English Language Learning Magazine

CONTACT

August 2016

The least you should know about Mandarin

Cognitive linguistics and language teaching

Reviewing language tests

VOLUME 42, NUMBER 3, AUGUST, 2016
ISSN # 0227-2938
Teachers of English as a Second Language Association of Ontario
www.teslontario.net/publication/contact-magazine

TESL Ontario
CONNECTION • COMMUNITY • EMPOWERMENT

IN THIS ISSUE

Contents

In this issue 2

Editor's Note 3

Contact Magazine 4

Articles

Exploring Rhyme and Reason in Vocabulary and Phraseology 5

Featured Articles: Tests

A Review of the Reading Section of the CELPIP-General Test 10

Telephone oral interview tasks in university admissions language testing 17

Steps to English Proficiency 26

Articles

Language on the orient express: A guide to Mandarin for English teachers 34

Supporting Blended Learning in an EFL Teaching Environment 40

Teaching critical thinking skills in a disciplinary context 46

LINC Field Trips in the Post-Funding World: Five Changes 52

Viva La Lingua Franca

Viva La Lingua Franca: On The Art of the Apology. 56

Calendar

September 30–October 1	BCTEAL: Teaching in Techno-colours: Imagine, Innovate, Inspire
October 5	International World Teachers' Day
November 9–11	Bringing IT together Conference
November 12	People for Education Annual Conference
November 19	Ontario Modern Language Teachers Association Fall Conference
November 20–26	ESL Week
November 24–25	TESL Ontario' 44th Annual Conference: "Recognizing Abilities and Possibilities"
November 16–18	York Region District School Board Conference
November 18–20	American Council on the Teaching of Foreign Languages; ACTFL Annual Convention and World Language Expo

EDITOR'S NOTE

Interest in generative and transformational grammar peaked in the 1970s, while interest in cognitive linguistics (CL) has been growing since the mid 80s. In our lead article, Frank Boers explains how CL can be applied to teaching vocabulary and phraseology.

We have three articles on language tests. Shayla Ahmad reviews the Reading Section of the CELPIP-General Test, an alternative for immigrants wishing to demonstrate English-language ability. Beverly Baker and her colleagues at the University of Ottawa report on their research into telephone oral interview tasks in university admissions. And Elizabeth Jean Larson & Clarissa Lau review the STEP test for elementary and secondary students.

With this issue, we begin a series of articles on the least you should know about various language. We begin with Eric Henry's introduction to Mandarin, one of the most common languages among international students and recent immigrants to Ontario. If there are languages you'd like to learn more about, please, write to suggest an article.

We also include articles about blended learning, critical thinking, and planning field trips, along with our Viva la lingua franca column by Eufemia Fantetti.

As always, we welcome and depend on your contributions. I do hope to hear from you.

Brett Reynolds

editor@teslontario.org

CONTACT

Contact is published three times a year (May, August, and November) by TESL Ontario. May is our conference issue. It is published for the members of TESL Ontario and is available free online to anyone.

Contact welcomes articles of general interest to association members, including announcements, reports, articles, and calls for papers.

Personnel

Editor	Brett Reynolds
EAB members	Hedy McGarrell
	David Wood
	Hanna Cabaj
Webmaster	Kevin O’Brien
Design	Yoko Reynolds

Legal

ISSN # 0227-2938

The statements made and expressed in articles are those of the authors and do not necessarily reflect the policies of TESL Ontario.

Acceptance of advertising does not constitute endorsement by TESL Ontario nor guarantee accuracy of information therein.

Copyright for all articles published in *Contact* rests with the authors, Copyright © 2016. Any reprints require the written permission of TESL Ontario and must clearly state *Contact* as the source along with the original date of publication. For permission, contact: rtilson@teslontario.org.

Cover image

https://commons.wikimedia.org/wiki/File:Dundas_Street_West_at_Huron_Street_Toronto_2010.jpg

TESL ONTARIO

TESL Ontario is a supportive community empowering educational professionals to help English language learners to thrive.

Contact TESL Ontario

TESL Ontario #405 - 27 Carlton St.
Toronto, ON M5B 1L2
Phone: 416-593-4243 or 1-800-327-4827
Fax: 416-593-0164
http://www.teslontario.net

Enquiries regarding membership or change of address should be addressed to the TESL Ontario Membership Coordinator at membership@teslontario.org.

Enquiries regarding advertising rates and reservation of advertising space should be addressed to the Office Coordinator at administration@teslontario.org.

Board

Chair	James Papple
Vice-Chair	Bernice Klassen
Treasurer	Jennifer MacKay
Secretary	Cheryl Fretz
Members-at-large	Sharon Deng
	Bhupinder Gill
	David Hazell
	Geoff Lawrence
	Melita Vrakela

Executive director

Renate Tilson

Affiliates

TESL Durham, TESL Hamilton-Wentworth, TESL Kingston, TESL London, TESL Niagara, TESL North York/York Region, TESL Northern Region, TESL Ottawa, TESL Peel/Halton/Etobicoke, TESL Toronto, TESL Waterloo-Wellington, TESL Windsor

EXPLORING RHYME AND REASON IN VOCABULARY AND PHRASEOLOGY

By Frank Boers, Victoria University of Wellington, New Zealand

No-one will dispute that language learning is to a large extent a matter of mastering myriads of words and phrases and that it therefore relies heavily on memory. Many pedagogy-minded applied linguists concur that a word or phrase is more likely to be remembered if the learner consciously "engages" with it in one way or another (Schmitt, 2008). This, then, raises the question of what kinds of cognitive engagement with lexical items are relatively fruitful, and how teachers (or materials writers) can prompt students to give these a try. This article considers a handful of proposals for stimulating engagement with words and phrases that are in broad agreement with a school of thought known as *Cognitive Linguistics*.

Cognitive Linguistics (CL) emerged in the 1980s as an alternative to the then dominant Chomskyan-style descriptions of language, and has been attracting a growing number of followers. Space constraints prevent me from doing justice to CL as a school of thought here (but see, e.g., Littlemore & Taylor, 2014, for extensive coverage). The essence of CL thinking is that language is an integral part of human cognition rather than a separate module in the mind. Accordingly, language should be described with reference to general cognitive phenomena which operate outside language as well (Lakoff, 1987; Langacker, 1987). Another corollary is that, as a manifestation of general cognitive phenomena, perhaps language need not be considered as a system of arbitrary signs[1], a system without rhyme or reason for the way concepts and messages are packaged. Instead of being totally arbitrary, several facets of language use may be "motivated" (Radden & Panther, 2004). In what follows we explore this tenet that linguistic phenomena may be motivated instead of arbitrary, and how this can serve as a channel for second language learners' engagement with (aspects of) vocabulary and phraseology.

Polysemy and figurative language

Many of the early investigations into the merits of applying CL to language teaching and learning concern polysemy. Polysemy refers to words or expressions which have more than one meaning. High-frequency words, such as prepositions (e.g., *over* and *under*), are particularly prone to polysemy (see Lindstromberg, 2010, for a reader-friendly treatment

1 The idea of an arbitrary relationship between signs and meaning was put forward by Ferdinand de Saussure over 100 years ago and underpins much linguistic theory.

of English prepositions). However, as a quick glance at a handful of dictionary entries will reveal, polysemy is the norm rather than the exception in language. Very often, words have a basic, literal meaning (e.g., *hurdles* referring to the obstacles to be jumped over in a running race) and also figurative meanings derived from it (e.g., *hurdles* referring to problems or difficulties to be overcome). One can wield a weapon but one can also wield power. One can tackle a player in sports such as soccer and one can tackle problems. Plants may flourish but so can businesses. It is not difficult to see how the more abstract, figurative uses of these words have extended from their literal uses through metaphor. It is not unusual, however, for a learner (perhaps especially in ESP/EAP contexts) to encounter a word used in a more abstract sense first. Of interest for teaching practice is the finding that making learners aware of the literal meaning of those words makes them more memorable (e.g., Boers, 2000a; Verspoor & Lowie, 2003). This is not so surprising, since it is well known that concreteness of meaning facilitates retention, probably thanks to the mental imagery that comes with it (cf. Dual Coding Theory, e.g., Paivio, 1986).

This also applies to figurative idioms. Although the idiomatic meaning of *being on the ropes* is abstract, re-connecting it to the context in which the expression is used in a literal sense (i.e., boxing) makes it more memorable. Recognition that *taking a back seat* literally means taking the role of passenger in a car may help a learner infer the figurative meaning of the expression—allowing others to take control and responsibility—that is associated with not being in the driver's seat. Giving attention to idioms may seem trivial, but studies have shown that even advanced second language learners often fail to comprehend idioms despite the presence of contextual clues (Boers, Eyckmans, & Stengers, 2007), and that this can cause serious communication problems in ESL contexts (Littlemore, Chen, Koester, & Barden, 2011). This is because, far from being the icing on the cake, idioms fulfill important pragmatic functions in discourse (O'Keeffe, McCarthy and Carter, 2007, pp. 80–99).

More impetus for endeavours to make words and phrases more memorable through the use of mental imagery has come from Conceptual Metaphor Theory, a strand of CL which began with Lakoff and Johnson's seminal book *Metaphors we live by* (1980). According to this theory, metaphor is not just a trope, confined to literary genres where it serves the purpose of embellishment. Instead, it is a fundamental cognitive ability which helps us to talk and think about intangible, abstract domains of experience. As a result, everyday language abounds with figurative expressions, although we are seldom aware of their figurative nature. For example, expressions such as *a chronic budget deficit, a financial injection, economic recovery*, and *a healthy economy* suggest that one of the ways in which people have come to understand issues in economics is by seeking analogies with human health and illness. One and the same 'metaphor theme' will be manifested in various expressions, and this provides a way of organizing lexis that may facilitate learning (e.g., Boers, 2000b). Expressions such as *being hot under the collar, being hot-tempered, losing one's cool, adding fuel to the fire* and *blowing off steam* can all be 'motivated' as instantiations of a general metaphor which likens anger to heat. The existence of this general metaphor itself is not arbitrary either; it is grounded in physical experience (a rising body temperature

is symptomatic of agitation). In a similar vein, *setting the stage, waiting in the wings, taking centre stage, playing to the gallery, behind the scenes* and *the curtain is down* all have their origin in the theatre, and the inclusion of these expressions in the English idiom repertoire is not surprising given the cultural significance of this domain. Raising learners' awareness of metaphors may benefit their engagement with phrasal and prepositional verbs as well. For instance, the use of *out* in phrases such as *find out* and *figure out* may reflect the general metaphor 'knowing is seeing' (also manifested in a expressions such as *I see what you mean* and *being in the dark about something*); if something is taken out of a container it becomes visible and thus "knowable".

Words of a feather flock together

The majority of the empirical investigations into the merits of applying CL to second language vocabulary and phraseology teaching have been concerned with the aforementioned motivated connections between the literal meanings of words and expressions and their extended, figurative uses (Boers, 2013, for a review of these investigations). However, interest in other types of motivation, including motivation of form is picking up. One recent trend is to explore potential reasons for the precise lexical makeup of standardised phrases where substitutions with synonyms would convey the same message but would sound unnatural. *Time will tell* sounds right whereas **time will say* does not. Statistical evidence has emerged that phonological (and orthographic) similarity very often plays a part, such that a large proportion of the English phrasal repertoire displays catchy sound patterns, especially alliteration (e.g., *bunk bed, fast food, slippery slope, peer pressure*) but also (near-) rhyme (e.g., *hot spot, high five, small talk, brain drain*; Boers & Lindstromberg, 2009: pp. 106–125). Some types of phrases, most notably binomial phrases (e.g., *part and parcel, wear and tear, spick and span*) and similes (e.g., *good as gold, right as rain, thick as thieves*) are particularly prone to this 'words-of-a-feather-flock-together' phenomenon (33% and 54%, respectively). Of interest for teaching practice is the finding that simply drawing learners' attention to the presence of alliteration or (near-) rhyme in the lexical phrases they encounter strongly enhances the mnemonic effect of these sound patterns (e.g., Boers, Eyckmans, & Lindstromberg, 2014; Eyckmans, Boers, & Lindstromberg, 2016). This is good return for minimal investment.

Iconicity

Iconicity refers to instances where the form of a word or expression reflects the characteristics of the concept being denoted. Iconicity is the most obvious in onomatopoeia (as in *hiss* and *splash*), but it extends beyond that. Evidence for this was furnished by experiments where participants were presented, for example, with the pseudo-words *bouba* and *kiki* and were asked to match these with drawings of shapes (Ramachandran & Hubbard, 2001). Regardless of language background, participants tended to match *bouba* with round shapes and *kiki* with jagged shapes. While "sound symbolism" is far from abundant in language, where it does occur, it can facilitate learning (e.g., Kantartzis, Kita, & Imai, 2011). In fact, simply asking learners to evaluate whether the form of a word somehow matches

its meaning fosters retention of both form and meaning of the word (Deconinck, Boers, & Eyckmans, 2010).

General implication and further reading

Inquisitive students will occasionally ask their teachers the "*why* question": "why do we say it like that in English?" The answer to this may often be, "That's is just the way it is." However, once we recognize that not everything in language is arbitrary, and that, instead, plausible explanations may be available, a more attractive option is to embrace (instead of dread) the *why* question, as an opportunity for engagement. The proposal to make the best of such opportunities when they present themselves also aligns with practices informed by socio-cultural theory (Lantolf, 2011).

Of necessity, we have in this short article considered just one facet of cognitive linguistics and have confined our examples of its pedagogic applicability to the realm of vocabulary. More comprehensive accounts of applied cognitive linguistics include Littlemore (2009) and Tyler (2012).

References

Boers, F. (2000a). Enhancing metaphoric awareness in specialised reading. *English for Specific Purposes, 19*, 137–147.

Boers, F. (2000b). Metaphor awareness and vocabulary retention. *Applied Linguistics, 21*, 553–571.

Boers, F. (2013). Cognitive Linguistic approaches to second language vocabulary: Assessment and integration. *Language Teaching: Surveys and Studies, 46*, 208–224.

Boers, F. & Lindstromberg, S. (2009). *Optimizing a Lexical Approach to Instructed Second Language Acquisition*. Basingstoke: Palgrave Macmillan.

Boers, F., Eyckmans, J., & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research*, *11*, 43–62.

Boers, F., Lindstromberg, S., & Eyckmans, J. (2014). Is alliteration mnemonic without awareness-raising? *Language Awareness, 23*, 291–303.

Deconinck, J., Boers, F., & Eyckmans, J. (2010). Helping learners engage with L2 words: the form-meaning fit. *AILA Review, 23,* 95–114.

Eyckmans, J., Boers, F., & Lindstromberg, S. (2016). The impact of imposing processing strategies on L2 learners' deliberate study of lexical phrases. *System, 56*, 127–139.

Kantartzis, K., Kita, S., & Imai, M. (2011). Japanese sound symbolism facilitates word learning in English speaking children. *Cognitive Science*, *35*, 626–630.

Lakoff, G. (1987). *Women, fire and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar, Volume* 1*: Theoretical prerequisites*. Stanford: Stanford University Press.

Lantolf, J. P. (2011). Integrating Sociocultural Theory and Cognitive Linguistics in the second language classroom. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning, Vol. 2* (pp. 303–318). New York: Routledge.

Lindstromberg, S. (2010). *English prepositions explained* (revised ed.). Amsterdam: John Benjamins.

Littlemore, J. (2009). *Applying Cognitive Linguistics to second language learning and teaching*. Basingstoke, UK: Palgrave Macmillan.

Littlemore J. & Taylor, J. (Eds.). (2014). *The Bloomsbury companion to Cognitive Linguistics*. London: Bloomsbury.

Littlemore, J., Chen, P. T., Koester, A., & Barnden, J. (2011). Difficulties in metaphor comprehension faced by international students whose first language is not English. *Applied Linguistics, 32*, 408–429.

O'Keeffe, A. M., McCarthy, M., & Carter, R. (2007). *From Corpus to Classroom: Language Use and Language Teaching*. Cambridge: Cambridge University Press.

Paivio, A. (1986). *Mental representations: A Dual Coding Approach*. New York: Oxford University Press.

Radden, G. & Panther, K-U. (Eds.). (2004). *Studies in linguistic motivation*. Berlin: Mouton de Gruyter.

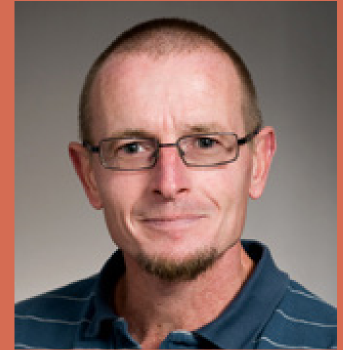
Ramachandran, V.S. & Hubbard, E.M. (2001). Synaesthesia: A window in to perception, thought and language. *Journal of Consciousness Studies, 8*, 3–34.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research, 12*, 329–363.

Tyler, A. (2012). *Cognitive Linguistics and Second Language Learning: Theoretical Basics and Experimental Evidence*. New York: Routledge

Verspoor, M. & Lowie, W. (2003). Making sense of polysemous words. *Language Learning, 53*, 547–586.

Author Bio



Frank Boers' initial research interests were in the fields of lexicology and semantics. Most of his more recent research interests, however, were sparked by his long experience as a language teacher and teacher trainer. He now publishes and teaches mostly on matters of second or foreign language teaching, often with a focus on phraseology. Frank is co-editor of the journal Language Teaching Research.

A REVIEW OF THE READING SECTION OF THE CELPIP-GENERAL TEST

By Shayla Ahmad, George Brown & Humber Colleges

The Canadian English Language Proficiency Index Program-General (CELPIP-G) Test aims to assess the general English language functional proficiency of individuals for adapting to life in Canada. The distinctiveness of the CELPIP-G Test, in contrast to other leading proficiency tests in the industry, arises from its design and use for Canadian immigration purposes. Apart from IELTS (International English Language Testing System), it is the only test that is accepted by the Department of Citizenship and Immigration Canada (CIC) for permanent resident and citizenship purposes ("Citizenship and Immigration Canada", n.d.). In addition, the test uses the English variety spoken in Canada. Accordingly, it exposes individuals to Canadian English rather than other varieties of English used in other proficiency tests. Using Bachman and Palmer's qualities of tests (1996) as a framework, I will critically examine the purpose, strengths, and limitations of the reading section of the CELPIP-G Test.

Background and Users

Paragon Testing Enterprises, a subsidiary of the University of British Columbia (UBC), launched the CELPIP-G test in response to the success of its placement test known as the Language Proficiency Index (LPI; L. Barrows, personal communication, October 29, 2015). The LPI measured beginning university and college level English reading and writing skills only. Therefore, there was an increasing demand to assess all four language skills of reading, writing, speaking and listening for general and academic purposes. In order to meet this demand, the CELPIP-G Test was developed in 2002 (L. Burrows, personal communication, October 29, 2015). In the same year, the test was recognized by CIC as a proof of general language proficiency for the Federal Skilled Workers Class; in 2012, it was accepted for the streams of the Canadian Experience Class of Immigration, and in 2013, it was established for all current federal economic immigration programs (Federal Skilled Worker Program, CEC, and Federal Skilled Trades Program), the Provincial Nominee Program, and the Business Program (L.Barrows, personal communication, October 29, 2015). In addition to the CELPIP-G Test, an academic English language proficiency test, CELPIP-Academic, was introduced in 2005 to fulfill language requirements for admission purposes in universities giving it parity with proficiency tests like IELTS, TOFEL (Test of English as a Foreign Language), and MELAB (Michigan English Language Assessment Battery). However, the CELPIP-Academic was retired in August, 2015 (L. Barrows,

personal communication, October 29, 2015). As a result of limitations identified in the validity and reliability research, a revised version of the CELPIP- G was introduced in April 2014, which presented changes to the structure and scoring template, time allotment to different sections, and the number of items in each skill appropriate to everyday situations (McKenzie, 2014).

Although the test was originally developed as a general and academic proficiency test, it is now generally used for the purpose of immigration in Canada. In addition to Citizenship and Immigration Canada, the test results are used by Canadian immigration associations and consultants for directing applicants for immigration to Canada and also by professional associations and employers who need evidence of English language proficiency of their members and prospective employees ("CELPIP General", n.d.)

In the previous version of CELPIP-G test, the scores of an individual needed to be transformed to the Canadian Language Benchmark (CLB) criterion. This proved to be problematic as the CLB had a scale of 1–9+ whereas the CELPIP-G was scored on a scale of 1–5 (high-low). The revised version CELPIP-G is now aligned with the CLB and the scores are given on a scale of 1-9+ (McKenzie, 2014). The following table shows each CELPIP-G level and its corresponding description.

Table 1

CELPIP-General Levels and Corresponding CLB Levels

CELPIP Level	CELPIP Descriptor	CLB Level
12	Advanced proficiency in workplace and community contexts	12
11	Advanced proficiency in workplace and community contexts	11
10	Highly effective proficiency in workplace and community contexts	10
9	Effective proficiency in workplace and community contexts	9
8	Good proficiency in workplace and community contexts	8
7	Adequate proficiency in workplace and community contexts	7
6	Developing proficiency in workplace and community contexts	6
5	Acquiring proficiency in workplace and community contexts	5
4	Adequate proficiency for daily life activities	4
3	Some proficiency in limited contexts	3
M	Minimal proficiency or insufficient information to assess	0, 1, 2

Note. From Paragon Testing Enterprises, n.d.

Purpose

The purpose of CELPIP-G is to assess an individual's general language ability or functional competency to communicate in everyday and workplace situations. The test takers include both English-first and English as second language speakers, and the test provides a score based on the CLB 2000 (Canadian Language Benchmarks). Aligned with its purpose, the test contains "Canadian English Language ... that accurately assesses listening, reading, writing, and speaking skills in typical everyday situations" ("*CELPIP- G Brochure",* n.d.). The construct of functional reading proficiency in English, the focus of this review, measures the test taker's ability to "engage with, understand, interpret, and make use of written English texts to achieve day-to-day and general workplace communicative functions" (Wu & Stone, 2015, p. 5).

Test Format and Scoring of the Reading Section

Offering plausible benefits of being computer-delivered, the CELPIP-G Test is completed in one three hour sitting. The reading section, which is one hour long, is devised in a multiple choice item format for the feasibility of quick marking, low expense, and objectivity (Weir, 1990, 1993). Individuals get one point for correct answers and a zero for incorrect answers. All test items are checked automatically by the computer with a pre-specified answer key. The reading tasks of the test are outlined to elicit responses reflecting the understanding of regular day-to-day and work-related written English, and the sections become gradually harder with the progression of the test (Wu & Stone, 2015). The multiple choice item formats of all sections of the reading test, apart from section three, have four distracters for one key. Section three has five distracters for one key. The first section, Reading Correspondence, is based on social or work related email communication and entails an understanding of the underlying "sociolinguistic and pragmatic" features of written communication along with vocabulary, sentence structure, and grammar. Analyzing graphical information (presented in an email) is the focus of Reading to Apply a Diagram section. The third section, Reading for Information, presents factual information for interpretation. Finally, in the form of an "editorial genre", the fourth section, Reading for Viewpoints, requires test takers to understand critical and abstract ideas (Wu & Stone, 2015). In addition, in order to ensure item quality, which is a standard procedure for tests of this nature, some unscored items are included in the reading test. These occur in any four sections of the reading test and are indistinguishable from the other items ("*CELPIP- General Test format and scoring"*, n.d.).

Validity and Reliability

Validity and reliability are the two most critical traits of language assessment, especially for large-scale tests that involve important consequences for a wide range of people. (Bachman & Palmer, 1996). Validity is established through "a long-term process of accumulating,

evaluating, interpreting, conveying, and refining evidence from multiple sources about diverse aspects of the uses of an assessment to fulfill particular purposes to its guiding construct/(s)" (Cumming, 2013, p. 1), and reliability is "considered to be a function of the consistency of scores from one set of tests and test tasks to another" (Bachman & Palmer, 1996, p.19). Research by Paragon Testing Enterprise provides substantial evidence on the reliability and validity of the CELPIP-G Test (Chen, 2014; Wu, 2011; Wu, et al., 2012). Focusing on aspects such as measurement error, gender bias, convergent and criterion validity, construct validity, and dimensionality, these studies concluded that despite some irregularities, the CELPIP-G test meets the standards of the industry and is analogous to other leading proficiency tests. Other studies (Wu & Stone, 2015; Wu, Stone, & Lui, 2015), on the other hand, raised questions about some aspects of the validity and reliability. The research conducted by Wu et al. (2012) looked at the suitability and effect of categorizing test takers into the different proficiency levels which determine their eligibility for immigration. The proficiency levels are established by three cut scores or thresholds that classify test takers into four proficiency levels: no proficiency, basic proficiency, moderate proficiency, and high proficiency. The study found the reading section to be less consistent and less accurate at the high proficiency level. This suggested the developers add items to the reading section at the third cut point to distinguish between high proficient and moderate proficient test takers. The revised scoring system of CELPIP –G accounted for this problem. A study carried out by Wu and Stone (2015) that focused on the reading test taking strategies of individuals, also provides evidence for validity. They found that the test takers gave priority to comprehension skills in analyzing tests rather than "test-wiseness", a factor that would undermine construct validity.

As presented on its website, Paragon Testing Enterprise asserts that its mission is "to develop and produce innovative, valid, fair, accurate, and high quality assessments through collaboration with the unit's members and with its many stakeholders". ("About research at Paragon", n.d.). Moreover, it "is dedicated to sharing the research on its English language testing program with users of its tests". ("About research at Paragon", n.d.). Accordingly it provides information regarding the items, users and test takers, sample tests, procedure, and outlines appropriate uses and warnings about misuses ("Paragon Testing Enterprise", n.d.).

Limitations

An important feature of language tests mentioned by Bachman and Palmer (1996) is test fairness. "Fairness implies that every test taker has the opportunity to prepare for the test and is informed about the general nature and content of the test, as appropriate to the purpose of the test." (Code of Fair Testing Practices in Education, 2004). The information regarding the Canadian English spoken variety is limited; it does not give test takers a clear description of the "Canadian variety" included in the test, which brings construct validity into question. The official CELPIP test website mentions "Canadian English is the English

variety spoken in Canada. It contains elements of British English and American English in its vocabulary, as well as many distinctive Canadianisms...both British and American English spellings are accepted on the CELPIP Test." (What is Canadian English?, n.d.).

The link provided by Paragon Testing Enterprise to a Wikipedia source is mostly theoretical in nature, and although it gives some examples of Canadian English Language norms, it poses difficulty for test takers to relate the information to test preparation. The available sample test is also not enough to gain a definite insight of the nature of Canadian variety of English. Therefore, the purpose of the test becomes a concerning factor as to what is being assessed: general reading comprehension skills, familiarity with Canadian English, or both?

The second drawback of the test is its use of multiple choice items as the only form of assessment in the reading section which raises issues of content validity. Although multiple choice items have the advantage of convenience and high reliability, it is "an unrealistic task, as in real life one is rarely presented with several alternatives from which to make a choice to signal understanding" (Weir, 1990, p. 44). In many cases it might lead to guessing, or the test taker can be trained to choose the correct answer. It also tends to place more emphasis on a test takers' familiarity with the linguistic system (Shin, 2012). It is, therefore, advisable to "use more than one test method for testing any ability" (Alderson, Clapham, & Wall, 1995, p. 45) because "by using a variety of different task types, the test is far more likely to provide a balanced assessment" (Buck, 2001, p. 153). Given the purpose of the CELPIP-G Test, assessing general English functional competency, items "which appear integrative, authentic, communicative, and pragmatic from real world language use situations" (Shin, 2012, p. 239) can be included.

Finally, an important issue that emerges from the discussion is the washback effects of the test. As the score of the CELPIP-G is often used in high-stakes situations, such as immigration, it exerts considerable influence on how test-takers prepare and take the test, the feedback they receive, and the decisions made on the basis of the test (Bachman & Palmer, 1996). So far, one study (Wu, et al., 2012) outlined unintended consequences in terms classification consistency and accuracy, but other areas of washback effect have not been investigated. However, one potential direct effect of the test is the test taker's investment of time and money to buy materials or enroll into programs to get a clear idea of the "Canadiniasms" the CELPIP-General aims to represent.

Conclusion

To conclude, despite the doubts expressed regarding general construct specification and item choice in the reading section, the CELPIP-G test is important, especially in Canada for assessing English language skills for permanent resident and immigration purposes. In comparison to IELTS, the alternative test accepted by CIC, CELPIP-G can be considered more feasible as it is cheaper. The computer-based delivery makes it less time consuming,

and the results are also available in a shorter time. International students and non-resident workers living in Canada who intend to become permanent residents can take the CELPIP-G Test to provide proof of English language ability. However, issues related to washback effects and fairness need to be revisited to account for test validity and reliability as the use of CELPIP-G Test expands in Canada.

References

About research at Paragon. (n.d.). Paragon Testing Enterprise. Retrieved from https://www.paragontesting.ca/research/

Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.

Bachman, F.L., & Palmer, A.S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Buck, G. (2001). *Assessing Listening*. Cambridge, UK: Cambridge University Press.

CELPIP- G Brochure. (n.d.). Paragon Testing Enterprise. Retrieved from https://www.celpiptest.ca/wp-content/uploads/2015/09/CELPIP-G-Brochure-WEB.pdf

CELPIP- General Test format and scoring. (n.d.). Paragon Testing Enterprise.Retrieved from https://www.celpiptest.ca/about-celpip-g/test-format-and-scoring/

CELPIP General. (n.d.). Paragon Testing Enterprise. Retrieved from https://www.paragontesting.ca/english-language-tests/celpip-test/

Citizenship and Immigration Canada.(n.d.). Retrieved from http://www.cic.gc.ca/english/informationapplications/guides/CIT0002ETOC.asp#CIT0002E4

Code of fair testing practices in education. (2004). Washington, DC: Joint Committee on Testing Practices, American Psychological Association. Retrieved from
http://www.apa.org/science/programs/testing//fair-code.aspx.

Cumming, A. (2013). Validation of language assessments. In C. Chapelle (Ed.)*The Encyclopedia of Applied Linguistics*. Malden, MA: Wiley-Blackwell. DOI:10.1002/9781405198431.wbeal1242

Mckenzie, T. (2014, April 15). Recent changes to the CELPIP test. Retrieved from https://*www.immigrationnation.ca/blog.aspx?entry=1892*

Paragon Testing Enterprise. (n.d.).Retrieved from https://www.celpiptest.ca/about-celpip-g/

Shin, D. (2012). Language assessment for immigration and citizenship. In G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing*. (pp.237-248). New York, USA: Routledge.

Weir, C. J. (1990). Communicative language testing. New York: Prentice Hall.

Weir, C. J. (1993). Understanding and developing language tests. New York: Prentice Hall.

What is Canadian English? (n.d.). Paragon Testing Enterprise. Retrieved fromhttps://www.celpiptest.ca/faq/

Wu, A. D. (2011).Report on the evaluation of the CELPIP-G Test scores. Paragon Testing Enterprises. Retrieved fromhttps://www.paragontesting.ca/wp-content/uploads/2012/06/Report-on-Evaluation-of-CELPIP-G-Test-Scores.pdf

Wu, A. D., & Stone, J. E. (2015). Validation through understanding test-taking strategies: An illustration with the CELPIP-General reading pilot test using structural equation modeling. *Journal of Psychoeducational Assessment*. Advance online publication.DOI: 10.1177/0734282915608575

Wu, A. D., Stone, J. E., & Lui, Y. (2015). Developing a validity argument through abductive reasoning with an empirical demonstration of the latent class analysis. *International Journal of Testing*. Advance online publication DOI:10.1080/15305058.2015.1057826

Wu, A. D., Wehrung, D., & Zumbo, B. D. (2012).The validation of the CELPIP-G Test for Canadian immigrants: Classification consistency and accuracy. Paper presented at the 2012 Annual meeting of the American Educational Research Association, Vancouver, Canada. Retrieved from https://www.paragontesting.ca/wp-content/uploads/2012/06/The-Validation-of-the-CELPIP-G-Test-for-Canadian-Immigration.pdf

Author Bio

Shayla Ahmad holds an M.A. in Applied Linguistics. She teaches EAP at George Brown College and English Communication Skills at Humber College. Her professional interests include assessment and second language acquisition. shayla.ahmad@hotmail.com

TELEPHONE ORAL INTERVIEW TASKS IN UNIVERSITY ADMISSIONS LANGUAGE TESTING

Interactive functions and reports of anxiety

By Beverly Baker, Joselyn Brooksbank, Irina Goundareva, Valerie Kolesova, Jessica McGregor, & Mélissa Pésant, University of Ottawa

We report here on research we undertook to examine the entrance tests for applicants to second-language teaching programs in either English or French. In particular, we were interested in examining the speaking portion of this test, a one-on-one telephone interview. Individual one-on-one interviews are still the most common method of assessing speaking for high stakes contexts such as these (Luoma, 2004), and they are often conducted by telephone for cost and time savings.

This oral interview task follows a standard format, consisting of a warm-up, two separate tasks (a role play with the interviewer and a single long turn discussing an opinion on an issue), and a wind-down. The complete interview lasts approximately 10 to 15 minutes. As with many current speaking tests (see Luoma, 2004), this oral interview was designed to capture information not just on language accuracy but on the ability to engage in a variety of interactional language functions, like being able to interrupt appropriately and politely, introduce topics, and manage turn-taking. However, sometimes oral tests do not succeed in doing this, instead being much less interactive: simply a series of answers to interviewer questions. So, we were interested in examining whether a variety of language functions were actually being elicited. If they were, then the test is more useful in making decisions about these future teachers' interactive abilities in the language that they hope to be teaching.

In addition, we were also interested in how the telephone format affected the level of anxiety of the test candidates. Language test anxiety is a well-known phenomenon (see Horwitz; 1996; Horwitz, Horwitz, & Cope, 1986) and anxiety has been shown to affect both the quantity and quality of linguistic output (e.g., Toth, 2012). There is some evidence that test candidates may have a preference for a face-to-face format, as it is more like authentic interaction (Qian, 2009), but there is little work on test-taker reports of anxiety with this particular test delivery method.

Therefore, our research questions were the following:

1. What is the nature of the interactive functions elicited by the telephone tasks?
2. How has the telephone format of the test affected feelings of anxiety in the test candidates (the applicants to the teacher training program)?

Method

In our study, we were informed by Weir's sociocultural model for test validation (Weir, 2005; see Taylor, 2011, for a slightly adapted version of the framework for the conceptualisation of speaking validity specifically). This framework allows for the collection of evidence to support the claims made for the use of this particular test.

Recorded Telephone Interviews

Audio recordings of the admissions telephone interview were obtained from the language testing administrators following university ethical approvals. The full interview database consisted of three versions of both the role-play and the opinion tasks. While further details cannot be provided for reasons of test security, the role play that was selected was a mock job interview, and the opinion task selected was the candidates' position on an educational topic.

A sample of 32 recordings was then chosen for qualitative analysis: 16 each for the role-play and the opinion tasks. Four recordings were chosen of English speakers being tested in English (their native language), four of French speakers being tested in English, four of French speakers being tested in French, and four of English speakers being tested in French. To control for the language proficiency of test takers in our analysis, the 16 task recordings that were chosen from among those receiving a grade of 4/5 (a clear pass) on the rating scale developed for the assessment. It was decided that recordings receiving a clear pass would be most suitable as the object of analysis because a variety of interactional functions is more likely to be elicited by successful candidates.

After the final 32 recordings were selected, each was analysed with an adapted version of the observational checklist used to validate speaking assessment tasks developed by O'Sullivan, Weir, and Saville (2002). Since that time, a variety of forms of this observation checklist have been used in test validation applications, including investigations of test form comparability (Weir & Wu, 2006) and for the collection of *a priori* validity evidence in speaking test development (Nakatsuhara, 2014).

This practical checklist was developed specifically for use while listening to oral responses, without the need for transcription. The checklist was adapted for our purposes in a three-step process. First, one researcher initially listened to all recordings and noted which functions were used or could conceivably be applied. This formed the initial list. Secondly, the test specifications documents were consulted and all interactive functions that were

cited as the focus of the tasks in this document were included. This required adding some functions to the original list (such as *changing topic*), but it was important to look specifically for the functions referenced in the test development documents. Finally, all researchers engaged in a group piloting session with the tool, listening to full interviews, to determine whether all utterances in the interviews could be assigned a function. In this manner, we were able to combine some categories and come up with as parsimonious a list as possible for practical and efficient use. The functions included in this adapted checklist can be found in Table 1 below.

Questionnaire

To respond to our second research question we made use of an adapted version of the Foreign Language Classroom Anxiety Scale (FLCAS; Horwitz, Horwitz, & Cope, 1986). The FLCAS uses a Likert scale to assess the agreement participants report on statements regarding their anxiety related to interacting in a foreign language instructional setting. This scale has proven to be valid for making useful inferences of anxiety in language students. The team examined all items and chose those which related to language assessment. We then translated the questionnaire from English to French in order to allow participants to complete the questionnaire in their choice of language.

The first three questions requested details regarding participants' language preferences. This allowed us to gain insight into perceived comfort levels while speaking English or French, or both. Participants were left space where they could explain in more detail which language they felt more comfortable speaking. For the purposes of this study, we did not include demographic information about other languages, only reported comfort in the languages of the test—English or French.

Questions four and five were designed to activate the participants' memories of the test event. Participants were asked to think back to when they completed the test, the topics they discussed, and the friendliness of the interviewer. The participants then responded to the anxiety-related items, and opinions about how the telephone test delivery format affected these anxiety levels. As a final open-ended question, we asked participants to provide us with any other information that they felt they would like to share with us regarding their telephone interview experience. The full questionnaire is included in Appendix A.

In April, 2015, upon permission from the instructors, we visited two first-year language teaching methodology classes, one in French and one in English, explained the purpose of our study, and requested that students complete the questionnaire. Thirty-one students voluntarily completed the questionnaire in the two classes.

Informational Functions	*personal information; expressing opinion; justifying opinion; summarizing; suggesting; comparing; speculating; describing; suggesting; expressing preferences*
Interactional Functions	*Agreeing; disagreeing; qualifying; asking for information; persuading; repairing breakdowns; negotiating meaning*
Managing Interaction	*Initiating; changing topic; terminating*

Table 1: Interactional functions for analysis (adapted from O'Sullivan, Weir, & Saville, 2002).

Results and Discussion

We found that the two tasks (the role play and an opinion long turn) differed in the interactional functions they elicited. As might be expected, *informational functions* dominated the opinion task. The opinion task was worded in such a way that candidates were able to engage in *comparing* (20 instances), *expressing opinion* (83 instances), *justifying opinion* (61 instances), and *speculating* (44 instances). These functions were either very rare or non-existent in the role-play task (where, for example, *justifying opinion* occurred only 6 times). On the other hand, only the role-play task elicited *interaction management functions* such as *initiating* and *terminating* interactions. Overall, the role-play task elicited a greater variety of functions, with at least a few instances of *informational, interactional,* and *managing interaction* functions. Therefore, the use of two tasks in the assessment is beneficial as it increases the coverage of language functions produced by the applicants.

An important finding also is the number of interactive functions that are NOT performed: many functions are elicited rarely or not at all by either of the tasks, including *agreeing* and *disagreeing*, *changing topic*, and *suggesting*. If we judge the quality of the task by the variety of interactive functions in which the candidate is able to engage, then there is improvement to be made. The interviewer might be able to encourage more use of the above interactive functions by mentioning them explicitly (e.g., "persuade me;" "what parts do you agree and disagree with?").

There is a challenge in eliciting a variety of functions due to the unequal power relationship between the test taker and interviewer: the interaction is managed primarily by the interviewer. This power differential, sometimes called *interviewer dominance*, has been well established for decades (see Lazaraton, 1992; Young & Milanovic, 1992). However, Kormos (1999) suggests that while non-scripted interviews are characterised by dominance of the interviewer, these roles can be reversed in role-play tasks where there can be interaction that is more "symmetrical". So there is potential for the role-play task to be designed to overcome challenges in the power differential as well as the variety of functions elicited. For example, it would be very interesting to explore the use of a role-play task where the test taker has the role with the higher position of relative power. If the test taker

were the person *conducting* a job interview, for example, then more of the responsibility for interaction management would fall on his or her shoulders. This loss of control for the interviewer means greater authenticity as well as more even interaction, but represents other challenges for both reliability and validity. How, for example, can similar test lengths for all candidates be ensured, if the candidates decide when to wrap up the interaction? The challenge is in creating tasks that allow for the power sharing that is indicative of the target domain, while providing reasonably equivalent testing conditions for all candidates.

Regarding our second research question related to anxiety, we found that the telephone format did not appear to have created more anxiety amongst test takers. The average response to Statement 6d, "I wish that the interview had been face-to face instead of over the telephone," was 2.5 (SD 1.4). Note that 5 equals strong agreement, 3 is neutral, and 1 equals strong disagreement with the statement, so while there is some divided opinion there is a slight preference to the telephone interview over the face-to-face form. In addition, most respondents were in agreement with Statement 6g, "It was less stressful to do the interview over the phone," for both tests and all the groups of respondents (M=3.48; SD 1.02). There was no significant difference between English and French test takers on either of these questions. Figures 1 and 2, below, summarise the individual responses to this question.

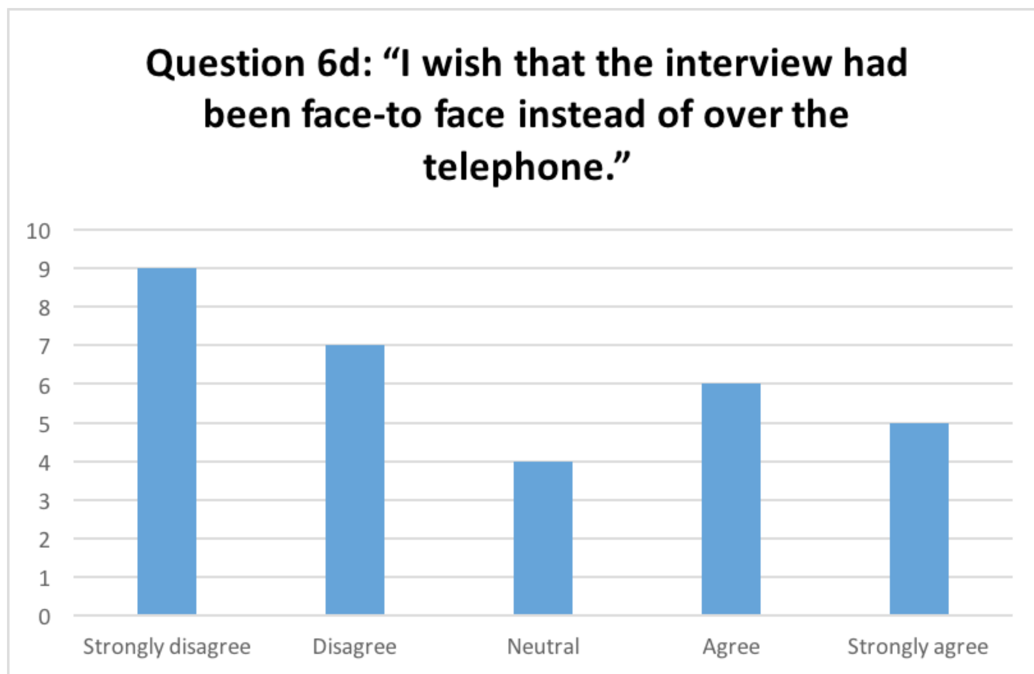


Figure 1. Responses to question 6d.

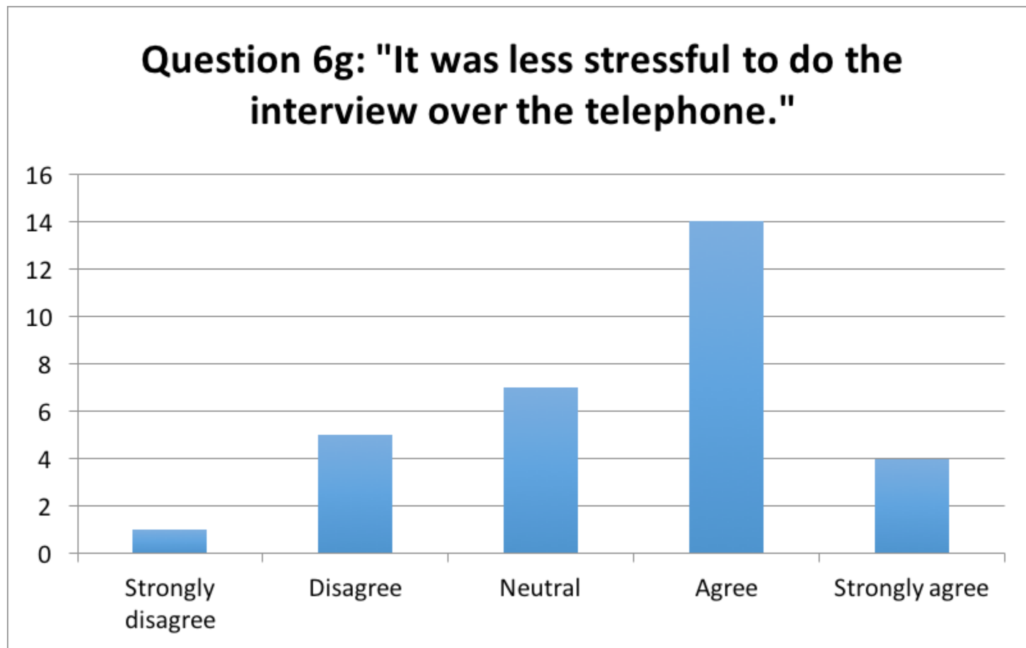


Figure 2. Responses to question 6g.

There was mild anxiety reported that did not have to do with the telephone format. The mean response for all anxiety-related questions is 2.6 (SD=1.25). This sentiment can be illustrated with the following comment by one student: "Au début, cela était un peu plus difficile avec le stress de parler, mais à fur et à[sic] mesure que l'entretien avançait, mieux c'était. La personne m'a aider [sic] à diminuer le stress un peu." (*At first, it was a little more difficult with the stress of having to talk, but as the interview went along, the better it got. The* [*interviewer*] *helped to reduce the stress a little.*)

In addition, English native speaker respondents taking the French test reported significantly greater anxiety than French native speakers taking the English test (M= 2.91; SD=1.2; p=.007). One explanation for this finding is that there were some difficulties with the delivery of the French test: several respondents who were tested in French indicated difficulty with finding the right phone number, whereas there were no issues reported with conducting the English version of the test. We might therefore question if both versions of the test would initially cause the same (relatively low) level of anxiety among test takers in both languages, but these additional logistic obstacles could have increased stress levels.

Another explanation is related to language proficiency. English applicants to the French teaching program have a lower relative proficiency level (see Table 1 below):

	English Speaker average score	French Speaker average score
English Test	4.96 (n=13)	4.81 (n=8)
French Test	3.72 (n=20)	4.86 (n=8)

Table 1. Average scores on each test by reported first language.

In fact, almost all the students who failed this particular exam and who were not admitted were English native speakers doing the French exam Literature has demonstrated that greater anxiety is associated with the lower proficiency levels (e.g., Liu, 2007; Tsai, 2013; Zhang & Liu, 2013). One English respondent made the suggestion that more information about the test content and format ahead of time would be helpful in reducing stress, as illustrated by the comment by E13: "If examples of what we would be interviewed on were provided prior, it would be less stressful. As I didn't know I would be doing a pretend skit I was worried wondering what it is I would be asked."

Conclusions and Future Work

We addressed the nature of interactive functions elicited by the telephone tasks and found that the role play task elicited a greater variety of such functions compared to the opinion task. Some functions were notably absent, and those related to interaction management were still limited. These functions can only be more present if the test candidates were given more control of the interaction. The fact that the role play enabled more opportunity for interaction management was sensed very clearly by student E5, who mentioned "I think doing a role play made it easier to take the lead and answer. However, that does not make it any less stressful!" This anxiety, in general, was not found to be extreme. In general, the participants did not show any particular concern about the telephone format of the test. Based on these results, we made the following recommendations to the test developers:

- In designing the written script for interviewers, it may be possible to increase the variety of certain interactional functions by naming them explicitly (e.g., with *persuade, summarise,* and *agree/disagree*).
- Test designers may want to design a role play task where the candidate is given a "character" to play who has relatively higher power in the interaction, in order to investigate the effects of this change on the nature of interactive functions produced.
- It does not seem problematic at this time to continue with telephone formats for interviews, in terms of anxiety. While stress levels are not exaggerated, this stress may be mitigated even further if additional information was provided to test takers about the test format and content, to the extent possible. This is in keeping with best practice in the field of language assessment (ILTA, 2007).

This validation study was small in scope, but was still found to result in useful evidence for increasing the quality of the test that is offered. This is a high-stakes test, playing a role in the selection of future members of the language teaching profession in Ontario and elsewhere. Efforts must be taken to develop tasks that allow university decision makers to draw meaningful inferences about these applicants' interactive abilities. Therefore, speaking-assessment tasks need to shake the interviewees out of their passive roles, where they can demonstrate their rhetorical skills and their abilities to initiate and manage interaction.

References

Horwitz, E. (1996). Preliminary Evidence for the Reliability and Validity of a Foreign Language Anxiety Scale. *TESOL Quarterly*, *20*(3), 559–562.

Horwitz, E. K., Horwitz, M. B., & Cope, J. A. (1986). Foreign language classroom anxiety. *Modern Language Journal*, *79*(1), 125–132.

ILTA (International Language Testing Association; 2007). Guidelines for Practice. Retrieved from: http://iltaonline.com/images/pdfs/ilta_guidelines.pdf

Kormos, J. (1999). Simulating conversations in oral proficiency assessment: A conversation analysis of role plays and non-scripted interviews in language exams. *Language Testing, 16*(2), 163–188.

Lazaraton, A. (1992). The structural organization of a language interview: A conversation analytic perspective. *System, 20*, 373–386.

Liu, M. (2007). Anxiety in oral English testing situations. *ITL: International Journal of Applied Linguistics, 153*, 53–76.

Luoma, S. (2004). *Assessing Speaking*. Cambridge: Cambridge University Press.

O'Sullivan, B., Weir, C., & Saville, N. (2002). Using observation checklists to validate speaking-test tasks. *Language Testing*, *19*(1), 33–56.

Qian, D. D. (2009). Comparing direct and semi-direct modes for speaking assessment: Affective effects on test takers. *Language Assessment Quarterly*, *6*(2), 113–125.

Sturges, J. E. & Hanrahan, K.J. (2004). Comparing telephone and face-to-face qualitative interviewing: A research note. *Qualitative Research*, *2*(1), 107–118.

Taylor (2011). In Taylor, L. (Ed.), Examining Speaking: Research and practice in assessing second language speaking. Studies in Language Testing 30. Cambridge: Cambridge University Press.

Toth, Z. (2012). Foreign language anxiety and oral performance: Differences between high- vs. low-anxious EFL students. *US-China Foreign Language*, *10*(5), 1166–1178.

Tsai, C. (2013). The impact of foreign language anxiety, test anxiety, and self-efficacy among senior high school students in Taiwan. *International Journal of English Language and Linguistics Research, 1*(3), 1–17.

Weir, C. (2005). *Language testing and validation: An evidence-based approach*. Palgrave MacMillan, Basingstoke.

Young, R., & Milanovic, M. (1992). Discourse variation in oral proficiency interviews. *Studies in Second Language Acquisition, 14*, 403–424.

Zhang, W. & Liu, M. (2013). Evaluating the impact of oral test anxiety and speaking strategy use on oral English performance. *The Journal of Asia TEFL, 10*(2), 115–148.

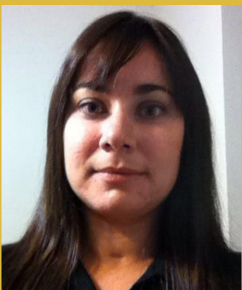
Author Bio

Beverly Baker is an Assistant Professor and Director of the Language Assessment Sector at the Official Languages and Bilingualism Institute at the University of Ottawa.

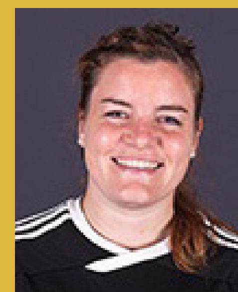
Valerie Kolesova is an ESL instructor and an MA student in Bilingualism Studies at the University of Ottawa, focusing on acquisition and assessment of sociopragmatic competence.

Joselyn Brooksbank is an MA student in Bilingualism Studies at the University of Ottawa, with an interest in family language policy, early childhood bilingualism, and Montessori education.

Jessica McGregor is an MA student in Bilingualism Studies at the University of Ottawa, focusing on the secondary course selection process in French as a second language, with a secondary interest in language identity and ideologies.



Irina Goundareva recently completed her PhD in Spanish Linguistics from the University of Ottawa, specialized in foreign language acquisition and Spanish language teaching.



Mélissa Pésant completed her MA in Bilingualism Studies in 2015. Her major research paper was on the resources used in second language textual revision. She is currently working in England as a French second language teacher.

STEPS TO ENGLISH PROFICIENCY

A review of the framework and its writing continua

By Elizabeth Jean Larson & Clarissa Lau,

Ontario Institute for Studies in Education, University of Toronto

Abstract

Given the current increased presence of English Language Learners (ELLs) in our Ontario elementary and secondary classrooms, teachers are faced with a challenging task of assessing each student's language ability, placing students in suitable programs, and tracking progress. The Steps To English Proficiency is a tool developed by the Ministry of Education, team of educators, and experts in assessment and content that will support classroom teachers to make appropriate decisions and continue to track their students' performance in order to make the most suitable choices for their students. This article will specifically review the writing portion of the framework and hopes to provide insight to teachers who may be interested to adopt this framework by highlighting the strengths and weaknesses of this framework.

Ontario's elementary and secondary classrooms have been quickly evolving as their student populations have become increasingly diverse. To properly address the particular needs of each student, classroom teachers need the best knowledge and tools to properly assess and track each student's progress. The Steps To English Proficiency assessment framework (STEP) is a tool designed for teachers to evaluate their students' language development across oral language and literacy (reading & writing) skills and offer in-depth insights into best assessment practices and teacher development in the classroom setting. To our knowledge, there have been no objective reviews of this unique assessment framework. Therefore, this paper will provide an overview of the history, content, and educational purposes of STEP in relation to the contextual needs of teachers in the classroom. As we believe that writing is a critical skill that can propel students toward better academic success, this review will place a specific focus on the writing portion of the STEP, looking at the continua in all four-grade clusters (Grades 1–3, 4–6, 7–8, and 9–12).

History

In Ontario, elementary and secondary classrooms are largely composed of students whose first language is not English. These students are referred to as English Language Learners (ELLs). These students can be new immigrants or students born in Canada and living in

multilingual environments (Jang, 2014). The initiative to address the language needs of the growing population of ELLs began in 2005 from a response to an Auditor General for Ontario's report (Office of the Auditor General of Ontario, 2005) regarding the status of English as a second language (ESL) and English as literacy development (ELD) in Ontario classrooms (Jang, Wagner, & Stille, 2010). The Auditor General report noted that there was a lack of information regarding how ELLs' English language ability improved and suggested the need for a tracking tool to understand both where the allocated funding was distributed and how ESL and ESD programs were implemented. To address the concerns in the report, from 2008 to 2011 the STEP proficiency scales were developed, validated, and field-tested with various stakeholders such as teachers, assessment experts, and ESL specialists (Jang, Wagner, & Stille, 2011). Following this, the STEP framework was introduced in 2012 and integrated within various school boards across Ontario.

The main purpose of the STEP assessment framework was to address the increasing achievement gap of ELLs with a focus on the classroom context, targeting both oral language and literacy skills (reading and writing). Results from the 2012 Ontario Secondary School Literacy Test (OSSLT) alarmingly identified that while 82% of first-time eligible students passed the test, only 66% of the ELL population did so (Zhang, Shulman, & Kozlow, 2014), suggesting a critical demand for support. Thus, an assessment framework like STEP is a useful tool, and the first in Ontario that allows teachers to track ELLs' language trajectory. To do so, STEP is constructed with more skills-based descriptions of student abilities and competencies. It is a descriptor-based proficiency scale, a current, global, assessment trend (Stille, Jang, & Wagner, In press). Unlike other descriptor-based proficiency assessments (e.g. Canadian Language Benchmarks (CLB), the American Council of the Teaching of Foreign Languages (ACTFL), and World-Class Instructional Design and Assessment (WIDA) English Language Proficiency), STEP is embedded in the Ontario curriculum and is specifically used for classroom-based assessment (Stille, Jang, & Wagner, In press).

Defining Writing

Any valid assessment measure must be grounded within a theoretically and empirically sound definition of the thing that is being assessed (i.e. a construct definition). Within the STEP framework, writing is seen as a *process* (rather than a mere product) that moves from "pre-writing and organization of ideas, to writing and editing" (Ontario Ministry of Education, 2012, p. 13). Jang, Wagner, & Stille (2011) note that this view of writing is supported by theories from Seow (2002) and Tribble (1996). It is important to note here that the content for the entire STEP framework was negotiated amongst three major groups of stakeholders: teachers in Ontario, the Ontario Ministry of Education, and a team of external researchers and content experts. The specific language chosen for all of the Observable Language Behaviour (OLB) continua (both descriptors and elements) resulted from these negotiations. For example, the wording of the STEP descriptors is linked to curriculum standards.

Continua Content

The STEP OLB writing continua consist of OLBs, particular language behaviours that are strongly tied to curriculum standards and can be observed in students' daily classroom activities across all subjects. The OLB writing continua for each grade cluster is composed of four "Elements"' (i.e. dimensions) and six "Steps". The elements and their sub-dimensions are listed in Table 1.

Table 1

Writing Elements of the STEP continua

Developing and Organizing Content	Engage in pre-writing, to generate ideas and information
	Organize ideas and information
Form and Style	Incorporate a variety of text forms and features in writing
Language Conventions	Choose words that convey meaning and add interest to the writing
	Write with fluency using a variety of sentence structures
	Use grammatical structures appropriate to the purpose
	Spell familiar words using a variety of strategies
Revising	Revise for content and clarity

Each OLB writing continuum consists of positively worded descriptors, which highlight students' abilities, rather than their deficiencies. Table 2 contains the descriptors for steps 1–6 for the grade 1–3 cluster in one section of the Developing and Organizing Content element and serves to further illustrate this point.

Table 2

Developing and Organizing Content Element – Observable Language Behaviours

	Observable Language Behaviours (OLB)					
Element: Developing and Organizing Content	Step 1	Step 2	Step 3	Step 4	Step 5	Step 6
Engage in prewriting to generate ideas and information	Generate key ideas using L1, English, and/or visuals by answering simple questions about personal experiences	Generate ideas by brainstorming with peers and teachers in L1 and English about personally relevant topics	Generate ideas with peers, using familiar strategies	Generate ideas about a topic, using a variety of strategies and key academic vocabulary	Generate ideas, using a variety of strategies and resources and academic vocabulary	Locate and select information for a writing topic, using resources

Note. From the *STEP Observable Language Behaviours Continua-Writing (Grades 1–3; Ontario Ministry of Education, 2012).*

The descriptors for each grade cluster are tied to curriculum demands and teacher expectations and are, as such, quite distinct from each other. Recent research by Jang, Cummins, Wagner, Stille, and Dunlop (2015) provides empirical evidence of how distinct the descriptor definitions are across the three modalities (speaking, reading, and writing), as well as how similar these descriptors are within proficiency levels. It is essential to empirically confirm that descriptors are distinct from each other to ensure that each step sequentially builds upon the previous step in order to track student progress. In other words, if the steps are too similar, then it may become unclear what a "1" means versus a "2", and the steps will lose their meaning.

However, there are differences regarding the students' use of their first language (L1) across age clusters. All of the OLBs for writing anticipate students to use their L1 in Steps 1 and 2, but there are differences among the Elements in which L1 use is expected. In the 1–3 cluster, L1 use is expected in 3 of the 4 Elements (except Revising), while in the 4–6 and 7–8 grade clusters, L1 use is expected only in the Developing and Organizing and Form and Style Elements. In the 9–12 grade cluster, use of L1 is only anticipated in the Developing and Organizing Element. The different treatment of students' use of their L1 in different grade cluster indicates a potential disadvantage for older ELLs. While the curriculum expectations are certainly more cognitively challenging as grade level increases, it is not clear why students' limited use of their L1 could not be accepted in the same elements in steps 1 and 2 across age clusters. Denying students' initial use of their L1 as they develop

their English skills may have detrimental effects and prevent students from achieving their full potential. Thus, it would be valuable to add descriptors that include the use of the L1 in steps 1 and 2 in three of the four elements (of the writing continua).

Implementation

The continua are frameworks and are therefore meant to be used on a continual basis to inform lesson plans and other pedagogical decisions, as well as other assessments of students' work and classroom behaviours. They are to be used by all classroom teachers, regardless of discipline, as well as ESL/ELD teachers to assess and track the language development of ELLs. *The STEP Users Guide* (2012) suggests using the OLBs for tracking purposes at least once a year. It can be assumed that students will progress differently in different elements. For example, it would be possible for a student to be at step 2 for Revising, but at step 3 for Developing and Organizing. However, students can only progress to a "higher" step if teachers have observed all OLBs in all four elements of the current step. Regarding the example above, tracking records would only indicate that student had overall completed step 2. Practical implementation of the STEP would be limited when teachers have dividing opinions of whether a student has achieved all Elements of a Step or not. This poses difficulty in the tracking procedures of STEP and teachers need to be more aware of their own interpretations of the step descriptors.

Limitations and Suggestions

Each teacher offers their own unique perspectives of their students and as such there should be measures in place to ensure accountability regarding teacher use and interpretation of the continua. There are currently no such measures in place and this calls the reliability (particularly inter-rater reliability) of the continua into question. One way to ensure a reliable use of the continua would be to have teachers from different disciplines and grades share perspectives on each student's language development with each other. This would have the added benefit of strengthening teachers' own classroom activities and assessment tasks.

While the STEP framework is flexible in its application and interpretation, it is quite inflexible regarding the process (for tracking purposes) through which a student can be said to have completed a particular step. A student must demonstrate *all* aspects of the four elements in order to complete that step. It could be argued that this is a rather narrow view of language development that focuses more on what students still cannot do (i.e. their deficiencies), rather than what they can do (i.e. their abilities). If a student has mastered one of the elements in a step, that should be acknowledged and they should be given more challenging tasks informed by the descriptors in the next step. It may be more useful, then, to have each element and its accompanied step in its own continua. This would provide a clearer picture of students' language development for whomever is reading the student's tracking record.

Go Back to Top

Strengths and Suggestions

A major strength of the STEP is that the descriptors have been developed through discussion with various experts such as teachers, education professionals, and assessment experts. While different stakeholders' perceptions of each construct can lead to validity concerns, it could be argued that these descriptors are better representations of the construct definitions most relevant to classroom-based assessments. Through the validation studies, teachers' inputs were crucial to appropriate the language used in the descriptors. Given that teachers are the primary users of STEP, the assessment framework utilizes descriptors that are applicable to teachers' classroom context so the framework can be broadly interpretable.

Another strength of STEP is its flexibility as a framework meant merely to guide teachers. But, this flexibility could also be seen as a burden for some teachers who may not have received education or training on how to prepare a range of materials that target specific descriptors of each element. In order for STEP to be easily accessible and pertinent to teachers' preparation of materials, the *STEP User Manual* (2012) should include examples of activities and assessment tasks that would elicit the skills in each element. This suggestion is not meant to restrict or direct teachers' work but rather is intended to support and optimize teachers' preparation.

Moving forward from the validation studies and field-testing, framework developers and experts should follow-up with teachers who are currently utilizing STEP in their classrooms. Factors such as the practicality and applicability of the framework in the classroom context should be further investigated. Also considering other contexts, we would recommend other educational systems consider using a framework such as STEP for their own classrooms. It is important to note that STEP is highly context-specific, so its applicability to other contexts should stem more from its theoretical perspective on language development rather than its particular descriptors, elements, or continua. Supporting ELLs' language development in any school should not only concern language, ESL, or English teachers, but teachers in all subjects and classroom contexts. The more teachers understand the unique needs of each student, the better they can provide essential supports and activities to create a truly equitable learning environment for all students.

References

Jang, E. E. (2014). Assessing English language learners in K-12 schools. *Education Matters*, *2*(1), 72–80.

Jang, E. E., Cummins, J., Wagner, M., Stille, S., & Dunlop, M. (2015). Investigating the homogeneity and distinguishability of STEP proficiency descriptors in assessing English Language Learners in Ontario schools. *Language Assessment Quarterly, 12,* 87–109.

Jang, E. E., Wagner, M., & Stille, S. (2010). A democratic evaluation approach to validating a new English language learner assessment system: The case of Steps to English Proficiency. *English Language Assessment, 4*, 35–50.

Jang, E. E., Wagner, M., & Stille, S. (2011). Issues and challenges in using English proficiency descriptor scales for assessing school-aged English language learners. *Cambridge ESOL Research Notes*, 45, 8–14.

Office of the Auditor General of Ontario. (2005). 2005 annual report of the office of the auditor general of Ontario. 9–10. Retrieved from http://www.auditor.on.ca/en/reports_en/en05/en_2005%20AR.pdf

Ontario Ministry of Education. (2012). *STEP: Steps to English proficiency: A guide for users*. Ontario: Queen's Printer for Ontario. Retrieved from http://www.edugains.ca/resourcesELL/Assessment/STEP/STEPUserGuide_January2012.pdf

Seow, A. (2002). The writing process and process writing. In J. C. Richards and W. A. Renandya (Eds), *Methodology in language teaching: An anthology of current practice* (pp. 315–320). Cambridge: Cambridge University Press.

Stille, S., Jang, E., & Wagner, M. (In press). Building teachers' assessment capacity for supporting English Language Learners through the implementation of the STEP language assessment in Ontario K-12 schools. *TESL*.

Tribble, C. (1996). *Writing*. Oxford: Oxford University Press.

Zhang, S., Shulman, R. & Kozlow, M. (2014). Tracking student achievement in literacy over time in English-language schools: Grade 3 (2005) to grade 6 (2008) to OSSLT (2012) cohort. (EQAO Research: Research Report).

Elizabeth Jean Larson (jeannie.larson@mail.utoronto.ca) has taught ESL for 10 years in Japan, Germany, and Canada. She has taught students of all ages in contexts varying from private junior and senior high schools, to private conversation schools, to business courses conducted in-company. She completed her Masters in Education in the Language and Literacies Program at OISE in the spring of 2016 and began her PhD in the same program at OISE that following September. Her research focuses on language education and language assessment, particularly regarding oral language.

Clarissa Lau is a third year doctoral student in the Developmental Psychology and Education program at OISE. Her teaching experiences have included high school and adult ESL learners as well as individuals with special needs. She has also developed curriculum materials for high school and adult ESL learners. Her main research interests involve issues of emotional regulation, assessment, and language development. She is involved in various research initiatives investigating cognitive, metacognitive, and affective factors that influence literacy development. She is currently involved in a provincial initiative to develop a language framework to assess and track oral language in the kindergarten years

LANGUAGE ON THE ORIENT EXPRESS: A GUIDE TO MANDARIN FOR ENGLISH TEACHERS

By Eric Henry, Saint Mary's University

Foreigners living in 1920s Shanghai apparently counselled new arrivals that "those who learn Chinese go mad" (Kane, 2006, p. 17). Certainly there has been a long tradition of outsiders looking on the Chinese language as a confusing mass of chicken-scratch writing and syncopated bursts of rhyming syllables. In turn, Chinese people themselves were often portrayed by Westerners as aloof, inscrutable and incomprehensible to outsiders. The reasons for surmounting such racist representations are pressing for both English-speaking students of Chinese and teachers of English to Chinese, particularly in light of the crucial role China will play on the global stage in the twenty-first century. With this in mind, this article provides a general overview of Mandarin Chinese geared towards teachers and students who have had limited exposure to the language. This overview is, by necessity, cursory and in several areas I have had to simplify what in other contexts would be questions of great complexity, but I hope this article can at least offer an entry point to those new to the language.

Mandarin Chinese, called *putonghua* (common speech) in mainland China and *guoyu* (national language) in Taiwan—although differences exist between the two national varieties—is the most widely spoken language in China today. Although many consider Mandarin and Cantonese to be dialects of the same language, this is analogous to calling French and Italian dialects of "European." Most linguists argue that we should describe the linguistic situation in China as consisting of a host of spoken Chinese *languages*, including at the very least Mandarin, Cantonese, Wu, Hakka, Fujianese and a few others (Ramsey, 2002). At the same time though, all more-or-less share a common written script.

The Script

The modern Chinese script has its origins in a logographic system of writing that was incised on turtle shells and animal bones during the Bronze Age as a means of divination (DeFrancis, 1984). The earliest oracle bone writing consisted of pictograms representing objects and animals, but these eventually came to represent words and concepts. Some modern Chinese characters derive from these early forms, such as *ma* (马) "horse", which is a stylized drawing of the animal, and *ri* (日) "day", representing an image of the sun. The majority of characters today, however, are not pictographic representations but can be more accurately described as

picto-phonetic; that is, they typically feature both semantic and phonetic components (Yin and Rohsenow, 1994, p. 21). The character *bo* (波) "wave", for instance, is composed of two parts (or radicals). On the left side of the character is a semantic component, three strokes representing "water", which is also found in words like *piao* (漂) "float" and *jiang* (江) "river". On the right side is a phonetic component, indicating that this word is pronounced like other characters that share its form, including (跛) "lame" and (菠) "spinach." These similarities originally emerged as rebus associations among similar sounding words, but as certain pronunciations have changed over the centuries, phonetic radicals now often represent a group of related sounds; related to the characters above is the word *po* (坡) "slope".

There are several different Chinese scripts currently in use. Hong Kong, Taiwan, Macao, and many overseas communities use a form of Chinese characters called *fantizi* ("complicated script"), which are unchanged from the imperial period, while in Mainland China a script reform movement in the late 1950s resulted in the adoption of *jiantizi* ("simplified script"). The other main scripts are the various systems of romanization such as Wade-Giles for Cantonese and the standard *pinyin* for Mandarin. The lack of congruity between these ways of representing Chinese languages alphabetically often leads to several different renderings of the same word, such as the name of China's capital city: it is Peking in earlier systems of romanization and Beijing in *pinyin*.

Phonology

English teachers may wish to note that the phonemic system of Mandarin does not include many common English sounds, so these often present difficulties for Chinese students. Notable among these are the vowel in words like *ship* that for Chinese students is often articulated more like *sheep,* the *v* which is often realized as a *w,* and the voiced and unvoiced *th* in words like *there* and *think* respectively. None of these are part of the Mandarin phonemic system. Mandarin also has a relatively limited set of vowel sounds: five in comparison to English's fourteen. There are also, however, a few semi-vowels (or glides) and many of the vowels combine into diphthongs and triphthongs.

Nevertheless, the Mandarin Chinese phonological system and its *pinyin* orthography are remarkably similar to English, with a few exceptions. First, whereas paired English consonants are typically differentiated on the basis of whether they are voiced or unvoiced (compare, for instance, the initial consonants in *tin* and *din*), in Chinese the key difference is whether they are aspirated or not (Wiedenhof, 2015, p. 34). The Chinese word *da* features an unaspirated initial consonant while *ta* features an aspirated one; both consonants are unvoiced, meaning that they do involve vibration of the vocal cords.

Three consonant minimal pairs also present difficulty for native English speakers, and their orthographic representation is not entirely obvious to the uninitiated. The three contrasts as they appear in *pinyin* are:

- *ch – q*
- *zh – j*
- *sh – x*

The three sets of consonants contrast in terms of place of articulation. *Ch-, zh-*, and *sh-* sounds are all unvoiced retroflex consonants, produced by placing the tip of the tongue just behind the alveolar ridge (ie. the ridge behind the upper teeth). *Ch-* and *zh-* contrast as aspirated and unaspirated affricates, while *sh-* is a fricative. The *q-, j-*, and *x-* sounds are all unvoiced palatal consonants, produced with the blade of the tongue touching the hard palate (ie. the roof of the mouth). *Q-* and *j-* contrast as aspirated and unaspirated affricates, while *x-* is a fricative. To put it more simply, *ch-* and *sh-* are pronounced much as they would be in English. In contrast, *x-* and *q-* represent palatalized versions of the same sound, pronounced slightly further back in the mouth. *J-* is pronounced much as it would be in English, while *zh-* is a retroflex version of the same sound. Many novice learners of Mandarin have trouble differentiating these contrasts as they are not perceived as salient in English phonology; thus the initial consonants in *ju* (居) "reside," and *zhu* (猪) "pig" may be perceived as allophones by a non-speaker but are meaningfully differentiated by the palatal-retroflex contrast in Mandarin. One other phoneme not shared with English is represented orthographically with the letter c, indicating an unvoiced aspirated alveolar affricate pronounced much like the *ts* in *tsunami* or *cats*.

Mandarin has a relatively limited syllable system, generally consisting of either consonant + vowel (ie. *bu, ge, pai*) or consonant + vowel + nasal consonant (*lan, rong, jiang*).[1] It is therefore common for Mandarin speakers to add a schwa after a final consonant or to insert a vowel in long consonant clusters when speaking English. Because of the limited set of syllables Mandarin provides for constructing words, tonality allows the language to differentiate four forms of each syllable. The four tones, often described as high (1st tone), rising (2nd tone), falling-rising (3rd tone) and falling (4th tone), can be represented orthographically in *pinyin* by placing a diacritic over the main vowel in each syllable.[2] Each diacritic is an iconic representation of the tone (ie. 1st marked by a bar above the vowel, 2nd by a rising slash, 3rd by what looks like a *v*, & 4th by a falling slash). For example, the word *lao* (pronounced similar to the first three letters in the word *loud*) can have the following four meanings according to tone:[3]

- *lāo* (捞) – "dredge"
- *láo* (劳) – "labour"
- *lǎo* (老) – "old"

1 A few syllables in Mandarin lack an initial consonant, such as the words ai 'love' and en 'to press with the finger,' but these are relatively rare.

2 Although common for pedagogical purposes, in many texts the tone diacritics are omitted and I have included them here only where relevant.

3 Actually, my basic Chinese dictionary lists ten meanings, because each of these tonal syllables has multiple homophonous characters (see the next paragraph) – I have merely chosen four here as representative examples.

- *lào* (烙) – “bake”; “iron”

Despite the ability of the tone system to expand the potential semantic identities of Mandarin syllables, each character still usually has at least a few homonyms, and many jokes rely as much on soundplay as they do on wordplay. It also allows Chinese internet users to locate information on banned topics by searching for homonyms, essentially substituting other sound-alike words for sensitive names or topics. The word *dāng* (裆), for instance, means “crotch” or “seat of the pants,” but is a popular online search string because it sounds like *dǎng* (党) meaning “political party” or, more specifically, “the Communist Party of China.” Another example is the seemingly innocuous phrase “grass-mud-horse” (*cǎonímǎ*) that sounds like one of the crassest insults in the Mandarin. When the dissident artist Ai Weiwei posted a picture of himself online dancing naked with a stuffed horse covering his crotch, it was widely understood to be an insult to the Chinese government.

Morphology

A common assumption among non-Chinese speakers is that characters are monosyllables and thus can act as a kind of mystical shorthand for broad ideas or feelings. For instance, although *ai* (爱) can be used as a verb meaning “love,” as in *wo ai ni* (“I love you”), when referring to the concept of “love,” *ai* is combined with *qing* (情), meaning “emotion,” to form a compound word. Most Chinese “words” are, in fact, compound words composed of two characters that may have related or relatively broad meanings in isolation. In its simplest form, this can consist of a syllable followed by *zi* (子), indicating an object or noun. Dictionaries define the character *mao* (帽) as “hat”, but in actual speech a hat is referred to as *maozi* (帽子). Most words, however, combine two separate morphological syllables to generate exact meanings, such as *huanjing* (环境) “environment”, which brings together *huan* “encircle” and *jing* “area”. Note that to use “encircle” as a verb in actual spoken Mandarin, one would typically say *huanrao* (环绕), where *rao* means “coil” or “move in a circle”. In each case, the monosyllabic characters are semantic elements on their own, but most words exist as disyllabic compounds.

Mandarin can be classified as an isolating language, meaning one that avoids inflectional affixes to express pluralization, tense, gender, and so forth. Many linguists characterize the language as having almost no morphological structure (Li & Thompson, 1989). Instead, Mandarin employs a class of function words that indicate these qualities at the level of sentence structure. For instance, questions are not formed through the rearrangement of words as in English (e.g., *I am dreaming* → *Am I dreaming*?), but through the addition of the interrogative auxiliary *ma* at the end of a sentence. The statement *ta you* (“he has [it]”) can be transformed into a question in this way: *ta you ma?* (“does he have [it]?”). We can also see in this example how, where they can be inferred from context, Mandarin speakers frequently omit the subject or object of a sentence.

Mandarin does not possess a tense system, but speakers can indicate the temporal quality of actions in two main ways. One is through the use of function words to mark grammatical aspect, such as completion (the auxiliary *le*) or continuation (the auxiliary *zhe*) of actions. In practice these operate much as past and progressive tenses do in English. The other way time is indicated is through topic-fronting, which positions the timeframe of the action at the beginning of a sentence. Teachers will note that Chinese students often begin English sentences this way, using words such as *yesterday* or *in the future* to start a sentence rather than placing them at the end as is more common in spoken English.

Discourse and Social Aspects of Use

Mandarin relies heavily on idiomatic sayings and expressions called *chengyu*. These pithy phrases, oftentimes drawn from literature or historical sources, act as evaluative statements on the situation at hand, allowing speakers to refer to their interpretations obliquely rather than directly. "East one sentence, west one sentence," refers to incoherence. "A donkey's lips do not match a horse's mouth," means an answer that is not relevant to the question. One *chengyu* dictionary from Taiwan contains over thirty thousand entries (Allen 2011, p. 73). *Chengyu* arise regularly in conversation and often laminate contemporary language with classical expression. By alluding to this past and to Chinese culture as a whole, *chengyu* are not merely poetic aphorisms but also tools of persuasion and argument. Consequently, Chinese students often overuse idioms as a rhetorical technique when writing in English.

Turning to names, in Mandarin surnames come before given names; asking a Chinese for his or her "first name" can therefore elicit confusion. Names usually take the form of a single character (and thus syllable) surname followed by a one- or two-character given name. Almost all surnames and given names are characters with other semantic referents. It is common for people to have many different names that facilitate everyday social interactions. Parents often refer to their children with a diminutive or "small name," such as *Lili* or *Baobao*. Friends may employ a nickname or append a descriptor to a person's real name, such as *Little Wang* or *Old Peng*. Similarly, at work people are often referred to by their position, such as *Teacher Xu* or *Department Head Zhang*. And many people adopt kinship terms to refer to particular types of relationship in addressing others; the most senior female teacher at a school, for instance, might be addressed by her co-workers as "elder sister". I have heard Canadian teachers wonder why their Chinese students adopt English "nicknames" or "aliases," and it is often assumed that this is motivated either by a desire to fit in or a fear that their "real" names cannot be pronounced. In effect though, English names are an extension of this logic. An English name represents a Chinese person's global persona and is no less authentic than any of the other names they carry.

Finally, although I have focused largely on standard Mandarin, it should be noted that there is a great deal of linguistic diversity across China. Every region, province and city has

its own idiosyncratic variety of the language, and these can diverge almost to the point of mutual incomprehension. The difference between urban and rural speakers is especially marked, and urbanites often told me that they had to read subtitles when rural citizens were interviewed on television. Local dialects are a matter of some pride to the area, and are generally used informally to develop feelings of belonging and camaraderie. Nevertheless, there is a heavy emphasis on correctness, and consequently a stigma on local dialects, in formal settings such as at work or school. Many people will shift between local and standard Mandarin as they travel from home to the office or from a casual lunch to the classroom.

Conclusion

Mandarin Chinese is a rich, complex language that offers challenges for English speakers who are often more familiar with European languages as a whole. Chinese characters, in particular, present a radically different approach to literacy than an alphabetic script. Nevertheless, I hope this outline of the language can serve to demystify some of the common elements of Mandarin and make it more accessible to a new generation of learners.

References

Allen, J. (2011). From literature to lingerie: Classical Chinese poetry in Taiwan's popular culture. In M. Moskowitz (Ed.), *Popular culture in Taiwan: Charismatic modernity* (pp. 65–85). New York: Routledge.

DeFrancis, J. (1984). *The Chinese language: Fact and fantasy*. Honolulu: University of Hawai'i Press.

Kane, D. (2006). *The Chinese language: Its history and current usage*. Singapore: Tuttle

Li, C. & Thompson, S. (1989). *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.

Ramsey, R. (2002). The languages of China. In S. Blum & L. Jensen (Eds.), *China off center: Mapping the margins of the Middle Kingdom* (pp. 45–64). Honolulu: University of Hawai'i Press.

Wiedenhof, J. (2015). *A grammar of Mandarin*. Amsterdam: John Benjamins.

Yin, B & Rohsenow, J. (1994). *Modern Chinese characters*. Beijing: Sinolingua.

SUPPORTING BLENDED LEARNING IN AN EFL TEACHING ENVIRONMENT

By John Allan & Jennifer Wicks, Centre for Teaching Excellence,

College of the North Atlantic - Qatar

Supporting blended learning at a college with a face-to-face classroom model can be a challenge, particularly in an EFL learning environment. Faculty need clear guidelines about the institutional definition of blended learning, a strong sense of why they should "blend" technology and face-to-face methods, as well as the know-how for tools and resources amongst a range of other considerations. In order to better support faculty, our Learning Technologies team offered a Blended Learning Workshop Pilot Series to support a culture of learning starting with basic technology tools and resources freely available to faculty. The following describes how we went about creating and delivering the series, and provides insights into what we learned from the experience and feedback from faculty.

Blended Learning

We have decided to use the definition from the University of Central Florida's Blended Learning Toolkit for the purposes of our pilot series, which describes blended learning as, "courses … where a portion of the traditional face-to-face instruction is replaced by web-based online learning." This definition is in keeping with our institutional mandate to promote and support blended learning without reducing face-to-face instructional time for faculty and students.

Background

As members of a faculty development department serving approximately 500 instructors at a college in the Middle East region, we have been tasked with supporting the implementation of a Learning Management System (LMS) and facilitating the training and ongoing support for faculty using technology in their teaching and learning.

An examination of the existing LMS training (technology-focused), faculty mentoring (pedagogy-focused), requests for additional training from various schools and departments, and a wealth of anecdotal evidence provided us with a starting point for our support of blended learning on our campus.

This starting point was to help faculty add to their toolbox of educational technologies and strengthen their comfort level with these resources. We decided to approach this through a series of workshops designed specifically from faculty feedback and demand. We decided

that face-to-face workshops delivered with a consistent schedule, style, and format was the most practical way for us to reach faculty at this point in time. The following details provide some insight into the planning and delivery of our Blended Learning Pilot Workshop Series, which was offered from January to April 2015.

Target Audience

To begin, we proposed the pilot series to the School of Language Studies and Academics on campus as it is a large school and is known to have a high number of early adopters and individuals who are inclined to participate in professional development opportunities in general. The majority of instructors in this department are also career educators with experience teaching in EFL environments; their experience is useful in preparing the content and delivery of the workshops with consideration for appropriate pedagogical approaches as there is already some common ground in terms of the basic tenets of teaching and learning and an appreciation of the EFL perspective. The dean of this school is also supportive of assisting his faculty in professional development for the use of learning technologies in the classroom.

Scheduling

Having received departmental approval, we worked with the chair to establish times when most faculty were not teaching. We used the Doodle scheduling tool in our meeting announcements. The chair sent a weekly email announcement containing the Doodle link for registration. Registration was capped according to the number of functioning computers in the lab, as we wanted each participant to have their own computer. In addition to the Doodle sign-up, participants indicated it would be helpful for us to send a follow-up meeting reminder using Outlook. Registration thus became a two-step process, and depended on us creating reminders for ourselves each week for the extra step.

The sessions were scheduled during a 45-minute block twice a week, but given the constraints of busy faculty schedules, helping participants log onto their computers, ensuring everyone had the necessary workshop files open, and providing a few minutes at the end for feedback surveys, we found that actual instructional and hands-on time was closer to 35–40 minutes.

Session Format

To help ensure efficiency during the session, we employed the following model:

1. Provide short introduction to the topic and objectives of the session
2. Distribute content such as image, text, or audio files to the participants' workstations through the LANschool resource (later using USB drives)

3. Provide written step by step instructions for most activities (soft and hard copies)
4. Demonstrate and/or provide visual displays for most actions performed during the workshop
5. Distribute and collect feedback from participants

This model helped us provide a consistent format for the benefit of the participants, as well as to establish a framework for materials development. Two to three facilitators helped guide participants through the activities, answered questions, and helped troubleshoot technical problems.

Workshop Topics & Materials

To narrow our menu of possible workshop topics, we interviewed instructional developers, instructors, departmental coordinators and students. We felt it was important to identify Web 2.0 resources for this workshops series as these freely available social and interactive tools offer countless opportunities for teaching and learning, and can be accessed easily both on- and off-campus by faculty and students.

One example of a Web 2.0 resource is Quizlet's free study tools and apps. Quizlet was one of the most popular events in our series. After the workshop, instructional developers and instructors requested the workshop materials, and we received several follow-up requests for further support. The enthusiasm of participants at the current capabilities of this particular tool was apparent, and served as an important reminder that attendees of the workshop sessions come with widely varying experience and knowledge of the technologies we were working with, even those individuals who have used particular tools in the past.

Once the offerings were established, we began development of workshop support materials. The template for materials, including digital and paper media, were based on in-house resources developed for training over the past year. Workshop materials were prepared in advance of each workshop and tested for errors or omissions, a process taking approximately six hours per workshop. Handouts with step-by-step instructions have been shared on our Blended Learning Portal where all faculty can now access tutorials for each of the workshop topics.

Faculty Input

After the initial three workshops (Tutela, Document Accessibility, Image Optimization), the department chair sent a poll to all faculty members in the department to ask for input on upcoming workshop topics. We presented a list of possibilities and provided space for additional suggestions as follows:

1. Copyright-friendly images
2. CSS (Cascading Style Sheets) Overview

3. LMS Quiz Basics
4. Embedding Web 2.0 Objects in an LMS
5. Hot Potatoes
6. HTML Basics 1
7. HTML Basics 2
8. Image optimization 1
9. Image optimization 2
10. Khan Academy
11. M-Reader
12. Padlet
13. Phet
14. Prezi
15. MoboSurvey
16. Quality Matters Overview
17. Quizlet
18. Slideshare
19. WebQuests
20. Other – open to their suggestions

There were 36 respondents; to our surprise, requests were for more technical topics than "out-of-the-box" tools. With this result, we started to create and refine our workshops and began running them the following week. Timing was critical as we did not want to lose any momentum from the first few offerings.

Workshop Feedback

We asked that all participants complete a feedback form at the end of each session. This gave us valuable data on participation rates. Feedback included suggestions for follow-up topics and future workshops, requests for specific information, suggestions for supporting materials, and requests regarding the workshop facilitation.

We incorporated this feedback where possible to improve our workshops and to appeal to the faculty who were attending regularly. We are compiling all feedback to inform future offerings and discussion amongst our team for planning purposes.

Attendance

Workshop attendance generally fluctuated between zero and 12 attendees. Low participation is always a risk when offering workshops, and particularly when piloting workshops. We wanted to experience the process of facilitating each session no matter how few attendees since we knew this would allow us to refine our workshops for the future. However, there were instances where the two facilitators outnumbered the participants, which is clearly not a desired scenario in terms of resource allocation.

The highest-attended session was Quizlet with a total number of 14 participants, and the two HTML Basics workshops drew a total of 22 participants. The two Image Editing sessions had 12 participants in total. Document Accessibility reached a total of eight participants, Image Optimization had six, LMS Quiz Basics had four, and Copyright-friendly Images had only three attendees (there were no attendees at one of these two offerings). It is important to note for future planning that HTML Basics (the highest attended session) was one of the faculty-requested topics from our survey. It is not surprising that tailoring session topics to faculty requests tends to boost attendance.

Attendance data may be valuable as an indicator of interest, participation rates, and popularity of topics. However, our data do not address contextual factors such as semester schedules, faculty commitments, and the potential influence of other institutional events or affective variables that affect faculty availability or the desire to engage in professional development.

Observations & Recommendations

It is apparent that faculty are keenly interested in technology workshops, and faculty input on topics helps facilitators develop practical and relevant offerings. While faculty input does not guarantee attendance, it is clear that offering relevant training opportunities can go a long way in terms of developing an ongoing conversation about blended learning and timely supports.

However, we recognize the need for a more comprehensive framework in which to place these workshops to provide holistic support for faculty with varying competencies with technology in the classroom. As with any support department, it is important to determine how our resources are best spent and to ensure that any support initiatives are followed-up with additional support for classroom application and institutional implementation.

The pilot also helped us understand the need for continued support for faculty with the institution's LMS. Without clear guidelines for use of the LMS, and access to learning analytics and data on actual usage, it is difficult to create a common basis for which to launch further blended learning supports, a concern echoed by many of our participants. Establishing a foundation for ubiquitous use of the LMS campus-wide should be a priority moving forward.

Go Back to Top

In summary, our recommendations (for our institution) are as follows:

1. Utilize feedback from participants in the pilot series to inform future offerings
2. Elicit faculty feedback on technological resources of interest to them through the use of surveys and data collected from instructional technologies staff
3. Host a discussion group with faculty and managers to discuss the outcomes of our pilot series and generate ideas for fostering an institutional "culture of blended learning"
4. Help administration establish a foundation (including clear guidelines and expectations) for ubiquitous use of the LMS campus-wide
5. Offer continuous support and training for faculty to use the LMS more effectively as a platform for blended learning, including faculty-led sessions on features of the LMS that are working for them
6. Purchase and utilize the LMS Learning Analytics tools to understand how faculty and students are currently using the LMS in order to make more informed decision about training and development
7. Make workshop materials available to all faculty through the LMS Portal
8. Continue to work on a Blended Learning Framework in order to devise a clear strategy for implementation institution-wide

Author Bio

With 15 years of rich and diverse experience in Education, leadership and innovation, **Jenn Wicks** is passionate about professional development and coaching. She is currently a Faculty Development Mentor at the College of the North Atlantic in Qatar and a certified executive coach. https://qa.linkedin.com/in/jennwicks

John Allan is currently an English as a foreign language Instructor at the College of the North Atlantic in Qatar. He is also a TESL Ontario guest blogger and a member of the Social Media Committee. When time allows he helps language instructors improve their blended learning events through learning objects and facilitating professional development workshops. https://ca.linkedin.com/in/johnharoldallan

TEACHING CRITICAL THINKING SKILLS IN A DISCIPLINARY CONTEXT

By Percival Santos, London Academy, London, UK.

Critical thinking (CT) is a fundamental defining concept of a Western university education (Barnett, 1997) and it lies "at the heart of EAP" (de Chazal, 2014, p. 12). However, there is a discrepancy between the findings of several experimental studies that show CT to be particularly suited to a content or theme-based approach, and how it is currently conceptualized and taught in many English for Academic Purposes (EAP) programs. Some EAP foundation programs tend to regard it as a discrete transferrable skill- one not necessarily embedded within a particular theme or subject.

This article aims to answer one question: how can CT best be taught to EFL/EAP students? It will argue that integrating the teaching of CT within a discipline or subject is particularly effective for EFL learners. It will also propose a discipline-specific model for teaching CT in the EFL classroom.

Should CT be Taught as a Discrete set of Skills?

Although many agree on the importance of CT, there are diverse opinions regarding what the concept is. Moon (2008, p. 33–34) asserts that its characteristics are assessment of evidence, critical appraisal, reflection, understanding, analysis, synthesis, and evaluation. Judge, et al. (2009, p. 1–2) emphasize the objective to subjective process and the necessity of providing supporting evidence. Ennis argues that it "is a form of reasonable, reflective thinking, focused on deciding on what to believe or do" (1987, p. 10).

Notwithstanding its exact meaning, let alone its definition, CT has often been conceptualized in EAP as a component of critical literary or critical reading skills. In this view, students are taught to "critically" read a text, looking for such things as assumptions and flaws. They are asked "to understand what it is saying" (de Chazal, 2014, p. 12), and they are taught to ask questions such as:

- What is the writer's main argument?
- How relevant and reliable is the evidence that the writer presents in support of their argument?
- What is their stance, and what meanings lie implicit in the text?
- What weaknesses are there in the material? (de Chazal, 2014, p. 12)

Nevertheless, one fundamental assumption some EAP programs make regarding the nature of CT is that it is composed of several generic, transferrable skills. However, the literature

manifests a difference of opinion on the topic, with the research falling into three broad camps. Some researchers (Ennis, 1989; Davidson, 1998; Gieve, 1998; Hawkins, 1998) assert that CT can be taught as if it were a set of generalizable and transferable skills—skills that the students can learn and apply to other literacy contexts. Others (Atkinson, 1997; Pennycook, 1999; Willingham, 2007) argue that CT is not an autonomous, cognitive skill that can be learned out of context; domain or subject-specific knowledge and practice are essential for its development. Alternatively, Jenicek & Hitchcock (2005) take the middle position, asserting that CT can be taught as an independent program or alternatively, it can be integrated into a subject matter course.

A Discipline-Specific Focus in CT

Bean (2011, p. 9) proposes that students need to understand the different ways various disciplines use evidence to make knowledge claims if they want to grow as critical thinkers:

> Some disciplines derive their evidence from observations of natural or cultural phenomena, sometimes converted to numbers, subjected to statistical analysis, and displayed in graphs and tables. Other disciplines use qualitative data from ethnographic observations, focus group transcripts, or interviews… What new students don't see is how these different kinds of data function as evidence in support of a claim. Teachers can accelerate students' understanding of a field by designing assignments that teach disciplinary use of evidence or that help students analyze the thinking moves within an evidence-based argument.

Under this view, the starting point for an effective CT lesson would have to be a focus on a discipline-specific problem-solving task. The lesson should focus on an academic problem that is relevant to a given discipline. That is, teachers should be asking questions or requesting students to solve problems that an economist or biologist would be asked to think about or solve.

Next is the role of evidence. Students already know how to formulate a thesis or argument. They are also familiar with using evidence to help strengthen their thesis and any assertion they make that relates to their thesis. Like a lawyer in a jury trial, a writer must convince her audience of the validity of her argument by using evidence effectively. A discipline-specific approach would show them that the types of evidence used changes from discipline to discipline—they might use quotations from a poem or a literary critic, for example, in a literature paper; they might use data from an experiment in a lab report.

Last is analysis. Students already know how to analyze data, that is, the process of interpreting *evidence* in order to support, test, or refine a *claim*. A strong thesis requires solid *evidence* and analysis to support and develop it because, without them, a claim is merely an unsubstantiated idea or opinion. A discipline-specific approach would show how different disciplines can use different kinds of analysis (although there will be many similarities).

Go Back to Top

Teaching Sociological Critical Thinking Skills

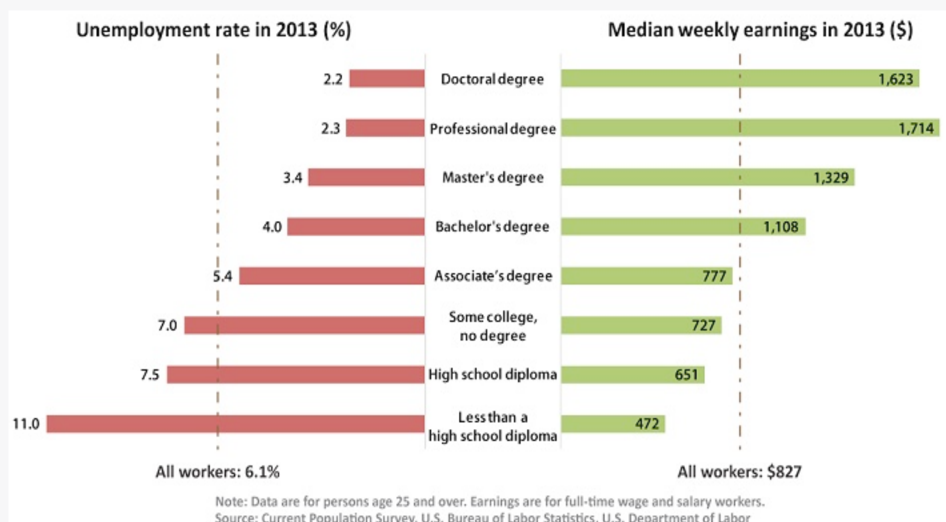
How might we teach a discipline-specific CT skills lesson? Let us take sociology as an example. What follows are two reading/speaking lessons based on extracts of a sociology power point presentation and an academic article. For the first lesson the class would first discuss the warm-up questions (Extract 1):

Extract 1: Education and Income

1. Is there a relationship between education and income; in other words, do better-educated people earn more than less educated ones?
2. What kind of evidence would be useful to answer this question?
3. What analytical tools, concepts or theories could we use to find a solution to this problem?

The class will then examine the information in the table below (Figure 1).

Figure 1: Earnings and Unemployment Rates by Educational Attainment



They will read the power point extract (Extract 2):

Extract 2: Analysis of Education, Income and Unemployment

From Figure 1 you can conclude that educational level is extremely important for increasing the income of all men and women. Statistical analysis of the earnings and unemployment data reveal a strong correlation between educational attainment and earnings, one the one hand, and an inverse relationship between educational attainment and unemployment, on the other hand. This means people with professional degrees make more, on average than those with only bachelors degrees and so on. Also, the higher your educational attainment, the lower the probability your will be unemployed.

In this extract, the author wants to show that people's educational level has an important effect on their income. Sociology frequently uses statistical tools such as mean and median, and utilizes earnings and unemployment level as evidence, and it breaks down these data according to educational level and presents them in visual form like the horizontal bar graph in Figure 1 in order to emphasize the link between education on earnings and unemployment.

What isn't very visible is the kind of reasoning that sociology uses. Sociological reasoning often tries to establish a causal relationship between two or more phenomena (Extract 2) and then it establishes a mechanism that explains how one causes or is connected to the other. It establishes a causal relationship between two variables, in this case education and income and next proposes a mechanism or theory that can explain why A causes B. The class can brainstorm reasons, theories, and mechanisms that can explain this phenomenon and then compare them with the relevant sociological theories and concepts at a later lesson.

Extract 3: Children's Educational Success and Parental Income

4. Is there a relationship between children's success at school and their parent's earnings?
5. What kind of evidence would be useful to answer this question?
6. What analytical tools, concepts or theories could we use to find a solution to this problem?

The class will then examine the information in the table below (Table 1).

Table 1: College Board Data for the Scholastic Aptitude Test

SAT	Test-Takers		Critical Reading		Mathematics		Writing	
	Number	Pct	Mean	SD	Mean	SD	Mean	SD
Family Income								
$0 - $20,000	105,680	14	433	105	461	119	428	104
$20,000–$40,000	125,982	17	463	103	481	111	453	101
$40,000–$60,000	109,444	14	485	102	500	108	473	101
$60,000–$80,000	97,649	13	499	102	512	106	486	101
$80,000–$100,000	83,659	11	511	103	525	106	499	102
$100,000–$120,000	72,776	10	523	103	539	107	512	104
$120,000–$140,000	38,556	5	527	103	543	106	517	104
$140,000–$160,000	29,437	4	534	103	551	106	525	105
$160,000–$200,000	35,474	5	540	105	557	108	534	108
More than $200,000	57,487	8	567	107	589	107	566	110
No Response	908,335		496	117	515	119	490	116

The data show that SAT scores go up in perfect tandem with $20,000 family income amounts. Willingham (2012. p. 33) asserts:

On average, kids from wealthy families do significantly better than kids from poor families. Household wealth is associated with IQ and school achievement, and that phenomenon is observed to varying degrees throughout the world. With a more fine-grained analysis, we see associations with wealth in more basic academic skills like reading achievement and math achievement. And the association with wealth is still observed if we examine even more basic cognitive processes such as phonological awareness, or the amount of information the child can keep in working memory.

For another related lesson the class would first discuss the warm-up questions (Extract 3).

In this article extract, the author wants to show that family or household wealth has an important effect on children's educational success. Sociology frequently uses numerical data such as household earnings, test scores, etc. as evidence, and it presents them in visual form like the chart in Table 1 above. The class can brainstorm reasons, theories and mechanisms that can explain this phenomenon and then compare them with the relevant sociological theories and concepts at a later lesson. The class can later focus on the language of correlation (Extract 4).

Extract 4: The language of Positive and Negative Correlation

Positive Correlation

- The higher _________, the more _______
- A is strongly connected to B
- A is associated with B
- There is a strong relationship between A and B
- Negative or Inverse Correlation
- A is inversely correlated with B
- There is a negative relationship between A and B
- The more __________, the less ________

Conclusion

CT has often been conceptualized in EFL as a component of reading comprehension skills. Nevertheless, one fundamental assumption EAP makes regarding the nature of CT is that it is generic and transferable, that it can be taught as a disembodied skill. This article argued that integrating the teaching of CT within a discipline or subject is particularly effective for EFL/EAP learners. It also proposed a model for teaching CT in the EFL/EAP classroom.

References

Atkinson, D. (1997). A critical approach to critical thinking in TESOL. *TESOL Quarterly*, *31*(1), 71–94.

Barnett, R. (1997). *Higher education: A critical business*. Buckingham: Society for Research into Higher Education and Open University Press.

Bean, J. (2011). *Engaging ideas: The professor's guide to integrating writing, critical thinking, and active learning in the classroom*. San Francisco, CA: Jossey Bass.

Davidson, B. (1998). A case for critical thinking in the English language classroom. *TESOL Quarterly*, *32*(1), 119–123.

de Chazal, E. (2014). *English for academic purposes*. Oxford: Oxford University Press.

Ennis. R. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher*, *18*(3), 4–10.

Ennis, R. (1987). A taxonomy of critical thinking dispositions and abilities. In J. B. Baron, & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9–26). New York: Freeman.

Gieve, S. (1998). A reader reacts: Comments on Dwight Atkinson's "A Critical Approach to Critical Thinking in TESOL." *TESOL Quarterly*, *32*(1), 123–129.

Hawkins, M. (1998). Apprenticing nonnative speakers to new discourse communities. *TESOL Quarterly*, *32*(1), 129–132.

Jenicek, M. & Hitchcock, D. (2005). *Evidence-based practice: Logic and critical thinking in medicine, XV. Chicago, IL: American Medical Association Press.*

Johns, A. (1988). The Discourse Communities Dilemma: Identifying Transferable Skills for the Academic Milieu. *English for Specific Purposes*, *7*, 55–60.

Judge, B., Jones, P., & McCreery, E. (2009). *Critical thinking skills for education students*. Exeter: Learning Matter.

Moon, J. (2008). *Critical thinking: An exploration of theory and practice*. Abingdon & New York: Routledge.

Pennycook, A. (1999). Introduction: Critical approaches to TESOL, *TESOL Quarterly*, *33*(3), 329–348.

Willingham, D. (2007). Critical thinking: Why is it so hard to teach? *American Educator*, *Summer*, 8–19.

Willingham, D. (2012). Ask the Cognitive Scientist Why Does Family Wealth Affect Learning? *American Educator*, *Spring*, 33–39.

Author Bio

Percival Santos (PhD, London School of Economics) was Assistant Professor of English for Academic Purposes at Dongbei University of Finance and Economics, Dalian, China. He currently teaches geography, sociology, history and religious studies at London Academy, London, UK. He has published several journal and magazine articles on critical thinking, research methods, academic writing, and the teaching of the social sciences.

LINC FIELD TRIPS IN THE POST-FUNDING WORLD: FIVE CHANGES

By Joseph Ng, adapted from his TOSCON 16 Presentation

Arms laden with Food Basics groceries, we trudged south west across snowy Victoria Park Avenue and Sheppard Avenue East and took shelter kitty corner in the iconic Johnny's Burgers in Scarborough before heading for Round Two of grocery shopping, a multicultural one at Hong Tai Supermarket.

Just one of dozens of field trips in recent years to places far and near to help my students not only to learn and use the language but also to settle in their new home of Canada. Looking back, it's hard not to notice how different field trips have become in the post-funding world. Gone are end-to-end, fully paid for rides in yellow school buses and padded seats to subsidized tourist destinations up the CN Tower or over at the zoo. Out they went with roundtrip TTC token reimbursements and all the rest.

We watched in learnt helplessness as those school trips, cushy on hindsight, went from three or four a year to two and one great big one, and then nothing at all. Thanks to the likes of the Toronto Region Conservation Authority's multicultural program, we still get the occasional and limited subsidy for a school bus to be shared by all classes, but we have to pay the entrance fee at the destination.

In short, it has been a belt-tightening time, and for those of us who would commend the educational glories of field trips, it was also time to walk the talk. And some walk it was, having to step out in at least five areas.

Risk Waivers and Photo Releases

"I hereby release and forever discharge" the organizers "of all actions, injury, damage," quoth the intake paperwork every new student has to sign, adult and child participants included, before it goes into their dossier. From then on until the participants leave, permission has been granted for their photos to be taken for internal sharing and publicity purposes. And the amount of permission papers saved is substantial.

As a caveat, every organization is different, and how and how often a teacher organizes field trips will depend much on the decision that organization makes on risks in this regard. But having the paperwork out of the way does free the teacher to other instructional mischief.

Openness to Transit and Car Pooling

Free from dependence on school buses, we need to decide on modes of travel, which throws up some interesting challenges.

The first is the concept of students meeting, not in the car park of the LINC school but by prior appointment with classmates at a designated TTC station, in our case Don Mills. Depending on the level of the challenge, students are sometimes told to go directly to the event downtown, a free concert for example. On a recent trip to the Toronto Reference Library, my new Yemeni student, who had been holed up in Fort McMurray, Alberta, for her first three years before coming to Toronto, excitedly told me upon arrival at the library entrance that the subway trip was for her a maiden voyage, and she'd stayed up the night before loading screenshots of the route map onto her smartphone. More typically, the students would have studied the route and exchanged phone and Whatsapp details in the days leading up to the trip.

The availability of private cars provided by volunteers or students varies greatly from one cohort to another. Some batches drive a fleet of higher-end SUVs. Others can't afford to take the bus and so walk to school, but often they're mixed. Liability and insurance, like school board field trips, would be covered by the car owners' insurance. But having cars and drivers to car pool simply opens up places beyond walking and TTC range. Daniel Ren heard our call for drivers and cars through our class alumni page on Facebook and responded. Having left my LINC 4–5 level the year before and moved to Mississauga, where he works as a computer programmer from home, he nonetheless made the journey back to his alma mater in North York, picked up a carload of new classmates to drive them to and from Niagara Falls, and finally headed back home on the congested 401.

Up Front Freedom to Find Free Fields

We finally made Niagara Falls after years of calculating travel times by coach or school bus, contacting the casino bus, and otherwise talking about it. It was the Harper era, and we really dug in to the War of 1812 lore, standing atop the cemetery at Drummond Hill and recalling that fateful summer battle with that Great Invader Enemy Down South before soothing our nerves on the winery route on the way back.

Or we'd skip the Falls on some trips to learn about navigating the Great Lakes at the Welland Canal Museum at Lock 3 or to check out agricultural products at Great Mountain Ginseng Farm.

But come August, we'd hit the 404 and head north to Stouffville to Farintosh Farms to pick corn. Owner Guy Farintosh would teach us how to eat corn right off the cob right off the stalk. We just had to pay for what we took home, bunches of hand-picked corn and beans and loads of watermelons and squashes from the barnhouse. The photos and memories came free.

Within the City of Toronto, government and nongovernment destinations abound for field trippers. City Hall, Nathan Phillips Square, and Queens Park are fair game in all seasons, even more when Council and subcommittees are in session, in outdoor skating season, or during Caribana or Tasty Thursdays. Toronto's two well-hidden ski centres offer discounted lessons with rental equipment, but always make sure to offer a free option for those not paying, where they can not only take pictures of classmates in funny poses but also learn about Ski Patrol and the technology of snow sports, touching skis, snowboards, and bindings, watching demos, sitting by the fireplace over a mug of hot chocolate, digging into their picnic lunch, etc.

The Art Gallery of Ontario has fabulous neighbourhood and community access programs that welcome needy individuals and LINC classes to view their favourite Group of Seven paintings or Henry Moore sculptures for free. For us, what was even more intimate were visits to the Durdy Bayramov Art Foundation, a unique museum created out of a private home along Bayview just north of York Mills. The founder and president of the foundation herself, Keya Bayramova, would be there to show us her late father's paintings and photos, which added a multicultural dimension through the genius of a Soviet-era Turkmen artist. The museum graciously accepts requests for LINC visits, and I'm immensely gratified to have met Keya when she first walked into my LINC class a few years ago.

Challenge to Integrate the Curriculum

The LINC curriculum is theme-based, with topics spanning public transit, geography, culture & society, government, and environmental issues to getting employment and starting a small business. The better-integrated the field trip with the theme, the greater the learning and traction of activities before and after. So it's no serious loss when we can't afford to visit Casa Loma or CN Tower these days as they are difficult to fit into the curriculum anyway. And they can probably be accessed for free on the one-year pass that comes when students get their Canadian citizenship and have to leave the LINC program.

The purpose of this sharing is not, of course, to be exhaustive in listing out all the field trip destinations in Toronto. That would futile as nobody knows all the possibilities out there, and things keep changing. However, if LINC teachers care to integrate field trips into their curriculum, there is a platform on which to pool their secrets consisting of a Facebook page and an online spreadsheet.

Catalyst to Pedagogy

Technology can be an aid to field-trip management. My students are told to check the class blog before heading out on any field trip. If they forget the meetup or destination time or venue or if there are any last-minute cancellations or changes, they can find out before it is too late. On the trip itself, they may have to take selfies or wefies of the things and places they see. At Toronto Reference Library, the pictorial artefacts they had to pose with and

upload to their blogs included various newly renovated features, the 3D printer, and the ESL section. Of course, they would typically also have to write a trip report on their blogs when they get back to class on Friday.

Yet technology is not the only catalyst to successful field trips. On trips to the AGO, for instance, we have used Dictation Triptychs, three-column handouts that marry the intense interactivity of jigsaw classrooms with the intense use of the four skills through dictation, all to accelerate the use of target functions and forms that are natural and real. So having studied the life of Canadian artist William Kurelek, we practised and simulated a gallery dialogue about two highly contrastive pieces of his works before playing it out in that very location itself in AGO, the Kurelek Room.

So there we were in the Hong Tai Supermarket, arms temporarily relieved as previously-bought groceries went into wheeled baskets, going around in pairs, Chinese student with non-Chinese, finessing the treatment and recipe of raw jellyfish and other merciless delicacies. Which seems as good a place as any to end my rambling journey.

VIVA LA LINGUA FRANCA: ON THE ART OF THE APOLOGY

By Eufemia Fantetti

My father is a fan of an Italian-dubbed, German soap opera called *Tempesta D'Amore*, saga television at its very best. Set in a five-star hotel with characters who cheat, lie, and connive like cousins at a Trump family-reunion picnic, it would occasionally pull me into the vortex of daytime viewing, spending the entire episode trying to figure out the Byzantine plotlines of complicated relationships.

"I don't get who these people are."

"They brother and sister-in-law," my father would explain, "but his wife, her sister, was die. Now she's love him and she's pretends she sick so he's no leave."

"But why is she 'remembering blood'? I don't get what they're arguing about."

"They was kill and bury somesbody together."

"I can't believe you called this a 'family' show."

In the end, my effort to improve my Italian through the technique of TV watching was a waste. After all, it is highly unlikely that I will ever need to know how to say, "The man you call father is not your real dad."

The Germans are paragon folk to my Teutophile father, hard-working and level-headed: "You can see from the shows how they talks with each other, they no so much boom-boom, yelling each other, upsets but no making sense. If someones makes mistake, they ask for the sorry, they make apologize."

Sure, *Sturm der Liebe* as a standard for etiquette is a bit of a stretch, but I did think my dad was right. There is an Art to the Apology, and how one approaches the task matters—it is fundamental to our progress as people on this planet. Canadians are considered mass producers of the term—though truth be told, if saying "sorry" instead of "excuse me" is what earned us this reputation as *über*-polite, is it really deserved?

Acts of atonement are headline-worthy in Canada. Remember Elbowgate? Trudeau offered three apologies for losing his patience. My favourite story along these lines is the fellow who sought out a classmate sixty-five years after punching him on a playground by placing an ad in the Canadian Jewish News. The *Toronto Star* documented the meeting of Thomas Caldwell and Howard Rosen and every time I read it, I get that tight-squeeze feeling in my

ribcage that my heart is expanding and rubbing up against the bones.

In sharp contrast, I get a shrinking feeling of disappointment when I think back to the first ESL class I ever taught. Of course, mistakes were expected but my botched attempt at apologizing remains with me. How does one translate an apology? Is "I'm sorry" understood the same way in every language?

Years ago, when I was travelling through Pushkar, India, I kept thanking the young man at the internet café who helped me figure out what was happening as I blogged my adventure. After a week of this "thank you" and "sorry" stuff, he told me to knock it off. Thank-you-sorry was the realm of the British, he said, a strange politeness that he insisted created an awkward and unpleasant distance—as if I thought I was better than him.

I sputtered, "I'm sorry—*what*?"

Reflecting on his response, I understood what he meant; his attitude was my Italian family's way of thinking as well. If I express thanks when my aunts offer me anything, they both answer "For what?" with an edge of hostility in their voices—as if to say I'm insulting them. *Sorry* is similarly dismissed with a shrug, "You're human, you'll make mistakes."

Then there's a differing opinion on the apology offered by one of my uncles: *Sorry* was the sucker born every minute and you were better off keeping your head down and never admitting anything. This comes from an old belief that silence was better than confessing culpability, probable cause, or any other words that could get tricky.

My father disagrees. "Is important to know how asking forgive," he says. "Maybe the person will be say 'No way, go away,'' but you shoulds be ask."

This brings me back to my first ever ESL class. The moment I feel worst about is one I couldn't have predicted or prepared for. Two weeks into the job, on the Friday after a long week of teaching, meeting new people, memorizing names, speaking too fast for literacy level learners, and bringing in every manner of self-created handouts or adjusted materials, a student put up his hand and said, "Teacher, ex-squeeze me, but—"

I heard nothing else. It started as a giggle and then, the more I tried to contain it, my reaction turned into a full blown, holding-my-stomach laugh. I apologized immediately, in between breaths. After I corrected his pronunciation, I attempted to explain why the mispronounced word was funny. I acted out why he should never say *ex-squeeze me* and how it could be mistaken as insincere or false. I was so exhausted that every attempt to clarify made me break out in fits of laughter. He laughed as well but his face was bright pink. There were only a few other students in the class, many had left at the break. We all practiced the pronunciation and I apologized for laughing.

He nodded, and never came back to class. Each day that he wasn't there, I silently cursed myself as I took attendance. Finally, after a week of berating myself, I called him, reaching him on his cell with the number I had in the class list. It took a few minutes for him to

figure out who I was, not because he'd forgotten me but all the confusion of being literacy level is amplified over the phone. I told him school was closed for a holiday on the following Monday. He understood, and said he'd found work. In the pause before hanging up I wondered why I hadn't looked up how to say "I'm sorry" in Turkish.