however small the holes, you will notice the passage from night to day and back again. You will gradually become aware of other patterns: the sounds and smells the arrival of certain phenomena: lightening, predators, companionship.

Second, on the basis of such patterns, it has learned to make predictions about what will happen next. Having a good idea what to expect is essential to the redistribution of resources (limbs, glucose, hormones, electrolytes, oxygen supply, even the location of the box itself – after all, our particular box has arms and legs). Without this redistribution, the organism will not be positioned to take advantages of the opportunities and to avoid the threats that the brain predicts from its environment (Damasio, 1994). The predictions are made by comparing (incomplete) patterns sensed through the holes with previously sensed (also incomplete) patterns. Just like the child who is able to complete the sequence "1, 2, 3, __, __," the brain learns to predict the next parts of a pattern based on its previous experiences of similar patterns.

Third, because the predictions are based on necessarily incomplete information about the world outside the skull, they will at first be crude and unreliable, so a mechanism is needed whereby new data, forming new patterns, can be used to update the basis on which future predictions are made. Thus, a heavy rain shower does not always occasion the sound of thunder. If I had predicted, based on limited experience, that the sound, smell, and feel of rain would soon be followed by the rumble of thunder, further experience would help me to form more nuanced predictions: sometimes thunder, sometimes no thunder. With further experience, I might then become aware of the sensory antecedents of a thunderstorm as opposed to those of a downpour without thunder. Further experience would continue to refine the basis of my pedictions from ever more subtle sensory signals that allow me to make ever more accurate predictions about whether that sounds, looks, and feels a certain way is likely to be accompanied by thunder.

Predictive Processing

This, in essence, is the Predictive Processing model of human cognition (Clark, 2015) first developed to account for how we thrive despite having access only to incomplete information (Feldman Barrett, 2020), but now coming to be recognized as the preeminent theory of how the brain makes sense of the world around it (Campbell, 2016). An efficient pattern detection routine sifts through whatever data is available. The

patterns are used to construct a series of “priors,” basically distillations of experience: an understanding, from previous encounters, of the sensory patterns occasioned by changes in our environment. The priors are used to make specific predictions about subsequent sensory information. The predictions are used, in turn, to set up the most appropriate balance of bodily resources to deal with the developing situation. Successful predictions lead to efficient use of resources. Unfulfilled predictions generate a “prediction error,” a signal that all is not as expected and that the priors for such a situation need to be modified, using the (unpredicted) data that generated the error signal, to allow for more accurate predictions and, hence, more efficient use of resources the next time this situation is encountered.

The Predictive Processing model has turned our understanding of human cognition upside down (Feldman Barrett, 2018). While classical models of cognition begin with sensory input, which the brain reacts to and tries to make sense of, the new model begins with the sense that our brain has made of the world so far (our priors), and uses it to generate predictions about the sensory data that is about to be perceived. These predictions then prime the organism to react. The reaction occurs, unless data coming from the senses suggests that the predictions are incorrect (prediction error), necessitating a change in the reaction and a modification of the priors for next time.

This is also, it should be noted, a model of learning (Campbell, 2017). The modification, by experience, of our expectations about the world is exactly what we mean by learning. It can be conceptualized in two ways: as attempting to make a new piece of the puzzle fit our existing understanding of the world; or as allowing the new information to modify our understanding. In reality, the two processes are one: mutual adjustment between priors and new sensory information to sharpen our predictive abilities.

The importance of priors in this process needs to be stressed. They are central both to predicting and to learning. They are formed by experience, or, more precisely, by repeated exposure to similar experiences (similar patterns perceived outside the skull/box). On the neurophysiological level, such experiences occasion the repeated firing of similar groups of neurons. From our example above, repeated exposure to rainstorms would cause the firing of a certain group of neurons. The neurons fired would not be exactly the same, as no two rainstorms are exactly the same. However, there is sufficient similarity between rainstorms for a core group of neurons to fire together during each rainstorm. These neurons that become accustomed to being fired together (Hebb, 1949) constitute our core concept of a rainstorm and thus are the “priors” on which our brain bases predictions about future rainstorms.

There is, however, another kind of learning. There are, after all, other brains in other boxes all around us. We can learn not only from our own experience but from that of others, too, thanks to our ability to communicate our experiences to others across time and space. Language (verbal or non-verbal) allows us to do this, liberating us from the solitude of the box. Learning from the experience of others increases our ability to generate accurate predictions about the world around us exponentially. It is the

superpower that has made human beings the most (ecologically) successful species ever to exist on this planet, able to adapt to just about any environment, to survive there and to thrive. Thanks to language, the brain is a social organism, and learning an essentially social process.

It should be noted that what we get from others is not experience; it is knowledge of experience. The predictive processing model suggests that knowledge and experience are treated quite differently. While knowledge is something that we may store in memory and later call on when it is useful, it is experience that directly builds, modifies, and fine-tunes our priors. As we shall see, knowledge is not the same as experience, but it can be converted into experience.

Language Learning

So, how do humans learn language? The same way they learn anything else: by using priors to generate predictions which either match the data accessed by our senses or, when they do not, are refined or refocused by it (Kuhl, 2000). This process of attention to and statistical sampling of patterns is how a baby learns about its physical, emotional, and social world. It is how we learn language, along with everything else.

It comes as no surprise that language is a series of patterns, whether those patterns are phonological, orthographical, semantic, grammatical, discourse-level, or socio-pragmatic. The codification of these patterns as "Rules of Language" is often mistaken by learners as a prescriptive process (especially in the case of grammar "rules") but, in fact, what the linguistic is trying to do, just like the new-born baby in the previous paragraph, is to *describe* the patterns that make up language in a way that is useful for predicting, understanding, and producing comprehensible patterns in the future. Without patterns, language would not be meaningful, since a shared understanding of the patterns is essential to communication. Thanks to the patterns, a language is both predictable and learnable.

Unlike the baby, searching for and codifying linguistic patterns in its environment, the learner of a second (or foreign) language has access to another learning resource: the experience of others. Teachers, material writers, and course designers attempt to use their greater knowledge of the language being learned to design experiences for the learner which will be helpful in learning the language more efficiently.

The question of how best to help somebody learn a second or foreign language has stretched human ingenuity for centuries (Kelly, 1976). Our latest understanding of cognition, and of learning, should be able to shed some light on the issue.

Like a Baby: Exposure to the Language

In order to generate and refine priors about the language and how it is used, the learner needs to be exposed to the language. The more experience they have of the language in use, the more accurate the priors will become, and the closer their predictions about the language will be to how it is actually used. For reception, this

means they will be able to make accurate predictions about what they will read or hear next. For production, it means that what they say or write will more accurately incorporate norms which will allow their meaning to be understood by others.

One way to gain exposure to the language is through Extensive Reading, reading for enjoyment and relaxation, rather than for study, at a language level that does not greatly tax the reader's ability to understand. Learners who read extensively in the foreign language process large numbers of examples of language use that can be sifted through and refined into priors by their (unconscious) brains. This is in strong contrast to the handful of grammar points and example sentences that might be encountered in a more traditional reading lesson.

Although the practice of Extensive Reading has taken root in language teaching practice over the last 40 years (Bamford & Day, 2004), its equivalent for other skills, Extensive Listening, Oral Fluency Practice, and Written Fluency Practice (sometimes known as journaling) are not as widely used. The Predictive Processing model suggests that they could have an equally important role to play in providing the wide and deep experience of language use that will help establish accurate priors.

Also Like a Baby: Natural Language Use

The Predictive Processing model also suggests that, as far as possible, the language use to which the learner is exposed should be natural. Obscure and archaic language forms and usages may offer a convenient way to test the limits of a learner's knowledge, but they are not helpful in building priors that will guide the learner to useful predictions about natural, contemporary language use. Similarly, the use of simplified texts should be encouraged only to the extent that the language used in them is natural and follows the conventions of texts written for more proficient users.

One of the defining features of natural language use is that it has a context. In making and fine-tuning priors, the learner's brain uses data not only about language form and collocation but also the social and linguistic context in which the language is appropriately used. Decontextualised example sentences and discreet-point vocabulary learning (such as lists of "3,000 words you need to know for the TOEIC," Exam World, 2021) are of limited use. As far as possible, language should be presented and studied within a context in which it would naturally be used.

Nudging the Process Along: Noticing

As we have seen, second/foreign language learners, like babies, can become aware of the patterns of language through extensive exposure. However, the teacher, or materials designer, can help the process along. Approaches to language teaching which emphasise the importance of having learners notice regularities in language form and usage (Schmidt, 1990) are very much in tune with the pattern recognition which is at the heart of the formation of priors.

One way to do this is to design texts and language learning tasks so that language

patterns are more salient than in naturally occurring texts. For example, when teaching about past tense forms, the teacher may select materials in which people recount past experiences in a way that deliberately contrasts the use of the various forms. Without specifically having their attention directed to these forms and their usage, learners gain a more concentrated exposure to these patterns than they would from undoctored extensive reading or listening texts, thus hastening the pattern recognition process.

Another Nudge: Awareness Rising

A further step in the direction of "helping" students to discover the natural patterns of language comes when a teacher or materials designer directs the learner's attention towards a linguistic pattern (Rost, 2020). This can take the form of; a) pointing up the context in which the patten occurs; or b) explicitly encouraging the student to focus on the pattern itself. So, continuing the previous example of a text about past experiences, the learner could be asked a) What was John doing when he was bitten by the jellyfish?" or b) "What tense formed do you see when somebody talks about an activity in the past that was interrupted?" Clearly, example b) offers a much bigger nudge that example a), but both types of awareness raising have their place, and they are often used in combination.

Many modern coursebooks now contain elements of noticing and awareness raising among their learning activities, but one book in particular sets out to teach grammar entirely through these processes. *Impact Grammar* (Ellis & Gaies, 1998) consists of a series of units built around grammatical contrasts (such as the past continuous v. simple past in our jellyfish example). The series editor, Michael Rost (2020) describes the structure of each unit in this way: 1) examples of typical learner errors involving the contrast; 2) a listening text, with comprehension questions, featuring the contrast prominently; 3) a written version of the listening text with blanks to fill in, drawing attention to the contrast; 4) an invitation for learners to formulate "rules" about the contrast; 5) a written text in which listeners are encouraged to find errors in the use of the contrast; 6) a speaking or writing activity giving students the opportunity to use the contrast; and 7) an explicit statement of the grammar contrast, hidden away in an appendix at the back of the book.

Learning by Using

As priors are formed from experience, the experience of using the language should be at the heart of language learning. Since the 1980s and the advent of a Communicative Approach, activities involving student interaction in the language (pair-work, group-work) have been a feature of many classrooms. More recently, Task-Based Language Teaching, in which learners, often working together, attempt to solve non-linguistic tasks by communicating in the target language, has also brought learner interaction in language to the fore.

Predictive Processing suggests that such activities have a place not only as a means

of practicing already-learned or recently-taught language forms, but also as a means of exploring the language, feeling out what patterns of use are meaningful and useful in communicating with others. In fact, the logic of Predictive Processing supports the use of interaction in the language to replace more traditional teacher-fronted presentations of new language. The learners are given a task which involves using a specific language feature, which they may or may not have experienced using previously. However, thanks to exposure to large amounts of language in use, the particular language pattern they need is unlikely to be completely unfamiliar to them. By using, for example the past continuous/simple past contrast in a task that involves sharing stories about childhood accidents ("I was crossing the road, when a car hit me."), the learners gain experience of using the language to communicate, just the kind of experience that shape their priors and hone them for future predictions.

Non-Interactive Language Practice Activities

There appears to be a case for language practice activities that are not interactive or exploratory. The criterion for experience to be translated into a prior appears to be the "importance" of the experience (Clark, 2015): experiences that are important are distilled into priors, those that are not are processed at a lower cognitive level and are often discarded (= forgotten). This explanation, of course, begs the question of how the brain in a box judges whether something is important or not. On an evolutionary level, anything that reduces or increases our chance of survival is important; we tend to experience this importance as an emotion (fear, lust, hunger, pleasure, etc.) and teachers have long known that associating an experience, or even a fact, with a strong emotion is a sure-fire way of having their students remember something. However, we cannot really build a language learning model on repeatedly inducing strong emotion.

On a more prosaic level, our brain in a box appears to work on the assumption that a pattern that is repeated has more importance than one that is not (Clark, 2015). So, one way we can raise the chances that a particular language feature will be incorporated into a learner's priors is to arrange for them to encounter that feature repeatedly. On the most basic level this connects with the language classroom staple "Repeat After Me." Pattern practice, narrow reading, and parallel writing are more sophisticated versions of this. Testing and re-testing of the same language point will achieve the same effect. In fact, learning scientists (Kang, 2016) recommend spaced repetition as one of the most effective of learning techniques: repeating the encounter with the new data with increasing intervals as time goes by.

Explanation/Presentation

What consequences does the model have for a more traditional kind of teaching, where a teacher explains things and the learner takes note? The advantage of an explanation is that it outsources the work of gaining knowledge to another brain in another box. It does away with the need to figure out how to integrate new information

into old patterns by handing the learner a ready-made solution.

In the field of Applied Linguistics, the question of whether conscious learning, as a result of being told something, can be transformed into acquisition (the ability to use a language form automatically, without conscious processing) has been widely and hotly debated and, at times, researched (Long, 1983). The conclusion is not yet clear. Predictive Processing, including the formation of priors, is considered by most scholars to be an almost completely unconscious process (Clark, 2015), with consciousness engaged only if the prediction error is egregiously large (a phenomenon we usually call "surprise").

What can be said of conscious awareness of language patterns, whether resulting from teacher explanation or not, is that it allows the patterns to be used for practice activities, the kind of activities described above, designed to give experience of the language in use, which we know to generate and refine priors. Receiving information is a poor substitute for experience. In the language classroom, the teacher's explanation of new material is a staple, enshrined as the first "P" in the "Presentation, Production, and Practice" approach (SeeTEFL, 2021). An explanation is useful in drawing conscious attention to a language point but, by itself, by no means guarantees incorporation of the point into an unconsciously held prior. For this, the others two "P"s are essential.

Moreover, explanation of a language point tends to be decontextualised, or rather presented in the context of language learning rather than as a real-world experience. It is experiences, rather than lectures, that inform priors. As a result, the explanation of a language point is most effectively done in the context of a specific example occurring in the learner's exposure to the language (a reading text, a search for a way to communicate an idea) rather than as a discrete-point lesson.

Testing

In many school subjects, testing is a way to gauge what and how much knowledge a learner has retained. As such, it is often more about memory than about learning. In foreign language learning, however, where the goal is ostensibly to learn to use the foreign language, accumulating knowledge, as we have seen, is not enough. What we need to measure is how far the learner has developed appropriate priors to use the language in a variety of situations. As a result, we should be aiming to test not what the learner knows about the language but what they can do with it.

A good language test will simulate actual language use by causing the learner to produce predictions and then allowing us to check on the accuracy of those predictions. This does not mean that testing should be limited to questions in which learners are overtly asked to generate predictions (as in "Fill in the blanks: 'fish and _____'" or "What do you think Tom will say when he hears the good news?"), although such questions have their place. Just as all cognition is based on drawing on priors to make predictions, acting on those predictions and then modifying priors if the predictions turned out to be in error, so, too, is all language processing. Traditional language tests, such as reading comprehension questions, translation, writing in response to a prompt,

listening for gist or for specific information, all rely on the learner having accurate priors on which to base their answers, and so are legitimate tests of the accuracy of their priors. In other words, as long as the testing task simulates real-life language use, it is testing learning rather than just knowledge.

A further use of testing should also be apparent: regular, low-stakes, classroom-based tests that prompt learners to use and re-use language patterns (typically vocabulary or pronunciation) in order to increase the importance of these patterns from the brain's perspective. Again, the more the testing situation can be made to simulate real-life language use the more effective such tests will be in generating priors.

Summary

Consideration of foreign language teaching in the light of the Predictive Processing model of cognition leads us to the conclusion that learning comes from experience. The more realistic the experiences that learners have in using the language, the more useful their learning is likely to be to them. This conclusion would lead us to de-emphasise some traditional language-teaching practices, such as explaining the language, forced memorisation, and tests that stimulate only the reproduction of knowledge, in favour of a focus on activities that allow learners to experience the language.

Large-scale exposure to natural usage of the language; activities that promote noticing of recurring language features or overtly guide students to focus on them; contextualized practice with the language; opportunities to encounter new language repeatedly; pair- and group-work; and tests that require learners to show they can use the language rather than just reproduce knowledge, are all part of the language teacher's list of learning activity. The Predictive Processing model suggests that it is on these kinds of activities that we need to focus our students' attention for maximum learning.

References

Bamford, J., & Day, R. R. (2004). *Extensive reading activities for teaching language.* Cabridge, UK: Cambridge University Press.

Campbell, G. (Host). (2016, January 28). Andy Clark on prediction, action, and the embodied mind (No. 126) [Audio podcast]. In *Brain science with Ginger Campbell.* https://play.acast.com/s/brainsciencepodcastneuroscienceforeveryone/bsp-126-andy-clark-on-prediction-action-and-the-embodied-mind

Campbell, G. (Host). (2017, July 31). Lisa Barrett on how emotions are made (No. 135) [Audio podcast]. In *Brain science with Ginger Campbell.* https://brainsciencepodcast.com/bsp/2017/135-emotions-barrett

Clark, A. (2015). *Surfing uncertainty: Prediction, action, and the embodied mind.* Oxford, UK: Oxford University Press.

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain.* New York, NY: Penguin.

Ellis, R., & Gaies, S. (1998). *Impact grammar: Grammar through listening.* London,

UK: Pearson Education.

Exam Word (2021). 3000 TOEIC vocabulary. https://www.examword.com/toeic-list/3000-vocabulary-1

Feldman Barrett, L. (2018) *How emotions are made: The secret life of the brain.* Boston, MA: Mariner Books.

Feldman Barrett, L. (2020, November 20). *The power of prediction: An emerging paradigm for studying brain, body, and mind* [Video]. YouTube. https://www.youtube.com/watch?v=X7jOjLimOlM

Hebb, D. (1949). *The organization of behavior: A neuropsychological theory.* New York, NY: Wiley.

Kang, s. H. K. (2016). Spaced repetition promotes efficient and effective learning: Policy implications for instruction. *Policy Insights from the Behavioral and Brain Sciences, 3*(1), 12-19.

Kelly, L. G. (1976). *25 centuries of language teaching*. Ilford. UK: Newbury House.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America, 97*(20): 11850-11857.

Long, M. (1983). Does second language instruction make a difference? A review of research. *TESOL Quarterly, 17*(3), 359-382.

Rost, M. (2020). Did you notice . . .? The role of noticing in the development of grammar. *MindBrainED Think Tank, 6*(11), 10-21.

Schmidt, R. W. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

SeeTEFL. (2021). PPP TEFL teaching methodology. https://seetefl.com/ppp-tefl-teaching-methodology/

Zimmerman, M. (1986). Neurophysiology of sensory systems. In R. F Schmidt (Ed.), *Fundamentals of sensory systems* (pp. 68-105). Heidelberg, Germany: Springer.