Korea TESOL Journal

KOTESOL
대한영어교육학회

Korea Teachers of English to Speakers of Other Languages

LOOKING FOR FLEXIBLE STUDY OPTIONS?

The University of Birmingham's Department of English Language and Applied Linguistics offers flexible personal development opportunities for professionals wishing to develop their skills and expertise. Our distance learning Masters programmes are delivered part-time over 30 months, to fit around your existing commitments.

APPLIED LINGUISTICS MA

This programme is for professionals wishing to further their personal development, and those who are interested in learning more about possible applications of language research. You will study topics including corpus linguistics, sociolinguistics, lexis, functional grammar, spoken and written discourse, and multi-modal communication.

TEACHING ENGLISH TO SPEAKERS OF OTHER LANGUAGES (TESOL) MA

This programme is for practising teachers of English as a second or foreign language who wish to develop their knowledge of classroom methodology and materials design. You will study topics such as language teaching methodology, second language acquisition, syllabus and materials, pedagogic grammar, lexis, and teaching young learners.

KEY FACTS

- Start in February, April, July, October or December
- Study entirely online
- All modules assessed by written assignment
- Pay per module

Find out more and apply:
www.birmingham.ac.uk/elal-dl

Korea TESOL Journal

KOTESOL

Korea Teachers of English to Speakers of Other Languages

Korea TESOL Journal

About Korea TESOL

Korea TESOL (KOTESOL; Korea Teachers of English to Speakers of Other Languages) is a professional organization of teachers of English whose main goal is to assist its members in their self-development and to contribute to the improvement of ELT in Korea. Korea TESOL also serves as a network for teachers to connect with others in the ELT community and as a source of information for ELT resource materials and events in Korea and abroad.

Korea TESOL is proud to be an Affiliate of TESOL (TESOL International Association), an international education association of almost 12,000 members with headquarters in Alexandria, Virginia, USA, as well as an Associate of IATEFL (International Association of Teachers of English as a Foreign Language), an international education association of over 4,000 members with headquarters in Canterbury, Kent, UK.

Korea TESOL had its beginnings in October 1992, when the Association of English Teachers in Korea (AETK) and the Korea Association of Teachers of English (KATE) agreed to unite. Korea TESOL is a not-for-profit organization established to promote scholarship, disseminate information, and facilitate cross-cultural understanding among persons associated with the teaching and learning of English in Korea. In pursuing these goals, Korea TESOL seeks to cooperate with other groups having similar concerns.

Korea TESOL is an independent national affiliate of a growing international movement of teachers, closely associated with not only TESOL and IATEFL, but also with PAC (Pan-Asian Consortium of Language Teaching Societies), consisting of JALT (Japan Association for Language Teaching), ThaiTESOL (Thailand TESOL), ETA-ROC (English Teachers Association of the Republic of China/Taiwan), FEELTA (Far Eastern English Language Teachers' Association, Russia), and PALT (Philippine Association for Language Teaching, Inc.). Korea TESOL is also associated with MELTA (Malaysian English Language Teaching Association), TEFLIN (Indonesia), CamTESOL (Cambodia), and ACTA (Australian Council of TESOL Associations), and most recently with ELTAM/Mongolia TESOL, MAAL (Macau), HAAL (Hong Kong), and ELTAI (India). Korea TESOL also has partnership arrangements with numerous domestic ELT associations.

The membership of Korea TESOL includes elementary school, middle school, high school, and university-level English teachers as well as teachers-in-training, administrators, researchers, material writers, curriculum developers, and other interested individuals.

Korea TESOL has nine active chapters throughout the nation: Members of Korea TESOL are from all parts of Korea and many parts of the world, thus providing Korea TESOL members the benefits of a multicultural membership.

Korea TESOL holds an annual international conference, a national conference, workshops, and other professional development events, while its chapters hold monthly workshops, annual conferences, symposia, and networking events. Also organized within Korea TESOL are various SIGs (special interest groups) – e.g., Reflective Practice, Social Justice, Christian Teachers, Research, Women and Gender Equality, People-of-Color Teachers – which hold their own meetings and events.

Visit *https://koreatesol.org/join-kotesol* for membership information.

KOTESOL

Table of Contents

Brief Reports

Book Review

Korea TESOL National Council 2019-20

National Officers

President: Lindsay Herron, Gwangju National University of Education
First Vice-President: Bryan Hale, Yeongam High School
Second Vice-President: Rhea Metituk, University of Ulsan
Treasurer: Phillip Schrank, Chosun University
Secretary: Lisa MacIntyre-Park, I-Can! English Academy
Immediate Past President: Dr. David E. Shaffer, Gwangju International Center

Committee Chairs

International Conference Committee Chair: Michael Free, Kangwon National University
Nominations & Elections Committee Chair: Mike Peacock, Woosong Culinary College
Publications Committee Chair: Dr. David E. Shaffer, Gwangju International Center
Membership Committee Chair: Lindsay Herron, Gwangju National University of Education
International Outreach Committee Chair: Allison Bill, Jeonju University
Technologies Committee Chair: John Phillips, Independent Computer and Network Specialist
Publicity Committee Chair: Wayne Finley, Woosong University
Research Committee Chair: Dr. Mikyoung Lee, University of Munich
Financial Affairs Committee Chair: Daniel Jones, Dong-eui University
Diversity Committee Chair: Travis Frank, Banana Kids English Academy

Chapter Presidents

Busan-Gyeongnam Chapter: Sunil Mahtani, Kyungnam University
Daegu-Gyeongbuk Chapter: Kimberley Roberts, Gyeonghye Girls' Middle School
Daejeon-Chungcheong Chapter: Mike Peacock, Woosong Culinary College
Gangwon Chapter: Reece Randall, Gangneung-Wongju Natl. University, Gangneung
Gwangju-Jeonnam Chapter: Bryan Hale, Yeongam High School
Jeonju-North Jeolla Chapter: Allison Bill, Jeonju University
Seoul Chapter: James Tardif, Gimpo Academy Instructor
Suwon-Gyeonggi Chapter: Tyler Clark, Gyeongin Natl. University of Education, Gyeonggi Campus
Yongin-Gyeonggi Chapter: James G. Rush II, Luther University

Korea TESOL Journal

The *Korea TESOL Journal* is a peer-reviewed journal, welcoming previously unpublished practical and scholarly articles on topics of significance to individuals concerned with the teaching of English as a foreign language. The Journal focuses on articles that are relevant and applicable to the Korean EFL context. Two issues of the Journal are published annually.

As the Journal is committed to publishing manuscripts that contribute to the application of theory to practice in our profession, submissions reporting relevant research and addressing implications and applications of this research to teaching in the Korean setting are particularly welcomed.

The Journal is also committed to the fostering of scholarship among Korea TESOL members and throughout Korea. As such, classroom-based papers, i.e., articles arising from genuine issues of the English language teaching classroom, are welcomed. The Journal aims to support all scholars by welcoming research from early-career researchers to senior academics.

Areas of interest include, but are by no means limited to, the following:

> Classroom-Centered Research
> Teacher Training
> Teaching Methodologies
> Cross-cultural Studies
> Curriculum and Course Design
> Assessment
> Technology in Language Learning
> Language Learner Needs

For call-for-papers information and additional information on the *Korea TESOL Journal*, visit our website:

https://koreatesol.org/content/call-papers-korea-tesol-journal

Research Papers

A Bourdieusian Interpretation of English Language Learning: The Case of Korea

Michael D. Smith

Kwansei Gakuin University, Nishinomiya, Japan

In recent years, educational research describing the sociological impact of the English language has drawn increasingly on the theories of Pierre Bourdieu to account for the mechanisms by which ELT imbricates in social stratification. Accordingly, this critical study takes as its analytical focus the Bourdieusian concepts of "habitus," "capital," and "field" in an effort to illustrate the structural and cognitive pressures that drive English language education and thus intergenerational social inequality. Specifically, Bourdieu's model is employed to foster a theoretical comprehension of the post-globalization developmental strategies of the Republic of Korea during a period of sustained political and social reform. It has been shown that the interplay between Korea and internationalization has resulted in the identification of English as a resource crucial to the accumulation of capital within the transnational arena. This conflation of internationalization and *Englishization* acts not only as an instrument for responding to global pressures but a vehicle for elite privilege reinforcement, sustaining circular forms of socioeconomic inequality on the basis of language proficiency – to the advantage of the agentive forces behind the local dissemination of English and the disadvantage of broader subaltern populations. As a consequence, EFL instrumentalization within the Korean sociolinguistic field is illustrative of the measures by which dominant classes propagate self-aggrandizing values and norms via the manipulation of cultural capital, thereby achieving the hegemonic subjugation of subordinate groups via structural and ideological mechanisms.

Keywords: English language education, social reproduction, stratification, Bourdieu

BACKGROUND

The determinative effects of what Bourdieu (1989) defines as *habitus*, *capital*, and *field* in the structuring and maintenance of social reproduction have long held the attention of academia. Within the field of applied linguistics, for example, Bourdieusian interpretations of the interrelations between English as a foreign language (EFL) and the structuring of social classification have emerged as a fundamental component of sociolinguistic criticism (see Lee, Han, & McKerrow, 2010; Silver, 2004). At the core of these inquiries is an attempt to uncover English language adoption as a dominant contributor to, and indeed, *driver of*, socio-educational inequity within non-native speaking locales. Given the observable association between neoliberal, globalized policy enactments and English language spread, some (Phillipson, 1992) have sought to categorize this dynamic as an all-too-deliberate Western-driven imposition that aims to enhance the social, political, and financial dichotomy between the Global North and South.

Contesting interpretations, however, attempt to account for the sociocultural features distinct to each setting and the potential for domestic actors to hegemonically structure local educational initiatives in an effort to mediate self-determined outcomes (Crystal, 2003). Accordingly, while macro-level frameworks describing the impact of global English are useful in interpreting the causal factors associated with the language's ideological structuring and sustained dominance, it is advisable that inquiries account for micro-level, individually situated context if they are to lay legitimate claim to comprehensiveness. In this regard, Bourdieu's interpretation of *habitus* offers a compelling lens through which to analyze and understand "the relation between structure and agent in the context of practice" (Chandler, 2013, p. 469) as situated within both the local and global domains. Moreover, an appreciation of Bourdieu's anastomotic concepts of *capital* and *field* is beneficial when attempting to illustrate the mechanisms by which EFL policy discourses impact upon the process of linguistic hierarchization – or, to be more specific, the manner by which numerous non-native-speaking locales have agentively assigned asymmetrical economic, social, and cultural capital to English, thereby framing their particular linguistic fields in overt ways (Silver, 2004).

The environment that has shaped foreign language education in the

officially monolingual Republic of Korea (hereafter, Korea), for example, has borne witness to a near-unparalleled pursuit of English. Specifically, English language learning (ELL) within Korea has been institutionalized under the guise of the nation's self-insertion into globalization and its accompanying shift toward neoliberalist political-financial practices. As a consequence, the capability to speak English is now regarded by many Koreans as a legitimate conduit for social and economic advancement, and a "major criterion in education, employment, and job-performance evaluation" (Song, 2011, p. 35). By way of example, the nation's *yeongeo yeolpung*[1], or "English fever," is so ingrained into the public consciousness that Koreans are recorded to have spent $16 billion annually on education (Kim, 2015). Lawrence (2012, p. 71) noted that private language institute tuition fees accounted for 73 percent of total domestic outlay, compared to the 24 percent provided by government-funded schooling.

As this study will illustrate, the mechanisms behind Korea's fascination − perhaps fixation − with English are both complex and anchored to features of local culture that have remained socially encoded for centuries. That Koreans are willing to learn English is self-evident; nevertheless, given the cognitive and financial costs borne the learner, there are ethical as well as academic motivations for this investigation. It should be noted, however, that despite the noticeable impact of globalization − and thus, the political machinations of the Global North − on Korean educational policy, it is not the purpose of this investigation to provide a post-colonial description of global-level Centre−Periphery hegemonic oppression (Phillipson, 2014). Rather, *locally agentive* educational policies will be critically examined in conjunction with socio-historical features consistent to the Korean setting to provide a micro-level account of the contemporary positionality of EFL education and its impact on the Korean EFL learner. As a consequence, this paper proposes a Bourdieusian investigation of Korean foreign language policy intentions in an effort to expose the degrees of interaction between institutionalized linguistic habitus and the scope of Korean EFL capital within that specific sociolinguistic field.

BOURDIEU'S HABITUS, CAPITAL, AND FIELD

Bourdieu frames his employment of habitus as "a set of historical relations 'deposited' within individual bodies in the forms of mental and corporeal schemata of perception, appreciation, and action" (Bourdieu & Wacquant, 1992, p. 16). Put simply, habitus is a conceptual tool fundamental to understanding the innate behaviors, competencies, and dispositions that one inherits and cultivates as a result of their codified social grouping and continued life experiences. Consequently, habitus is central not only to the successful navigation of familiar social environments; it also extends to an individual's intuitive predilection towards a range of cultural objects, including the means why which they absorb, process, and utilize language. Through the lens of *linguistic habitus*, Bourdieu (1977) characterizes language as one's embodied system of "objective structures" (Bourdieu, 1977, p. 662), determined initially by one's socially accumulated linguistic competencies, before assimilation into a distinct cognitive and somatic disposition. Due to this process, habitus is inherently linked to an individual's ability to anticipate strategically − and thus exploit − their accumulated linguistic resources, such as pronunciation, diction, lexicon, and foreign language competencies.

Thus, all forms of discourse act not only as a mechanism of communication but as an agency of dominance (Bourdieu, 1982) imbricating in "power struggles over what is and what is not regarded as acceptable and valuable" (Zotzmann, 2013, p. 253). Accordingly, language exhibits analogous behavior with that of an economic market (Bourdieu, 1977, p. 662), involving the transaction of linguistic capital that is subject to variation in market worth, depending on its specific sociolinguistic context (British Received Pronunciation possessing higher prestige over Estuary English, for example). As a consequence, the presence of the market of linguistic exchange not only facilitates *but is central to* the continual reconditioning of individual habitus. Bourdieu (1977) reasons:

> Situations in which linguistic productions are explicitly sanctioned and evaluated, such as examinations or interviews, draw our attention to the existence of mechanisms determining the price of discourse which operate in every linguistic interaction (e.g., the doctor–patient or lawyer–client relation), and more generally in all social relations. It follows that agents continuously subjected to the

sanctions of the linguistic market, functioning as a system of positive and negative reinforcements, acquire durable dispositions which are the basis of their perception and appreciation of the state of the linguistic market. (p. 654)

While the connection between habitus and one's first language (L1) is overt, compelling are the mechanisms by which linguistic habitus contributes to second (L2) or foreign language acquisition. One is afforded a higher likelihood of winning what Bourdieu terms "the games of culture" (Bourdieu, 1984, p. 54) when their habitus exhibits commonality with cultural markers laid by broader society (thereby obeying the specific logic of the *field*). Thus, when a community elects to allocate symbolic resources to a foreign language, one's linguistic practice must be modified to accommodate the emergent dynamics of the market. Fundamentally, linguistic habitus represents the corporeal embodiment of linguistic capital, itself an individual feature of the broader forms of capital.

Bourdieu (1986) develops the concept of capital beyond its narrow economic interpretation to incorporate mutually constitutive social and cultural features. Accordingly, while economic capital represents those resources that may be converted directly into financial assets, such as money and property (Cho, 2017), an individual's social capital is determined proportionally by their cumulative social network – and the opportunities and obligations that those connections facilitate. Cultural capital, meanwhile, is described as those symbolic assets, both tangible and intangible, that are acquired by virtue of one's distinguishing level of social stratification – thereby representing an implicit currency that is traded during the navigation of culture. Specifically, cultural capital is structured around a triumvirate of the institutionalized, objectified, and embodied forms.

The formal measurement of one's cultural competence or authority, typically in the form of academic or professional credentials, is representative of the institutionalized state. In its objectified form, cultural capital is exemplified by the hierarchized material objects, or "cultural goods" (Bourdieu, 1986, p. 243) that are used to distinguish one's social class or degree of capital – which may include items such as clothing, automobiles, or art. Finally, the embodied state of cultural capital is denoted by the forms of knowledge that reside within an individual in the shape of enduring dispositions, both physical and

cognitive (Bourdieu, 1986), which are inherited via the process of socialization. Crucially, as an individual consolidates embodied forms of cultural capital, those items metamorphosize into a form of habitus, and thus lack the capacity for instantaneous transmission (Bourdieu, 1986, pp. 244–245). In this regard, one of the earliest forms of capital in the embodied state is that which is acquired via language, such as an individual's L1 or regional dialect.

Significant in Bourdieu's description of capital is the role of scholastic achievement in the reproduction of social stratification via the "hereditary transmission of cultural capital" (Bourdieu, 1986, p. 244). Cultural capital is, like all forms of currency, an item that can be retained, invested, or consumed during the procurement of desired resources (Kingston, 2001). That the social return of education is presumed to be dependent both on the qualification's institutionalized worth and the social capital that one has accrued via networked credentials (Bourdieu, 1986) is thus unsurprising. This dynamic facilitates the conversion of cultural capital into economic capital via desirable forms of employment, which in turn, permits a user to aggrandize further and hereditarily transmit cultural capital. Consequently, the educational institution reflects, and is indeed responsive to, "the cultural orientation of the dominant class" (Kingston, 2001, p. 89). This interlinked process of conversion, therefore, represents a significant determiner of social inclusion and exclusion (DiMaggio, 1982).

All forms of capital are transformed automatically into symbolic capital when an actor enters a specific social environment or field. Subsequently, each field possesses a distinct and unquestionable orthodoxy or *doxa*, according to which the group at an aggregated level will evaluate an individual to ascribe him or her their social position within that specific environment. Given that individuals inhabit differentiated positions within distinct fields, it is consistent that the mechanisms by which they utilize language are similarly differentiated (Zotzmann, 2013). As a consequence, fields act as spatial arenas in which capital in all of its forms is accrued and, crucially, contested between accumulating actors. Thus, it is prerequisite that one observes the broader processes by which power is engendered and regulated between fields, agents within those fields, and the linguistic and broader habitus of the agents involved if one is to comprehend the distinct linguistic dynamics of a locale (Greenfell, 2012).

THE HISTORICAL CONTEXT OF ELL IN KOREA

When assessing Korea's complicated relationship with ELL, one must first be appreciative of the significance of both East Asian traditionalism and the opening of the peninsula to outside influence in 1876. Prior to this, Korea's political and societal values had remained effectively unchallenged for centuries, manifested in the saturation of Neo-Confucianist ideologies that were embedded culturally during the late Goryeo period (918–1392). Most notably, filial piety, dedication to the nation-state, strict adherence to hierarchy, and cultivation of the self via scholarly pursuit (Han, 2007). The fate of the peninsula was to be irreversibly altered in 1876, however, following the endorsement of the Japan–Korea Treaty of Amity[2], which opened the peninsula to foreign diplomatic relations. Agreements with Western nations soon followed, including the United States in 1882 and the British Empire a year later (Kim, 2005), with the ensuing influx of foreign service personnel, traders, and, crucially, Christian missionaries proving instrumental in forging the country's early English language teaching (ELT) institutions.

In Bourdieusian terms, this period was fundamental to the reshaping of various Korean fields and the forms of capital contained therein. Nevertheless, while the influences of external cultural ideologies on the reorganization of feudal Korea are apparent, the mechanisms with which they were employed were connected initially to pre-existing social stratification dynamics. For instance, while a dedication to academic pursuits and the Neo-Confucian ideal of the learned gentleman was, and indeed remains, an emblematic component of local cultural discourse, access to elite education was highly restricted. The civil service examinations that acted as the gateway for feudal-era prestige and power were tightly controlled, with students typically drawn from Korea's powerful ruling aristocracy (Palais, 2014, p. 138). For instance, the alumnus of Korea's first state school by modern standards[3], the *Yugyong Kongwon* (Royal College), consisted exclusively of the sons of high-ranking officials (Yi, 1984), who were schooled as English language interpreters in an effort to ease diplomatic negotiations with foreign dignitaries. Given that these roles were dependent upon requisite heritage and correlative social capital, the utilization of ELT in the maintenance and replication of asymmetrical social reproduction during this period is manifest.

Nevertheless, English rapidly became a symbol of egalitarianism and democracy within Korea (Shin, 2007), with several non-government affiliated mission schools delivering education to the underprivileged: recruiting students irrespective of heritage, wealth, or gender. Indeed, mission schools were the only source of education for women at this time, with Korea's first female school, the *Ewha Haktang*, still thriving today in the form of Ewha Womans University, one of Asia's largest and most prestigious female-exclusive educational institutes (Kim, 2007). English would subsequently develop into a significant component of local tertiary education, with fellow elite universities, including Seoul National, enhancing ELL's institutional capital by employing English language testing as a means of screening prospective students (Kim, 2011) – a trend that continues to this day.

Following the Korean War and the signing of the Korean Armistice Agreement in 1953, the country entered a period of extreme poverty that was to last beyond the deposition of the corrupt Lee Seung-man government and subsequent coup of May 16th, 1961. Nevertheless, by the late 1960s, Korea had entered a cycle of modernization that was, in part, manifested by President Park Chung-hee's commitment to human capital development via uniform basic scholarship. Specifically, Park manipulated the Confucian ideals of communalism and scholarly progression during the *Saemaul Undong* (New Community Movement – Korea's compressed modernization period), providing unparalleled levels of social cohesion and the highly literate workforce that was to be instrumental in facilitating the dramatic phase of Korean economic development commonly referred to as the "Miracle on the Han River" (Kleiner, 2001, p. 254).

The post-Korean War period is further characterized by sustained neo-colonial exposure from the West, particularly from the US, and gradual assimilation into Western ideologies, including globalization, neoliberalism, democracy, and consumerism (Kim, 2000) – with EFL positioned as a mechanism fundamental to the realization of these concepts (Lee, Han, & McKerrow, 2010). Continued exposure was, in time, facilitated by a newly emergent power stratum, consisting of repatriated Koreans who had fled the previously war-ravaged peninsula or received US-funded Western educations as part of the agreement that ensured Korea's participation in the Vietnam conflict. By the early 1980s, returnees were positioned in highly influential government roles that allowed them to advocate the perceived benefits of EFL and foreign

schooling. Yim (2007) and Jeon (2009) also note that the announcements that Korea was to host both the 1986 Asian and 1988 Olympic Games provided further encouragement. Specifically, the opportunity to showcase an educated and sophisticated society coincided with the acute period of Korean internationalization known locally as *segyehwa*[4], providing the impetus for a boom in ELL that has witnessed EFL emerge as a fundamental component of Korean globalization discourse.

KOREA'S POST-*SEGYEHWA* ELL FIELD

The overt connection between education and social stratification, while far from unique to the Korean setting, has been recognized as a sustained feature of local social reproduction. As noted by J. Lee (2010), EFL, in particular, is "consumed as a symbolic measure of one's competence and is associated with job success, social mobility, and international competitiveness" (J. Lee, 2010, pp. 246−247). Thus, the interplay between Korea and internationalization has resulted in the identification of English as a resource crucial to the accumulation of capital within the transnational arena. For example, *segyehwa* policy is commonly referred to by local linguists (Jeon, 2009; Song, 2011) as being central in propagating, from the governmental level to the public, the significance of EFL capital − and thus, the language's value within the market of linguistic exchange. This was illustrated directly in contemporary educational reform, specifically the Korean Ministry of Education's Sixth (1995) and Seventh (2000) National Curricula (cited in Chang, 2009, p. 88), which as described by Chang (2009), emphasized the belief that "if Korea is to function effectively as a nation in the era of globalization, then Korean people must be able to communicate effectively in English" (p. 94). As a consequence, EFL functions locally as an indicator of internationalization and modernity, and, by these measures, its acquisition defines the means by which social agents are legitimized and integrated within an industrialized *global* Korea.

Subsequently, English has been ideologically positioned as requisite not only to the maintenance and enhancement of national competitiveness but as a "tool for Korea to *survive*[5] in the international community" (Yim, 2007, p. 37). As noted by Lee, Han, and McKerrow (2010), this modality was particularly prevalent during the presidency of

Lee Myung-bak. Specifically, Lee called upon Social Darwinist allegory to depict English as a "weapon" within the "battlefield" of globalization (cited in Lee, Han & McKerrow, 2010, p. 338), thereby dogmatically manipulating *survival of the fittest* rhetoric in an effort to enhance the symbolic worth of EFL's institutional-cultural capital and, as a consequence, conserve the dominant linguistic habitus.

Given the context described thus far, it is wholly unsurprising that ELL has been embraced as a mechanism for Korean self-advancement. English's association with Korean modernization discourse and initial circumvention of the cultural-social capital features necessary for participation in education has left an indelible mark on the habitus of local agents and, indeed, the broader Korean field. Due to these dynamics, several local academics (J. Lee, 2010; Shin, 2010) have interpreted Korean EFL instrumentalism as being structured in such a manner that it serves to conserve and replicate a distinctly asymmetric social-relation structure. Song (2011), for example, notes that

> English has been recruited, in the guise of globalization, to exploit the meretricious ideology of merit to the advantage of the privileged classes and the disadvantage of the other classes of the society. English in South Korea cannot be understood fully unless it is recognized that its importance has not been as much engendered by globalization, as it has been resorted to as a subterfuge to conceal where the responsibility for inequality in education lies within the society. (Song, 2011, p. 35)

Likewise, Shin (2010) explicitly calls upon Bourdieu to describe the elite-driven symbolic capital attached to EFL – notably "authentic" forms of English acquired at prestigious Western universities – as reflecting the desire to reproduce pre-existing social positionality via the creation of a "new capital of distinction" (Shin, 2010, p. 11). This is achieved not only by the aforementioned forms of inner-circle[6] education that are accessible to only the wealthiest of Koreans but the demand by local educational and industrial institutions for standardized EFL assessment mechanisms, such as those provided by the Test of English for International Communication (TOEIC) and Test of English Proficiency (TEPS)[7]. According to Jambor (2012), Korea's two most prestigious technical universities, KAIST (formally the Korea Advanced Institute of Science and Technology), and POSTECH (Pohang University of Science

and Technology) both employ EFL as the sole means of communication whilst, as mentioned previously, almost the entirety of elite-level local educational institutions utilize standardized EFL testing measures as a means of filtering prospective students.

Moreover, Song (2011), notes that "over 90 percent of employees in manufacturing and export industries are continuously assessed for their English competence" (p. 42). Thus, EFL in the context of the Korean socio-educational field is embedded firmly within the structures of institutional capital and competition, with its acquisition representing a collective imperative, rather than agentive determination (Piller & Cho, 2015). The hegemonic ideologies that serve to benefit both the language testing industry and Korean establishment simultaneously act to constrain the material and symbolic choices of local agents rather than, as they posit, expand them – a process consistent with Bourdieu's (1998) acerbic description of the neoliberal utopia.

This model is reinforced directly by the socioeconomic mechanisms associated with local ELL. A study by Kim (2012), for example, found that "seventy percent of students from families earning 5 million won[8] or more a month received private English education in 2010, fully 3.5 times the 20 percent from those earning less than 1 million won[9]" (Kim, 2012, p. 3). Thus, the attainment of the forms of capital requisite for executing local success is limited amongst the financially disadvantaged, thereby representing a distinct English divide. In doing so, it may be argued that the positionality of EFL within the Korean socio-educational-industrial fields embodies the discursive authority of predetermined meritocracy, reinforced on the institutional and ideological levels by the hereditary transmission of economic, social, and cultural capital.

EFL CAPITAL AS A PRIMARY AGENT OF DISTINCTION

Recognizing as a basis that each social field constitutes an arena of practice in which context-specific doxa are applied, reinforced, and, crucially, internalized by actors, capital represents those resources that may be strategically exploited to establish positionality within the framework of a specific field, thereby contributing to the reproduction of its sphere of influence (Bourdieu, 1986). Thus, each social arena facilitates contestation over not only the economic forms of capital but

"instruments for the appropriation of symbolic wealth socially designated as worthy of being sought and possessed" (Bourdieu, 1986, p. 73). Within the Korean sociolinguistic field, for instance, institutionally-driven foreign language ideologies have resulted in the explicit attribution of symbolic capital to EFL, which, according to the framework laid by Bourdieu (1977), yields enhanced legitimacy within the market of linguistic exchange. Due to the relative position of English within the global economy, local EFL users are afforded significant opportunities for gaining access to desirable forms of education and employment, and thus, the economic capital inherent to these features.

As a consequence, the conjunction between material wealth and the primacy of EFL instrumentalism represents a visible manifestation of cultural-economic capital conversion. This process is achieved via the propagation of language ideologies that serve to rationalize pre-existing social structure and "justify social inequality as an outcome of linguistic difference" (Piller, 2015, p. 1) − a mechanism consistent with local cultural dynamics. Specifically, social capital within Korea is tied heavily to the concept of *yonjul* (approximately: ties, or connections), which manifests itself based on an individual's ascribed status − commonly via ancestry, regional origin, wealth, or education. As described by Horak (2015), "*yonjul* ties exist for a purpose, often to secure personal gains and benefits" (p. 78), while also serving as a barrier to those who do not share the particularistic relation. However, while *yonjul* dynamics may act to preserve Centre−Periphery socioeconomic exclusion, the "densely knit network creates pressure for an individual to conform to the group ... exerting enormous pressure for conformity to the members" (Yee, 2000, p. 342). Thus, the constraints and pressures regulating sociolinguistic reproduction within such groups function by both positive and negative measures.

Considering one's capacity for social network exploitation depends broadly on an actor's accrued cultural and economic capital, Korea's "*yonjul* society" (Lee, 2000, p. 369) offers a convincing exemplar of Bourdieu's (1986) capital regulation-reproduction process. Deliberately structured by those possessing the corresponding linguistic competences and, thus, "vested interests in the practical and symbolic meanings of English" (Cho, 2017, p. 170) to function as a gatekeeper to the various forms of capital. Accordingly, the standardized testing mechanisms by which local EFL "productions are explicitly sanctioned and evaluated" (Bourdieu, 1977, p. 654) by local institutions represent certificates of

cultural competence (Bourdieu, 1986), resulting in "graduates from privileged backgrounds" (Smith & Kim, 2015, p. 342) gaining overwhelmingly favorable access to preferred forms of employment – aided by their command of EFL-related cultural capital and the social capital accrued via dynastic privilege and alma mater. English within Korea, therefore, must be understood as an overt manifestation of power dynamics between actors (Bourdieu, 1977), specifically embodying a skewed distribution across socioeconomic positionality due to "the scholastic yield from educational action depend[ing] on the cultural capital previously invested by the family" (Bourdieu, 1986, p. 244).

Consolidated as a product to be consumed during social reproduction (Phillipson, 2008), EFL within Korea facilitates cultural hegemony from a Bourdieusian perspective. Specifically, those possessing high degrees of cultural, social, and economic capital have established EFL as a dominant form of embodied capital, or *taste*, to function as a prestige code by those who speak English, and to act as a barrier to social mobility to those who do not. Via metaphors that depict English as a transnational tool of development, ELL interlocks with symbolic and material systems, legitimizing existing class fractions and asymmetrical social distinction due to the internalization of ideologies by those retaining lower volumes of capital. The elite-driven cultural capital of local EFL, in naturalizing the misconception that English is necessary to the enhancement of Korea's global positionality, exemplifies what Bourdieu terms "symbolic violence" (Bourdieu, 1984, p. 358). Moreover, considering that the social conditioning governing dominant forms of taste is internalized during the early stages of cognitive development, the modification of such behaviors may prove acutely difficult, if not impossible. Based on these assertions, one's embodied forms of cultural-linguistic capital – reflective of the systematic corporeal and affective structuring of a social actor – represents, perhaps, the most visible demonstration of habitus.

THE KOREAN LINGUISTIC HABITUS

As noted by Crossley (2001), field and habitus are intertwined in an openly circular dynamic in which participation within a specific field frames the habitus, which, as a result, influences the actions that

reproduce the field during the accumulation of capital. Thus, one's linguistic habitus is constituted, reinforced, and modified via a continual process of positive and negative sanctioning, occurring across a range of linguistic markets (Shin, 2014). In describing the interplay between Korea and the features of power associated with English language appropriation, for example, it has been shown that the linguistic habitus of local agents are influenced by a notable interaction between agency and structure (Cho, 2017), specifically via institutionally endorsed neoliberal ideologies that explicitly link ELL to the enhancement of both individual and national development, modernity, and mobility (J. Lee, 2010). In this regard, negotiations between Koreans and local sociolinguistic norms are connected broadly to the meretricious notion of advancement via demonstrable English language competency and achievement – determined by attending financially prohibitive private ELL academies, and both local and inner-circle-based elite education institutions (Shin, 2010), which are, in turn, endorsed by standardized language assessment measures.

Thus, linguistic habitus within Korea is characterized by an imbalance between objective and subjective pressures, specifically "between recognized norms and the capacity to produce" (Bourdieu, 1977, p. 658). As noted by Cho (2017), this process is particularly relevant to the analysis of local EFL power structures "in which recognized acceptability and legitimacy in English remains elusive due to the ever-fiercer competition over English and resultant ever-rising standards of English" (Cho, 2017, p. 21). Nevertheless, that EFL-related capital is exploited locally as a screening mechanism for industry, and education is reflective of endemic academic inflation within the Korean education system. D. Y. Lee (2010) notes that, in 2005, the university acceptance ratio of secondary school graduates reached 82.1 percent, resulting in local tertiary education alone proving insufficient to the generation of the academic capital (both symbolic and tangible) required for social advancement (D. Y. Lee, 2010). As a result of this process, the significance of EFL within the Korean linguistic habitus is particularly heightened amongst the middle and upper classes (Park & Lo, 2012), who possess the requisite economic capital to access ELL. This dynamic has resulted in the widely described "English divide" (Cho, 2017; J. Lee, 2010; Shin, 2010), which functions to enhance polarized forms of socio-educational stratification and, thus, cultural distinction.

Considering that said partition is characterized by the unequal

distribution of EFL-related linguistic capital "between children of wealthy parents and those from lower-income families" (Jeon, 2012, p. 407), which, given the intensely competitive nature of the Korean industrial and educational domains (Byun, Schofer, & Kim, 2012), is essential to generating socioeconomic success, one must inquire whether the criterion that would distinguish Korea as an authentic meritocracy have, in fact, been achieved. Nevertheless, ideologies describing the enhancement of social mobility via individual competition and rigorous dedication to scholarly achievement are so ingrained within the Korean habitus that they act to mask the structural inequalities governing local social reproduction, which, in turn, sustains the division of EFL proficiency based on dynastic capital.

Additionally, it has been noted that Social Darwinist rhetoric has intersected with ideologies advocating extreme forms of patriotism in an effort to enhance the dominant positionality of EFL within the Korean sociolinguistic field (Lee, Han, & McKerrow, 2010). Thus, language learning by way of a social actor's distinct linguistic capital matters not only in the context of social mobility but fuses with other social categories, including citizenship (Shin, 2014, p. 99). In connecting these concepts, the Korean habitus is responsive to the pervasive ideology that one must participate in (preferably desirable forms of) ELL if one is to distinguish oneself as a productive and patriotic citizen. Subsequently, this dynamic emphasizes the potential cognitive friction in positioning culturally extrinsic products as mechanisms vital to the enhancement of national prestige. Moreover, given that financially disadvantaged Koreans are typically excluded from ELL, the upper and middle classes are, by extension, afforded enhanced opportunities of being recognized as "patriotic" citizens, potentially enhancing affective disharmony within those excluded from ELL.

In consequence, the core expectations by which EFL is rationalized locally favor the wealthy, to the exclusion of those constrained by circular socioeconomic stratifications and a resultant dearth of professional advancement. The presence of this dynamic within the Korean habitus may result in the acceptance of exclusion based on predetermined wealth, further increasing social class frictions. Given both the professional and social requirements for EFL competency and the seemingly immovable presence of English within the Korean habitus, one must pose the question whether local language learners do, in fact, agentively elect to participate in ELL.

CONCLUSIONS

In drawing on Bourdieu's concepts, this investigation has attempted to critically appraise the underlying structures that secure the hegemonic role of English within the Korean sociolinguistic sphere. Subsequently, it is apparent that EFL dominance within this setting is maintained and strengthened in accordance with "habitus that forms Koreans' dispositions that determine how they are supposed to act and respond in dealing with and making sense of English in their everyday lives" (Park, 2009, p. 27). In this context, the integration of Korean EFL with the various forms of capital illustrates that pervasive language ideologies continue to circulate and evolve via the sociolinguistic practices of local speakers (Cho, 2017), thereby lending credence to Bourdieu's (1982) assertion that "the language of authority never governs without the collaboration of those it governs, without the help of the social mechanisms capable of producing this complicity" (Bourdieu, 1982, p. 113). It has therefore been demonstrated that homologies between sociolinguistic fields, and the relations between actors within those fields, are anchored in specific forms of power, expressed here by a distinctly circular transmission of EFL-related capital.

As a consequence, EFL instrumentalization within the Korean sociolinguistic field is illustrative of the procedures by which a dominant class propagates self-aggrandizing values and norms via the manipulation of cultural items, thereby achieving the hegemonic subjugation of subordinate groups via ideological mechanisms. Due to the discursive realities of English adoption within Korea, it is thus paramount that future research provides ethnographic comprehension of the sociological impact of ELL and the procedures by which its constituent processes are effectuated within this distinct field. Indeed, an appreciation of the micro-level dynamics by which social mobility, citizenship, identity, and ELL converge in the lives of EFL users, and the means by which their linguistic practices are driven via the habitus formed through social reproduction, is invaluable in understanding the *authentic* processes by which the uneven power structures inherent to linguistic hierarchization are sustained and, in turn, internalized in the form of unquestionable orthodoxies and habits that act to fundamentally preserve and circulate "the very structure that has given rise to them in the first place" (Malik & Mohamed, 2014, p. 72).

THE AUTHOR

Michael D. Smith is an instructor in English as a foreign language at Kwansei Gakuin University, School of International Studies, Japan. Currently enrolled as a doctoral student at the University of Bath, he holds a teaching license and MA in applied linguistics, and is an alumnus of University College London Institute of Education, where he gained an MA in technology and education with distinction. Michael's research interests include sociolinguistics, World Englishes, international education, and CALL. Email: Michael.Dean.Smith@Kwansei.ac.jp

REFERENCES

Bourdieu, P. (1977). The economics of linguistic exchanges. *Social Science Information*, *16*(6), 645−668. https://doi.org/10.1177/053901847701600601

Bourdieu, P. (1982). *Language and symbolic power*. Harvard University Press.

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. Harvard University Press.

Bourdieu, P. (1986). The forms of capital. In J. G. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241−58). Greenwood.

Bourdieu, P. (1989). Social space and symbolic power. *Sociological Theory*, *7*(1), 14−25.

Bourdieu, P. (1998). The essence of neoliberalism. *Le Monde Diplomatique*. https://mondediplo.com/1998/12/08bourdieu

Bourdieu, P., & Wacquant, L. (1992). *An invitation to reflexive sociology*. University of Chicago Press.

Byun, S. Y., Schofer, E., & Kim, K. K. (2012). Revisiting the role of cultural capital in East Asian educational systems: The case of South Korea. *Sociology of Education, 85*(3), 219−239. https://doi.org/10.1177/0038040712447180

Chandler, B. (2013). The subjectivity of habitus. *Journal for the Theory of Social Behaviour*, *43*(4), 469−491. https://doi.org/10.1111/jtsb.12025

Chang, B. M. (2009). Korea's English education policy innovations to lead the nation into the globalized world. *Pan-Pacific Association of Applied Linguistics*, *13*(1), 83−97.

Cho, J. (2017). *English language ideologies in Korea*. Springer.

Crossley, N. (2001). The phenomenological habitus and its construction. *Theory and Society*, *30*(1), 81−120.

Crystal, D. (2003). *English as a global language* (2nd ed.). Cambridge University Press.

DiMaggio, P. (1982). Cultural capital and school success: The impact of status

culture participation on the grades of U.S. high school students. *American Sociological Review*, *47*(2), 189−201. https://doi.org/10.2307/2094962

Grenfell, M. (2012). *Bourdieu, language, and linguistics*. Continuum.

Gunz, H., & Peiperl, M. (2007). *Handbook of career studies*. SAGE.

Han, K. K. (2007). The archaeology of the ethnically homogeneous nation state and multiculturalism in Korea. *Korea Journal*, *47*(4), 8−31.

Horak, S. (2014). Antecedents and characteristics of informal relation-based networks in Korea: Yongo, Yonjul, and Inmaek. *Asia Pacific Business Review*, *20*(1), 78−108. https://doi.org/10.1080/13602381.2013.791567

Jambor, P. Z. (2012). *English language influence on THE-Reuters 2010 University Rankings; The evidence: A South Korean model in a global context* [ERIC Number ED528167]. https://eric.ed.gov/?id=ED528167

Jeon, M. (2009). Globalization and native English speakers in English Program in Korea (EPIK). *Language, Culture, and Curriculum*, *22*(3), 231−243.

Jeon, M. (2012). English immersion and educational inequality in South Korea. *Journal of Multilingual and Multicultural Development*, *33*(4), 395−408. https://doi.org/10.1080/01434632.2012.661438

Kachru, B. B. (1985). Standards, codification, and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds), *English in the world: Teaching and learning the language and literatures* (pp. 11−30). Cambridge University Press.

Kim, S. S., (2000). *Korea's globalization*. Cambridge University Press.

Kim, C. K. (2005). *The history of Korea*. Greenwood Press.

Kim, E. M. (2007). The first footsteps across the frontier. In H. A. Willoughby (Ed.), *Footsteps across the frontier* (pp. 25−30). Ewha Womans University Press.

Kim, E. G. (2011). English educational policies of the U.S. Army Military Government in Korea from 1945 to 1948 and their effects on the development of English language teaching in Korea. *Language Policy*, *10*(3), 193−220.

Kim, M. (2012). Equity and efficiency of English education spending. *KDI Policy Forum*, *245*, 2−3.

Kim, B. R. (2015). The English fever in South Korea: Focusing on the problem of early English education. *Journal of Education and Social Policy*, *2*(2), 117−124.

Kingston, P. (2001). The unfulfilled promise of cultural capital theory. *Sociology of Education*, *74*, 88−99. https://doi.org/10.2307/2673255

Kleiner, J. (2001). *Korea, a century of change*. World Scientific.

Lawrence, C. B. (2012). The Korean English linguistic landscape. *World Englishes*, *31*(1), 70−92.

Lee, J. H. (2000). Society in a vortex? Yonjul network and civil society in Korea. *Korea Journal*, *40*, 366−391.

Lee, D. Y. (2010). Cultural habitus in the everyday life of Koreans. *Munhwa*

Gwahak, *61*, 169−195.

Lee, J. (2010). Ideologies of English in the South Korean "English immersion" debate. In M. T. Prior, Y. Watanabe, & S. K. Lee (Eds.), S*elected Proceedings of the 2008 Second Language Research Forum* (pp. 246−260). Cascadilla Proceedings Project.

Lee, J. H., Han, M. W., & McKerrow, R. E. (2010). English or perish: How contemporary South Korea received, accommodated, and internalized English and American modernity. *Language and Intercultural Communication*, *10*(4), 337−357.

Malik, A., & Mohamed, A. (2014). English as cultural capital: EFL teachers' perceptions − A cross-cultural study. *Journal of Sociological Research*, *5*(2), 63−74. https://doi.org/10.5296/jsr.v5i2.6413

Palais, J. B. (2014). *Confucian statecraft and Korean institutions: Yu Hyongwon and the late Choson Dynasty*. University of Washington Press.

Park, J. S. (2009). *The local construction of a global language: Ideologies of English in South Korea*. Mouton de Gruyter.

Park, S. Y., & Lo, A. (2012). Transnational South Korea as a site for a sociolinguistics of globalization: Markets, timescales, neoliberalism. *Journal of Sociolinguistics*, *16*(2), 147−164. https://doi.org/10.1111/j.1467-9841.2011.00524.x

Phillipson, R. (1992). *Linguistic imperialism*. Oxford University Press.

Phillipson, R. (2008). Lingua franca or lingua frankensteinia? English in European integration and globalisation. *World Englishes*, *27*(2), 250−267. https://doi.org/10.1111/j.1467-971X.2008.00555.x

Phillipson, R. (2014). *English, the lingua nullius of global hegemony − The politics of multilingualism: Linguistic governance, globalisation, and Europeanisation*. Université de Genève.

Piller, I. (2015). Language ideologies. *The international encyclopaedia of language and social interaction* (pp. 1−10). https://doi.org/10.1002/9781118611463.wbielsi140

Piller, I., & Cho, J. H. (2015). Neoliberalism as language policy. In T. Ricento (Ed.), *Language policy and political economy* (pp. 162−186). Oxford University Press.

Shin, H. J. (2007). English language teaching in Korea: Toward globalization or glocalization? In J. Cummins & C. Davis (Eds.), *International handbook of English language teaching* (pp. 75−86). Springer.

Shin, H. J. (2010). *"Gireogi gajok": Transnationalism and language learning* [Unpublished doctoral dissertation]. University of Toronto.

Shin, H. J. (2014). Social class, habitus, and language learning: The case of Korean early study-abroad students. *Journal of Language, Identity, and Education*, *13*(2), 99−103. https://doi.org/10.1080/15348458.2014.901821

Silver, R. E. (2005). The discourse of linguistic capital: Language and economic policy planning in Singapore. *Language Policy*, *4*, 47−66.

Smith, M. D., & Kim, D. H. (2015). English and linguistic imperialism: A Korean perspective in the age of globalization. *Mirae Journal of English Language and Literature*, *20*(2), 331−350.

Song, J. J. (2011). English as an official language in South Korea: Global English or social malady? *Language Problems and Language Planning*, *35*(1), 35−55.

Yee, J. Y. (2000). Social networks of Koreans. *Korea Journal*, *40*(1), 325−352.

Yi, K. B. (1984). *A new history of Korea*. Harvard University Press.

Yim, S. (2007). Globalization and language policy in South Korea. In A. Tsui & J. Tollefson (Eds.), *Language policy, culture, and identity in Asian contexts* (pp. 37−53). Lawrence Erlbaum.

Zotzmann, K. (2013). Bourdieu, language and linguistics. *Language and Intercultural Communication*, *13*(2), 252−255. https://doi.org/10.1080/14708477.2013.771518

FOOTNOTES

[1] All Romanization of Hangeul employs the Revised Romanization of Korean system.

[2] Referred to in Korean as *Ganghwa-do Joyag* (The Treaty of Ganghwa Island).

[3] In this instance, a "modern" school is one that provides a Western-style curriculum.

[4] Literally: globalization. President Kim Young-sam's 1994 drive for internationalization, which produced significant reforms of the Korean political, cultural, and social economies (Kim, 2000).

[5] Italics added for emphasis.

[6] "Inner circle" is a spatial metaphor denoting the conventional bases of the English language, such as the UK, US, and Australia (Kachru, 1985).

[7] Initially designed by Seoul National University in 1992, the TEPS English proficiency test is the primary method of evaluating local English language skills.

[8] & [9] Approximately 4,186 USD and 837 USD, respectively, employing historically accurate rates of currency exchange.

Achievement Emotions in Foreign Language Learning Between German and Korean Students

Mikyoung Lee

University of Munich, Munich, Germany

This study examined cross-cultural differences in independent/ interdependent self-construal and achievement emotions regarding foreign language learning (FLL). It also investigated the relationships between self-construal and emotions in the FLL context. To investigate these issues, German ($N = 200$) and Korean ($N = 250$) students completed a questionnaire assessing self-construal and achievement emotions. The results showed that both independent self-construal and interdependent self-construal were higher in Germany. In terms of achievement emotions, enjoyment, hope, pride, anger, and boredom were higher in Germany, whereas anxiety, hopelessness, and shame were higher in Korea. Furthermore, structural equation modelling showed that independent self-construal was correlated with positive emotions, while interdependent self-construal was correlated with negative emotions. This suggests that culture plays a key role in the process of emotional experiences in FLL. The present study is meaningful, given that it integrated psychological perspectives from culture and emotion research into the FLL field to contribute to developing a more comprehensive understanding of students' emotions in this domain.

Keywords: independent self-construal, interdependent self-construal, achievement emotions, foreign language learning (FLL)

INTRODUCTION

Educational contexts are formed by social and cultural variables (e.g., self-construal); thus, it is important to involve the sociocultural perspective in learning and motivation research (Tanaka & Yamauchi, 2004). Since culture plays a fundamental role in shaping individuals' cognitive and affective processes (Elliot, Chirkov, Kim, & Sheldon,

2001), the consideration of cultural variables may help to explain cultural transmission to individuals' emotional experiences. To date, there has been little attention to cultural variables that may influence achievement emotions, although the constructs of emotions have become popular in academic domains. In addition, most cross-cultural research concerning self-construal has been conducted mostly with samples of North American (Western) and Asian (Eastern) adults (Heine, Lehman, Peng, & Greenholtz, 2002). The present research tapped into this issue with European (German) and Asian (Korean) high school adolescents in an academic context, particularly in foreign language learning (FLL), which is closely related to culture (Hinkel, 1999).

The English language is playing an important role as a global tool of communication as the world is becoming globalized (Lee, 2014). Crystal (2003) reported that approximately only 25% of English speakers in the world are native English speakers, which means that most communications through English are taking place among non-native English speakers (Seidlhofer, 2005). Consequently, many people in the world invest a lot of effort to speak English fluently as a foreign language. Studying English as a foreign language is challenging and takes a lot of time in order to reach a certain level of competency (Lee, 2014). Considering that learners may suffer from a variety of emotional experiences while learning English, it is worthwhile to investigate achievement emotions in FLL.

Most importantly, the present study made an effort to integrate psychological perspectives from culture and emotion research into the FLL field. To this end, the present study examined (a) cross-cultural differences in independent/interdependent self-construal and achievement emotions in FLL in German and Korean students and (b) the relations between self-construal and achievement emotions in FLL.

LITERATURE REVIEW

Self-Construal

Numerous cultural studies have suggested that culture plays an important role in establishing individuals' perspectives of themselves as well as others, and their relationships with others (e.g., Markus &

Kitayama, 1991; Singelis, 1994). They differentiated independent and interdependent self-construal, indicating that people might possess self-construals specific to the culture. In Western individualistic cultures, people are regarded as holding independent self-construal that emphasizes the separateness and uniqueness of the individual, and their main task is to become distinguished from others through personal accomplishments. In comparison, in Eastern collectivistic cultures, people are expected to possess interdependent self-construal that focuses on connectedness and relationships with others. Their main task is to maintain group harmony (e.g., Markus & Kitayama, 1991). This self-construal might heavily influence individuals' cognition, emotions, and motivations (Markus & Kitayama, 1991).

Singelis (1994) argued that independent and interdependent self-construal could coexist in a higher or lower degree within each culture. In fact, some studies have indicated that conventional theories of culture might have been losing their applicability over time. For example, Li et al. (2006) found that although the Chinese were more interdependent than Canadians, they were less interdependent than Indians. This indicates that China has become more individualistic, given that the Chinese had been more collectivistic and less individualistic than Indians (Hofstede, 2001).

ACHIEVEMENT EMOTIONS

Pekrun (2006) defined achievement emotions as "emotions tied directly to achievement activities or achievement outcomes" (p. 317); they are differentiated based on object focus (activity vs. outcome) and valence (positive vs. negative). Crossing these two dimensions produces a 2-by-2 taxonomy of achievement emotions: positive activity emotions such as enjoyment, negative activity emotions such as boredom and anger, positive outcome emotions such as hope and pride, and negative outcome emotions such as anxiety, hopelessness, and shame. Previous research on achievement emotions has typically emphasized outcome emotions, but activity emotions have not gained much attention (Pekrun, 2006). Achievement emotions here include not only outcome emotions such as hope, pride, anxiety, hopelessness, and shame but also activity emotions like enjoyment, boredom, and anger in the FLL field.

According to Pekrun's (2000) social-cognitive, control-value theory of achievement emotions (see Figure 1), emotions are induced by the individuals' control and value appraisals. Specifically, cultural norms and values (i.e., distal factors) may influence the parents', teachers', and peers' expectations, feedback, and consequences of achievement; and induction of values in the English subject (i.e., proximal factors). These proximal factors determine different levels of control and value appraisals in students, which in turn influence the students' emotional experiences differently across cultures.

FIGURE 1. Basic Assumptions of the Control-Value Theory of Achievement Emotions (Adapted from Pekrun, 2000).
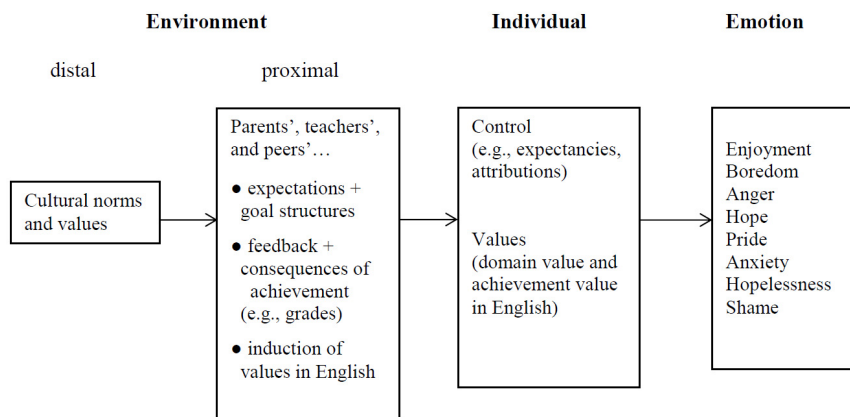
The FLL process is involved with comprehensive emotions (Imai, 2010), and particularly in foreign language classrooms, considering emotions is more important than other subjects. This is because the students' self-image becomes fragile if they do not possess language skills necessary to express themselves (Arnold, 2011). In fact, emotion has been considered a critical factor influencing FLL and achievement (Arnold & Brown, 1999; Schumann, 2001). Dewaele (2005) argued that emotions cannot be separated from language. Therefore, the emotional dimensions of FLL should be investigated in foreign language research. Since language researchers have considered language anxiety as the most significant emotional element (Oxford, 1999), several empirical studies on language anxiety have been conducted (e.g., Aida, 1994; Gardner, 1985; Horwitz, 2001; Truitt, 1995). The results imply that language

anxiety influences language learning and performance negatively. Little research has been implemented regarding comprehensive emotions, except for anxiety (Imai, 2010). It is important to consider a broader range of emotions to understand the complexity of emotions in FLL and ultimately improve the quality of foreign language instruction.

Meanwhile, numerous studies have demonstrated that Easterners tend to express lower life satisfaction, less pleasant emotions, and more negative emotions compared to Westerners (e.g., Diener & Diener, 1995; Kitayama et al., 2000). For example, negative emotions (i.e., anxiety, guilt, and pessimism) tend to be prevalent in Eastern collectivistic cultures, whereas positive emotions (i.e., optimism, pride, and joy) are promoted in Western individualistic cultures (e.g., Chang, 2001; Eid & Diener, 2001). Other theorists have proposed that individuals endorsing independent self-construals would show anger and pride more frequently, whereas individuals endorsing interdependent self-construals would experience shame, sympathy, and embarrassment more frequently (e.g., Kitayama et al., 2006; Markus & Kitayama, 1991). Regarding boredom, cross-cultural studies have been rare and produced conflicting results. There has been a tendency for boredom to be explained by higher levels of independent self-construal and lower levels of interdependent self-construal among German and Chinese students (Zirngibl, 2004). However, Americans and Australians evidenced less boredom proneness than Lebanese and Hong Kong students (Sundberg et al., 1991).

In academic contexts, however, little research has investigated cross-cultural differences in emotions, except for test anxiety. The study by Frenzel et al. (2007) is an exception, revealing that Chinese students experienced more enjoyment, pride, shame, anxiety, and less anger than Germans in mathematics. Studies on test anxiety showed that Koreans experienced higher anxiety than Germans (Schwarzer & Kim, 1984), and that Iranians and Indians presented higher anxiety than U.S. students (Sharma et al., 1983). Based on Pekrun's (2000) social-cognitive, control-value approach to emotions (Figure 1), students' anxiety is formed by their individual control and value appraisals (i.e., lack of controllability and high value of achievement), which are influenced by proximal factors such as expectations from family and teachers. Since people in Eastern collectivistic cultures emphasize harmonious interdependence with others, it is natural for students from collectivistic cultures to consider the opinions of important figures like parents and teachers. Thus, they may be more eager to live up to the parents' and

teachers' expectations and may be more afraid of disappointing them than their Western counterparts are. These proximal factors are affected by distal factors such as different culture norms and values, suggesting that culture plays a role as a potential antecedent of emotions.

Aims and Hypotheses

The aims of this study were to examine cross-cultural differences in self-construal and achievement emotions in the FLL context in German and Korean high school students, and to investigate the relationships between self-construal and emotions in FLL. The research questions and hypotheses are as follows:

RQ1. What are the cross-cultural differences in self-construal and achievement emotions in the FLL context in German and Korean high school students?

Hypothesis 1. Independent self-construal is higher in German students; interdependent self-construal is higher in Korean students.

Hypothesis 2. Enjoyment, hope, pride, anger, and boredom[1] are higher in German students; anxiety, shame, and hopelessness are higher in Korean students.

RQ2. What are the relationships between self-construal and achievement emotions in FLL?

Hypothesis 1. Independent self-construal is positively related to enjoyment, hope, pride, anger, and boredom.

Hypothesis 2. Interdependent self-construal is positively related to anxiety, shame, and hopelessness.

METHODS

Participants and Procedure

A total of N = 200 tenth- to twelfth-grade high school students in Germany (age M = 16.14, SD = .98, 60.5% girls) and N = 250 eleventh- to twelfth-grade students in Korea (age M = 17.54, SD = .62, 60.1% girls) participated. Both German and Korean students came from a private academic school located in the suburbs of a large city. One colleague who was familiar with the school systems in both countries helped choose these schools, considering the similar academic level and socio-economic background across the two schools. Concerning school size, the number of enrolled students was 845 in the German school and 930 in the Korean school. Both schools were located in residential and non-commercial districts. All in all, it was assumed that students in both schools belonged to a similar middle class relative to the overall socio-economic status in both countries.

A brief introductory letter was provided to students, assuring them of the confidentiality of their responses. Students then voluntarily completed the questionnaire that assessed their self-construal and achievement emotions during their regular English class under the supervision of their English teachers. It took about 20 minutes to complete the questionnaire.

Measures

Self-Construal

To examine independent self-construal and interdependent self-construal, the Self-Construal Scale (SCS, Singelis, 1994) was used. Sixteen items were selected from the original 24 items of the SCS, which considered contents and higher factor loadings (independent self-construal, e.g., "I enjoy being unique and different from others in many respects"; interdependent self-construal, e.g., "It is important for me to maintain harmony within my group"). Participants rated each item on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The Cronbach's alphas (Germany/Korea) were .54/.64 for the independent self-construal scale and .64/.51 for the interdependent self-construal scale. The alphas might have been lower for high school

students because the original SCS was developed based on university students. Also, the present teenager participants might not have fully developed cultural constructs compared to adults. Moreover, the lower reliabilities may reflect the measure itself. In fact, the alphas for the SCS in the present study were similar to the ones for this measure found in previous studies (Lam & Zane, 2004; Okazaki, 2000). As Lam and Zane (2004) suggested, the SCS might need to be reexamined and supplemented to include items that could enhance the reliabilities of the measures. Despite relatively low alphas, the SCS measure was included for further analysis, as it is widely used and its construct validity and reliability are well established (e.g., Hannover et al., 2000; Singelis, 1994).

Achievement Emotions

The Achievement Emotions Questionnaire – Language (AEQ-L), a modified version of the Achievement Emotions Questionnaire – Mathematics (AEQ-M, Pekrun et al., 2011), was used to assess eight emotions of students in English classes: enjoyment, hope, pride, anger, anxiety, shame, hopelessness, and boredom. The words "math class" in the original items were substituted with "English class." This measure consisted of items regarding three different academic situations (class, learning, and test-related) and reflected different components of emotions (affective, cognitive, motivational, and physiological components). Regarding the activity emotions, the enjoyment scale included ten items (e.g., "I look forward to my English class"; $a = .85/.83$), the boredom scale included five items (e.g., "I can't concentrate because I am so bored"; $a = .83/.85$), and the anger scale included seven items (e.g., "I get angry because the material in English is so difficult"; $a = .80/.83$). For the outcome emotions, the scales contained five hope items (e.g., "I have an optimistic view toward studying English"; $a = .78/.72$), six pride items (e.g., "I am proud of my contributions to the English class"; $a = .77/.74$), eleven anxiety items (e.g., "I worry if the material is much too difficult for me"; $a = .87/.86$), six hopelessness items (e.g., "I keep thinking that I will never get good grades in English"; $a = .89/.85$), and eight shame items (e.g., "I am embarrassed about my lack of knowledge in English"; $a = .86/.81$), where a = Germany/Korea. The alphas for each emotion scale in both countries were acceptable to good (Cortina, 1993). The participants responded on a scale from 1 (strongly disagree) to 5 (strongly agree).

Translation Procedure

Systematic and thorough translation procedures of all measures were undertaken for the equivalence among English, German, and Korean versions. Existing German versions of the SCS (Hannover et al., 2000) and the AEQ-L (Pekrun et al., 2011) were utilized. For the Korean sample, two English–Korean bilinguals first translated the English versions of all measures into Korean after considering potential cultural differences. Then in order to check consistency, another bilingual blindly back-translated the translations into English. The back-translated versions were compared with the English versions to confirm their accuracy. The translators discussed all items considerably until they agreed with clarity and precision. Regarding the SCS, the existing Korean version (Youn, 2000) was used with a slight modification with the word "coworker" being substituted with "friend" in the items. Finally, a German–Korean bilingual compared the German and Korean versions to ensure semantic equivalence across the two languages.

Data Analysis

Correlational analysis was used to examine the relations among the study variables (i.e., self-construal, achievement emotions, basic demographics). Multi-group confirmatory factor analyses (CFAs; Byrne & Campbell, 1999)[2] were conducted to evaluate the cross-cultural equivalence of the self-construal and achievement emotions' measures. Multi-group CFA is the most widely used to evaluate cross-cultural equivalence of measures rigorously (Byrne & Campbell, 1999). Models were assessed by the comparative fit index (CFI; Bentler, 1990), the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), and the root-mean-square-error of approximation (RMSEA; Steiger & Lind, 1980). The rules of CFI > .90, TLI >. 90 (Lance, Butts, & Michels, 2006), and RMSEA < .080 with $N < 250$ (Hu & Bentler, 1999) were adopted as indicating an adequate fit. Finally, the study hypotheses on the relations between self-construal and achievement emotions were tested using multi-group structural equation modelling (SEM) with Mplus 7 (Muthén & Muthén, 1998-2012).

RESULTS

Preliminary Analyses

Table 1 presents the intercorrelations among all the study variables in Germany and Korea. For instance, the positive emotions of enjoyment, hope, and pride were positively intercorrelated, as were the negative emotions of anger, anxiety, shame, hopelessness, and boredom in general. The emotions with differing valences were negatively related with one another. Regarding self-construal, there was a positive correlation between independent and interdependent self-construal in Korea.

TABLE 1. Pearson Product-Moment Correlations for the Study Variable

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1. Enjoy-ment	–											
	–											
2. Hope	.74**	–										
	.66**	–										
3. Pride	.76**	.65**	–									
	.74**	.64**	–									
4. Anger	-.68**	-.64**	-.46**	–								
	-.54**	-.49**	-.42**	–								
5. Anxiety	-.54**	-.67**	-.50**	.60**	–							
	-.33**	-.46**	-.36**	.62**	–							
6. Shame	-.35**	-.47**	-.38**	.45**	.72**	–						
	-.25**	-.39**	-.34**	.55**	.79**	–						
7. Hope-lessness	-.52**	-.66**	-.47**	.58**	.84**	.64**	–					
	-.47**	-.59**	-.53**	.70**	.81**	.72**	–					
8. Boredom	-.59**	-.44**	-.37**	.62**	.22**	.12	.26**	–				
	-.62**	-.46**	-.43**	.76**	.43**	.38**	.53**	–				
9. IND	.13	.14*	.19**	-.03	-.08	-.01	-.08	.13	–			
	.21**	.23**	.22**	-.07	.01	.00	-.04	-.07	–			
10. INT	.02	-.06	-.09	.04	.22**	.22**	.29**	-.14*	.13	–		
	.23**	.09	.19**	-.04	.06	.03	.02	-.09	.22**	–		
11. Gender[a]	.01	.05	.09	.02	-.15*	-.23**	-.19**	.02	-.06	-.17*	–	
	.02	.07	.05	-.15*	-.19**	-.23**	-.22**	.01	.06	.16*	–	
12. Age	-.17*	-.07	-.07	.10	.03	.01	.08	.23**	.07	-.01	-.12	–
	-.09	.07	.09	-.09	-.27**	-.23**	-.20**	-.18**	-.01	-.03	.02	–

Note. IND = independent self-construal; INT = interdependent self-construal. Upper row Germany; Lower row Korea. [a]Gender is coded 1 = boys and 2 = girls. *$p < .05$, **$p < .01$.

Research Question 1: Cross-Cultural Differences in Mean Levels

Table 2 presents the results of independent group t-tests regarding differences in self-construal and achievement emotions between German and Korean students. The means and standard deviations are also displayed. Both independent self-construal and interdependent self-construal were higher in Germany: $t(420) = 8.60$, $p < .01$ for independent self-construal, and $t(421) = 3.66$, $p < .01$ for interdependent self-construal, which was unexpected.

TABLE 2. Mean Levels for Self-construal and Achievement Emotions Across Countries

	Mean[a]		SD		t	df	d
	Germany	Korea	Germany	Korea			
IND	3.81	3.36	0.5	0.58	8.60**	420	0.83
INT	3.53	3.35	0.51	0.51	3.66**	421	0.35
Enjoyment	2.92	2.79	0.58	0.61	2.32*	420	0.22
Hope	3.33	3.24	0.71	0.73	1.35	423	0.12
Pride	3.06	2.97	0.67	0.7	1.33	411	0.13
Anger	2.25	2.04	0.73	0.73	3.02**	422	0.29
Anxiety	2.18	2.55	0.7	0.73	-5.20**	422	-0.52
Shame	1.93	2.56	0.77	0.79	-8.20**	415	-0.81
Hopelessness	1.95	2.44	0.82	0.9	-5.90**	423	-0.57
Boredom	2.42	2.36	0.77	0.82	0.67	423	0.08

Note. IND = independent self-construal; INT = interdependent self-construal. [a]possible range 1–5. *$p < .05$. **$p < .01$.

Concerning achievement emotions, enjoyment and anger were significantly higher in Germany than in Korea: $t(420) = 2.32$, $p < .05$ for enjoyment, and $t(422) = 3.02$, $p < .01$ for anger. Hope, pride, and boredom were slightly higher in Germany. In contrast, anxiety, shame, and hopelessness were significantly higher in Korea: $t(422) = -5.20$, $p < .01$ for anxiety, $t(415) = -8.20$, $p < .01$ for shame, and $t(423) = -5.90$, $p < .01$ for hopelessness.

Research Question 2: Relationships between Self-Construal and Achievement Emotions

Structural equation modelling (SEM) using Mplus 7 (Muthén &

Muthén, 1998-2012) was conducted in a total merged sample to examine the relations between self-construal and achievement emotions. Scale items were used as manifest indicators of latent variables for the self-construal and emotions.

FIGURE 2. **Structural Parameter Estimates of Self-Construal and Achievement Emotions**
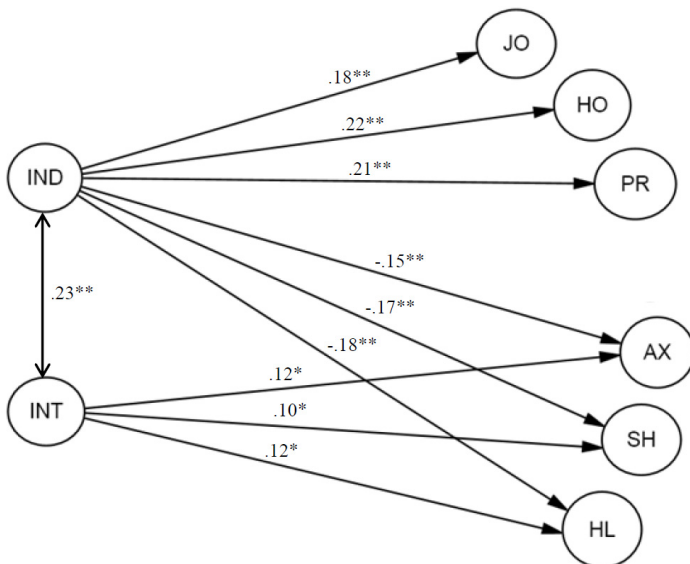
Note. Only significant path coefficients are presented. Indicator variables, error variables, and correlations between error variables are omitted for simplification. IND = independent self-construal; INT = interdependent self-construal; JO = enjoyment; HO = hope; PR = pride; AX = anxiety; SH = shame; HL = hopelessness. *p < .05, **p < .01.

Figure 2 displays the results of the SEM for the relations between independent /interdependent self-construal and achievement emotions. It displays correlations between independent and interdependent self-construal as well as significant path coefficients between independent/interdependent self-construal and emotions. The model fit was; χ^2 (4172) = 8322.67, CFI = .87, TLI = .85, and RMSEA = .072. Although CFI and TLI were slightly below the conventional standard, this model was used for analysis, considering that multi-group CFAs of these variables demonstrated better fit compared to the ones in the previous studies[2]. Independent self-construal was positively related to positive emotions such as enjoyment (β = .18, p < .01), hope (β = .22,

$p < .01$), and pride ($\beta = .21$, $p < .01$), but negatively to negative emotions such as anxiety ($\beta = -.15$, $p < .01$), shame ($\beta = -.17$, $p < .01$), and hopelessness ($\beta = -.18$, $p < .01$). Interdependent self-construal was positively related to negative emotions such as anxiety ($\beta = .12$, $p < .05$), shame ($\beta = .10$, $p < .05$), and hopelessness ($\beta = .12$, $p < .05$). However, anger and boredom were related to neither independent self-construal nor interdependent self-construal.

Discussion

Based on previous cross-cultural studies, it was expected that German students (Western culture) would demonstrate higher independent self-construal, whereas Korean students (Eastern culture) would show higher interdependent self-construal. However, the hypothesis was not completely supported. Both independent self-construal and interdependent self-construal were higher in German students. The reason for this result may be that the present study compared European (Western) with Asian (Eastern) adolescents, whereas previous studies mostly compared North American and Asian adults (Heine et al., 2002). Moreover, independent self-construal and interdependent self-construal could coexist within each culture (Singelis, 1994), which is reflected in the result.

Due to globalization and the prevalence of the internet, teenagers may be more influenced by other cultures now than in the past, which could have led to a change in independent/interdependent self-construal among both Western and Eastern youths. Particularly, since Korean students focus on English learning, they might be influenced by the related Western culture, which eventually could make them culturally adapted. Another explanation may be the way the constructs were measured. Heine et al. (2002) mentioned a reference effect, which happens when people from a certain cultural group evaluate themselves by comparing themselves to individuals of a different reference group. The participants may have compared themselves with other members of their society and concluded that they possess a relatively higher or lower independent/interdependent self-construal compared to others. Moreover, the present finding might have been influenced by acquiescence bias (Schimmack et al., 2005). In fact, the self-construal measures have been reported to be inclined to acquiescence bias, which refers to the tendency of respondents to agree or disagree with a survey item, regardless of the

actual content of the item. As a consequence, this renders the actual relationship to appear differently (Kam et al., 2012). Finally, the self-construal construct might not have been completely conceptualized for adolescents compared to adults.

The findings on mean level differences of achievement emotions can be interpreted using Pekrun's (2000) social-cognitive, control-value theory of achievement emotions as well as previous findings on Eastern and Western cross-cultural differences in emotional experiences (e.g., Eid & Diener, 2001; Heine & Lehman, 1995; Markus & Kitayama, 1991). In Asian cultures, academic success is very important for the individuals' successful future. Thus, the parents force and expect their children to perform well, which influences the students' control and value appraisals in school subjects. Since Korean students highly value the English domain, and maybe more so compared to German students (Lee, 2014), these appraisals might have determined how students experience emotions differently in the English domain in both countries.

The findings show that enjoyment and anger were higher in Germany. In contrast, anxiety, hopelessness, and shame were higher in Korea, confirming that these negative emotions are frequent in Eastern collectivistic cultures, while positive emotions such as enjoyment are promoted in Western individualistic cultures. For example, a higher achievement anxiety in Korea is in line with evidence showing that people from Eastern countries generally show higher anxiety (e.g., Eid & Diener, 2001). Less hopelessness in Germany and higher hopelessness in Korea corroborate earlier findings that Westerners tend to be more optimistic compared to Easterners (e.g., Chang, 2001). Since Korean students are under achievement pressure from parents and teachers and highly value achievement in English, they put a lot of effort into studying English. Nevertheless, many Korean students may feel their performance is never commensurate with their parents' and teachers' expectations. This might have made the Korean students experience more anxiety, shame, and hopelessness, but less enjoyment. Moreover, the Korean students might have experienced and reported more shame than the German students due to the encouragement of modesty in Asian cultures (Schoenhals, 1993). Assuming that Korea and Japan share cultural characteristics, the results can be further interpreted based on Kitayama et al.'s (2006) finding that Japanese values promote socially engaging emotions (e.g., friendly feelings, guilt, and shame), whereas North Americans endorse socially disengaging emotions (e.g., pride and

anger). In particular, anger may be avoided in Eastern collectivistic cultures, but accepted in Western individualistic cultures (e.g., Frenzel et al., 2007; Kitayama et al., 2006), which may explain higher levels of reported anger in German students.

The country differences were small and not significant for hope and pride. When considering the influence of culture, hope and pride should be higher in Germany. Also, Asian students might be less confident than Western students due to the extreme emphasis on performance with high pressure (e.g., Stevenson & Lee, 1990). However, if Korean students highly value FLL, they should also show high levels of hope and pride based on Pekrun's (2000) social-cognitive, control-value theory of achievement emotions. Whereas the influence of the cultural context would imply higher levels of hope and pride among German students, a lower subjective value of English would reduce these emotions, resulting in non-significant cross-cultural differences in the mean levels of these emotions.

The lower tendency of boredom in Korean students is likely due to their learning environment in Korea. Since performance in English has a heavy influence on admission to university, regardless of specific majors, Koreans often possess higher achievement values for English. Therefore, they would need to study English, no matter how much genuine interest they have. However, in Germany, outstanding English performance is important, mainly for the students who want to major in the language. This implies that German students might hold lower achievement values in English on average. As a result, this could have led to a slightly higher boredom result for German students since some students might lack interest in English.

Overall, the present findings on the cross-cultural difference in emotions are not surprising, especially when considering the unique situation of the English domain in Korea where performance in English is extremely crucial for students' academic careers. Under this stress and pressure from parents and teachers, learning English in Korea could induce negative emotions rather than positive ones.

In terms of the relations between self-construal and achievement emotions, the results partially confirm earlier findings. Independent self-construal related positively to enjoyment, hope, and pride, and negatively to anxiety, shame, and hopelessness, while interdependent self-construal related positively to the negative emotions. This indicates that independent self-construals tend to experience more positive

emotions, while interdependent self-construals might promote negative emotions. For example, pride and enjoyment were more strongly expressed in independent self-construals, while anxiety and shame were higher in interdependent self-construals (e.g., Eid & Diener, 2001). Moreover, Chang (2001) suggested that independent self-construals tend to be more optimistic, while interdependent self-construals tend to be more pessimistic, explaining the result on the links between self-construal and hope/hopelessness.

Unexpectedly, independent self-construal was not related to anger and boredom. This might be due to the fact that emotions were examined particularly in the FLL context and due to the positive correlation between independent and interdependent self-construal (see Figure 2). Furthermore, the concept of self-construal might not have been fully cognized in the teenager participants, which might have concealed the theoretically postulated relations between self-construal and emotions. Despite some inconsistent findings, the main results of this study indicate that culture still plays a key role in the process of individuals' emotional experiences in FLL. Therefore, foreign language teachers should be aware of students' cultural aspects in their classrooms so that they can have a positive influence on students' emotional experiences in learning a foreign language. On the other hand, researchers could also deliberate on the idea that culture might be changing over time, and conventional theories might be losing their applicability, considering some inconsistent findings. It might be time to reconsider the main theories of self-construal in cultural psychology, as some researchers have argued (e.g., Levine et al., 2003; Oyserman et al., 2002).

The present research has a few limitations, and accordingly, some directions for future research are suggested. This study was conducted in one high school in each country with teenagers to examine cultural differences in self-construal and achievement emotions, particularly with the English subject, assuming that these samples would represent Western and Eastern countries. However, we must be careful to generalize the results for Western and Eastern cultures from the present samples in an academic setting. Since language and culture are closely related (Hinkel, 1999), it would be interesting to see whether self-construal has the same influence on emotions in other subjects. To explore cultural differences and the relations of culture with achievement emotions more thoroughly, future studies should expand the number of schools and subjects and include various samples of both Western and

Eastern cultures.

This study utilized self-report data; the participants might not have accurately seized actual psychological phenomena such as cultural variables (i.e., self-construal) and achievement emotions. Particularly, it is difficult to exactly evaluate whether the cross-cultural comparisons of achievement emotions are differences in intensity, duration, or frequency of the emotions or not. Also, it might have been possible for the participants from each country to have simply considered that it was socially acceptable to articulate certain emotions, but not others (Zirngibl, 2004). More qualitative studies should be designed to describe cultural differences in psychological phenomena in future research to complement this issue.

CONCLUSIONS

Given that studies on emotions in FLL are so few, except for language anxiety, the present research contributes to developing a more comprehensive understanding of students' emotions in this domain. It is also expected to motivate other researchers to conduct more studies in this field. Considering the relationship between emotions and students' learning is much more important, especially in a foreign language classroom than for other subjects, since learners' self-perception becomes more susceptible when they do not have language skills to express themselves effectively (Arnold, 2011). Therefore, it is imperative that educators take the affective side of language learning into account in their classrooms. This study has looked into a range of positive and negative emotions in the FLL classroom. The results of the present research will contribute to scientific knowledge about language learning, given that this investigation produced new insights on students' emotions when learning a foreign language. These insights will help develop effective interventions that will enhance positive emotions, learning, and academic performance (Astleitner, 2000). In addition, the present research urges that foreign language teachers to be aware of students' cultural aspects in class for a positive influence on students' emotions in FLL and performance. This study also expands the realm of cross-cultural psychology by comparing achievement emotions in academic contexts between European (Western) and Asian adolescents

(Eastern), considering that most cross-cultural studies have been conducted with North American (Western) and Asian adults (Eastern). Most of all, the current research contributes to interdisciplinary investigation in the FLL field by integrating psychological perspectives from culture and emotion research into this field.

THE AUTHOR

Mikyoung Lee is a guest researcher in the Department of Educational Psychology at the University of Munich, Germany. Her research interests are achievement goals and emotions, teachers' emotions and emotion regulation, and foreign language acquisition. Email: mikyoung.lee@psy.lmu.de

REFERENCES

Aida, Y. (1994). Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. *Modern Language Journal*, *78*, 155–168.

Arnold, J. (2011). Attention to affect in language learning. *Anglistik. International Journal of English Studies*, *22*, 11–22.

Arnold, J., & Brown, D. H. (1999). A map of the terrain. In J. Arnold (Ed.), *Affect in language learning* (pp. 1–24). Cambridge University Press.

Astleitner, H. (2000). Designing emotionally sound instruction: The FEASP-approach. *Instructional Science*, *28*, 169–198.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238–246.

Byrne, B. M., & Campbell, T. L. (1999). Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure. *Journal of Cross-Cultural Psychology*, *30*, 555–574.

Chang, E. C. (2001). Cultural influences on optimism and pessimism: Differences in Western and Eastern construals of the self. In E. C. Chang (Ed.), *Optimism and pessimism: Implications for theory, research, and practice* (pp. 257–280). American Psychological Association.

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*, 464–504.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98.

Crystal, D. (2003). *English as a global language*. Cambridge University Press.

Dewaele, J. (2005). Investigating the psychological and emotional dimensions in

instructed language learning: Obstacles and possibilities. *The Modern Language Journal*, *89*(3), 367–380.

Diener, E., & Diener, M. (1995). Cross-cultural correlates of life satisfaction and self-esteem. *Journal of Personality and Social Psychology*, *68*, 653–663.

Eid, M., & Diener, E. (2001). Norms for experiencing emotions in different cultures: Inter- and intranational differences. *Journal of Personality and Social Psychology*, *81*, 869–885.

Elliot, A. J., Chirkov, V. I., Kim, Y., & Sheldon, K. M. (2001). A cross-cultural analysis of avoidance (relative to approach) personal goals. *Psychological Science*, *12*, 505–510.

Frenzel, A. C., Thrash, T. M., Pekrun, R., & Goetz, T. (2007). Achievement emotions in Germany and China: A cross-cultural validation of the Academic Emotions Questionnaire – Mathematics (AEQ-M). *Journal of Cross-Cultural Psychology*, *38*, 302–309.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitude and motivation*. Edward Arnold.

Hannover, B., Kühnen, U., & Birkner, N. (2000). Interdependent and independent self- knowledge as a determinant of assimilation versus contrast in contextual priming. *Journal of Social Psychology*, *31*, 44–56.

Heine, S. J., & Lehman, D. R. (1995). Culture, self-discrepancies, and self-satisfaction. *Personality and Social Psychology Bulletin*, *25*, 915–925.

Heine, S. J., Lehman, D. R., Peng, K., & Greenholtz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales? The reference-group effect. *Journal of Personality and Social Psychology*, *82*, 903–918.

Hinkel, E. (1999). *Culture in second language teaching and learning*. Cambridge University Press.

Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations*. SAGE.

Horwitz, E. K. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics*, *21*, 112–126.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1-55.

Imai, Y. (2010). Emotions in SLA: New insights from collaborative learning for an EFL classroom. *Modern Language Journal*, *94*, 278–292.

Kam, C., Zhou, X., Zhang, X., & Ho, M. Y. (2012). Examining the dimensionality of self-construals and individualistic-collectivistic values with random intercept item factor analysis. *Personality and Individual Differences*, *53*, 727–733.

Kitayama, S., Markus, H. R., & Kurokawa, M. (2000). Culture, emotion, and well-being: Good feelings in Japan and the United States. *Cognition and Emotion*, *14*, 93–124.

Kitayama, S., Mesquita, B., & Karasawa, M. (2006). Cultural affordances and emotional experience: Socially engaging and disengaging emotions in Japan and the United States. *Journal of Personality and Social Psychology*, *91*, 890–903.

Lam, A. G., & Zane, N. W. S. (2004). Ethnic differences in coping with interpersonal stressors: A test of self-construals as cultural mediators. *Journal of Cross-Cultural Psychology*, *35*, 446–459.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The sources of four commonly reported cutoff criteria: What did they really say? *Organizational Research Methods*, *9*, 202–220.

Lee, M. (2014). *Achievement goals, emotions, and foreign language performance in German and Korean students* [Unpublished doctoral dissertation]. The University of Munich.

Levine, T. R., Bresnahan, M. J., Park, H. S., Lapinski, M. K., Wittenbaum, G. M., Shearman, S. M., Lee, S. Y., Chung, D., & Ohashi, R. (2003). Self-construal scales lack validity. *Human Communication Research*, *29*, 210–252.

Li, H. Z., Zhang, Z., Bhatt, G., & Yum, Y. O. (2006). Rethinking culture and self-construal: China as a middle land. *The Journal of Social Psychology*, *146*, 591–610.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*, 224–253.

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (6th ed.). Muthén & Muthén.

Okazaki, S. (2000). Asian American and White American differences on affective distress symptoms: Do symptom reports differ across reporting methods? *Journal of Cross-Cultural Psychology*, *31*, 603–625.

Oxford, R. L. (1999). Anxiety and the language learner: New insights. In J. Arnold & H. D. Brown (Eds.), *Affect in language learning* (pp. 58–67). Cambridge University Press.

Oyserman, D., Coon, H. M., & Kemmelmeier, M. (2002). Rethinking individualism and collectivism: Evaluation of theoretical assumptions and meta-analyses. *Psychological Bulletin*, *128*, 3–72.

Pekrun, R. (2000). A social-cognitive, control-value theory of achievement emotions. In J. Heckhausen (Ed.), *Motivational psychology of human development: Developing motivation and motivating development* (pp. 143–163). Elsevier Science.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*, 315–341.

Pekrun, R., Goetz, T., Frenzel, A. C., Barchfeld, P., & Perry, R. P. (2011). Measuring emotions in students' learning and performance: The achievement emotions questionnaire (AEQ). *Contemporary Educational*

Psychology, *36*, 36−48.

Schimmack, U., Oishi, S., & Diener, E. (2005). Individualism: A valid and important dimension of cultural differences between nations. *Personality and Social Psychology Review*, *9*, 17−31.

Schoenhals, M. (1993). *The paradox of power in a People's Republic of China middle school*. Sharpe.

Schwarzer, C., & Kim, M. (1984). Adaptation of the Korean form of the test anxiety inventory: A research note. In H. M. van der Ploeg, R. Schwarzer, & C. D. Spielberger (Eds.), *Advances in test-anxiety research* (Vol. 3, pp. 277−285). Erlbaum.

Schumann, J. H. (2001). Learning as foraging. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 21−28). Second Language Teaching and Curriculum Center, University of Hawaii at Manoa.

Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal*, *59*(4), 339−341.

Sharma, S., Parnian, S., & Spielberger, C. D. (1983). A cross-cultural study of the test anxiety levels in Iranian and Indian students. *Personal Individual Differences*, *4*, 113−120.

Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, *20*, 580−591.

Steiger, J. H., & Lind, J. C. (1980). *Statistically based tests for the number of common factors* [Paper presentation]. Annual Spring Meeting of the Psychometric Society, Iowa City, IA, United States.

Stevenson, H. W., & Lee, S. (1990). Context of achievement: A study of American, Chinese, and Japanese children. *Monograph of the society for Research in Child Development*, *55*, 1−2.

Sundberg, N. D., Latkin, C. A., Farmer, R. F., & Saoud, J. (1991). Boredom in young adults: Gender and cultural comparisons. *Journal of Cross-Cultural Psychology*, *22*, 209−223.

Tanaka, A., & Yamauchi, H. (2004). Cultural self-construal and achievement goal. *Hellenic Journal of Psychology*, *1*, 221−237.

Truitt, S. N. (1995). *Anxiety and beliefs about language learning: A study of Korean university students learning English* (ID 304246027) [Doctoral dissertation, The University of Texas at Austin]. ProQuest Dissertations and Theses Global database

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *3*, 1−10.

Youn, I. (2000). The culture specificity of epistemological beliefs about learning. *Asian Journal of Social Psychology*, *3*, 87−105.

Zirngibl, A. (2004). *Students' emotional experiences in mathematics. Examining the impact of culture, classroom climate and gender from a social-cognitive perspective* [Unpublished doctoral dissertation]. The University of Munich.

FOOTNOTES

[1] Due to a lack of research on boredom, a hypothesis was generated based on Pekrun's (2000) social-cognitive, control-value theory. Given that English performance is critically relevant to all types of academic careers in Korea, it was expected that students would show high engagement in English, no matter how much genuine interest they had. In Germany, however, English achievement is not universally important across all types of academic careers. As such, Korean students may value English more than German students. This situation might lead more students in Germany to be bored during English classes. Thus, I predicted higher average boredom scores for German students.

[2] According to the results of multi-group CFAs, all measures except the interdependent self-construal scale demonstrated configural and metric invariance. This confirms that similar latent constructs are measured and that the underlying factors have the same units across the two groups (Chen, 2007). For the interdependent self-construal scale, it demonstrated only configural invariance. However, a liberal rule was applied for this scale, considering that insufficient fits in self-construal scales might be acceptable because these scales measure a broad range of characteristics compared to other psychological scales (e.g., Levine et al., 2003; Singelis, 1994). For example, Levine et al. (2003) reported low CFIs (.44–.64) and high RMSEAs (.08–.27) in the Korean, Japanese, and American samples. The fit indices in self-construal scales in the present study are in fact better than the ones in earlier studies.

An Examination of a CA-Informed Test of L2 Oral Pragmatics

F. Scott Walters
University of Illinois at Urbana-Champaign, Illinois, USA

The theoretical basis of the relatively new field of second-language pragmatics testing (SLPT), speech act theory (Searle 1969), and the test method usually employed, the discourse completion task (DCT; e.g., Hudson et al., 1995) are called into question by findings from conversation analysis (CA; e.g., Pekarek Doehler, 2018; Sacks, Schegloff, & Jefferson, 1974). While CA has been useful in post hoc L2-test validation (Lazaraton, 2002; Okada, 2010; Ross, 2007), its use as a resource for a priori test-task design is still in its infancy; however, research (e.g., Youn, 2015) suggests some promise. The present study presents data from an L2 oral pragmatics test-development project that builds on earlier work (Walters, 2007, 2009, 2013, in press) by widening the sample of learner proficiencies and the range of EFL pragmatic targets, the latter derived from CA empirical findings (Pomerantz 1978, 1984; Schegloff 2007; Wong & Waring, 2010) and by employing two raters differentially trained in CA, one an English native speaker (NS), the other a Korean L1 speaker with advanced English proficiency. The test protocol involved low-intermediate to advanced L2 learners engaging in oral interactions with a NS tester. Response data were audio-recorded and then analyzed by the two raters according to CA techniques to determine (a) whether the test tasks generated interactions that could be used to validly infer intermediate L2 oral pragmatic ability and (b) the impact of differential levels of CA training on rater performance. Qualitative and quantitative results are given and implications for the Korean EFL educational context discussed, along with suggestions for further research.

Keywords: language assessment, conversation analysis, L2 pragmatics, rater training

INTRODUCTION

L2 Pragmatics Testing, Interaction, and Conversation Analysis

The concept of *interactional competence* (IC) has established itself in the applied linguistics literature over the last few decades. Perhaps the earliest mention of the term is by Kramsch (1986), who pointed out that talk is co-constructed. Other researchers working in the tradition of Hymes (1972), such as Hall (1993) and Young (2000), have proposed analytical frameworks to articulate the features of spoken interaction, such as knowledge of speech-act sequences, turn-taking strategies, and sociolinguistic knowledge. Still others, working in the discipline of *conversation analysis* (CA; e.g., Hall et al., 2011; Markee, 2006; Pekarek Doehler, 2018; Pekarek et al., 2015; Sacks et al., 1974; Schegloff et al., 1977; Schegloff & Sacks, 1973), have contributed to development of the concept of IC by analyzing the mechanisms through which speakers co-construct talk, such as opening or closing a conversation, engaging in turn-taking, or repairing breakdowns in talk. In short, the notion of IC posits that in communication, meaning is *jointly* created (McNamara, 1996) by speakers in specific social contexts, whether by native speakers (NSs) of a given language or by L2 learners. This perspective has implications for the teaching of English as a foreign language (EFL) in the Republic of Korea. However, one may note that the Korean national School Curriculum, while it lists "English conversation" among its subjects (Ministry of Education, Science, and Technology, 2008, p. 11) and breaks down that overall skill into canonical subfields of listening and speaking (e.g., pp. 56–57). the articulation of an IC-focused model appears somewhat tentative (e.g., p. 71). Regardless, for the purposes of this paper, one may note further that the notion of IC also has implications for L2/EFL assessment.

Among the various subfields of L2 assessment pertaining to IC, is *second language pragmatics testing* (hereinafter, SLPT). Research bases for SLPT include interlanguage or cross-cultural pragmatics (Blum-Kulka et al., 1989) and studies into implicatures (Bouton, 1988; Roever, 2006, 2013) and conversational routines (Bardovi-Harlig, 2019). Much SLPT research has been rooted in speech act theory, the origin of which has been credited to Searle (1969, 1975) and Austin (1962), who pointed out that people use language not only to assert the truth or falsity of a

proposition but also to perform intentional acts of one kind or another (1962, p. 3), such as requests, suggestions, invitations, refusals, and apologies (see Kasper, 1992, for a longer list).

However, various researchers have pointed out limits to speech act theory that may undermine its utility in creating assessments of L2 interactional competence. For example, while the theory focuses on the language behavior of individuals, certain speech acts, such as promises, cannot be performed by a single speaker (Hancher, 1979; Mey, 2001; Rose, 1992). Also, Richards and Schmidt (1983) and Kasper (2006) point out that speech acts are chiefly defined in terms of the intentions and beliefs of rational actors, whereas examination of conversational data seldom provides direct evidence of such beliefs, which are often assumed a priori by the investigator/tester. Such assumptions are problematic in terms of test validity (Chapelle, 1999; Kane, 2006; Messick, 1989). That is, inferences about test-taker ability perhaps should be grounded in hard test-data and not in meta-pragmatic, intuitive, observer beliefs about what is normative.

Limitations have also been found with regard to the test method usually employed in speech act theory-related studies, the discourse completion task (DCT). Employed in interlanguage and second language pragmatics research (e.g., Blum-Kulka et al., 1989), its use in SLPT was pioneered at the University of Hawaii by Hudson et al. (1995). A typical DCT prompt consists of a short, written paragraph outlining some interactional setting involving two speakers – a hypothetical one and the test-taker – the latter of which is asked to write a response in a blank below the paragraph. An example prompt from Blum-Kulka et al. (1989, p. 14) is given below:

At the University
Ann missed a lecture yesterday and would like to borrow Judith's notes.
Ann: _____
Judith: Sure, but let me have them back before the lecture next
        week.

As Roever (2001, 2011) points out, the DCT is a very efficient data-collection tool. Standardized, situational prompts can allow for relatively controlled comparisons across subjects of different L1s, different L2s, or proficiency levels, and between native and non-native

speakers. Tests can also be given to large groups at once. Perhaps the greatest advantage in terms of content-related validation (Bachman, 1990) is that the format constrains the data such that the investigator is likely to collect examples of precisely the types of linguistic/pragmatic forms that are of interest.

Despite these advantages, the DCT and its variants (e.g., Bardovi-Harlig & Su, 2019; Brown, 2001; Hudson et al., 1995; Yamashita, 2001) have limitations for the assessment of IC. For example, subjects may elaborate on the prompt-context in ways not anticipated by the investigator (Blum-Kulka & Olshtain, 1986; Faerch & Kasper, 1989; J. Walters, 1979). The DCT also under-represents the construct of IC/pragmatic competence (Grabowski, 2007; Roever, 2011). Most crucially, both the native-speaker norming procedure used in constructing DCTs (e.g., Hudson et al., 1995), and L2 test-taker responses to them, are intuition-based and show signs of inauthenticity when compared with empirical conversational data (Golato, 2003; Walters, 2013). Thus, the theoretical and practical implications of employing speech act theory and traditional SLP test methods are significant.

These implications arguably hold for curricula and educational practices that may not explicitly employ DCT formats in teaching or assessment. For example, in the Korean EFL teaching context, indications of linguistic inauthenticity appear in the Korean national School Curriculum (Ministry of Education, 2008), an appendix of which contains example phrases offered for oral communication (pp. 65–85). However, the examples are decontextualized and likely based on NS intuition. Further, the theoretical framework appears to be notional-functional in approach (page 44 of the 2008 curriculum guide specifically mentions "language functions"), implying a firm connection between form and function that, according to empirical data (e.g., Golato, 2003), may not always exist. Therefore, if L2/EFL test users – from EFL teachers in public schools and colleges to government agencies and businesses – need their instructors and learners to have adequate washback from IC/SLP tests, then speech act-based DCTs, or any assessments of oral skills with decontextualized EFL functions, will only provide information regarding a learner's meta-pragmatic knowledge, not a valid picture of online pragmatic skill. Hence, there is arguably a need for a more valid assessment approach to IC/SLP.

A proposed alternative, as a theoretical basis for SLPT, is conversation analysis (CA). Founded by Sacks et al. (1974), CA

constitutes an approach to the study of naturally occurring talk that eschews a priori categorizing of utterances (e.g., Searle, 1969). Such categories are not imposed on the data (Heritage, 1984, p. 243) but emerge, "bottom-up," in the course of the analysis. The raw data of CA consists of close transcriptions of video- and/or audio-recorded interactions. By avoiding a priori, meta-pragmatic abstractions, CA can offset the theoretical shortcomings of speech act theory with regard to the L2 pragmatics construct. Indeed, CA findings have arguably enriched the construct of IC, an example of which being the model proposed by Celce-Murcia et al. (1995) with its inclusion of phenomena such as turn-taking, self-initiated repair (p. 28), and the use of adjacency pairs (p. 14).

Oral-Tester Behavior and Validity

The above points notwithstanding, there is the practical question of how CA is to be applied to IC/SLPT, and what its impact on test validity might be (e.g., Kane, 2006; Messick 1989). Application of CA to tests of L2 IC can be divided into two general tracks: one consisting of post hoc CA studies that are relevant to L2-test validation, the other being a priori L2 test development.

Along the post hoc CA research track, some research findings suggest that the institutional nature of some L2 oral interview tests, standardized for reasons of test reliability and fairness, is sometimes violated by the interviewer-tester or the test-taker, either of whom may use non-institutional or unanticipated interactional strategies in the course of co-managing the talk. For example, Lazaraton (2002) found that testers sometimes went beyond "neutral" tester roles, providing assistance to examinees in a paired discussion task, to the extent that test results were affected. Tester interventions included rewording test instructions, supplying vocabulary, and correcting examinee responses. Similarly, Ross (2007) performed contrastive CA analyses on transcripts of oral proficiency interviews (OPIs) of an EFL examinee who had backslid to a lower, numerical score since his initial interview, and found that differences in interactional style between that examinee's two OPI interviewers had affected the respective scores. More recently, Kasper (2013) employed CA to determine what interactional strategies testers used on the OPI to facilitate oral-task uptake by L2 examinees. She found that two devices were used, the first being, non-controversially,

the *first pair part* (i.e., initial half of a dyadic sequence that makes expectable a certain response or *second pair part*) that conveyed the task instructions to the test-taker. The second device was *third-position repair*, in which after the tester delivers the prompt and the test-taker acknowledges that information. The tester then either interrupts the test-taker if the latter's ensuing response is not congruent with the expected response. As another example, Okada and Greer (2013) used CA to examine transcripts of 71 English language OPI role-plays in Japan, focusing on tester "pursuit" strategies, which included asking the examinee multiple questions containing possible response alternatives in order to elicit examinee responses.

In sum, this non-exhaustive review gives a sense of the variety of interactional strategies employed by L2 oral-test interviewers. A relevant point is that such interventions may either strengthen a test-designer's *validity argument* (Chapelle, 1999; Kane, 2006), namely, that the test results reflect examinee ability if the tester's action is a legitimate means to avoid, for example, a conflation of pragmatic competence with listening competence in the rating procedure. Alternatively, such tester behavior might weaken the argument if an intervention is analyzably an inappropriate "hint" to the examinee, allowing him or her to display tokens of mastery that would not be manifest absent the intervention.

The other SLPT research track involves a priori application of CA to test design. At this writing there have only been a few such studies. For example, Youn (2015), while investigating the validity of using an open role-play in an English for academic purposes classroom context, developed rating criteria by applying CA to audio-recorded examinees' oral performances during role-play tasks. From these performances, the investigator derived five rating categories: the ability to utter a conversational turn with fluency, degree of appropriate use of language, sociopragmatic sensitivity, degree of interactive engagement and turn organization. These categories were used to create a rubric used by raters to evaluate the recorded interactions. Notable in this study is the attention paid via the scoring rubric to turn organization (Schegloff, 2007), in particular, *adjacency pairs* (i.e., two-turn interactions consisting of a first pair part and a second pair part), the rubric criteria for relative proficiency using such adjacency pairs focusing on completeness and fluency. However, exactly how "completeness" was operationalized is not entirely clear. Nonetheless, Youn's (2015) study is remarkable for the care taken to develop the rating criteria, for the statistical analysis

performed to determine the degree of consistency of rating, and for the use of Kane's (2006) argumentation model to appraise the validity of the procedure.

Another example of an a priori CA-informed test (CAIT) is by Walters (2007, 2009, 2013, in press). This study constitutes groundwork for the present study, but for space reasons only a short summary will be given. (See the sections on delimited target pragmatic domain and test method for details about the testing procedure.) This first CAIT study involved 70 ESL participants, each engaged in a one-on-one interaction with a native English-speaking tester. In the course of a 15- to 20-minute audio-recorded conversation, the tester deployed one of three categories of conversational actions as oral prompts, to which the examinee responded; the audio-recorded responses were later scored independently by two raters trained in CA. The oral prompts given by the tester were based directly on CA research findings. (See the section on test method, below, for more on how CA findings informed the construction of the test procedure.)

Analysis of results partially validated the operationalized pragmatic construct and provided evidence that a CAIT protocol was practical and useful. However, there were certain limitations in this earlier study (hereinafter referred to as "Phase I," the present study being "Phase II"): (a) low reliabilities attributable to homogeneity (advanced ESL proficiency) of the sample (Kunnan, 1992); (b) limited content representativeness (Bachman, 1990; Roever, 2011) in that only three pragmatic actions were tested; (c) the small number of tasks, also arguably affecting low reliabilities; and (d) problems with examinee uptake in one of the three oral test-tasks targeting pre-sequences (see the section on delimited target pragmatic domain). Finally, a question arose concerning (e) the effectiveness of CAIT raters with widely differential CA training: The two raters in Phase I, both formally trained in CA, possessed a high level of CA-rater expertise. However, since such expertise takes time to acquire, it was of interest to discern a lower limit of training required for a rater to engage in a CA-informed rating protocol.

All these issues indicated the need for further research toward answering the overarching question as to the validity, reliability, and practicality − that is, the overall usefulness (e.g., Bachman & Palmer, 2010) − of using a priori-CAIT approaches, as opposed to non-CAIT test methods, in the assessment of IC/SLP.

RESEARCH QUESTIONS

Given the above-mentioned concerns regarding shortcomings of non-CAIT, speech act-theoretic SLPT procedures − as well as the limitations arising from Phase I − two research questions for a "Phase II" (i.e., the present study) were formulated.

RQ1. Would CAIT prompts, targeting an expanded operational domain of pragmatic skills and administered not only to advanced- but also to intermediate-level L2 learners, generate interactive strategies by the tester-interviewer that would positively affect validity of inferences of IC/SLP skills?

RQ2. What might the implications of employing raters with "minimal CA expertise" be for CAIT usefulness?

PARTICIPANTS

Thirty adults participated in this study, all non-native speakers (NNSs) of English either attending or residing near a major Midwestern university in the United States. The L1s of the NNS participants were Chinese (16), Spanish (3), Indonesian (3), Korean (2), and one each of Turkish, Brazilian Portuguese, Afrikaans, Arabic, Tamil, and Estonian. The age range was 24−66 with an average of 35.8 years. Sixteen participants were female and fourteen male. Their length of residence in the U.S. ranged from 2 weeks to 60 months (average 34.8 months), and the reported length of English study and/or use, ranged from 3 to 50 years (average 17.97 years). Participants' educational statuses consisted of undergraduate (3), graduate (13), and post-graduate (14). Given that one of the research questions focused on the efficacy of a CAIT measure with intermediate L2 learners, care was taken to attempt to recruit participants at that proficiency level. The basic criterion was oral-proficiency sub-score on the TOEFL, IELTS, or some other large-scale measure, where available. For the purposes of this study, "intermediate" was defined as a speaker possessing any score below the official minimum score for full-status admission by the University. This resulted in 10 participants, or one third of the total dataset, being so classifiable.

DELIMITED TARGET PRAGMATIC DOMAIN

Assessment Responses and Compliment Responses

The operationalized L2 pragmatic norm for this CAIT consisted partly of targets from the earlier study, which were categories of adjacency pairs (Schegloff, 2007; and described above). Common to both studies were *assessment responses* (Pomerantz, 1984) and *compliment responses* (Pomerantz, 1978). As for the former, there were a number of assessment-response subtypes in the operational norm, for example, *agreements with upgrade*:

A: T's- tsuh beautiful day out isn't it?
B: Yeh it's just gorgeous. ←

Another subtype, *weakened agreements* (or half-agreements), were used:

A: I know but I, I- I still say thet the sewing machine's quicker.
B: Oh it c'n be quicker but it doesn' do the job. ←

As for compliment responses, a number of subtypes were also included in the norm, for example, *reference shifts*, which deflect attention away from the recipient of a compliment to some other relevant, contextualized feature:

A: You're a good rower, Honey.
B: These are very easy to row. Very light. ←

Justification for employing these prompts in the Phase II study was threefold. First, three of the targets were employed in Phase I, and given the new focus on intermediate learners, consistency for comparison purposes across the studies was deemed necessary. Second, the conversational actions of assessments, compliments, and pre-sequences are arguably common to a wide variety of interactional settings − this is apparent from the data in Pomerantz (1978, 1984) and Schegloff (2007) − including informal university-community contexts, whether in classes or in off-campus socializing by adult students, their dependents, and those who are non-students. Third, while pre-sequence responses had

been included in the Phase I study, most of the delivered prompts did not produce uptake (i.e., they failed to elicit responses) because the pre-sequences had inadvertently been delivered at the end of tellings (Schegloff, 2007, p. 215) given by the tester, which diminished the prompts' salience as pre-sequences. Revisions to the test specifications for this second study were intended to eliminate that turn-constructional/prompt pitfall.

Modified and New Test Tasks

The operationalized norm also included *pre-sequence responses* (Schegloff, 2007). These responses can constitute a range of conversational actions/subtypes, for example, *go-aheads* (which allow an interlocutor to complete a proffered verbal action) and *blockings* (which prevent completion of a conversational action by an interlocutor):

A: Hey I got sumpin thet's wild
B: What.   ← [go-ahead]
A: Ya know one a' these great big red fire alarm box thet'r on the corner? I got one.

                    *   *   *

A: Didju hear about thee, pottery and lead poisoning
   [(   )
B: [Yeah Ethie was just telling us ← [blocking]
                  (Shenkein, cited in Schegloff, 2007)

As mentioned in the Introduction, the Phase I study revealed problems with uptake when deploying the pre-sequence prompts − specifically, pre-suggestions from the tester regarding ways to overcome cultural-adaptation issues. Lack of uptake (about 30% of all responses) occurred for two general reasons. First, real-life situations of some examinees did not admit of a life-difficulty about which the tester could proffer simulated advice; hence, no pre-sequence could be constructed and delivered. A second reason occurred when an intended pre-sequence was inadvertently embedded by the tester within a telling (Schegloff, 2007, p. 203) − that is, a narrative told by the tester − causing the intended pre-sequence to lose salience; thus, it did not elicit a go-ahead or blocking action in the examinee. To avoid either of these pitfalls, in

the present study the specifications (and perforce the tester) took care not to preface any of the pre-suggestion prompts with tellings (Schegloff, 2007, p. 215). Also, multiple attempts at eliciting pre-sequence responses pertaining to acculturative or academic challenges were made in the course of some recorded conversations. (In the present study, only four pre-sequence prompts out of 30, or about 13 percent of the total, failed to produce uptake.)

An additional target for this operationalized norm involved *repair moves*, conversational actions that indicate, and repair, a breakdown in understanding. These also were deemed appropriate for this academic test setting since clarifications and corrections necessarily arise in various discoursal contexts, on campus or off it:

A: What happen for the boat?
B: What?   ←
A: What's wrong with the boat?
<div align="right">(Gass and Selinker, 2008)</div>

           * * *
A: Well did 'e ever get married 'r anything?
B: Hu:h?   ←
A: Did jee ever get married?
<div align="right">(Schegloff et al., 1977)</div>

TEST METHOD

Important to the present study is how and to what extent the proposed SLPT measure is "CA-informed." While a more detailed discussion of the methodological and theoretical rationales for this approach to CAIT can be found in Walters (in press), two facets can be briefly presented here. One is the development of the interviewer protocol; the other is in the rating procedure (see CAIT rating method section below). Instead of relying on the intuitions of native-speakers in the creation of DCTs (e.g., Blum-Kulka et al., 1989; Hudson et al., 1995) – which essentially provide not speakers' online practices but their meta-pragmatic beliefs about what constitutes correct behavior (see the Introduction) – the development of this CAIT relied on native-speaker norms as found in CA empirical research (see the section above on

delimited target pragmatic domain). Preliminary focus was on the adjacency-pair dyads in that data (e.g., assessments and assessment responses). Samples of relevant first pair parts were incorporated into the test-item specifications and thus into the tester procedure. In an actual, administered test protocol, these first pair parts appeared realistically in relatively informal, audio-recorded interactions in English between an NS tester and ESL examinee. These interactants conversed on three topics dealing with ordinary aspects of life in a college town: some object or location in the community; an academic- or job-related personal trait or acculturative skill; and a current life-challenge faced by the examinee (new job, academic challenges, etc.). This topic-order was generally observed, though not always. The topics, mandated in the task specifications, were not known by the examinee beforehand. As discussion of each topic proceeded, the tester without warning delivered an oral prompt (e.g., a first pair part) to elicit a pragmatic target (e.g., a second pair part). Each target was associated in the test specifications with one of the conversational topics, for the sake of consistency across administrations. For example, during the discussion of life in a college town, an assessment first pair part was deployed by the tester; during the discussion of a job-related trait, a compliment (intended as a first pair part of a compliment sequence) was given by the tester; and during the discussion focused on a life-challenge, a pre-suggestion was given by the tester. The prompt, intended to elicit repair moves, was not associated with any topic per se and could appear at any time during the recorded conversation.

While it might be convenient to style these interactions as role-plays, that term is inaccurate here since the "roles" of the tester and examinee were not peculiar to social-interactive dyads such as doctor–patient, lawyer–client, customer–clerk, teacher–student, and so forth. Rather, the relative informality of the test procedure was intended to ameliorate the power differential intrinsic to protocols such as the Oral Proficiency Interview (OPI), and thus the participants "acted" within their real-world "roles" as members of the same university community. To establish a degree of informality, all oral prompts (intended first pair parts or other prompts) were embedded meaningfully in the preceding discussion so that the prompts would appear as authentic outgrowths of the emergent talk and not take on the salience of a "test prompt." If there was no examinee uptake with a given prompt, the discussion continued and the tester delivered a second or third prompt later on. For example, where

possible given the trajectory of the conversation, such multiple prompt-deliveries (e.g., sometimes two or possibly three compliments) were delivered by the tester in an attempt to elicit different sub-types of pragmatic responses from the examinee and thus collect a greater sample of pragmatic actions upon which to base an overall judgment as to a portion of L2 oral pragmatic ability. This was unlike the procedure given in the Phase I study, which usually accepted only one example of uptake as a sufficient sample of pragmatic behavior for a given target. The Phase II study conversations ranged from 12.5 to 26.5 minutes and the average was 18.8 minutes.

In sum, those factors which arguably contribute to the test protocol being "CA-informed" are basing the oral prompts in CA empirical data rather than in NS intuitions; those prompts' organic delivery in a test administration approximating the informality of ordinary conversation; and the exposure of the raters to CA findings in order to accustom them to the operationalized norm. This third criterion will be discussed immediately below.

CAIT RATING METHOD

As mentioned above, the other "CA-informed" facet of the procedure was the rating protocol. Similar to that of the Phase I study, this consisted of several components: two independent raters, a scoring rubric with scale and descriptors, a tabular set of rater guidelines giving examples of pragmatic normative behavior (see Appendix B), a sheet for CA-style transcriptions and commentary by the raters, and a series of post-rating hermeneutic dialogues between the raters. Each is discussed below.

Raters

Whereas, in the Phase I study, two raters trained in CA were employed to make post-hoc judgments of NNS oral-production performance, in Phase II only one rater (a native speaker of English) with 10 years' experience studying CA was employed (hereinafter, the "experienced rater"). The other rater, an advanced Korean–English bilingual NNS with multiple years of L2 teaching experience (the

"novice rater"), had no CA training prior to the onset of the study. (Despite inherent interest of rater L1 background as a topic [e.g., Carey et al., 2011; Gass & Varonis, 1984; Winke et al., 2012], rater L1 was not an independent variable in this study.) As mentioned earlier, the rationale for investigating behavior of raters with highly differential CA experience was that CA training is time-consuming, and at this writing, the pool of those experienced in CA might be expected to be small or nonexistent in most classroom-based oral-L2 test settings. Hence, there is potential usefulness in determining to what extent an "on-the-fly" regime of an abbreviated, "applied-CA" training for raters immediately prior to oral-test rating would be sufficient to achieve inter-rater reliability and to support validity of the CAIT method. In the present study, the experienced rater was also the tester. Neither rater engaged in online rating of examinee behavior during test administrations but did so afterward, from audio-recordings, and independently of each other.

Rating Scale and Guidelines Booklet

Rating sheets for each examinee were used by the two raters. Each sheet contained a scale based on Bachman and Savingon (1986). For example, rating descriptors for interactions involving assessment responses were as follows:

> Overall, the examinee shows
> 4 = evidence of control of assessment responses
> 3 = more evidence of control than evidence of no control of
> assessment responses
> 2 = more evidence of no control than evidence of control of
> assessment responses
> 1 = no evidence of control of assessment responses

For the other three pragmatic targets – compliment responses, pre-suggestion responses, and other-initiated repair moves – the descriptors were identical aside from the name of each target. As with the earlier study, the vagueness of the descriptors in the rating scale was intentional. This was because this application of CA to SLPT rating was an exploratory one, and since how CA-trained raters might behave was of interest, it was felt that looseness in the descriptors would give raters some flexibility in applying their CA expertise to the rating task and, in

conjunction with the raters' written commentaries and the hermeneutic dialogues (see below), would also reveal something of their thought processes while rating.

The scoring rubric, containing rating criteria and rating scale, was supplemented with a rater guidelines booklet, which presented examples of the operationalized pragmatic norms in tabular form; the examples were taken directly from empirical CA data: Pomerantz (1978), Pomerantz (1984), Schegloff (2007), and Wong and Waring (2010). The Phase II version of the booklet also included examples of the new, fourth target of other-initiated repair moves. While there were slight edits to the rubric instructions for the sake of clarity, glosses to the guidelines document were not necessary since the NNS rater was an advanced user of the L2.

CA Transcription and Comment Sheet

The reverse side of the scoring sheet had space dedicated to comments by the raters, who were asked to justify their ratings. The sheet also provided space for raters, if they wished, to transcribe passages relevant to the targets while listening to the audio-files; such transcriptions were not required but encouraged. As with the earlier study, any rater transcriptions were to be compared to more careful, post-hoc, CA-style transcriptions made by the principle investigator (PI). Both these types of data, following Fulcher (1993) and Pollitt and Murray (1996), were intended to provide grist for later development of the rating scale.

Hermeneutic Dialogues and Rater Training

An additional aspect of the CAIT rating process involved a hermeneutic post-rating procedure, in which rater agreement was attempted via a contextualized, dialogic process of assessment (Linn et al., 1991; Moss, 1994, 1996; Schwandt & Jang, 2004). In this process, raters jointly revisited differently rated examinee responses as well as the rating materials mentioned above. They then engaged in a dialogue to resolve differences in interpretation. In both studies, this procedure not only helped raters to achieve consensus but also provided information regarding the extent to which the raters were relying on their CA

training. In addition, for the present study, given that one of the raters was a novice CA rater, the dialogues took on the character of ongoing rater-training sessions, of which there were five, one a preliminary training period using recordings of Phase I responses, and four conferences that addressed Phase II data. (See Walters, 2007, for more information on the background and rationale for hermeneutic conferences.)

RESULTS

Tester Interventions

The literature review above provided examples of tester interventions in interview-style L2 oral assessments that sometimes went beyond "neutral" tester roles, strategies such as the use of third-position repair to correct examinee responses, providing assistance to examinees, and "pursuit" strategies to elicit targeted examinee responses. Analysis of the data from this CAIT administration revealed several examples of the tester providing assistance to examinees and employing pursuit strategies but only one example of third-position repair, which is shown in Extract 1. ("T" indicates the tester, and "E" the examinee. See Appendix A for CA transcription conventions.)

```
Extract 1 (Participant 11, OIR, Chinese L1, F)
1. T:      so tell me about your your hometown Nanjing.
2.         (1.0)
3. E:      uh: (0.5) my <hometown>. (1.0) °How to s[ay°(   )
4. T:  →                                          [Nanjing
5.         is- is is is a nice city °like you said°.
6.         (0.5)
7: E:      uh actually my hometown is called Zhengzhou,
```

Here the pragmatic target is other-initiated repair moves (coded OIR). The examinee has earlier in the conversation mentioned the name of her hometown; in line 1 the tester has intentionally misnamed it in an effort to elicit an other-initiated repair move from the examinee. After a one-second pause, in the second turn the examinee evidences thinking about the question by engaging in a delay move (*uh:*) and a repetition of a key noun phrase (*my hometown*), followed by a pause and then

sotto voce a word search (*how do you say*). However, she does not give a sign of having heard the proffered error, (*Nanjing*). In the third turn, line 4, the tester deploys third-position repair repeating the intentional error, which the examinee, orienting (as intended) to the error *Nanjing* as a trouble-source, then responds with an other-initiated repair move in line 7 that includes the correct city name.

Examples of the tester providing assistance to examinees are shown in Extract 2, below. The conversational topic is a variant of the mandated topic of acculturation challenges, namely, the livability of an apartment complex. The pragmatic target is assessment responses (coded ASR).

```
Extract 2 (Participant 5, ASR, Chinese L1, F)
1. T:    So: Orchard South is very very crowded.
2.       (0.5)
3. E:    Oh. crowded,
4. T:→ Crowded means uh there are m- so many people.
5.       (0.8)
6. E:    N:no, I think of- uh compared with Beijing (0.5)
7. T:→ uh-huh
8. E:    the people ((in my neighborhood)) is not very
9.       ((laughter))
```

Here the tester deploys an assessment in line 1. After a half-second silence, the examinee in line 3 initiates a repair, inquiring as to the meaning by repeating the trouble-source (*crowded*) with an intermediate intonation. Note that since examinee understanding of the meaning of the adjective *crowded* is essential for the success of the task prompt, the examinee's attempt at repair is appropriate and authentic. The tester intervention appears in line 4 in the form of a short definition of the term, which rounds out the repair sequence. The stage has now been organically set for the examinee possibly providing the targeted assessment response, which begins in line 6. A second tester intervention can also be seen in line 7, in the form of a continuer (*uh-huh*) deployed after a half-second pause in the middle of the examinee's assessment-disagreement in line 6. After this, the examinee completes her disagreement in lines 8 and 9 (minus the adjective *crowded*, which, given the evidence in lines 2 and 3 that she was unable to comprehend the word, may still be problematic for her in terms of productive competence at this point).

An example of the third tester-intervention type, the pursuit strategy, is given in Extract 3. Here, the topic is the livability of a house and an

apartment complex, respectively, in adjoining college towns. As in the preceding extract, the pragmatic target is assessment responses. The tester delivers the assessment in line 2; however, the examinee does not clearly answer that utterance with an assessment response: the word *yeah* in line 3 is indeterminate, and the examinee continues her story, begun in line 1, about living in a U.S. college town. In line 11, the tester deploys another assessment, which succeeds in eliciting an assessment response in lines 13−15, as does another assessment in line 18, which is answered by an assessment response in line 21.

```
Extract 3 (Participant 13, ASR, Korean L1, F)
 1. E:    I lived in Champaign just (.) six months before,
 2. T:    Oh. Champaign's a a terrible town.
 3. E:    yeah I lived in a <house> like- (.) because I
 4.       wan- I was curious about how to- what it means to
 5.       live in a house in U.S. but it was terrible
 6.       mistake. They had a: rat inside the house.
 7. T:    Oh rats inside the house.
 8. E:    Yeah.
 9. T:    So you got out of there.
10. E:    Yeah.
11. T: →  Yeah. (.) But- Urbana really is equally bad,
12.       (1.0)
13. E:    Yeah it's um .hhh ((nasal exhalation))(0.2) I
14.       think at least like uh uh where I <live now>,
15.       (0.2) I- at least I don't have mice. ((smile))
16. T:    ((laugh)) So that's good,
17. E:    Yeah.
18. T: →  Yeah. Oh okay. And- and uh the neighbors are
19.       pretty (.) bad neighbors.
20.       (0.8)
21. E:    .hhh uhm? In Orchard Downs ((apartment complex
22.       name)) I think it's (.) better?
```

One may note that the interventions shown in Extracts 2 and 3 (as elsewhere in the data) facilitated both the delivery of the assessment prompt and authentic examinee responses, and therefore did not threaten validity.

Statistics

Descriptive statistics for this CAIT administration were obtained in two phases. The first was before the raters engaged in a series of post-rating, hermeneutic sessions; the second occurred after the raters had ended each hermeneutic post-rating session and revised scores were

recorded. Here, for simplicity, some tables give only the post-dialogic values. Table 1 provides descriptive quantitative results, values differing only slightly between the experienced rater and the novice rater. Table 2 gives indices of inter-rater reliability: Pearson's *r*, percent absolute agreement, and percent absolute plus adjacent agreement. For the post-hermeneutic session, the Pearson's values slightly exceeded those in the Phase I study, and the proportions of absolute agreement in this study were slightly higher (53% versus 40%; see Walters, 2007, p. 177). Also, coefficient alphas were calculated for both raters, with *r* = .56 for the novice rater and, interestingly, *r* = .40 for the experienced rater. The alpha values in the present study (*r* = .54 for the novice rater and *r* = .22 for the experienced rater) were significantly higher than the respective attenuated values obtained in the Phase I study (*r* = .21 and *r* = -.02; see Walters, 2007, p. 170), though still somewhat low. That they were still relatively low could be attributed either to the relatively small number of test tasks in the measure or to shortcomings of the rating scale (see Discussion).

TABLE 1. Oral CAIT Phase II Descriptive Statistics, Post-hermeneutics

Statistic	Novice CA Rater	Experienced CA Rater
Mean Score*	12.90	12.90
Median	13.50	13.50
Mode	14.00	14.00
Range	7.00	7.00
Variance	3.61	3.40
SD	1.90	1.84
Skewness	-0.98	-0.94
Kurtosis	-0.17	0.02

Note. *Possible total score: 16. *n* = 30.

TABLE 2. Oral CAIT Phase II Indicators of Inter-rater Reliability

Statistic	Session	
	Pre-hermeneutic	Post-hermeneutic
Pearson's *r*	0.55	0.82
Percent Absolute Agreement	0.33	0.53
Percent Adjacent Agreement	0.23	0.37
Percent Absolute + Adjacent	0.57	0.90

As shown in Table 3, item means (*p*-values) for the four test tasks were generally similar for each rater, with the task intended to elicit compliment responses emerging as being of greatest difficulty, the task targeting pre-suggestion responses of moderate difficulty, the task focused on assessment responses of average difficulty, and the task designed to elicit repair moves being of greatest facility.

TABLE 3. Phase II Oral CAIT Item Means (*p*) by Rater*

Pragmatic Target	Novice CA Rater	Experienced CA Rater
Assessment Response (ASR)	.56	.57
Compliment Response (CMR)	.16	.20
Pre-sequence Response (PSR)	.24	.28
Other-Initiated Repair (OIR)	.87	.87

Note. *Post-hermeneutic values only.

The following section will provide a few examples of discrepant rating that may suggest explanations as to why the inter-rater reliability and absolute agreement were relatively low.

Rater Training and Inter-rater Hermeneutic Dialogues

The rating sessions were preceded by a short training period in which the novice rater evaluated a set of audio-recorded responses from the Phase I study. This process enabled the new rater to become familiar with the scoring rubric (descriptors and scale), the transcription and comment sheet, the scoring rubric, and the general format of the recorded interactions before engaging with the new data sets. After rating the Phase I interactions, the PI gave verbal feedback on the ratings and clarified any unfamiliar aspects of the rating materials. The rating period then commenced and lasted four weeks.

At approximately weekly intervals, the two raters held a series of post-rating dialogues, which had two purposes. One was for the novice rater to share impressions about the rating process. For example, the novice rater reported that the initial rating runs were "hard" and required multiple hearings of the audio files as well as continual referencing of the rubric. She also reported that in some instances where the rubric did not seem helpful, she attempted to arrive at a rating by recourse to her intuitions about NS verbal behavior. A single recorded interaction took

from 20 minutes and, in the beginning, up to 45 minutes. She estimated an average rating time per recording of 30–35 minutes (compared to an average 25-minute reported time by the experienced rater). However, she further stated that rating became "easier as time went on" and that she felt "more confident" about rating the interactions toward the end of the rating period, reporting that the series of post-rating sessions were "helpful" in assisting her to "confirm what I'm doing is right or wrong or a little deficient." She also reported that the PI's clarifications of the rubric and explanatory glosses added to the rubric were also helpful.

The other purpose of the post-rating dialogues was for the two raters to attempt to achieve a "hermeneutic circle" (Moss, 1994) – that is, a resolution to discrepant ratings through joint examination of problematic examinee utterances analyzed in conversational context and with reference to the scoring rubric. The highest rate of discrepancies, perhaps predictably, appeared during the first dialogue; here the inter-rater reliability was relatively low, $r = .78$. In some of the succeeding three post-rating sessions, however, higher inter-rater reliabilities emerged (see Table 4), arguably due to the increasing experience by the novice rater. However, not all dialogues resulted in changes to individual ratings; if either rater felt that an original rating was valid based in his or her understanding of the operational norm as reflected in the rubric and scoring guide, the score was not altered.

TABLE 4. Inter-rater Reliability Coefficients Pre- and Post-hermeneutic Session

Session	Session Type		*n*
	Pre-hermeneutic	Post-hermeneutic	
Session 1	.13	.78	6
Session 2	.85	.95	4
Session 3	.49	.78	11
Session 4	.61	.72	9
All	.55	.82	30

Rater Discrepancies

The dialogues uncovered various apparent causes for the rating-discrepancies, which in general differed between the two raters. For the novice rater, one cause was simple unfamiliarity with applying the rubric to relevant turns and sequences of talk in the audio-recording. Another cause involved the novice rater occasionally missing phenomena that seemed non-nativelike but which were in fact normative, including, for example, *reference shifts* manifest in compliment responses (coded CMR), as shown in Extract 4.

```
Extract 4 (Participant 3, CMR, Chinese L1, M)
1. T: =oh volun- volunteering (.)
2. E: yeah volun[teer
3. T:            [volunteering is very nice. Uh:=
4. E: =yeah volunteering yeah [yeah
5. T:                        [you're- you're you-
6.    you're a good man.
7. E: yeah. I- yeah I- because I'm retired.
```

The focus of this conversation was the examinee's having volunteered to supervise gatherings of high-school children in a local church. After some hesitations in line 5, the compliment is delivered in line 6. In line 7, instead of responding with a token *thank you*, the examinee produces a reference shift, drawing attention away from himself; that is, he does not address the issue of whether or not he is *a good man*, but points to circumstantial reasons for his action of volunteering: *because I'm retired*. This latter part of his response in line 7 is arguably normative, according to CA findings as represented in the scoring rubric, but the novice rater did not recognize it as such until during the first hermeneutic session, when the normative phenomenon of the reference shift was pointed out in the scoring guidelines. The novice rater then revised that subscore from 2 to 3, which then matched the experienced rater's score. Both raters, it should be noted, declined to give this compliment response a full score of 4 as the first part of the examinee's turn in line 7, *yeah*, appeared non-nativelike relative to the examples in the scoring guide; that is, the *yeah* seemed an assessment response embedded in an overall compliment response. Thus, the response was determined to have both NNS- and NS-like elements.

Other examples of misperceiving native-like turn designs as non-native error were pauses and delays incorporated into assessment

responses, as shown, for instance, in Extract 5, below.

```
Extract 5 (Participant 16, ASR, Chinese L1, M)
1. T: So you've lived here for about a year, and uhm (.)
2.    uh Ur:bana: very boring city.
3.    (1.0)
4. E:  mm and I- I don't think so. (1.0). I like this city
```

In this example, the one-second pause in line 3 and the use of *mm* in line 4 are analyzably not indications of lack of ability at response, as the novice rater had first surmised, but rather valid pragmatic options normatively available to an interlocutor who wishes to disagree with an assessment by "cushioning the blow" (Pomerantz, 1984).

There were a few instances in which the experienced rater also overlooked a significant aspect of an interaction, whereas the novice rater was able to detect it as the latter gained more experience. One example is in Extract 6, below, targeting pre-suggestion responses (coded PSR). The examinee is recounting how he often has difficulty ascertaining the meaning of English words in some academic texts when the tester delivers a pre-advice item in lines 1 and 2:

```
Extract 6 (Participant 20, PSR, Arabic L1, M)
1. T: Well you know the way to solve these (.) semantic
2.    problems.
3.    (0.2)
4. E: yeah.
5. T: you: you can (.) look at a sentence and you: try to
6.    guess the meaning [from:
7. E:                   [okay but some[time
8. T:                                 [the context.
9. E: okay but- but sometime (0.2) it make difference.
10.   (within) (.) the same (.) context.
```

The experienced rater initially gave this examinee's utterance, shown above in line 4, a score of 3, feeling that it was more a non-native-like "continuer-style" go-ahead. However, the novice rater pointed out that the examinee's contradictory assertions (lines 7, 9, and 10) indicated that the utterance at line 4 was actually a blocking move unheeded by the tester/interlocutor, to which the experienced rater agreed. A re-examination of the audio data further revealed that the examinee had placed a noticeable increase in volume on that syllable (*yeah*), which supports the interpretation that it was a block, deserving of a full score of 4.

Another cause of rater discrepancy concerned severity of rating. One

example is shown in Extract 7.

```
Extract 7 (Participant 30, CMR, Spanish L1, M)
1. T: and you've told me a bit of your personal
2.    history. And um (0.2) uh: I think you're
3.    a very adaptive intelligent person.
4.    (0.2)
5. E: Yeah I am adaptive.[Eh
6.                       [uh-huh
7. E: like I told you I used to work on the coast it's-
8.    totally d- different conditions from my- from
9.    the city.
```

Lines 2 and 3 contain the compliment-prompt, which is followed by a short silence at line 4 into which a recipient-answer is expectable. Line 5 contains the beginning of the recipient's compliment response, *yeah I am adaptive*, which is followed by an account in lines 7 through 9. This response, however, is not a normative compliment response; it is not an acceptance token; nor is it a scaled-down agreement, diminishing the strength of the evaluative term in the compliment, nor is it a reference shift. In fact, it more closely resembles a direct agreement of an assessment (Pomerantz, 1978). Here both raters gave lower scores, the novice rater a 3 and the experienced rater a more severe rating of 2. However, hermeneutic re-analysis suggested that while the design of Examinee 30's turn resembled an assessment response, his further utterance in lines 7–9 could be interpreted as a partial reference-shift inasmuch as it seems to focus on the tester's adjectival compliment *adaptive* and not *intelligent*, suggesting that the examinee was orienting to the tester's utterance-prompt as a compliment. Thus, a compromise score of 3 was deemed appropriate.

A cause of discrepancy that was common to both raters, was the lack of salience of other-initiated repair (OIR) sequences. As noted above, prompts with assessments, compliments, and pre-sequences were delivered, per the item specifications, in generally the same order and with similar topical content. Hence, their appearance was more or less predictable to the raters, who could find them in the recording and rate the responses. However, trouble sources as grist for OIR moves could appear at any time in the interaction – the tester would deploy them at opportune moments as the conversation proceeded. Thus, raters, at least early in the rating sessions, sometimes failed to isolate the OIR sequences. As the novice rater reported, some passages in the recorded interactions "are very crowded and I get tired." After the first dialogic

session, as both raters gained experience with rating this new test task, OIR sequences were found and rated with regularity.

In any event, both raters agreed that applying the rating scale to oral responses was often problematic due to the imprecise wording of the scale. As the experienced rater put it during the fourth hermeneutic dialogue, "When is a compliment response a 3 and when a 2? How many instances of 'oddness' [i.e., examples of NNS-like utterances elicited via multiple prompts] qualify? What features of turn design − and what combinations of features − make for a 3 or a 2?" Accordingly, provisional expedients were arrived at through rater dialogue; for example, it was decided that a PSR that looked like a continuer but seemed to function as a go-ahead would receive a 3.

DISCUSSION AND CONCLUSIONS

The fundamental question driving this two-phase study is to what extent a CAIT protocol is preferred to a non-CA-informed approach to testing L2 pragmatics. A discussion of the results of this study, in light of the two research questions, is offered as a step in addressing that overarching inquiry. In addition, implications for EFL educational settings will be suggested.

The first research question asked whether CAIT prompts, targeting an expanded operational domain of pragmatic skills and administered not only to advanced but also to intermediate-level L2 learners, could generate interactive strategies by the tester-interviewer that would positively affect validity of inferences of IC/SLP skills. As seen in the Test Method section, earlier, extracts from the response data in this CAIT administration show that the tester made three types of interventions that aided the L2 pragmatics assessment process. In the first case, a third-position repair elicited a targeted pragmatic move on the part of the examinee; in the second, a tester-aid in the form of a vocabulary gloss (*crowded*) allowed the interactants to close a repair sequence and then the examinee to respond to a prompt with a targeted second pair part; and in the third case, a pursuit strategy allowed the tester multiple times to elicit a targeted second pair part. One may note that the interventions shown in Extracts 1, 2, and 3 (as elsewhere in the data) were embedded within evolving, online interactions, organic with

the emergent sequential contexts and relevant to the respective conversational topics, as well as being consonant with the testing purpose, which was to target select subskills of L2 pragmatics. The interventions also elicited responses that arguably were authentically embedded in those conversational sequences – such authenticity supported by evidence of interlocutor orientation toward previous turns in an analyzably logical manner. It is important to note that the authenticity of these responses and the organic effectiveness of the oral prompts, both unfolding between interlocutors in real-time, stand in contrast to the meta-pragmatic responses elicited by DCTs in other studies. (The focus on adjacency pairs in this study, which it shares with some DCT studies, should therefore not be a source of confusion.) Hence, one may conclude that this CAIT method allowed for tester interventions by which the tester and examinee co-created authentic sequences which were evidence supportive of valid inferences of the respective examinees' commands of a few delimited domains of IC/SLP.

The second research question asked what the implications of employing raters with "minimal CA expertise" might be for CAIT usefulness. It may be convenient to explore these issues with reference to three major, interconnected dimensions of testing: reliability, validity, and practicality (e.g., Bachman & Palmer, 2010). Regarding the issue of reliability, one may note, first, that the inclusion of L2 speakers with intermediate-level proficiency in the testing sample seems to have allowed for better reliabilities, unlike with the advanced-only examinee sample in the Phase I study. Second, it seems necessary to note also that the post-hermeneutic, inter-rater reliability value of $r = .82$, while more encouraging than the Phase I results, was still somewhat low, suggesting issues with the rating process (see below). Nonetheless, there is some evidence for test usefulness for the CAIT rating procedure given a general increase in reliability coefficients both before and after each hermeneutic dialogue, as shown in Table 4, from the first rating session ($r_{pre} = .13$; $r_{post} = .78$) to the fourth ($r_{pre} = .61$; $r_{post} = .72$). This seems indicative of the novice rater's increasing facility with the rating process (as also testified in the post-rating interviews) as well as of the efficacy of the hermeneutic dialogues in resolving (some) rater differences.

However, the relatively low rater self-consistency as evidenced in the Alpha values as well as discrepancies in rater severity suggest that refinements to the scoring rubric are indicated; hence, the vague descriptors, while necessary for this early, exploratory stage of CAIT

development, should be supplanted by more definitely worded rater guidelines (cf. Fulcher, 1993; Pollitt & Murray, 1996). Such refinements might assist raters, whether experienced in CA or novice, make judgments with greater accuracy and without depending on intuitions about verbal behavior, help resolve differences in rating, and facilitate better, individualized feedback to novice CA raters. Although inter-rater reliability indices of approximately r_{post} = .82 may be undesirable for high-stakes test settings, it may be that a CAIT protocol such as the one examined here may be useful for low-stakes settings such as classroom-based EFL assessments. However, with refinements to the scoring rubric, the potential for CAIT use in high-stakes settings may emerge with further trialling.

Regarding practicality of the CAIT method, three issues − rating time, time for administration, and rater training − are relevant. As for the first, one may note that at the beginning of the rating period, the novice rater reportedly took about 45 minutes per examinee recording, that rating time falling to about 30 minutes, whereas the experienced rater reported an estimate of 25 minutes throughout. Not only did the rating times fall for the novice, but the inter-rater reliability with the experienced rater increased overall. This suggests that with adequate time and training for raters the usefulness/practicality of the measure would increase. Regarding test administration time, the average test-response recording time in this study was 18.8 minutes, which is comparable to the 15−30-minute interview times for the ACTFL Oral Proficiency Interview (Language Testing International, 2019). While these times may be challenging in some high-intakes testing contexts, it seems reasonable to hypothesize that given a more precise, CA-data-driven scoring rubric, as suggested in the preceding paragraph, average rating time could be reduced and the practicality of a CAIT method thereby enhanced.

Finally, regarding the rater-training procedure: One may recall that the novice rater began training with select audio-recordings of Phase I interactions, a rating sheet, and rating guide, the outcomes of using which were discussed with the PI, followed by four weeks of "live" rating punctuated with near-weekly hermeneutic dialogues with the PI. Qualitative evidence for the effectiveness of this training procedure might be found in the post-rating interviews, in which the novice rater reported that the hermeneutic dialogues and the PI's clarifications of the rubric were helpful in addressing rating issues. However, objective evidence for the training procedure's effectiveness may also be found in

the increased post-hermeneutic, *r*-values obtained across the four dialogic sessions. Significantly, they may also be found in the post-rating interviews in which the novice, an advanced bilingual Korean L1 rater, reported instances in which the descriptors were insufficient and decided to award a score based on her intuitions of L1 behavior. While intuition-based scoring is not the goal of a program of CAIT development, it is notable that the rater was *consciously aware* of these two modes of rating − one based on objective CA findings, the other on intuitive judgment. That this novice had attained such self-reflective awareness after only a few days of training and exposure to a limited range of CA examples, speaks to the possible practicality of the protocol, even for novice raters new to CA. One might argue that this finding has implications for CA-informed rater-training in, for example, university EFL programs in Korea. As such, this study can be seen as generally resonating with the findings of Kondo (2010) and Kim (2009), who found evidence supporting the practice of training and employing NNS raters of L2 oral performance assessments.

In brief, it may be argued that, given the above quantitative and qualitative results collected in response to the two research questions, the present study adds to the evidence provided in Phase I of this line of research (Walters, 2007, 2009, 2013) which suggests possible usefulness of a CAIT protocol over intuition- and speech act theory-based SLP measures such as the DCT.

Nonetheless, there are some limitations to this study, which may have bearing on future CAIT development. One issue is the limited content representativeness of the measure (Bachman, 1990; Roever, 2011) in that only four aspects of IC/SLP were operationalized. Future versions of an oral CAIT should target other aspects of L2 pragmatic competence, such as use of apologies, requests, conversational openings and closings, and other forms of repair. Second is the problem of the vague rating descriptors, mentioned earlier. The process for refining these is beyond the scope of this paper (for suggestions, see Walters, in press), but suffice to say that any modifications to the scale must conform to empirical CA studies and not reflect meta-pragmatic beliefs as to the nature of IC/SLP. Moreover, relevant to this line of research are the above-mentioned facets of rating: rater variability (Weigle, 1998), self-consistency (Kim, 2009; Kondo, 2010; Lunz et al., 1990), rater severity (McNamara, 1996), and individualization of feedback on experienced versus novice raters (e.g., Elder et al., 2005; Knoch, 2011;

Lunt et al., 1994; O'Sullivan & Rignall, 2007; Wigglesworth, 1993). Given that several of these studies noted limitations on the effectiveness of rater training, and given that only a single novice rater was involved in the present CAIT study, it will be useful to replicate this study with a set of CA novices (and CA-conversant raters) assisted by a more finely crafted rubric. This may enable researchers to discover to what extent such a rater training-and-testing protocol can adapt to the multi-faceted challenges of IC/SLP rater performance.

The Author

F. Scott Walters is a lecturer in the Department of Linguistics at the University of Illinois at Urbana-Champaign (UIUC). He received his PhD from UIUC and has taught TESOL courses both there and at Queens College. His research interests include L2 assessment, L2 teacher training, and conversation analysis. Email: fwalters@illinois.edu

References

Austin, J. L. (1962). *How to do things with words*. Cambridge University Press.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice*. Oxford University Press.

Bachman, L. F., & Savingon, S. J. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL Oral Interview. *The Modern Language Journal*, *70*, 380−390.

Bardovi-Harlig, K. (2019). Routines in L2 pragmatics research. In N. Taguchi (Ed.), *The Routledge handbook of second language acquisition and pragmatics* (pp. 47−62). Routledge.

Bardovi-Harlig, K., & Su, Y. (2019). *Developing an empirically-driven aural DCT for pragmatics assessment* [Paper presentation]. The 21st Annual Conference of the Midwest Association of Language Testers (MwALT), Indiana University, Bloomington, IN, United States.

Blum-Kulka, S., & Olshtain, E. (1986). Too many words: Length of utterance and pragmatic failure. *Studies in Second Language Acquisition*, *8*, 47−61.

Blum-Kulka, S., House, J., & Kasper, G. (1989) (Eds.). *Cross-cultural pragmatics: Requests and apologies*. Ablex.

Bouton, L. F. (1988). A cross-cultural study of ability to interpret implicatures

in English. *World Englishes*, *7*(2), 183−196.

Brown, J. D. (2001). Pragmatics tests: Different purposes, different tests. In K. R. Rose & G. Kasper (Eds.), *Pragmatics in language teaching* (pp. 301 −325). Cambridge University Press.

Carey, M. D., Mannell, R. H., & Dunn, P. K. (2011). Does a rater's familiarity with a candidate's pronunciation affect the rating in oral proficiency interviews? *Language Testing*, *28*(2), 201−219.

Celce-Murcia, M., Dornyei, Z., & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics*, *6*(2). https://escholarship.org/uc/item/ 2928w4zj

Chapelle, C. (1999). Validity in language assessment. *Annual Review of Applied Linguistics*, *19*, 254−272.

Elder, C., Knoch, U., Barkhuizen, G., & von Randow, J. (2005). Individual feedback to enhance rater training: Does it work? *Language Assessment Quarterly*, *2*(3), 175−196.

Faerch, C., & Kasper, G. (1989). Internal and external modification in interlanguage request realization. In S. Blum-Kulka, J. House, & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 221−247). Ablex.

Fulcher, G. N. (1993). *The construction and validation of rating scales for oral tests of English as a foreign language* [Doctoral dissertation]. Lancaster University. http://languagetesting.info/articles/store/FulcherPhD.pdf

Gass, S., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). Routledge.

Gass S., & Varonis, E. M. (1984). The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, *34*(1), 65−87.

Grabowski, K. C. (2007). Reconsidering the measurement of pragmatic knowledge using a reciprocal written task format. *Columbia University Working Papers in Applied Linguistics*, *7*(1), 1−48.

Golato, A. (2003). Studying compliment responses: A comparison of DCTs and recordings of naturally occurring talk. *Applied Linguistics*, *24*, 90−121.

Hall, J. K. (1993). The role of oral practices in the accomplishment of our everyday lives: The sociocultural dimension of interaction with implications for learning of another language. *Applied Linguistics*, *14*, 145−167.

Hall, J. K., Hellerman, J., & Pekarek Doehler (Eds.). (2011). *L2 interactional competence and development*. Multilingual Matters.

Hancher, M. (1979). The classification of cooperative illocutionary acts. *Language in Society*, *8*, 1−14.

Heritage, J. (1984). Conversation analysis. In J. Heritage (Ed.), *Garfinkel and ethnomethodology* (pp. 233−292). Polity Press.

Hudson, T., Detmer, E., & Brown, J. D. (1995). *Developing prototypic measures of cross-cultural pragmatics*. University of Hawai'i Press.

Hymes, D. (1972). Models of the interaction of language and social life. In J. J.

Gumperz & D. Hymes (Eds.), *Directions in sociolinguistics: The ethnography of communication* (pp. 35−71). Holt, Reinhart, & Winston.

Kane, M. (2006). Validation. In R. Brennan (Ed.), *Educational measurement* (4th ed.; pp. 17−64). American Council on Education; Praeger.

Kasper, G. (1992). Pragmatic transfer. *Second Language Research*, *8*(3), 203−231.

Kasper, G. (2006). Speech acts in interaction: Towards discursive pragmatics. In K. Bardovi-Harlig, C. Felix-Brasdefer, & A. Omar (Eds.), *Pragmatics and language learning* (Vol. 2, pp. 281−314). University of Hawaii, National Foreign Language Research Center.

Kasper, G. (2013). Managing task uptake in oral proficiency interviews. In S. Ross & G. Kasper (Eds.), *Assessing second language pragmatics*. Palgrave Macmillan.

Kim, Y. (2009). An investigation into native and non-native teachers' judgments on oral English performance: A mixed methods approach. *Language Testing*, *26*(2), 187−217.

Knoch, U. (2011). Investigating the effectiveness of individualized feedback to rating behavior: A longitudinal study. *Language Testing*, *28*(2), 179−200.

Kondo, Y. (2010). Examination of rater training effect and rater eligibility in L2 performance assessment. *Journal of Pan-Pacific Association of Applied Linguistics*, *14*(2), 1−23.

Kramsch, C. J. (1986). From language proficiency to interactional competence. *Modern Language Journal*, *70*, 4, 366−372.

Kunnan, A. J. (1992). An investigation of a criterion referenced test using G-theory, and factor and cluster analyses. *Language Testing*, *9*, 30−49.

Language Testing International. (2019). *ACTFL speaking assessment: The Oral Proficiency Interview (OPI)*. https://www.languagetesting.com/oral-proficiency-interview-opi-2

Lazaraton, A. (2002). *A qualitative approach to the validation of oral language tests (Studies in Language Testing, 14)*. Cambridge University Press.

Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press.

Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria [CSE Technical Report 331]. *Educational Researcher*, *20*(8), 15−21.

Lunt, H., Morton, J., & Wigglesworth, G. (1994). Rater behaviour in performance testing: Evaluating the effect of bias feedback [Paper presentation]. The 19th Annual Congress of Applied Linguistics Association of Australia, University of Melbourne, Melbourne, Australia.

Lunz, M. E., Wright, B. D., & Linacre, J. M. (1990). Measuring the impact of judge severity of examination scores. *Applied Measurement in Education*, *3*(4), 331−345.

Markee, N. (2006). A conversation analytic perspective on the role of quantification and generalizability in second language acquisition. In M. Chalhoub-Deville, C. A. Chapelle, & P. Duff (Eds.), *Inference and*

generalizability in applied linguistics: Multiple perspectives (pp. 135−162). John Benjamins.

McNamara, T. (1996). *Measuring second language performance*. Addison Wesley Longman.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (pp. 13−103). Macmillan; American Council on Education.

Mey, J. L. (2001). *Pragmatics: An introduction*. Blackwell.

Ministry of Education, Science, and Technology. (2008). *The school curriculum of the Republic of Korea*. http://gangwonepik.weebly.com/uploads/1/3/8/5/13851570/national_school_curriculum-english2008.pdf

Moss, P. A. (1994). Can there be validity without reliability? *Educational Researcher*, *23*(2), 5−12.

Moss, P. A. (1996). Enlarging the dialogue in educational measurement: Voices from the interpretive traditions. *Educational Researcher*, *25*, 20−28.

Okada, Y. (2010). Role-play in oral proficiency interviews: Interactive footing and interactional competencies. *Journal of Pragmatics*, *42*, 1647−1668.

Okada, Y., & Greer, T. (2013). Pursuing a relevant response in oral proficiency interview role plays. In S. Ross & G. Kasper (Eds.), *Assessing second language pragmatics*. (pp. 288−310). Palgrave Macmillan.

O'Sullivan, B., & Rignall, M. (2007). Assessing the value of bias analysis feedback to raters for the IELTS writing module. In L. Taylor & P. Falvey (Eds.), *IELTS collected papers: Research in reading and writing assessment*. (Studies in Language Testing 19; pp. 466−476). Cambridge University Press.

Pekarek Doehler, S. (2018). Elaborations on L2 interactional competence: The development of L2 grammar-for-interaction. *Classroom Discourse*, *9*(1), 3−24.

Pekarek Doehler, S., & Pochon Berger, E. (2015). The development of L2 interactional competence: Evidence from turn-taking organization, sequence organization, repair organization, and preference organization. In T. Cadierno & S. Eskildsen (Eds.), *Usage-based perspectives on second language learning* (pp. 233−268). De Gruyter Mouton.

Pollitt A., & Murray, N. L. (1996). What raters really pay attention to. In M. Milanovic & N. Saville (Eds.), *Studies in language testing 3: Performance, testing, cognition, and assessment*. (pp. 74−91). Cambridge University Press.

Pomerantz, A. (1978). Compliment responses: Notes on the co-operation of multiple constraints. In J. Schenkein (Ed.), *Studies in the organization of conversational interaction* (pp. 57−101). Academic Press.

Pomerantz, A. (1984). Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action: Studies in conversational analysis* (pp. 79−112). Academic Press.

Richards, J. C., & Schmidt, R. W. (1983). Conversational analysis. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 117–154). Longman.

Roever, C. (2001). *A web-based test of interlanguage pragmalinguistic knowledge: Speech acts, routines, implicature* [Unpublished doctoral dissertation]. University of Hawai'i at Manoa.

Roever, C. (2006). Validation of a web-based test of ESL pragmalinguistics. *Language Testing*, *23*(2), 229–256.

Roever, C. (2011). Testing of second language pragmatics: Past and future. *Language Testing*, *4*(2), 463–481.

Roever, C. (2013). Testing implicature under operational conditions. In S. Ross & G. Kasper (Eds.), *Assessing second language pragmatics*. Palgrave Macmillan.

Rose, K. (1992). Speech acts and questionnaires: The effect of hearer response. *Journal of Pragmatics*, *17*, 49–62.

Ross, S. J. (2007). A comparative task-in-interaction analysis of OPI backsliding. *Journal of Pragmatics*, *39*(11), 2017–2044.

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest schematics for the organization of turn-taking for conversation. *Language*, *50*, 696–735.

Schegloff, E. A. (2007). Some variation in sequence organization. *Sequence organization in interaction: Volume 1. A primer in conversation analysis*. (pp. 220–230). Cambridge University Press.

Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, *8*, 289–327.

Schegloff, E. M., Jefferson, G., & Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, *53*(2), 361–382.

Schwandt, T. A., & Jang, E. E. (2004). Linking validity and ethics in language testing: Insights from the hermeneutic turn in social science. *Studies in Educational Evaluation*, *30*, 265–280.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press.

Searle, J. R. (1975). A taxonomy of illocutionary acts. In K. Gunderson, (Ed.), *Language, mind, and knowledge: Minnesota studies in the philosophy of science* (Vol. 7, 344–369). University of Minnesota Press.

Walters, F. S. (2007). A conversation-analytic hermeneutic rating protocol to assess L2 oral pragmatic competence. *Language Testing*, *24*(2), 155–183.

Walters, F. S. (2009). A conversation analysis-informed test of L2 aural pragmatic comprehensions. *TESOL Quarterly*, *43*(1), 29–54.

Walters, F. S. (2013). Interfaces between a discourse completion test and a conversation analysis-informed test of L2 pragmatic competence. In S. Ross & G. Kasper (Eds.), *Assessing second language pragmatics*. Palgrave Macmillan.

Walters, F. S. (in press). Some considerations regarding validation in CA-informed oral testing for the L2 classroom. In S. Kunitz, N. Markee, & O. Sert (Eds.), *Emerging issues in classroom discourse and interaction: Theoretical and applied CA perspectives on pedagogy*. Springer.

Walters, J. (1979). The perception of deference in English and Spanish. In C. A. Yorio, K. Perkins, & J. Schachter (Eds.), *On TESOL '79* (pp. 288–296). TESOL.

Weigle, S. C. (1998). Using FACETS to model rater training effects. *Language Testing*, *15*(2), 263–287.

Wigglesworth, G. (1993). Exploring bias analysis as a tool for improving rater consistency in assessing oral interaction. *Language Testing*, *10*(3), 305–323.

Winke, P., Gass, S., & Myford, C. (2012). Raters' L2 background as a potential source of bias in rating oral performance. *Language Testing*, *30*(2), 231–252.

Wong, J., & Waring, H. Z. (2010). *Conversation analysis and second language pedagogy: A guide for ESL/ EFL teachers*. Routledge.

Yamashita, S. (2001). Using pictures for research in pragmatics: Eliciting pragmatic strategies by Picture Response Tests. In T. Hudson & J. D. Brown (Eds.), *A focus on language test development* (pp. 35–56). University of Hawai'i Press.

Youn, S. J. (2015). Validity argument for assessing L2 pragmatics in interaction using mixed methods. *Language Testing*, *32*, 199–225.

Young, R. (2000, March). *Interactional competence: Challenges for validity* [Paper presentation]. Joint Symposium on Interdisciplinary Interfaces with Language Testing, Annual Meeting of the American Association for Applied Linguistics and the Language Testing Research Colloquium. Vancouver, British Columbia, Canada.

APPENDIX A

CA Transcription Conventions

Part A. Temporal Relationships

(1) Square Brackets [   ]

    [a vertically-arrayed pair of left square brackets on two adjacent lines indicates a

    [point of overlap onset between two speakers

    ]   a vertically-arrayed pair of right square brackets on two adjacent lines indicates a point

    ]   where overlap between utterances ends

    Example:
```
1        A:      yeah  volun[teer
2        B:                 [volunteering is very nice.
```

(2) Equal Sign – two meanings, (a) and (b):

Used when one person's utterance follows on to a second person's utterance with no discernable silence between the utterances; that is, when one utterance is "latched" onto another. Note the latching between lines 1 and 2 below:

```
1        A:      volunteering is very nice. Uh: =
2        B:      =yeah  volunteering
```

(3) Silence
    (a) Numbers in parentheses – e.g., (0.3) – represent silence in tenths of a second.
    (b) A dot in parentheses (.) indicates a micropause, hearable but not easily measurable.

Part B. Symbols denoting aspects of speech delivery

(4) Punctuation (used to indicate intonation, not grammaticality)

> . (period) − falling intonation contour
> , (comma) − continuing intonation
> ? (question mark) − rising intonation, not necessarily a question
> :: (colons) − prolongation of sound preceding the colon; the longer the sound, the
>   more colons used, e.g., "tough cou:rses" or "Oh: : :"
> - (hyphen) − after a word or word part, a self-interruption or cut-off

(5) Other Symbols

> > < (greater-than, less-than) − talk between signs rushed.
> < > (less-than, greater-than) − talk between signs slowed down.
>   < (less-than sign alone) − succeeding talk starts with a rush.
> .hh   inhalation (sometimes with a raised dot before the letter "h").
> <u>talk</u> − underlines indicate emphasis or stress, either by higher pitch or increased loudness.
> <u>: :</u> − underlined colons indicate rising intonation contour.
> (words) − empty parentheses indicate utterances that are inaudible and thus not transcribable; parentheses with words indicate conjectural utterance-parts.
> °words° − small circles indicate speech much softer than previous talk.
> ((words)) − double parentheses contain information on non-verbal events accompanying the interaction, such as interlocutor hand-gesture, direction of eye-gaze, or other actions performed by the speakers.

APPENDIX B

Rater Guidelines [excerpt]

Instructions: For quick reference, here are some examples of English pragmatic norms from the CA literature regarding assessment responses, compliment responses, pre-sequence responses, and other-initiated repair moves. This list of examples is *not* exhaustive; actual participant performance will of course differ from these examples in various ways: vocabulary, grammar, intonation, topic, etc. However, keeping these examples in mind may be useful when listening to the examinee responses. (Note: Some examples below have been edited slightly for reasons of space and clarity.)

Assessment Responses

Response Types	Provisional Normative Examples (sample only)
AGREEMENTS	
Upgrades	
Stronger evaluative terms	A: T's- tsuh beautiful day out isn't it? B: Yeh it's just gorgeous ←
Intensifier modifying prior term	A: She seems like a nice little [lady B: [Awfully nice little person ←
Same Assessments	
Addition of *too*	A: She was a nice lady – I liked her B: I liked her too ←
Pro-terms (*that was/he is*)	A: He's terrific! B: He is. ←
Downgrades (weak agreements)	A: Oh it was just beautiful. B: Well thank you uh I thought it was quite nice. ←

Can Teachers Thrive at *Hagwons*? Challenges and Possibilities in Private Language Schools in South Korea

Esther Ahn

Chungdahm Institute, Seoul, Korea

This paper argues that it is possible for *hagwon* teachers in South Korea to not only survive but also thrive both professionally and holistically. This standpoint contradicts the mainstream perception of learning and teaching in the private sector where neoliberal principles seem to predominate. I discuss both challenges and opportunities one can find in such a learning environment by critiquing critical pedagogy and suggesting relational pedagogy. This paper may be particularly relevant to the interests of English language teaching professionals as it questions the generally simplistic dismissal of *hagwon* teachers in the ELT world in Korea while little research has been done on the subject. With an emphasis on the complexity of human beings and the versatility of learning places, I hope to draw attention to the notion that no human places, including *hagwons*, are completely hopeless.

Keywords: *hagwon*, private education, relational pedagogy, critical pedagogy, literature in EFL/ESL classes, education in South Korea, neoliberalism

CONTEXT: AN OVERVIEW ON EDUCATION IN KOREA

South Korean parents spend over US$15 billion on private education annually, which is more than anywhere else in the world. The system of private institutes called *hagwons*, also referred to by some as "shadow education," are typically after-school programs that specialize in language learning, test preparation, math, and science, among other subjects. These private institutions for supplemental education are ordinary and commonplace in South Korea, but they are stigmatized and

regarded suspiciously in the Western world.

One important factor that has contributed to this *hagwon* culture in Korea is that Koreans value education more than anything else. For them, education is a gateway to progress, distinction, and proof of one's value. The importance of gaining admission into top universities makes competition for spots in them fierce and thus fuels the demand for supplementary lessons from private institutions. *Hagwon* culture is thus indicative of Koreans' fierce competition and determination that have shaped modern Korea with its phenomenal economic success since its independence from Japanese colonialism and its emergence from catastrophic postwar poverty.

As the private educational system has become a free market where supply and demand rule everything, many people have criticized and warned of legitimating *hagwons* at the cost of undermining public education. Indeed, permeated with the principles of neoliberalism, which can be summed up as "attaining the maximum financial profit," private education may lead to the exclusion of the economically disadvantaged from equal educational opportunities. However, with nearly 75 percent of Korean students currently attending *hagwons*, the scale and the scope of the private educational sector has reached a point where they cannot easily be dismissed as illegitimate places.

In this context, instead of focusing on delegitimizing them, this paper asks how we can reimagine the private education system and move toward it. As few reports of the private education sector as a context of teaching and learning exist, it is time we had more serious discussions about what is being taught and how it is being taught in these places. This paper is an attempt to examine the content and quality of education and draw attention to teachers working in private institutes. From the perspective of a teacher currently working at a *hagwon*, I will particularly focus on practices in the context of private English language education in *hagwons* in South Korea and delineate their challenges and possibilities. By suggesting relational pedagogy, I attempt to counter the mainstream ideas on teaching and learning at a *hagwon* and to give an affirmative answer to the question of whether teachers at *hagwons* can *thrive*, not just survive.

Like everything else in life, the reality of *hagwons* and the lives of the teachers working at those places are complicated. Individual teachers' motivations to work at *hagwons* are widely different. However, just like any other institution, *hagwons* are also a human space where both

challenges and hopes exist. In this light, the criticisms of *hagwons* in this paper come from a hope for revision and reform for the betterment of education and people's lives.

Korean students seem to be the unhappiest of students in a global context. After being put through long days at school, they go to various academies to study more. After dinner, they do homework from school and their academies. The intensity with which students' mothers drive them to study might seem quite alarming to a third person. Still, many students are appreciative of the linguistic growth that they have made through the aid of supplementary education. In this paper, I particularly emphasize the use of literature in language learning environments. When a language is taught critically and through good literature, not only students' critical thinking skills but also their moral imagination can be developed. Teachers at *hagwons* can teach and lead class discussions in such a way that students would gradually be able to see themselves in a globally social context and judge for themselves about the content and the quality of education that they are getting.

CHALLENGES

While *hagwons* offer educational and relational possibilities, the constraints presented in their system cannot be ignored. Not only are the material conditions of *hagwon* learning difficult for students, but also the conditions of work for teachers can be challenging. In particular, teaching hours are grueling. During students' vacation periods in summer and winter, there are extra classes called "intensives." In those periods, some teachers may be asked to teach 38 to 40 hours a week. Also, they are often asked to work on national holidays.

Teaching certificates or TEFL qualifications may not be so useful in this environment. This is because, in *hagwons*, those who make decisions about hiring and promoting teachers do not value them as much. Consequently, there has been an influx of untrained, inexperienced "native-speaking" teachers, which contributes to the degeneration of TESOL as a field. Many of the foreign teachers do not have any teaching certificate or a degree in education. More often than not, they are fresh out of college and more enthusiastic about being exposed to different cultures in Asia than teaching per se. Naturally, most of these teachers are only temporary residents in Korea. This temporary status

may have affected their seriousness in trying to change or challenge the existing ways of learning and teaching at *hagwons*. Due to this tendency, the profession of teaching English suffers as a whole from the lack of quality standards and a public loss of faith in English teachers as professionals.

Another factor that contributes to the degeneration of teaching quality is the evaluation system at private institutes. Teachers are evaluated based on students' responses to a few questions regarding each teacher's class. One may question the usefulness of the survey questions, which tend to be vague, let alone the validity of the marks on each teacher evaluated by 11- to 15-year-old students. More often than not, entertaining teachers get higher marks on these surveys. Because the managers take the results into account when negotiating the teacher's salary at the time of renewal of his or her contract, the pressure on the teachers to be fun in classes, above all other requirements, is considerable.

Another cost of *hagwons* is that students may lose interest and motivation in their public education and their stress may increase. Beyond furthering class inequalities, this system has a severe human impact. The fact that Korean high school students sleep only an average of four hours per night while attending class for up to eleven hours a day cannot simply be attributed to the existence of *hagwons*. Still, they do not alleviate the intensity of their competition. The negative effects of limited sleep have been linked to weight gain, high rates of depression, and higher suicide rates. These studies on the health effects of *hagwons* have called into question the value of this supplementary private education when it physically brings harm to students (West, 2017, p. 88).

These conditions make *hagwons* a difficult place to practice creative learning and an integrated life for a teacher. Teachers who make their living teaching English in a *hagwon* may undergo an inner conflict in that they work for an institution that allegedly harms both the individual students and society by deepening inequality through restricted access to linguistic capital. Some may argue that this contradiction is a condition of life in today's world where neoliberalism has become firmly entrenched. To have an integrated life, however, it is important to bridge the gap between personal and professional values. In his book *The Courage to Teach*, Palmer (1998) says that good teachers join self and subject and students in the fabric of life. He encourages teachers to

recover their identity and integrity and reclaim the wholeness of their lives in the very place where their work takes place: in the classroom. To overcome these challenges, it is important for teachers working at *hagwons* and other similar neoliberal education settings to carefully examine the status quo and ask what is possible in all educational contexts for better teaching and learning and to try to bring about change. Those individual teachers, as marginalized workers in difficult jobs with low benefits, have a stake in raising awareness and voicing the necessity for reforming the system.

In an article from *The New York Times*, Koo (2014) depicts *hagwons* as "soulless facilities, with room after room divided by thin walls, lit by long fluorescent bulbs, and stuffed with students memorizing English vocabulary" (para. 8). The article suggests that legislation be passed criminalizing excessive private education. Over the last few decades, some education scholars have warned teachers against any sort of engagement with private education, on the basis that this could potentially legitimize private education at the expense of supporting and improving government-sponsored education. However, in Korea, the enterprise of private English language education has simply become too large to ignore and to pursue policies of disengagement with. Reflecting this reality, there might be more useful and practical questions to pursue.

It is in this context that this paper raises the following questions: Are *hagwons* completely hopeless, or can they offer some benefit for society and individuals? Can they also become sustainable over the long run? What changes are desirable and needed? Is it possible to teach at a *hagwon* in a way compatible with one's personal or moral values and convictions? Can teachers at *hagwons* thrive with a sense of joy and fulfillment, not just survive? What are the possibilities that *hagwons* uniquely provide in understanding education?

A CRITIQUE ON CRITICAL PEDAGOGY

In light of the notion that everyone should have equal opportunities and access to education, making a living at a *hagwon* may create tension arising from the gap between personal and professional values. Because *hagwons* do not get any financial support from the government, their fees can be quite high, which means that only financially privileged

students can attend the classes with top-notch learning resources and advanced technology. For a teacher who believes that educators also have a responsibility to work for changes to amend injustices and inequalities in the world, the neoliberal principles by which most *hagwons* are run create inner conflicts. Is it possible, in this context, to adopt a practice that matches one's own personal moral and political convictions at a *hagwon*?

In Korea, the English language serves as a gatekeeping mechanism to both universities and employment through testing requirements. In his paper on *hagwons* in South Korea, West (2017) says that simply removing English language requirements to eliminate the inequalities of the system is not enough to address the inequalities that are exacerbated by the *hagwon* system in Korea. He sees *hagwons* as an almost perfect embodiment of neoliberal education policy. They are loosely regulated and run strictly according to market principles through which students get the education they can afford, rather than according to a right-to-education principle whereby all students are guaranteed an equal opportunity to education (p. 155).

In resisting the market principles in education, in some cases West suggests critical pedagogy as a teaching methodology, which particularly focuses on helping students question and challenge the existing power structure as a way to address social justice and democracy in a learning environment. In his paper, he examines the language learning and teaching in neoliberal spaces in Korea and asks if critical pedagogy is applicable in those environments. His definition of neoliberalism follows "financialization of everything" (Harvey, as cited in West, 2017, p. 86), and he compares its five principles to the private English education system in Korea: uncritically accepting markets, a concentration of wealth and power, deregulation of industry, privatization of services that were once provided by government, and an emphasis on individualism over social responsibility.

Tied to the ideas of deregulation, privatization, and individualism, neoliberalism is fundamentally about a certain ideal of freedom. *Hagwons* thrive on the ideas of privatization of learning; marketization and competition, which might come at a cost to the democratic process; the quality of learning and teaching; and the meaningfulness of relationships between students and teachers. West points out that, at private institutions in Korea, language teaching can be seen as a product, which is produced, distributed, and controlled by the private market.

While the value of the product has been increasing, its distribution has been increasingly unequal.

For a critical pedagogue, teaching is essentially a political act. Even choosing textbooks becomes a matter of controlling knowledge production and domination. Therefore, a critical pedagogue might suggest that the materials used for learning should be learner-generated wherever possible. Unless the textbooks contain some political ideology, however, at an actual learning place, what they contain might affect students less compared to how they are taught by teachers. That is why, when two different teachers use the same textbook, their students come out of the classroom with very different focus and understanding.

It is true that critical pedagogy provides an important view on education and developing critical consciousness. With its focus on human rights, it offers a valuable view for challenging repressive structures in the status quo and constructing a fairer society. However, encouraging students to develop a certain kind of lens so that they would suspect a power structure in everything they see might not be the best kind of education. Too much emphasis on its principles may be misleading. If neoliberalism is about gaining profit from everything, critical pedagogy demands a constant awareness of the existing power structure while politicizing everything, in which case it becomes yet another ideology. The view that sees teaching as an inherently political act presupposes an assumption that teachers try to indoctrinate the students with their political beliefs. It also assumes that students will naively adopt the ideology that is either consciously or unconsciously being taught by the teacher. When a class is run based on a political perspective that constantly asks who has the dominant power, however, it may lose its direction of learning. For these reasons, a teacher might want to first question the idea of putting ideology and radical politics at the center of their teaching.

Of course, students should be encouraged to motivate themselves and to make good, informed decisions about the direction or the field of their interest. However, it takes time and training for their minds to develop enough to be able to self-reflect, think critically, and make judgment calls judiciously. It might be a little premature to talk about democratic environments when students are not even fully aware that their freedom can be misused by their own erratic choices, by following instant gratification, when they cannot even control their urges.[1] Rather, a more subtle and nuanced attention should be paid to what sort of person is in

the position of authority in class, that is, the teacher. Depending on the teacher's character, the classroom may become interactive or oppressive, directionless or clear, and directed to a certain goal. Teachers have many roles to play in the classroom, and this includes the judicious exercise of legitimate authority. Here, teachers' own interest and personal character affect many intuitive decisions made in class. Teachers attuned to the issues of social justice, for example, may want to alter the balance of power in the classroom and help students imagine alternatives.

In guiding students' attention, a certain kind of language learning can help students develop critical skills so that they do not follow any ideologies blindly or compromise their personal values of the current system. In class, they should be encouraged to critique many different aspects of life, including their education system, and to envision alternative settings. Through their continuous self-examination and questioning of the structure and legitimacy of existing power in societal systems, they may picture free, yet responsible, individuals and a kinder and fairer society. For example, students sometimes talk about the intensity of education in Korea and express strongly adverse feelings against its education system. In response to their anger and frustration, the teacher can duly acknowledge their situation and, ideally, lead the discussion critically. As a teacher, it is important to listen to and legitimize students' concerns. But a teacher should be able to remind the students that they are developing adequate tools so that they could address the problems in constructive, not reactionary, ways. In this way, students' difficulties can be duly acknowledged while the teacher maintains the purpose of the class in the manner of persuasion rather than force.

Considering the prevalence of private education in South Korea, it might perhaps be more edifying to turn our attention to how we teach at *hagwons* and how teachers can fortify students' moral imagination even in those places. What is being taught and how well it is being taught in private institutions can affect students' critical thinking and develop their empathy for others. In this way, the private sector of education can supplement the formal one.

CREATING AN INTERACTIVE SPACE THROUGH LITERATURE

It is a questionable stereotype that "serious" learning can happen only in a highly structured, teacher-created environment. To talk about the problems underlying the assumption, we need to ask first what we mean by certain terms that are often used uncritically. When we ask what sort of teaching is better than others, we assume that there is a certain hierarchy in the quality of classes. But it is not easy to objectively measure the degree of effectiveness in teaching and to what extent teachers are accountable for students' progress in learning. This is because, in judging the nature and the quality of a class, one should consider a variety of factors such as class activities, the teacher's philosophy, and his or her focus in the process of learning. Given that no one wants to be "barely adequate" when facing the opportunity to excel, what does it mean to excel in teaching, particularly in language learning environments?

In business, it is important that the head of a company lead by example, set expectations, and push the boundaries. For educational institutions, the quality of teaching and of leadership itself heavily depends on the leaders' ability to discern good teachers. Also, managers at *hagwons* can create a work environment where teachers and administrators cooperate and communicate with each other to set the goals together, and individual performance is assessed objectively.

A banking style of teaching, which is based on John Locke's idea of *tabula rasa* that teachers deposit knowledge in the otherwise empty minds of the students (Locke, 2009), has been, for the most part, removed from contemporary educational discourses. In today's discourses on learning and teaching, it is normative to say that teacher-centered, test-focused, and academic learning may not be the most effective methodology. Influenced by Western theories, the English language curriculum in Korea has also moved toward principles of progressive educational practice, which emphasizes student-centered teaching and communicative competence. The basic assumption of this approach is that students are not passive, empty vessels. They have their own values, interests, and agency. Along with their own lived experience, their parents, families, friends, and teachers help them construct their worlds. For them, a language teacher at a *hagwon* is only one of many different

sources of learning. Nevertheless, it is not an unimportant role.

The most fundamental level of facts about a learning environment is that a teacher and many students are voluntarily in a room where learning takes place. In this teacher–student relationship formed in a classroom, the teacher can help students have the right frame of mind to engage with the language they are learning and use the opportunities before them so that they can build a good learning environment together and learn in a meaningful way. From a student motivational perspective, a part of the language teacher's job is to help them think about their reasons for learning a foreign language and the value it could offer to them and their future lives. In their learning process, students should feel connected not only to their teachers but also to their peers, their institution, and the language itself.

Even though democratizing the classroom sounds fair and desirable, having it as a top priority in class may create some unintended negative effects, such as losing a sense of purpose for their learning. When the teacher tries to accommodate a majority of students who are not mature enough to judge properly, the class may lose its direction. Nevertheless, one thing that a class can be democratic about is the way information is shared and valued.[2] In an effective language learning environment, students are expected to speak much more than the instructor. One way to realize this is for the teacher to ask engaging questions that may be personally meaningful to students.

In asking personally and socially meaningful questions, good literature becomes a particularly great tool in a private language institution. After all, education is about nurturing individuals to grow in thinking and empathizing with others while trying to make good changes for the world. But realizing these may not necessarily come from emphasizing politically motivated activities predominantly associated with social justice. Great literature introduces the richness of human experience and the complexity of a life in communities with subtlety and nuance. Discussing such books by asking good questions may teach students to be kind and generous towards the disadvantaged and the marginalized. In the end, however privileged they are, and however unfairly they have access to more opportunities in learning another language, when the lessons themselves are dealing with serious life questions such as injustice and the pain of others through literature, they help students to see themselves through a critical lens.

Admittedly, some of the novels read at private institutes seem to be

beyond the linguistic level of the students, and this causes a great deal of frustration both for teachers and students. But it is not impossible to have a meaningful class when the goal is certain. It is true that teachers might have to negotiate in terms of eliciting nuanced details and subtleties in the books. But many elements of the books can still be used for discussing personalized, relevant, and meaningful questions for the students. In particular, when the Socratic method is applied well, it can encourage students to think critically and creatively.

By raising questions for discussion, teachers can challenge the prevalent view that private educational institutions take: Students are customers. Rather than try to please or merely entertain them, teachers can use questions to persuade and reason with them. Teachers should be able to inform parents of a student's behavior and performance without fearing that they will withdraw their child from their *hagwon*. Not all *hagwons* mindlessly pursue their revenue no matter what. Such teachers, who resist mediocrity and work with a sense of accountability, do exist at *hagwons*. They continue to develop their teaching skills, to improve themselves, and to be alert to the changes happening outside the classroom.

For teacher motivation, the typical image of a ladder or a pay raise may sometimes help. But, for a teacher, there is no greater measurement of progress than a class filled with meaningfulness, where personally believed values are practiced and materialized.[3] In this sense, the system that negotiates teachers' payment based on how young learners simply "like" the teacher demoralizes those who continuously try to improve their lessons. Everyone who participates in education should know that teachers' likability does not always match the "goodness" of their classes.

RELATIONAL PEDAGOGY

In a politically polarized world such as ours, there is an increasingly loud voice to see teaching as a political act. But teaching is, first and foremost, a relational act. Learning of anything, but especially a language, is a social act involving interpersonal communication, as it inherently involves interaction with another person. In its current form, the statement sounds like it is saying "learning a language is a social act

because it is a social act." Language learning is about everydayness. Its learning environment can be a good place for practicing one's values, educational principles, and philosophies in the everyday context. A language class is not a separate world from each participant's other world. Before and after class, for both teachers and students, things happen beyond the classroom that have an effect on what happens in the classroom where a teacher and many students intersect and interact with one another. Focusing on this organic view of students and teachers as individual human beings, relational pedagogy underscores the inherently social and relational nature of language learning.

At *hagwons*, the number of students in each class fluctuates each term, but class sizes are usually capped at fifteen students, so teachers have an excellent opportunity to develop close rapport with students. Also, administrators and teachers pay a lot of attention to the potential dropouts who seem to have various behavioral and performance issues. They talk to the students at a personal level and ask themselves how to make their classes more engaging. I am hesitant to think that the reason behind this attentiveness is just the profit of the institute. Each individual who makes up the whole institution wants to make their work more meaningful and more effective for all the participants. From their firsthand experience, they learn that focusing on the human connection is a crucial way to encourage the students. This kind of quality teacher–student and peer relationship is considered one of the most valuable resources in education and has been shown to be linked to a range of desirable educational outcomes (Furrer et al., 2014).

It is curious, then, despite the importance of positive teacher–student relationships that has been explained through various theoretical frameworks, that research on teacher–learner relationships and rapport remains almost non-existent in language education. It is telling of our less than proper understanding of what constitutes a quality teacher–learner relationship in a language learning setting.

We humans are socially situated, relational beings. Relating is an active, ongoing process in a constant state of flux. And a large part of the nature and the quality of a relationship is dependent on the role of human agency. We can, to varying degrees, actively construct our relationships with things, peoples, and the world, and thereby potentially alter how we choose to relate to various aspects of our lives. We choose a certain perspective in relating to our contexts in our own personal ways based on our personal frames of reference. As Charon (2009)

succinctly puts it, "An environment may exist, but it is our definition of it that is important" (p. 28). A teacher's perception of the self is not entirely determined by a seeming environment, but instead it emerges from the interaction of a network of personally relevant relationships, from the way one relates to the present contexts and cultures based on his or her own interpretation of them.

The interaction of all the relationships in the classroom together generates collective qualities such as group dynamics, rapport, trust, and general atmosphere. And all of these are known to be very important for effective teaching and successful learning. Nonetheless, the hub of all relationships in the classroom is likely to be the teacher, who forms a relationship with each individual learner as well as the group collective as a whole. In this sense, Palmer (1998) defines good teaching as something that cannot be reduced to technique: "Good teaching comes from the identity and integrity of the teacher" (p. 10). He then adds, "Good teachers possess a capacity for connectedness. They are able to weave a complex web of connections among themselves, their subjects, and their students so that students can learn to weave a world for themselves."

From a relational perspective, it is problematic to set boundaries on the professional domain since each teacher's own perception of such boundaries is different from another's. Definitions of different domains will be highly personal and what is relevant for a teacher's professional domain will also vary in ways meaningful for an individual at any one point in time. Despite the popular notion of professionalism, however, it is likely that personal and professional spheres will interconnect strongly for teachers.

As the learners' relationships to parents, teachers, classmates, and administrators deeply affect their progress of learning, teachers' relationships to their schools, resources, tasks, colleagues, and administrators also influence their personal and professional lives. Also, how they relate to their own competencies and beliefs about teaching and learning shapes their teaching. A relational view thus suggests a holistic view in seeing teachers as having various cognitive, affective, and motivational dimensions. It attempts to connect mind, body, and individual experience in social contexts narrated across our pasts, presents, and futures. It sheds a new light on learning and teaching processes, drawing attention to inherently dynamic and complex ways we relate to the world and others around us.

TEACHERS AS COMPLEX HUMAN BEINGS

The aims of the teaching profession are to help learners to learn to the best of their abilities. Because learners exist, teachers are hired and educational institutions exist. Given learners' centrality to that process, it is natural that we strive to understand the main beneficiaries of the teaching and educational processes. However, Mercer (2018, p. 505) calls into question "a kind of inequality" at work in discussing language learning regarding the relative status of teachers compared to learners.

What teachers think, feel, and believe manifest in their practices in the classroom. There is a correlation between teachers' personal mental lives and their practices in the classroom and learners' lived experiences. In other words, their personal and professional wellbeing are directly connected to the quality of their teaching as well as student performance. According to insights from various strands of neuroscience, it is clear that teacher and learner psychology are intricately related, and their emotions and motivation can be contagious to learners and coworkers. In this regard, Mercer points out that there has been so little attention devoted to teachers in comparison to diverse attempts to understand learners. She argues that it is important also to pay close attention to teacher psychology with the same degree of complexity and diversity. Both teachers and students are active contributors to their own subjective experience in classroom.

Traditionally, teachers have been seen in terms of "the learners' environment" rather than as agents in their own right. A prevalent discourse that expects teachers to put their learners first, "sacrificing" themselves for the benefit of their learners has led to our being reluctant to talk about teacher needs and their professional wellbeing. This widespread notion of learner-centered mentality has contributed to the predominance of interest in the learner, drawing much attention to learner individuality and learner needs. However, as an unintended consequence, the emphasis on understanding learners might have neglected the teachers' needs and motives. This inequality becomes exacerbated particularly in a private educational setting where the focus on learners is heightened.

It is important to understand teacher motivation before we consider how teachers can foster the motivation of their learners. It is obvious that a teacher with high job satisfaction and positive morale is more

likely to teach creative and effective lessons. Teachers need to look after themselves and nurture their own motivational basis so that they can give something worthwhile to their students. However, teachers are increasingly under enormous stress, particularly those working at *hagwons* due to the lack of job security, despite the fact that they have extra responsibilities such as maximizing student retention rates − and thus having to "please" students to some degree − and minimizing language anxiety. This is especially true for teachers whose first language is not English.

Most people want to make their professional lives compatible with their personal values. Likewise, many teachers working in private educational institutions are asking how they can realize their personal value and meaning systems through their profession. Among the values teachers may prize, the quality of the relationship between teacher and student can be one of the most rewarding aspects of the teaching profession, but it can also be the source of emotionally draining and discouraging experiences.

Teachers at *hagwons* should be able to teach in a way that is compatible with their values and principles. Even the less-than-ideal class material can offer an opportunity for teachers to be creative by expanding the subject or by breaking it down into digestible pieces for the students. An environment, even when it just *feels* highly controlled and autocratic, does not require all participants to succumb to it. We need to ask what is really called for to truly resist the invisible force that dehumanizes individuals and undermines their resilience in holding onto their values and freedom. The question of resistance is an important one in all kinds of oppressed environments. But *hagwon* teachers should ask themselves whether there is absolutely no room for exercising their freedom creatively. As a teacher, any small act of choosing what s/he believes to practice in class in order to edify students' development in learning can be a meaningful form of resistance. More often than not, teaching in such a way does not directly oppose what *hagwons* as a supplementary educational place are trying to do.

As Mercer (2018) points out, a language teacher's identity is extremely comprehensive in that it is not an object but "something that is enacted, dynamic, and multifaceted" and "a process or way of being" (p. 507). People as individuals or groups can never be simply understood and explained away. A theory needs to be able to accommodate the real-life complexity of the human experience. It requires a holistic,

organic view of a human being. The novel, *Looking for Alaska*, by Green (2005), has an apt description of this complexity when a character portrays his friend, Alaska, after her death:

> ...we are greater than the sum of our parts. If you take Alaska's genetic code and you add her life experiences and the relationships she had with people, and then you take the size and shape of her body, you do not get her. There is something else entirely. There is a part of her greater than the sum of her knowable parts. (p. 220)

People are continuously *becoming*. Human experience is inherently dynamic and constantly evolving in response to either self-motivation or external circumstances. In particular, it is noteworthy that individuals are not passive recipients of the influence of their social contexts, but instead, they exercise their agency to differing degrees in making meaning out of their experiences and contexts, while both influencing and being influenced by them. Unlike the popular notion behaviorism purports, contexts and cultures are not "monolithic external objective variables affecting an internal inner world" (Mercer, 2018, p. 511). They are first subjectively interpreted and thus bear different meanings for individuals. Across time and space, individuals are being continuously shaped by multiple contexts, which are not externally or objectively defined. These multiple contexts of one's life are all part of what a person thinks, believes, and feels. In other words, contexts are within us.

CONCLUSIONS

When asked why Koreans care so much about education, novelist Min Jin Lee, who is working on a novel about *hagwons* in Korea, surmises that Koreans want power. She asks, "What will a mother do when she wants her child to have a little more than she has? And how is she different from any of us?" (as cited in Bolotnikova, 2019, para. 9). She points to the humanity of mothers, children, managers, and teachers who are all participants in making a *hagwon* culture. They are vulnerable and frightened humans who are very complicated. The system and culture of the *hagwons* may be harsh, but upon closer inspection, there may be a social purpose they are filling. Lee remarks that *hagwons*

are not just for studying, but instead they almost function like a community center providing daycare services. As such, I suggested in this paper that *hagwons* are, fundamentally, human places where empathy, intelligence, and a capacity for change exist.

Even at private institutes, teachers are one of the most valuable stakeholders in the language learning and teaching processes. It is, therefore, lamentable to have little understanding of what inspires teachers and makes them flourish in their professional roles. More importantly, teachers matter to their learners "not just because of what and how they teach, but because of who they are as people" (Day et al., 2007, p. 1). For teachers, a fuller sense of fulfillment comes not from a narrow concept of professionalism or being good at what they are paid to do, but from human relationships in which all available emotional, intellectual, and spiritual capacities are played out in very personal ways. Those teachers who are keen to the quality of their interaction with students try to develop classroom communities that promote academic, social, and emotional growth, and thus restore the language of meaning, values, character, and interest. And these teachers do exist in *hagwons* in South Korea.

There has been a voice asking for reforming our education system, not based on test scores but based on creativity and social/emotional capacities. Some of the classes that *hagwons* offer, such as literature classes, may contribute to developing such capacities within students. In a troubling era of political polarization, when we focus on the relational aspects of learning and teaching in classrooms, and good questions are raised through reading great literature, it is possible for the teachers to be both personally and professionally fulfilled and for the students to learn to think critically and creatively despite the constraints of a supplementary educational setting.

THE AUTHOR

Esther Ahn is a Cambridge CELTA-certified teacher at Chungdahm Institute in Seoul, South Korea, and a member of the National Coalition of Independent Scholars (NCIS). She received her MA from Yonsei University and has published and translated works notably in the interdisciplinary field of literature and theology. Email: esther.ahn@ncis.org

REFERENCES

Bolotnikova, M. N. (2019, February 13). Min Jin Lee on her new novel and writing about the Korean diaspora. *Harvard Magazine*. https://harvardmagazine.com/2019/02/min-jin-lee-on-her-new-novel-and-writing-about-the-korean-diaspora

Charon, J. M. (2009). *Symbolic interactionism: An introduction, an interpretation, an integration* (10th ed.). Pearson.

Day, C., Sammons, P., Stobart, G., Kington, A., & Gu, Q. (2007). *Teachers matter: Connecting lives, work, and effectiveness*. McGraw-Hill.

Furrer, C.J., Skinner, E. A., & Pitzer, J. R. (2014). The influence of teacher and peer relationships on students' classroom engagement and everyday motivational resilience. *National Society for the Study of Education*, *113*(1), 101−123.

Green, J. (2005). *Looking for Alaska*. SPEAK.

Koo, S.-W. (2014, August 1). An assault upon our children. *The New York Times*. https://www.nytimes.com/2014/08/02/opinion/sunday/south-koreas-education-system-hurts-students.html

Locke, J. (2009). *An essay concerning human understanding*. Oxford University Press. (Original work published 1689)

Mercer, S. (2018). Psychology for language learning: Spare a thought for the teacher. *Language Teaching*, *51*(4), 504−525.

Palmer, P. J. (1998). *The courage to teach: Exploring the inner landscape of a teacher's life*. John Wiley & Sons.

West, G. B. (2017). Doing critical language pedagogy in neoliberal spaces: A materialist narrative analysis of teaching young learners of English in a Korean hagwon. *Second Language Studies*, *36*(1), 75−165.

FOOTNOTES

[1] In this regard, not all resistances should be taken seriously. Needless to say, that a student does not want to do a certain task at a given moment does not mean that the task is not worthwhile.

[2] When technology is used appropriately, it can greatly benefit learning.

[3] In my personal teaching experience, I could see a potential for young learners to envision a different and better society from promoting the idea "without mercy, justice is impossible" ("fairness is impossible without kindness") in a classroom setting. Gradually, students began to see that the point of getting team points from doing teamwork was not to win, but to have more "fun" in the process of learning. When there was a big gap between different teams, they even volunteered to donate some of their team points to the team that was falling behind so that team members would not be too discouraged or give up altogether.

Writing Better Introductions and Conclusions for English Argumentation Essays

Carol Rinnert

Hiroshima City University, Hiroshima, Japan

Learning to write effective introductions and conclusions is an important goal for developing writers. This paper investigates how writers develop the ability to construct effective introductions and conclusions for their English argumentation essays. For more than a decade, we have analyzed English and Japanese writing by diverse groups, ranging from novice to advanced Japanese EFL writers and North American L1 writers (Kobayashi & Rinnert, 2012, 2013; Rinnert & Kobayashi, 2009; Rinnert et al., 2015). In this paper, we identify characteristics of stronger and weaker introductions and conclusions in English argumentation essays by these writers. Based on the analysis, we formulate some practical ways to help our students (and ourselves) improve the way to write these essential parts of every text.

Keywords: English argumentation essays, coherence, engaging readers, introduction, conclusion, components

INTRODUCTION: AIMS OF THE STUDY

The goal of this paper is to examine how writers can construct introductions and conclusions that engage their readers and contribute to higher quality texts. Introductions and conclusions are important parts of any text. They can be challenging for experienced writers in their first language (L1), and they often prove to be even more daunting for writers in a foreign language (FL). Therefore, it seems worthwhile to try to gain a better understanding of how writers develop the ability to construct effective introductions and conclusions.

Introductions and conclusions differ according to the genre and discourse type of the writing (e.g., research papers, reports, expository

essays). For this study, we looked at argumentation essays, one of the most common types of writing taught in formal classes. By argumentation essay, we mean that the writer takes a position for or against a given topic and supports the position with evidence that could persuade the reader that the position is reasonable.

METHOD

For the analysis, we examined a total of 53 English argumentation essays written by six groups of writers. Five of the groups were Japanese:

Novice 1 (Nov1): First-year university students who had received intensive preparatory training in both Japanese and English writing before taking university entrance essay exams;

Returnee (Ret): First- (or second-) year university students who had come back to Japan after spending 2.5 to 3 years at overseas high schools;

Experienced Group 1 (Exp1): Third-year university students who had never studied overseas;

Experienced Group 2 (Exp2): Fourth-year university students who had spent one year studying at universities overseas; and

Experienced Group 3 (Exp3): Graduate students and teachers who had spent at least 3 years studying and working overseas after their undergraduate degrees.

The last group were native English speaking writers:

North American (NA): Third- and fourth-year university students in the U.S. or Canada who were studying Japanese as a foreign language.

The English essays we analyzed were written under the same basic conditions: no time limit and use of dictionaries allowed. Each writer wrote on one of two topics:

For the Novice and Returnee writers:
1. University students living alone vs. living with family
2. University students traveling alone vs. in a group

For the Experienced and NA writers:
1. For or against elementary school students learning a foreign language
2. For or against elderly people living with family

The same writers also wrote a Japanese essay on the other topic. After they had completed both essays, we asked the writers to reflect on their English and Japanese writing, including their perceptions of similarities and differences between them. (For more information about the series of studies the essays were taken from, see Kobayashi & Rinnert, 2012, 2013; Rinnert & Kobayashi, 2009; Rinnert, et al., 2015.) To supplement our text analysis in this paper, we occasionally draw on some of those reflections, as well as our analysis of the Japanese essays on the same topics. We also refer to the results of an evaluation study of some of the essays by two highly experienced native English-speaking writing teachers in Japan.

In this paper, we first look at introductions; next, at conclusions; and then, at some of the ways introductions and conclusions work together to contribute to the quality of the whole essay. We conclude with a list of pedagogical implications drawn from the study.

ARGUMENTATION ESSAY INTRODUCTIONS

A text's introduction could be its most important part, at least in terms of attracting the reader. This is because an interesting introduction will encourage the reader to continue, whereas an uninteresting one will likely lead the reader to stop reading or to continue only reluctantly, which could result in a negative evaluation of the text. Because introductions are important, much effort has been devoted to identifying the features of effective introductions. The best known of these is probably Swales's (1990) characterization of research article introductions, the Create a Research Space (CARS) model. This model specifies the moves a research article writer can make to contextualize

and establish a niche for the research, as well as to identify a gap in the current literature that the article aims to fill.

Writing teachers are interested in how to help their students learn to construct introductions in a variety of genres, and many writing textbooks include instruction on how to create effective introductions. For example, one introductory English academic writing textbook (Davis & Liss, 2006, p. 8) identities three components of introductions to five-paragraph essays: (a) a "hook" (an opening sentence or two that grabs the reader's interest); (b) background information about the topic; and (c) a thesis statement (containing the specific topic and "controlling idea" for the essay). Another (Oshima & Hogue, 2006, p. 59) characterizes an introductory paragraph as having two parts: (a) general statements that introduce the topic and interest the reader and (b) a thesis statement that gives the specific topic and may also include a listing of sub-topics or the organization of the essay.

In order to explore how the writers in our study constructed the introductions in their argumentation essays, we identified the components of the introductions they wrote. We found five common components, which are listed and defined in Table 1.

TABLE 1. Salient Components of Argumentation Introductions

Component (abbreviation)	Definition/Explanation
Announcement (A)	Stating overall goal, procedure and/or structure for the essay, or raising the topic of the essay (can include rhetorical questions)
Position (P)	Stating or implying the writer's opinion (claim) for or against the issue addressed in the essay
Context (C)	Background: Presenting general, personal, and/or specific information about the topic Raising issue: Referring to contrastive sides of an argument, alternative view, or controversy on the topic
Focus (F)	Defining terms, establishing perspective for the argument, narrowing/clarifying the topic or issue, and/or setting conditions for the argument
Preview (Prev)	Introducing what is to be discussed: General content of the argument or specific content of one or more of the main points/reasons

To see an example of each of these components, let us look at the following introduction by a North American writer (NA-13), which was

the only one that contained all five components. (Throughout this paper, excerpts from the participants' writing is presented as it appeared in the original, with only spelling corrected if necessary.)

> **Sample Introduction 1.** NA-13, Announcement + Context + <u>Position</u> + Preview + Focus
>
> ---
> $^{(A)}$Should elderly people live or not live with family? $^{(C)}$This issue has many different viewpoints. $^{(P)}$<u>My opinion is that elderly parents should not live with family</u> $^{(Prev)}$for the following reasons: hardship on adult children, lack of freedom for both parties (adult children and elderly parents), and discouraging retirement planning. $^{(F)}$One caveat is that I will define family as adult children of an elderly person.
> ---

This introduction begins with an Announcement (A) of the topic in the form of a question; continues in the second sentence with a general Context (C) statement that characterizes the issue as having "different viewpoints"; states the writer's Position (P) in the following sentence (underlined), along with a specific Preview (Prev) of the reasons in the same sentence; and ends with a Focus (F) in the last sentence that defines the term "family" in this essay.

To compare the introductions across the groups, we computed their average length and number of different components. Table 2 shows these averages by group.

TABLE 2. Introduction Length and Number of Components: Averages by Group

	Nov1	Ret	Exp1	Exp2	Exp3	NA
English						
Words	20.0	56.4	62.4	53.9	77.3	92.4
Components	1.5	2.4	2.4	2.4	2.3	2.8

Note. Nov1 = Novice Group 1; Ret = Returnees; Exp1 = Experienced Group 1; Exp2 = Experienced Group 2; Exp3 = Experienced Group 3; NA = North Americans.

As we can see in the table, there appears to be a developmental trend toward longer and somewhat more complex introductions as writers gain more experience. The average number of words was less for Novice EFL, greater for Experienced Group 3, and even greater for North Americans. However, the number of components did not differ so much across the groups. Although the number of components was low for

Novices (1.5), it was very similar for Returnees and all the experienced EFL groups (2.3–2.4), and only a little higher for North American L1 writers (2.8).

We also looked at how often each component occurred to see which ones were the most popular across the groups. Table 3 shows an overview of how many writers across the groups used each component. As we can see, Position was used by a majority of writers in every group, and Context was almost as widely used. However, Announcements seemed to be less popular among more experienced writers, and there were big differences among groups in the use of Focus and Previews. Most notably, Focus was used often only by Experienced Group 3, but the use of Previews varied considerably across the groups. (We will look more closely at the use of Previews in the section below.)

TABLE 3. Frequency of Introduction Components by Group

Group	A	P	C	F	Prev
Nov1 (N = 8)	◎	◉	◎	•	–
Ret (N = 7)	◎	◉	◉	–	◎
Exp1 (N = 9)	•	◉	◉	–	•
Exp2 (N = 9)	◎	◉	◉	–	◎
Exp3 (N = 4)	–	◉	◉	◉	–
NAmer (N = 16)	•	◉	◉	•	◎

Note. A: Announcement; P: Position (explicit, implied or conditional); C: Context; F: Focus; Prev: Preview.
(N =): Number of argumentation essays for the group; ◉ : Feature of 50% or more of essays; ◎ : Feature of 29% to 49% of essays; • : Feature of at least one essay, but less than 29% of essays; – : Feature of no essays.

In addition to the above analyses, we also looked at how individual writers combined the components in their introductions. Looking at all the results, we found three main trends that may help us to understand how introductions to argumentation essays change with more writing experience. First, as pointed out above, introductions tend to become longer and more complex. Second, introductions become less formulaic and more individually crafted to fit particular audiences. Third, typical

placement of position statements inside EFL writers' introductions appears to change with more experience. Let us now look at each of these findings.

Length and Complexity

One obvious difference between inexperienced and experienced writers was whether or not there was a separate introduction paragraph. Half of the Novice EFL writers (4 out of 8) did not have separate introductory paragraphs in their English essays. In contrast, virtually all the other essays had distinct introductory paragraphs. It therefore seems that creating a separate introductory paragraph is a first developmental step toward writing an effective introduction.

Among the more experienced writers, Experienced Group 2 tended to have relatively short introductions. For example, one writer's (Exp2-6) English introduction, shown below, has three components, expressed in one long and one short sentence. It consists of a Position statement (underlined) and Preview in the same sentence, followed, in the second sentence, by an Announcement of the organization of the essay to come.

Sample Introduction 2. Exp2-6: <u>Position</u> + Preview + Announcement

> (P)<u>I think that old people should live with their family members</u> (Prev)in terms of four points, old people's loneliness, preventing senility, immediate help by family members and good effects for grandchildren's growth. (A)Now I will explain each of the points.

In this introduction, the writer makes her intent very clear and leads the reader smoothly into the body of the essay. However, if the writer added some Context, for example, some background information to situate the argument, the opening paragraph could draw in the reader more effectively.

In fact, providing extended Context may be the easiest way to lengthen an introduction and make it more interesting at the same time. For instance, let us examine the following introduction (see Sample Introduction 3) by an Experienced Group 1 writer (Exp1-1). This introduction begins with five and a half sentences of Context (C), consisting of a personal story related to the topic, followed by general background on the issue and mention of the opposing side at the beginning of the last sentence. The last sentence continues with a

Position (P) statement (underlined) and a Preview (Prev). In this introduction, the writer tried to attract the interest of the reader by pointing out both the relevance of the topic to our daily lives and the importance of the issue. At the same time, there is an acknowledgement of potential criticism of the writer's position (opposing viewpoints). Then, after a clear statement of the writer's position, there is an overview of the argument in the form of a Preview of the supporting reasons developed in the essay.

Sample Introduction 3. Exp1-1: Context + Position + Preview

> (C)Yesterday, I talked with one of my friends on campus. He is from Indonesia and he cannot speak Japanese well, so we always talk in English. These days English is coming more and more important. If you can use English, you can communicate with people from all over the world. However, it is said that Japanese people are not good at speaking English, so many professors and school teachers are trying to teach English with elementary school children. There are many criticism of this current of early English education, but (P)I strongly agree to this idea, (Prev)because it can help children to get not only skills of speaking English, but also skills of communicating with people from other countries of cultures.

Many of the Returnee and North American writers wrote introductions for their argumentation essays that were developed in similar ways to the one above by Exp1-1. Most commonly, besides a Position statement, these involved extended Context, often mentioning both sides of an issue; a Preview of the supporting reasons; or both Context and Preview. These components fit the textbook description of effective introductions, given earlier, and were reportedly learned by most of these writers in their secondary school English writing instruction outside Japan. In the case of the Experienced Group 1 writers, some of them were enrolled in an academic writing class taught by an American teacher at the time of the data collection. From the interviews, we learned that they tended to consciously apply what they had learned about introductions (and conclusions) in that class.

More Original Introductions by Advanced Writers

In comparison with the other experienced EFL and L1 English writers, the English introductions by the most experienced EFL group

(Exp3) were not only longer, but also more varied. In fact, no two introductions contained the same combination and ordering of components. This suggests that these writers were reshaping rather than simply reusing components. Unlike the introductions by the other groups, those by these writers were non-formulaic. They included no cases of Announcements or Previews, and they were characterized by frequent use of Context and Focus (limiting and clarifying the argument). In fact, these writers used Focus in 75% of their introductions. For example, the following introduction, by a member of Experienced Group 3 (Exp3-6), consists entirely of extended Focus and a Position (underlined). The essay begins with a question, which we considered as the title, and the introduction serves to answer the question.

Sample Introduction 4. Exp3-6: Focus + Position

Early foreign language education should start with elementary school children?

[F]I think it depends on how many hours are to be spent for foreign language classes. If it is longer than or the same as the time spent for national language classes, it is just too much. It also depends on whether foreign language education is compulsory or optional. If it is meant for all elementary school children, [P]I do not agree with it.

The writer carefully limits the topic according to the number of hours to be spent on foreign language teaching in the elementary school classroom (either more or less than for national language learning) and whether the instruction is required or optional, and then expresses her conditional Position against it. By limiting the topic and expressing the Position conditionally, the writer conveys a sense that she has thought carefully about the issue and is approaching it in both a cautious and realistic way, which could make her argument more credible to the reader.

Placement of Position Statements

One thing that has intrigued us for a long time is the location of Position statements in the introduction. Many of the introductions in our set of essays by EFL writers started out with a Position statement in the first sentence, as we saw in Sample 2 above. This is quite different from the English writing textbook advice referred to earlier and has been

noted in the literature. Basically, it has been observed that L1 English writers tend to place their main point (thesis) at or near the end of their introductory paragraphs, while many Japanese writers place their main point at the very beginning of the introduction (Hirose, 2001; Sato, 2014). For example, Hirose (2003) found that all but two of 15 writers (87%) in her study put their position statements in the first sentence of both their L1 Japanese and L2 English short argumentation essays.

We found a similar pattern in our study. A great many of the L1 English writers (75%) placed their position statement in the *last* sentence of their English introductions, whereas a large number of L1 Japanese writers (71%) placed their position statement in the *first* sentence of their Japanese introductions. In contrast, the experienced EFL writers fell in between these two L1 groups. In their English essays, equal numbers (42%) placed their positions in the first or the last sentence. That is, they did not follow the tendencies of either L1 group. These findings suggest that the Japanese EFL writers may have adopted features from both languages and incorporated them into an overlapping repertoire of writing knowledge that they could choose to apply in either language. (This suggestion fits the notion of a merged repertoire of knowledge, which we have discussed elsewhere, e.g., Kobayashi & Rinnert, 2012.) Nevertheless, it may be beneficial to tell L2 English writers that English audiences generally expect a position statement to come at or near the end of the introduction.

ARGUMENTATION ESSAY CONCLUSIONS

Like introductions, conclusions play an important role for the reader. It is a well-known psychological principle that endings − of words, sentences, and larger pieces of discourse − stand out and are more easily remembered than beginning or middle parts. This implies that strong endings can be memorable and ultimately make an argument more convincing for the reader. In fact, in our experimental study of English essay evaluation by different groups of readers in Japan (Rinnert & Kobayashi, 2001), we found that ratings of conclusions correlated highly with ratings of the overall quality of the whole essays. Moreover, a substantial number of the members of all groups of evaluators (28% to 38%) offered positive or, more often, negative comments on the

conclusions. Therefore, it seems clear that learning to write effective conclusions is an important goal for developing writers.

TABLE 4. Salient Components of Argumentation Conclusions

Component (abbreviation)	Definition/Explanation
Conclusion Marker (Mkr)	Explicitly signaling the beginning of the conclusion with a discourse marker (e.g., *In conclusion*, or *For the reasons above*)
Position (P)	Stating, restating, or implying the writer's opinion for or against the issue addressed in the essay; implied positions may be presented indirectly as part of another separate component; positions may also be conditional (dependent on specific conditions being met)
Concession (Cs)	Recognizing the other side of the issue and/or problems with the side being taken
Summary (Sum)	Condensing one or more of the points of the argument already made in the body (can be very general, e.g., *more merits than demerits*)
Extension Text-based (Ext-T)	Interpreting the content of the essay more deeply, offering a solution to a problem related to the argument, making a new proposal, or showing future concern
Writer-based (Ext-W)	Giving personal comments or an emotional appeal

To examine our writers' conclusions, we identified the six conclusion components listed above in Table 4. None of the conclusions contained all six components, and very few contained more than four components. Therefore, we chose the three typical sample conclusions below to illustrate the full range of components we found.

The first sample conclusion is by a member of Experienced Group 1 (Exp1-2).

Sample Conclusion 1. Exp1-2: Marker + Summary + <u>Position</u> + Text-based extension

> $^{(Mkr)}$In conclusion, $^{(P)}$<u>I strongly think that early foreign language education should start with elementary school children</u>. $^{(Sum)}$This is because younger children can learn easily and can have a lot of opportunities. $^{(Ext-T)}$Furthermore, I think which language required to take should discuss in any country.

This conclusion has four components. The Conclusion Marker (Mkr, "In conclusion") is followed by the Position statement (P, underlined). Then

there is a Summary (Sum) of the reasons given in the essay. Finally, there is a Text-based Extension (Ext-T) proposing that every country should discuss which foreign language to require, an idea that was not mentioned earlier in the essay.

The second sample is by a North American writer (NA-14).

Sample Conclusion 2. NA-14: Summary + <u>Position</u> (implied) + Writer-based extension

> $^{(Sum\ with\ P(imp))}$Due to both the greater number of strong language learning environments and the comparative strength of a youngster's mind when facing the task of language learning, <u>the answer to when the best time to begin foreign language education is clear.</u> $^{(Ext-W)}$I only wish that I had been given the opportunity to study more languages as a child.

This conclusion starts with a Summary of the writer's argument in favor of foreign language education in elementary school, which includes an implied Position statement (P(imp), underlined) that does not explicitly state his position. It ends with a Writer-based Extension (Ext-W) that expresses the writer's personal wish that he had been given the chance to learn a foreign language earlier.

The third sample is also by a North American writer (NA-2).

Sample Conclusion 3. NA-2: Marker + Concession + <u>Position</u> (conditional) + Summary

> $^{(Mkr)}$Ultimately, $^{(Cs)}$the decision to live with family or not is one that depends on the family and the elderly person in question. $^{(P(cond))}$<u>A reasonably healthy elderly person</u>, however, <u>should not have to live in-residence with family if</u> $^{(Sum)}$it compromises their sense of independence or causes them emotional distress.

This last conclusion starts with a Marker ("Ultimately"), continues with a Concession (Cs) to the idea that some families may have different needs, and ends with a conditional Position (P(cond), underlined) and Summary of the gist of the argument given in the essay against elderly living with family.

For the conclusions, like the introductions, we computed the average lengths and numbers of components across groups. We also counted the number of writers who used each of the components and examined how the different writers combined the components. The lengths of the

conclusions are reported in Table 5, and an overview of the distribution of the components across the groups in L1 and L2 is presented in Table 6.

TABLE 5. Conclusion Length and Number of Components: Averages by Group

	Nov1	Ret	Exp1	Exp2	Exp3	NA
Words	21.8	48.7	38.7	45.2	81.0	83.8
Components	1.8	2.1	2.9	2.7	3.0	2.8

As shown in Table 5, we found that the conclusions generally tended to get longer with more writing experience, even though the average number of components did not change much. We also found, as seen in Table 6, that Conclusion Markers, Position statements, and Summaries were very commonly used across the groups.

TABLE 6. Frequency of Conclusion Components Across Groups and Languages

Group	Mkr	P	Cs	Sum	ExT	ExW
Nov1 (N = 8)	◉	◉	–	◎	•	–
Ret (N = 7)	◉	◉	◎	◉	•	–
Exp1 (N = 9)	◉	◉	•	◉	•	•
Exp2 (N = 9)	◉	◉	–	◉	•	–
Exp3 (N = 4)	•	◉	◉	◉	◉	–
NAmer (N = 16)	•	◉	◎	◉	◎	•

Note. Mkr: Conclusion marker; P: Position (explicit, implied or conditional); Cs: Concession; Sum: Summary; ExT: Text-Based Extension; ExW: Writer-Based Extension. (N =): Number of argumentation essays for the group; ◉ : Feature of 50% or more of essays; ◎: Feature of 29% to 49% of essays; • : Feature of at least one essay, but less than 29% of essays; – : Feature of no essays.

In Table 6, we can see a possible tendency toward more use of Concessions with more writing experience. Like Counterarguments, Concessions bring in reference to the opposing side, so they can be a way of acknowledging possible objections from the reader.

Finally, by looking at the ways writers combined components, we found some interesting patterns in the use of Extensions across the groups. In the rest of this section, we would like to discuss these patterns and how they may relate to changes in writing instruction in recent years.

In our original analysis of the experienced EFL writers' conclusions (Kobayashi & Rinnert, 2007; summarized in Rinnert & Kobayashi, 2009), we reported that the less experienced undergraduates (Exp1) used many more extensions in Japanese (90%) than in English (20%), but the more experienced undergraduates (Exp2) did not use many extensions in either language. Interpreting our findings, we suggested that many of these EFL writers were applying learned L2 writing conventions and in some cases transferring them to their L1 writing. For example, those with overseas experience (Exp2) tended not to use extensions across languages, even though some of them perceived that Japanese conclusions should include some kind of extended or future perspective. We noted that the reported perceptions of many of these student writers appeared to match L2 English writing textbook advice about conclusions (e.g., Langan, 2000; Reid, 1988; Smalley & Hank, 1982), particularly the importance of concisely rephrasing the main points in different words and not adding any new ideas at the end.

When we expanded the analysis of conclusions to cover argumentation essays by the North American and L1 Japanese writers, we found that the L1 Japanese conclusions tended to include more extended ideas, as opposed to greater use of specific summaries in the L1 English conclusions. Thus, we observed that the L1 Japanese writers often ended their essays with a deeper interpretation of the content, while the L1 English writers frequently presented a final synopsis of the gist of their argument to appeal to readers (Kobayashi & Rinnert, 2012). We also noted that the most experienced EFL writers (Exp3) used more extensions in their English than in their Japanese essays, which we could not explain, other than to suggest that it may have been related to their L2 disciplinary training.

However, our current reanalysis of the conclusions reveals a more complex picture of the use of extensions in English conclusions. First, for all groups, text-based extensions were much more frequent than writer-based extensions; in fact, writer-based extensions were not used at all by Experienced Group 2 or 3 (see Table 6). Second, contrary to our earlier findings, when we added together the text-based and writer-based

extensions, we found that many L1 English writers (63%) used extensions. These findings, together with the fact that half of the most experienced writers (Exp3) also used extensions in their English conclusions, suggest that going beyond a simple restatement of the position and summary of the argument, especially by using a text-based extension, may be a productive way to conclude English argumentation essays.

This suggestion may reflect something of a shift from the earlier emphasis in writing instruction (in textbooks written in the 1980s and as late as 2000, referred to above) on not including any new ideas in English conclusions. In fact, the two more recent writing textbooks referred to earlier (in the section on introductions) give similar recommendations to include extended ideas in academic English essay conclusions. According to Oshima and Hogue (2006, p. 72), the conclusion has three purposes: (a) to signal the end of the essay by using a transition signal (such as, "In conclusion"); (b) to remind the reader of the main points of the essay; and (c) to leave the reader with a strong final message, for example, by making a prediction or recommendation, or suggesting results or consequences. Likewise, Davis and Liss (2006) say that the conclusion brings the essay to a close by restating "the thesis of the introduction in different words" to connect the conclusion to the introduction, and that it may also (a) give advice; (b) "make a prediction or ask as question"; or (c) "provide new insights or discoveries that the writer has gained through writing the essay" (p. 11). Thus, there is no mention in either textbook of avoiding any new ideas in the conclusion; instead, both suggest ending with an appeal to the reader in the form of a text-based extension (e.g., a prediction or a suggestion), or else, according to the second book, a writer-based one (i.e., what the writer has personally learned in the process of writing).

Further support for including extensions is seen in reflections the North American writers made on their L1 English conclusions. When asked about what they paid attention to when writing conclusions in their English essays, almost half (47%) referred only to restating their main point, e.g., "Summarizing the most important points succinctly and memorably" (NA-15). But a substantial number (37%) also mentioned going beyond what was already said to include something new. For example, one writer said that in both his English and Japanese conclusions, he was concerned with:

Restatement of my thesis/argument and possible applications/ prospects concerning the future. Also suggestions for further improvement or ideas for development of the subject of my argument. (Interview with NA-8)

In sum, then, we can advise that developing writers could be encouraged to write longer conclusions that include text-based extensions. For example, one of the more highly rated essays by an Experienced Group 2 student (Exp2-1) contains a conclusion, shown below, that is twice as long (96 words) as the average for that group (45.2 words).

Sample Conclusion 4. Exp2-1: P + Sum + Ext-T

> (P)Having foreign language classes with elementary school children will be a great opportunity for the children. (Sum)They can easily learn new languages and have chances to think global. The lack of other subjects class hours and the importance of the mother tongue are the things people who disagree worry about, but we cannot say that the education levels of other subjects will get down. This might be the charge to look over the curriculum of the subject. (Ext-T)Foreign language classes will be a good material for children to make progress for their knowledge and their heart.

This conclusion starts with a restatement of the Position (underlined). It next continues with a detailed Summary that includes the counterargument and refutations developed in the body of the essay. It then ends with a text-based extension that goes a bit beyond the argument in the body to encompass benefits not just for the children's knowledge but also for their hearts. Although the conclusion alone cannot account for the relatively high overall quality score given to this essay (5.25 out of 7), it can be said to leave the reader with a sense that the writer has presented a comprehensive and thoughtful argument.

So how do introductions and conclusions together relate to the overall evaluation of essays? In the next section, we would like to consider how these two parts, in combination with each other and the rest of the text, can contribute to, or detract from, the quality of an essay.

INTRODUCTIONS AND CONCLUSIONS WORKING TOGETHER

We would like to begin by focusing on Previews in introductions and Summaries in conclusions of essays. First, let us look at the similar roles they serve and why they might pose something of a problem for our EFL writers. Then, we can go a step further to reconsider some of our earlier interpretations of the use of these components in light of their functions and how they, together with other components, may relate to essay quality. To do this, we will look at two individual essays. In the process, we hope to draw some implications for ways to empower writers to construct more effective introductions and conclusions for their argumentation essays.

Looking at the definitions of Preview and Summary in Tables 1 and 4, we can see that they basically serve the same function of condensing the content of the argument. The only difference is that the Preview appears near the beginning of the essay, and the Summary, at the end. In essence, then, if a writer includes both a Preview and a Summary in an argumentation essay, the same argument is being repeated in three different places: the introduction, the body, and the conclusion. As pointed out many years ago by John Hinds (1987, p. 144), this format follows the traditional advice for making an American English speech: "Tell 'em what you're going to tell 'em, tell 'em, then tell 'em what you told 'em." This formula has definite advantages, especially for oral presentations, because it makes the main points clear and memorable for the audience. However, it can also run the risk of conveying a feeling of unnecessary repetition, especially in a short written text. In fact, in our evaluation study (2001), mentioned above, we found that this kind of redundancy in some of the essays was criticized by members of all three groups of Japanese raters (inexperienced and experienced student writers and teachers), though not by the L1 English speakers.

The above considerations suggest there may be some resistance to the use of both a Preview and a Summary in these essays, especially by Japanese writers. So, we decided to count how many writers used both components in the same essay, as compared to those who used just one or the other, or neither. Table 7 summarizes what we found. As we can see in the table, including both components in the same essay is not the most popular option for any of the groups, except the North Americans

(although at 38%, it still does not constitute a majority of writers in the group). Most strikingly, none of the Novice or Experienced Group 3 writers included both a Preview and a Summary in any of their essays. Moreover, for all the groups except the North American writers in English, the choice of Summary alone was more frequent than – or in a few cases, the same frequency as – the choice of both Preview and Summary together. Finally, we can see that in most cases, the use of a Preview, a Summary, or both, is more common than leaving the two components out of the essay. Overall, these findings suggest that it may not be necessary to include both a Preview and a Summary in the same essay, and that Summaries may be more highly valued than Previews when only one is included.

TABLE 7. **Use of Preview and Summary in English Argumentation Essays**

Group	Both (+Prev +Sum)	Preview Only (+Prev –Sum)	Summary Only (–Prev +Sum)	Neither (–Prev –Sum)
Nov1	0%	0%	**38%**	**63%**
Returnee	**43%**	0%	**43%**	14%
Exp1	**33%**	11%	**44%**	11%
Exp2	**33%**	0%	**44%**	22%
Exp3	0%	0%	**50%**	**50%**
NA	**38%**	25%	**31%**	6%

Note. +: present; –: not present; Prev: Preview; Sum: Summary; bolded entry: over 30%.

The concerns about redundancy discussed above, along with the findings reported in Table 7, may also help us to qualify some of our earlier findings regarding general versus specific Previews and Summaries. As we mentioned above, both the Preview and Summary can range from a general overview of the argument to a detailed synopsis of specific points discussed in the essay. In earlier publications (Kobayashi & Rinnert, 2007; Rinnert & Kobayashi, 2009; Kobayashi & Rinnert, 2012), we reported some differences across languages in the use of general versus specific Previews and Summaries. For example, we found proportionally more use of general than specific Previews in Japanese by our undergraduate EFL writers (Exp1 and Exp2), and we also found that the Japanese writers' L1 conclusions tended to include more general Summaries, as opposed to more use of specific Summaries in the North American writers' L1 English conclusions. However, it now seems clear that it is not enough to look at introduction and conclusion components

in isolation. In fact, it is highly likely that they are interrelated, and they may often be tightly connected. For example, in the present analysis we found that many of the more experienced writers who included an explicit Position statement in the introduction chose to use an implied Position in the conclusion. Similarly, some writers who presented a specific Preview in the introduction made their Summary more general in the conclusion, and vice versa. In sum, then, it appears necessary to look at both the introduction and the conclusion in the context of a whole essay to clarify their roles more precisely.

So how can introductions and conclusions work together to weaken or strengthen an essay? Let us now look at two essays that illustrate each of these cases (weakening and strengthening).

In the first essay, by an Experienced Group 1 writer (Exp1-5), the introduction and conclusion, along with poor language use (average 3.5 out of 7), detract from an otherwise relatively highly evaluated essay (with content scores averaging 5.2, and structure scores averaging 5.75 out of 7), yielding an overall quality score of five out of seven. In fact, one of the evaluators specifically pointed out the weaknesses of this essay as "background missing, intro and conclusion weak" (evaluator 2).

So why were the introduction and conclusion of this essay considered so weak? One obvious possibility is the length. Both are relatively much shorter than the averages for this group (introduction: 29 words, compared to 62.4 average for Exp1; conclusion: 27 words, compared to 38.7 average). However, in terms of the number of components, they are actually higher than the averages for the group (introduction: three components, compared to 2.4 average; conclusion: three components, compared to 2.9 average). The introduction starts with an explicit Position statement in favor of elderly living with family, continues with one sentence of Context that mentions the writer's personal background living with her grandmother, and ends with an Announcement of the writer's intention of using that background as a basis for the essay. The conclusion starts with a one-sentence general Summary of the writer's argument in favor of elderly living with family, which includes an implied Position (P(imp), underlined), and ends with a Writer-based Extension (Ext-W) referring to the writer's personal desire for a positive future outcome for many families.

These same components were used more successfully by other writers in their introductions and conclusions. For example, one North American writer effectively introduced personal experience in the

introduction as the basis to support the same argument in favor of elderly living with family. Therefore, it is safe to say that it was not the choice of the particular combinations of components that led to a low evaluation but rather the lack of development of these components (one short sentence for each). In the words of the same evaluator quoted above: "The longer paragraphs, with more than one or two predictable sentences, generally held my interest better than the formulaic short paragraphs" (evaluator 2).

Sample Essay 1. Exp1-5; Introduction: Position + Context + Announcement; Conclusion: Summary + Position (implied) + Writer-based extension

---

(P)I think that old people should live with their family members. (C)Actually, I have experienced life with my grandmother. (A)Based on this fact, I want to write this thesis.

Living with old people has two important advantages. First, when old people get serious sick, their family members can support and help them. If old people are alone, they will not be able to have a care when they get serious sick. This is very important fact, and this is a matter of life and death. For example, my family live with my grandmother now, and the other day my grandmother fell down suddenly. If she was alone, she may lost her life. However, my father noticed this accident immediately, he helped her. Thus, my grandmother narrowly escape death. Moreover, when old people get a sick, their family member can support them mentally. Mental support is rather important, and if old people are alone, they can't feel their relief. Now, my grandmother is bed ridden, but our conversation seems to play a important role in her health.

Secondly, if the family has a child, the child can have various experiences through living with old people who are different generation people. Old people have experienced various accidents and events, they can tell their grandchild it. The child can learn from grandparents' story. For example, I have heard various interesting stories from my grandmother. I can imagine the old days from it. When I was a child, I looked forward to hearing these stories. Furthermore, old people have different sense of values from the grandchild. Living with people who have different sense of values is valuable experience for each, that is, grandchild and grandparents. Through living with my grandmother, I can learn kindness for different generation.

It is true that living with old people have some difficulties. For example, my acquaintance often complain about living with his grandparent. He says that they young generation people can't understand old people any more. However, I think that they can understand each other someday. Living with old people is a big chance that we can understand people who have different sense of values.

(Sum with P(imp))Through living old people, their family members can grow up in many aspect. (Ext-W)I hope that many families accept their grandparents, and they lead a full life.

In contrast to the above essay, the following essay by an Experienced Group 2 writer (Exp2-7) appears to have benefited from a relatively stronger introduction and conclusion. The content scores for this essay were rather low (ranging from 4.25 to 4.5 out of 7), though the structure and language use scores were higher (5 to 5.5), and the overall score was 5.25, which was the second highest score for the group. One of the two reviewers was much more critical of this essay than the other one, but nevertheless described a strength of the essay as follows: "The introduction and conclusion do help the reader see the essay as a unified whole" (evaluator 1).

Sample Essay 2. Exp2-7; Introduction: Position + Preview; Conclusion: Marker + Position + Summary

---

(P)Early foreign language education for elementary school children has a lot of advantages especially for their listening and speaking skills. (Prev)Elementary school children, who are actively developing abilities to listen and imitate sounds in their developing process can be trained their listening and speaking skills of foreign language effectively by native speakers.

In elementary school class, students don't need to learn grammar or writing sentences, but should learn and use greetings or some simple phrases with classmates or teachers in order to be familiar with the sound of the foreign language. Elementary school children are pleased to imitate sounds more than junior high school students who are embarrassed to imitate the sound of foreign language and have accent of their own mother tongue. The younger they are, the more they like to imitate sounds. They tend to try to speak foreign language as if they are native speakers. The nature of children causes them to listen the foreign language with concentration and also enables them to develop listening skills. Therefore, such learning realizes development of tongue and ears of elementary school children for the foreign language.

Moreover, such listening and speaking trainings lead [to] children's interest in the language, because this training is not so much the kind of knowledge learnings with text book and exercise drill book as enjoyable play with friends and teachers. Many children would be interested in study of the language and even culture of the people who speak the language. It makes easier to study the language in following study stage, such as grammar, reading or writing.

(Mkr)Therefore, (P)people who arrange education programs should realize the great advantages to start early foreign language education for elementary school children. (Sum)It is important to learn and familiar with the sound of the foreign language while their abilities to listen and imitate are actively developed, besides enjoyable speaking of the language and broadened interest in the language study is helpful to study the language in following study in junior high school.

How do the introduction and conclusion of the essay specifically contribute to its quality? First, they are longer than those in the preceding essay. The introduction is close to the average number of words and components for the group (52 words vs. 53.9 average for Exp2; 2 components vs. 2.4 average), and the conclusion is much longer than the average (71 words vs. 45.2 average; three components vs. 2.7 average). Second, relatively long, complex sentences express these components in both parts. In the introduction, the Position in favor of foreign language education is elaborated by specifying "listening and speaking skills," and the Preview refers to specific supporting reasons and a suggestion that native speakers could be effective in teaching foreign language to children. Similarly, in the conclusion, the Position and Summary include specific contents of the gist of the argument. Third, even though there is both a Preview and a Summary, they do not create a feeling of redundancy. Although one key point was mentioned in both the Preview and Summary ("abilities to listen and imitate" being "actively developed" by young learners), the emphasis otherwise differs. In the introduction Preview, the emphasis is on teaching in effective ways, and in the conclusion Summary, it shifts to stressing benefits for the learner (i.e., enjoyment, interest, and productive future study).

Even though the introduction and conclusion of the essay appear to work together well to frame the essay, as pointed out by the reviewer above, one weakness should be noted. The introduction paragraph ends with a reference to children being "trained...effectively by native speakers." This could mislead the reader into thinking that the role of native-speaking teachers will be a main focus in the essay, which it is not. In fact, there is no direct reference to native-speaking teachers in the rest of the essay, although there is mention of children enjoying imitation of sounds and trying to speak "as if they are native speakers," which could imply a very indirect reference to native-speaking teachers. This case illustrates the problem of a potential gap between the introduction and the rest of the essay. In our study, this kind of gap often appeared when writers provided only a partial Preview of one or two points in the introduction, whereas other key points developed in the essay were not mentioned in the Preview.

Based on our findings in this and the preceding sections, it is clear that learning to write effective introductions and conclusions is a complex process. In the final section, we would like to suggest some ways to apply these findings.

PEDAGOGICAL IMPLICATIONS

In this last section, we present seven practical implications for the classroom based on the findings of this study. Let us look at each of these interrelated suggestions in turn.

1. *Make separate paragraphs for introduction and conclusion*: Writers can distinguish clearly between the opening, body, and concluding parts of an essay by learning to create separate introduction and conclusion paragraphs. This is an important step for novice English writers to acquire a basic schema for essay organization.

2. *Learn typical patterns*: Writers can be introduced to typical patterns of components for introductions and conclusions. For introductions, these would include Context + Position + Preview. For conclusions, they involve Conclusion Marker + Position + Summary. These standard components can prove useful for developing writers to orient their readers and make it easy to follow the argument through to the end of the essay. At the same time, we would like to point out that the use of such formulaic introductions and conclusions should not be considered an end point, but rather a step on the way to more effective writing that is individually designed to reach particular audiences in local contexts (see Suggestion 7 below).

3. *Write longer introductions and conclusions*: Writers should be encouraged to expand their introductions and conclusions. As we have seen, the best way to do this does not seem to be to add more components. Rather than employing four or five short, simple components, it appears that extending and elaborating two or three longer components is more effective.

4. *Give Context at beginning of introduction*: Writers can connect with their readers by starting with Context the readers can relate to. Drawing on various theories of argumentation, Sato (2014) convincingly showed that starting an argumentation essay with "a *shared context*" or "*common starting points*" between reader and writer (p. 9, italics in original) is usually the most persuasive way to lead the reader to accept the writer's viewpoint in both English and Japanese. Thus, it would seem

that starting an introduction with Context to establish shared understanding may be the most effective strategy for beginning an argumentation essay in any language. (A possible exception might be an argumentation essay test format that requires the writer to take a position for or against and support it convincingly within a short time period.)

5. *Maintain coherence between introduction, body, and conclusion*: In order to develop their control (writer agency) over the quality of the essay, writers need to learn to connect the parts of their essays to each other smoothly. To do this, there should not be any gaps in meaning between the introduction and body, or between the body and conclusion. In addition, it would be ideal to establish some relation between the introduction and conclusion.

6. *Avoid too much redundancy*: Writers should take care to eliminate too much repetition between the introduction and conclusion. In particular, if there is a Preview in the introduction and a Summary in the conclusion, they should contain different wording. It might also be beneficial to make them different in terms of their level of specificity. For example, a general Preview of the gist of the position could go well with a specific Summary of all the main points that support it.

7. *Move beyond formulaic patterns*: As they become more advanced, writers could be encouraged to experiment with unpredictable, non-formulaic patterns to enhance the quality of their writing. For introductions, this could involve adding a Focus component to clarify the terms and limits of the argument, as necessary, depending on what the writer assumes the imagined readers think about the topic. Similarly, in conclusions, it would include adding a Concession or an Extension, either text-based or writer-based. A Text-based Extension could be a future implication, broader perspective (e.g., prediction or recommendation), or suggested results or consequences (Oshima & Hogue, 2006). A Writer-based Extension could include what the writer has personally learned in the process of writing (Davis & Liss, 2006).

As we have seen in this paper, the relations between introductions, conclusions, and the essays they begin and end are complex. Therefore, it is not surprising that the ability to write successful introductions and conclusions takes a long time to develop. In fact, as pointed out by Kellogg (2011) and others, professional writing expertise takes at least 20 years to achieve and never stops growing as long as we keep writing. As writing teachers and writers ourselves, we can keep trying to find better ways for writers to connect to their audience effectively, especially in introductions and conclusions, which stand out most memorably for our readers and are probably the most essential parts of every written text.

ACKNOWLEDGEMENTS

THE AUTHOR

Carol Rinnert is Professor Emeritus of International Studies at Hiroshima City University in Hiroshima, Japan. She received her PhD from SUNY/Buffalo and has taught linguistics and English in the U.S., Yemen, and Japan. She has been researching multilingual writing and intercultural pragmatics with Hiroe Kobayashi for over 20 years. Email: rinnert@hiroshima-cu.ac.jp, rinny132001@yahoo.com

REFERENCES

Davis, J., & Liss, R. (2006). *Effective academic writing 3: The essay*. Oxford University Press.

Hinds, J. (1987). Reader versus writer responsibility: A new typology. In U.

Connor & R. Kaplan (Eds.), *Writing across languages: Analysis of L2 texts* (pp. 141−152). Addison-Wesley.

Hirose, K. (2001). Persuasive writing in L1 and L2: A look at Japanese EFL students' rhetorical organization strategies. *JACET Bulletin, 33*, 43−56.

Hirose, K. (2003). Comparing L1 and L2 organizational patterns in the argumentative writing of Japanese EFL students. *Journal of Second Language Writing, 12*, 181−209.

Kellogg, R. T. (2008). Training writing skills: A cognitive development perspective. *Journal of Writing Research, 1*(1), 1−26.

Kobayashi, H., & Rinnert, C. (2007) Transferability of argumentative writing competence from L2 to L1: Effects of overseas experience. In M. Conrick & M. Howard (Eds.), *From applied linguistics to linguistics applied: Issues, practices, trends* (pp. 91−110). British Association for Applied Linguistics.

Kobayashi, H., & Rinnert, C. (2013). L1/L2/L3 writing development: Longitudinal case study of a Japanese multicompetent writer. *Journal of Second Language Writing, 22*(1), 4−33.

Kobayashi, H., & Rinnert, C. (2012). Understanding L2 writing development from a multicompetence perspective: Dynamic repertoires of knowledge and text construction. In R. M. Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 101−134). De Gruyter Mouton.

Langan, J. (2000). *College writing skills* (5th ed.). McGraw-Hill.

Oshima, A., & Hogue, A. (2006). *Writing academic English* (4th ed.). Pearson Longman.

Reid, J. (1988). *The process of composition*. Prentice Hall Regents.

Rinnert, C., & Kobayashi, H. (2001). Differing perceptions of EFL writing among readers in Japan (with H. Kobayashi). *The Modern Language Journal, 85*, 189−209.

Rinnert, C., & Kobayashi, H. (2009). Situated writing practices in foreign language settings: The role of previous experience and instruction. In R. M. Manchón (Ed.), *Learning, teaching, and researching writing in foreign language contexts* (pp. 23−48). Multilingual Matters.

Rinnert, C., Kobayashi, H., & Katayama, A. (2015). Argumentation text construction by Japanese as a foreign language writers: A dynamic view of transfer. *The Modern Language Journal, 99*(2), 213−245.

Sato, R. (2014). *The first sentences and persuasion strategies in argumentative essays written by Japanese EFL students* [Unpublished master's thesis]. Hiroshima University.

Smalley, R. L & Hank, M. R. (1982). *Refining composition skills: Rhetoric and grammar for ESL students*. Macmillan.

Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge University Press.

Implications of Foreign Language Classroom Anxiety to Macau EFL Students

Ut Meng Lei
Beijing Foreign Studies University, Beijing, China
Ka Lon Chan
University of Macau, Macau, China

This study focuses on the influence of foreign language classroom anxiety (FLCA) on EFL learners' performance based on the context of Macau. The study was administered to English learners from three universities and four secondary schools in Macau. An adapted questionnaire from Horwitz et al.'s (1986) Foreign Language Classroom Anxiety Scale (FLCAS) was utilized, with the latent constructs about Communication Apprehension, Fear of Negative Evaluation, Test Anxiety, Language Classroom Anxiety, and Peer Influence. It focuses on both theoretical and practical aspects of the influence that FLCA has on EFL students' academic performance and teacher's pedagogy. The results indicate that learners exhibit a higher-than-average anxiety level that negatively correlates with their assessment. Peers were considered by the participants to be a factor as influential as assessments. Language instructors should hence reconsider how non-test-related factors like learning environment and peers can be better designed and utilized for refining effects in EFL learning.

Keywords: foreign language classroom anxiety, latent constructs, peer influence, five-construct model

INTRODUCTION

It is well known that foreign language classroom anxiety (FLCA) has been found to correlate with language learners' self-perceptions, beliefs, feelings, and behaviors related to the language learning process (Horwitz et al., 1986). Despite the enormous amount of research

conducted on the issue, no studies have been done in Macau, and neither are there studies that investigate the effects of FLCA for both final-year secondary school students and first-year university students. The current study, with respondents coming from the aforementioned two levels of students, aims at investigating how FLCA affects Macau's EFL (English as a foreign language) learning outcomes and performance as well as studying the correlative influence that different constructs of Horwitz et al.'s (1986) Foreign Language Classroom Anxiety Scale (FLCAS) have on each group of respondents.

Horwitz et al. (1986) were first to come up with the concept of foreign language classroom anxiety, which refers to the nervous, worried, and uneasy feelings that make language learners particularly pressured, and hinder students from performing smoothly, specifically in the classroom. They developed the FLCAS, which has since become popular among scholars in various constructs and contexts. The scale includes 33 items associated with scenarios of foreign language learning that fall into three categories: communication apprehension, test anxiety, and fear of negative evaluation. The scale has been modified by a number of scholars in different cultural contexts and to specific language aspects: listening, speaking, reading and writing (Aida, 1994; Cheng et al., 1999; Liu & Jackson, 2008; Mak, 2011; Matsuda & Gobel, 2004; Park, 2012; Young, 1994; Zhao, 2007). Concisely speaking, Aida (1994) adapted the FLCAS to examine the reliability of the constructions of FLCAS so as to investigate anxiety experienced by learners of non-Western languages, and she shed light on Japanese learning in her research. In her study, four constructs are categorized: speech anxiety and fear of negative evaluation, fear of failing the class, and comfortableness in speaking with native Japanese, as well as negative attitudes toward the Japanese class. Moreover, six items (Items 2, 6, 15, 19, 28, and 30) originally in the FLCAS were eliminated from Aida's research design, three of which belonged to the construct of Test Anxiety. She agreed with MacIntyre and Gardner's (1989) explanation that test anxiety could be a common psychological problem of learning and doesn't particularly influence Foreign Language Anxiety. Afterwards, Park (2012) reviewed Aida's study and doubted that the small number of respondents (98 Korean students) could produce fully reliable and a comprehensive interpretation of the data. He compared five different models with one to four constructs (Aida, 2004; Cheng et al., 1999; Liu & Jackson, 2008; Matsuda & Gobel, 2004; Toth, 2008) in his study to test their validity

and identify the most suitable model of FLCAS. In his study, 918 university students of an English conversation course participated. The result showed that a four-construct model gives better-suited indices than models with one to three constructs. The above studies together have given inspiration to the authors of this paper to investigate whether the affective filter could be different in the Macau context and which specific model is to be considered as most appropriate. In short, the authors have come up with different latent construct structures based on the FLCAS of Horwitz et al. (1986) for producing the best analyses in specific contexts.

Regarding various classifications of factors on this widely discussed topic, almost none of the studies specifically point out the effects contributed by peers of language learners. The authors of this study aim to reclassify performance anxiety and give deeper insight into the language anxiety derived from peer influence. In a previous study, Horwitz et al. (1986) noted in their findings that anxious students are worried about being left behind by other students and having their peers comment negatively about them. Therefore, anxious students are prone to being truant to avoid being called on in class to avoid embarrassment. They also stress over learning foreign languages. Shao (2014) stated that cooperative learning can help Chinese students overcome language anxiety by exchanging experiences and understanding learning materials. Based on the above, the authors of this study believe that peers bring about certain effects on foreign language learners. In that study, all the respondents completed an adapted questionnaire from Horwitz et al.'s (1986) FLCAS. Five constructs were investigated: Communication Apprehension, Fear of Negative Evaluation, Test Anxiety, Anxiety of English Learning, and Peer Influence. Among them, Communication Apprehension and Peer Influence were found to play a more vital role in affecting student performance at both secondary and tertiary levels. Additionally, this study reflects on what teachers can do to alleviate the anxiety of students in relation to the Macau context.

METHOD

Participants

There are two main groups of participants in this study: Macau EFL students from secondary schools and from tertiary institutions. There were 531 participants of F.6 (Grade 12 students) from four local EFL schools. They signed a consent form and completed the study's questionnaire survey. The participants consisted of 270 males (50.8%) and 261 females (49.2%). The male–female proportion is roughly 1:1. Their ages ranged from 15 to 20 years old. The four EFL schools are all traditional Chinese-medium schools in which Chinese (Cantonese) is the main lingua franca and medium of instruction. Like most other EFL schools in Macau, students in the four schools speak Chinese, Mandarin and some other dialects of Chinese, in their daily life. Their English competence ranged from A2 level (pre-intermediate) to B1 level (intermediate), with a small number of them reaching the B2 level (upper-intermediate) according to the Common European Framework of Reference for Languages (CEFRL). In their English classes, they study the four necessary skills of English, namely reading, listening, writing, and speaking; and other language features, predominantly grammar or English usage. As F.6 is the exit level of secondary school in Macau, most schools prepare their students for university entrance exams, and so, reading and writing are the two skills that gain the most attention. Besides usual teacher-fronted classes, English classes are sometimes conducted in groups, and students may be asked to give presentations based on a particular topic. Evaluation is mostly done with assignments and summative assessments like quizzes, tests, and end-of-term exams.

For the college participants, there were 74 first-year students from three tertiary institutions of Macau. Among them, 21 were males (28.4%) and 53 were females (71.6%). The large gender gap in the subjects reflects a similar phenomenon in tertiary institutions in Macau. According to the 2018 data (Higher Education Bureau, 2019), 43.5% of students attending a higher institution are males while 56.5% are females. In one of the target institutions of the study, the male:female ratio even reached 1:1.9, hence the large gap.

In the majority of higher education institutions in Macau, it is mandatory for first-year students to take courses related to English

learning. However, the purpose of the instruction varies. In the first two target institutions, students are expected to study English in a professional context, while in the third one, they learn English mainly for academic purposes. English classes are conducted in a wide variety of ways and are subject to the curriculum design and the style of the instructors. Like their secondary counterparts, the competence of the tertiary participants ranged from pre-intermediate (A2) level to upper-intermediate (B2) level.

Instruments

This study employed Horwitz et al.'s (1986) FLCAS as the instrument of the study. As mentioned in the Introduction, the FLCAS is a 33-item questionnaire survey (see Appendix) aiming at identifying respondents' anxiety level in English learning. A large number of studies have made use of different latent construct constitutions of the FLCAS in their studies, suggesting their own interpretation of the corresponding areas of anxiety of the 33 items. This study compared four different models with different constitutions of latent constructs: the original set of three latent constructs suggested by Horwitz et al. (1986), the four latent constructs suggested by Park (2012), and a hypothesized four-construct model as well as a hypothesized five-construct model suggested by the authors.

Park's (2012) latent construct division differs from Horwitz et al.'s (1986) in the way that Park included the additional latent construct of Foreign Language Classroom Anxiety. All nine items of this latent construct were originally categorized as Test Anxiety in Horwitz et al.'s model, leaving only three items of Horwitz et al.'s construct of Test Anxiety unchanged. The first model that this study suggests (Model 3, see Table 1), with four latent constructs, is similar to Park's model but several of the items originally categorized as Communication Apprehension have been re-classified as English Class Anxiety (which was called Foreign Language Classroom Anxiety in Park's model), making it a latent construct with 14 items, compared with 9 in Park's model.

The second model suggested in this study (Model 4 in Table 1) has the most latent constructs of all models with five, including one that has never previously been adopted for a study: Peer Influence. However, from daily classroom observation of EFL learners and from pilot selected interviews, the authors found peers to be a factor potentially influential

to a learner's feeling about learning English.

Constructs

The three latent constructs suggested by Horwitz et al. (1986) in their devised FLCAS are Test Anxiety, Communication Apprehension, and Fear of Negative Evaluation. *Test anxiety*, as defined by Sarason (1978), refers to "the tendency to view with alarm the consequences of inadequate performance in an evaluative situation." It expresses language learners' worry about failing to conform to the requirements set by a certain set of graded assessments. *Communication apprehension* can be understood as a person's level of fear or anxiety associated with either real or anticipated communication with another person or persons (McCroskey, 1978). In an English classroom, it is not hard to find learners who are very reluctant to speak up, most probably out of fear of not knowing how to express themselves clearly and accurately. Finally, *fear of negative evaluation* refers to the avoidance from others' evaluations, distress over their negative evaluations, and the expectation of negative evaluation from others (Watson & Friend, 1969).

Park's (2012) study attempted to divide items originally categorized as Communication Apprehension into two constructs: one still about anxiety arising from communication, and the other about learners' understanding of a foreign language class. For better understanding, the latter construct is labeled English Class Anxiety in this study.

Finally, this study suggests that a fifth construct, Peer Influence, should be added. From the authors' observation, *peer influence* can be understood as the contributing and debilitating factors that arise from the coexistence, interaction, and behavior of peers together with whom one learns.

RESULTS

Comparison of the Four Models

Internal consistency reliabilities using Cronbach's α were computed on the items in the constructs of all four models. As shown in Table 1, Cronbach's α for the items in the latent constructs ranged from .706 to .919. A Cronbach's coefficient α of higher than .600 can be considered

consistent (Landau & Everitt, 2004). The relatively lower values found in constructs such as Test Anxiety in Model 2, 3, and 4, and Fear of Negative Evaluation and Peer Influence in Model 4 can be attributed to the relatively small number of items in each construct (Park, 2012).

TABLE 1. Constructs of Models of the FLCAS

Model	Constructs	No. of Items	Cronbach's α
Model 1 (Horwitz et al. 1986)	Communication Apprehension	11	.882
	Test Anxiety	15	.916
	(Fear of) Negative Evaluation	7	.822
Model 2 (Park, G. P. 2012)	Communication Apprehension	15	.919
	Foreign Language Classroom Anxiety	9	.864
	Fear of Negative Evaluation	6	.789
	Test Anxiety	3	.740
Model 3 (Lei & Chan)	Communication Apprehension	8	.880
	English Class Anxiety	14	.901
	Fear of Negative Evaluation	7	.828
	Test Anxiety	4	.706
Model 4 (Lei & Chan)	Communication Apprehension	7	.862
	English Class Anxiety	13	.900
	Fear of Negative Evaluation	5	.757
	Test Anxiety	4	.764
	Peer Influence	4	.727

In order to determine if the hypothesized models, especially the construct of Peer Influence, can be a construct of an FLCAS, confirmatory factor analysis (CFA) was performed. Indices including the chi-square (χ^2) statistic, the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the Bentler-Bonett normed fit index (NFI) were computed. All results of the indices are listed in Table 2.

TABLE 2. Goodness-of-Fit Indices for the Constructs of the FLCAS

Model	Number of Constructs	χ^2	df	RMSEA	CFI	NFI
Model 1 (Horwitz et al. 1986)	3	2598**	492	.084	.807	.774
Model 2 (Park, G. P. 2012)	4	2481**	489	.082	.818	.784
Model 3 (Lei & Chan)	4	2545**	489	.083	.812	.779
Model 4 (Lei & Chan)	5	2584**	485	.085	.808	.775

Note. **$p < .01$. χ^2 = chi-square statistic; RMSEA = root mean square error of approximation; CFI = the comparative fit index; NFI = Bentler-Bonett normed fit index.

It can be found from Table 2 that all the chi-square statistics are found to be significant, which means the models are inadequate to fit the data (Park, 2012). However, the chi-square statistic is easily affected by the sample size and the complexity of the models, and a bigger correlation and a bigger sample size usually lowers the possibility for the chi-square statistic to accept a model (Bentler & Bonett, 1980; Kline, 2005; Marsh et al., 1988; Marsh & Hocevar, 1985; Schumaker & Lomax, 1996). As for RMSEA, a good model fit value is less than 0.05, and if the value is between 0.05 and 0.08, the model is said to be of a fair fit (Browne & Mels, 1990; McDonald & Ho, 2002; Schumaker & Lomax, 2004; Steiger, 1989). None of the values in Table 2 reaches this range, suggesting a relatively weaker fit in the model. For both CFI and NFI, a value of higher than .90 indicates a good fit. All of the values of CFI and NFI for the four proposed models were found to be close but falling short of the good-fit cut-off point. (Bentler, 1990; Bentler & Bonett, 1980; Park, 2012). Overall, it can be said that the goodness-of-fit indices presented in Table 2 indicated that the two original models and the two hypothesized models did not fit the data.

Considering the above outcome, this study analyzed the Pearson product-moment correlations among the latent constructs suggested in Model 4. The results, as presented in Table 3, indicate that the corrections among the five constructs are significant, and it might suggest the lower goodness-of-fit indices mentioned above.

TABLE 3. Pearson Correlations Among the Five Constructs in Our Hypothesized Model

	CA	ECA	FNE	TA	PI
Communication Apprehension (CA)	1	.799**	.738**	.714**	.734**
English Class Anxiety (ECA)		1	.761**	.811**	.769**
Fear of Negative Evaluation (FNE)			1	.674**	.657**
Test Anxiety (TA)				1	.687**
Peer Influence (PI)					1

Note. **$p < .01$.

From the results presented in Table 2 and 3, although Model 4 with a five-latent-construct constitution may not prove to be a good fit in indices, it comes close and rivals the other three hypothesized models. Thus, the researchers attempted to carry out additional data analysis with Model 4.

Data Analysis of the FLCAS with Model 4

The means of FLCA of each of the five latent constructs of Model 4, as well as the overall means of FLCA, were computed and found to range from 2.937 and 3.256. All the means are above the value of 2.5, which suggests a relatively higher level of FLCA exhibited by the respondents. In addition, among the top 6 questionnaire items with the highest means (see Table 4), two of them fall into the construct of Peer Influence, suggesting that peer influence plays a role in Macau learners' English learning. Among the overall means, Peer Influence ranked second highest among the five constructs, and it was found to be slightly more influential in secondary learners than in their tertiary counterparts (secondary mean: 3.2098 vs. tertiary mean: 2.9831).

TABLE 4. Top 6 Items with Highest Means of FLCA

Item	Construct	Mean	Std. Deviation
When I'm on my way to language class, I feel very sure and relaxed. (reversed)	ECA	3.52	0.987
I always feel that the other students speak the foreign language better than I do.	PI	3.48	1.107
I start to panic when I have to speak without preparation in language class.	CA	3.46	1.163
I get nervous when the language teacher asks questions which I haven't prepared in advance.	FNE	3.46	1.047
I am usually at ease during tests in my language class. (reversed)	TA	3.44	1.139
I keep thinking that the other students are better at languages than I am.	PI	3.42	1.124

Note. ECA = English Class Anxiety; PI = Peer Influence; CA = Communication Apprehension; FNE = Fear of Negative Evaluation; TA = Test Anxiety.

As for the comparison between secondary and tertiary learners, it can clearly be seen in Table 5 that for all five latent constructs and the overall FLCA, secondary learners exhibit a slightly higher level of anxiety compared to their tertiary counterparts. This might be due to the fact that secondary F.6 learners face more stress in learning because of their upcoming university admission challenge.

TABLE 5. Descriptive Statistics: Mean for Each Construct Based on Education Level

Construct	Education Level	Mean	Std. Deviation
Communication Apprehension	Secondary	3.2705	.79725
	Tertiary	3.1486	.75498
Fear of Negative Evaluation	Secondary	2.9456	.72805
	Tertiary	2.8757	.74371
Test Anxiety	Secondary	3.0904	.92623
	Tertiary	2.9020	.96321
English Class Anxiety	Secondary	3.0712	.75556
	Tertiary	2.9387	.70627
Peer Influence	Secondary	3.2098	.79650
	Tertiary	2.9831	.90785
Overall Anxiety	Secondary	3.1123	.71345
	Tertiary	2.9746	.69637

Table 6 shows the correlation between different latent constructs of Model 4 and the overall FLCA of Macau EFL learners. The analysis by means of the Pearson product-moment correlations showed that English Class Anxiety was found to have the highest correlation with the overall FLCA average. Such a result is consistent with both groups of secondary and tertiary respondents of the study. However, it is worth noting that Test Anxiety had the highest correlation with FLCA in Horwitz et al.'s (1986) study but in our study ranked behind English Class Anxiety and Communication Apprehension in its correlation with overall FLCA among the secondary respondents, and behind English Class Anxiety among all the tertiary respondents. It is even behind Peer Influence in two of the three institutions of the sample tertiary groups. This suggests that with a different latent construct of factors, the correlation between different constructs and foreign language learning anxiety may vary. Further research is hence needed to discover more possible relationships between the different constitutions of latent constructs within the area of FLCA.

TABLE 6. Pearson Correlations Among the Five Constructs in Our Hypothesized Model

			CA	FNE	TA	ECA	PI
S1	OAA	Pearson Correlation	.893**	.847**	.833**	.958**	.819**
		Sig. (2-tailed)	.000	.000	.000	.000	.000
		N	181	181	181	181	181
S2	OAA	Pearson Correlation	.899**	.846**	.864**	.953**	.814**
		Sig. (2-tailed)	.000	.000	.000	.000	.000
		N	187	187	187	187	187
S3	OAA	Pearson Correlation	.919**	.847**	.892**	.969**	.891**
		Sig. (2-tailed)	.000	.000	.000	.000	.000
		N	122	122	122	122	122
All Secondary	OAA	Pearson Correlation	.903**	.852**	.867**	.961**	.843**
		Sig. (2-tailed)	.000	.000	.000	.000	.000
		N	514	514	514	514	514
C1	OAA	Pearson Correlation	.838**	.887**	.910**	.942**	.847**
		Sig. (2-tailed)	.000	.000	.000	.000	.000
		N	31	31	31	31	31
C2	OAA	Pearson Correlation	.852**	.815**	.897**	.958**	.955**
		Sig. (2-tailed)	.000	.000	.000	.000	.000
		N	15	15	15	15	15
C3	OAA	Pearson Correlation	.930**	.747**	.816**	.976**	.879**
		Sig. (2-tailed)	.000	.000	.000	.000	.000
		N	28	28	28	28	28
All Tertiary	OAA	Pearson Correlation	.873**	.816**	.875**	.958**	.871**
		Sig. (2-tailed)	.000	.000	.000	.000	.000
		N	74	74	74	74	74
Overall	OAA	Pearson Correlation	.900**	.847**	.868**	.960**	.846**
		Sig. (2-tailed)	.000	.000	.000	.000	.000
		N	588	588	588	588	588

Note. **Correlation is significant at the 0.01 level (2-tailed). OAA = Overall FLCA Average; CA = Communication Apprehension Average; FNE = Fear of Negative Evaluation Average; TA = Test Anxiety Average; ECA = English Class Anxiety Average; PI = Peer Influence Average; S = Sample Secondary School; C = Sample Tertiary Institution.

DISCUSSION

As is shown in Table 5, both secondary students and their tertiary counterparts exhibit foreign language classroom anxiety with a mean of

higher than 2.5 for each item. This highlights the relatively high FLCA among Macau EFL learners, and suggests that some solutions to the situation are needed. Moreover, according to Table 6, the factor of English Class Anxiety influences students' overall level of anxiety the most, which is followed by Communication Apprehension, Test Anxiety, Fear of Negative Evaluation, and Peer Influence. One item depicting English Class Anxiety was found to show the highest mean of FLCA, namely the reversed item "When I'm on my way to language class, I feel very sure and relaxed." In other words, a majority of language-anxious students do not enjoy English class and feel perturbed when thinking of going to English classes. One point that is worth considering is that Peer Influence contributes to two of the six highest items that account for language anxiety. Although Peer Influence does not seem comparatively salient to overall FLCA based on the Pearson correlation, the effects derived from peers cannot be overlooked. When looking at the two corresponding items, we recognize that learners are inclined to compare English capability with their classmates, which can result in high anxiety if students cannot foster a positive attitude towards language learning. Recommendations should be given to teachers on how to modify and introduce methods to engage all students and boost the confidence of relatively anxious students.

In order to gain a deeper understanding of the anxiety of students towards learning English, a semi-structured interview was conducted following the questionnaire survey. Twelve consenting students from one of the secondary schools were selected as interviewees. They were evenly divided into high-anxiety and low-anxiety groups based on the result of their responses. Open-ended questions and scale questions were asked. Abstracting the main points from their responses, the following insights were found: first, both groups reflected having higher anxiety towards speaking in English than taking English tests. Most of them agreed that speaking demands greater mastery of English, and they were afraid of giving immediate responses in English. In contrast, English tests can mostly be prepared for in advance, and accordingly, anxiety can be effectively lessened with clearer test coverage. With regard to speaking English, almost all of the interviewees agreed that there is a lack of chances to speak up in English classes. The more motivated and less anxious interviewees can usually seek ways to practice after school. Some ideas like "I will talk with my friends in English after class," "there are some platforms on the internet to talk with foreigners," and

"I am really fulfilled when they [the native speakers] praise my accent" were reflected. However, students from the high-anxiety group are unmotivated in and out of the classroom. The lack of motivation in English class deprives them of a sense of achievement, which leads to anxiety when they are required to speak in English or take English tests.

IMPLICATIONS

Similarities in the Education and EFL Situations in Korea and Macau

The education frameworks and EFL teaching and learning environments in Korea and Macau are very similar. Both regions run a basic educational system of six years of primary education, three years of lower-secondary, and three years of senior-secondary (South Korea: Learning Systems, n.d.). Students in both places perform substantially better in academics than the global average: in PISA 2018, both regions ranked in the top 10 in all three areas of reading literacy, mathematics, and science, with a score of at least 20 marks higher than the OECD average (Macau Special Administrative Region Education and Youth Affairs Bureau, 2019). In addition, female students in Korea and female students in Macau outperform their male counterparts with a statistically significant difference of over 20 points in reading literacy (PISA, 2019a, 2019b).

Concerning the perspective of learning English, children in both Korea and Macau learn English at a young age, with Macau's children starting as early as pre-primary education (when children reach the age of 3) and Korea's starting at Grade 3 (South Korea Education, n.d.). English is not the official language of either region; however, as both regions put heavy stress on tourism, global trade, and the service industry, English is an international language that is valued in both places, especially in the educational systems. For secondary graduates of both places, English is a core section for the test that determines their admission of tertiary educational institutions. In Korea, the English section of the College Scholastic Ability Test (CSAT) is deemed to be "notoriously difficult" (Park, 2018). Because of this, secondary school graduates-to-be in both regions face enormous pressure for a place in

tertiary institutions, and the situation is especially serious in Korea, where students begin studying for CSAT as early as their first year of high school, attending extracurricular study academies and cram schools for hours each day after their regular classes, which can be up to 16 hours of studying each day (Tai, 2018). They aim high for prestigious universities, and the college entrance exam is believed to determine students' course of life and future professions in a South Korean society where graduating from a prestigious university is crucial to obtaining a successful job (Liu, 2019). The above phenomenon can also be reflected in the massive amount of spending on English learning in both regions. It is estimated that in Korea, a total of about 15 trillion won (US$15.8 billion), is spent on English learning per year (Jeon & Choi, 2006).

Concerning EFL instruction, it has been found that rote learning and a heavy emphasis on grammar is the norm of English instruction in both regions, causing EFL students to lack a real sense of the use of English in real life and as a world lingua franca (Hogan, 2015).

Similarities in Foreign Language Classroom Anxiety Between Students in Korea and Macau

Based on this study of Foreign Language Classroom Anxiety of Macau EFL secondary school and first-year university students, and Manley's (2005) study of Foreign Language Anxiety of first-year students in a Korean university, it has been found that EFL students of both regions exhibit a relatively higher anxiety level when it comes to Communication Apprehension (in the study by Manley, it is referred as "uncomfortableness when speaking"), accounting for 5 out of the 15 highest-ranked items in Macau, and 10 out of the 15 highest-ranked items in Korea. Despite a different latent construct, Peer Influence (referred to as "comparisons to peers" in the study by Manley) can be seen as one of the factors that may affect students' foreign language anxiety. Both studies suggest that students from Korea and from Macau face a relatively higher level of anxiety when learning English, and the factors that contribute to such anxiety are similar.

Implications of FLCA for EFL Teaching and Learning in Both Macau and Korea

Speaking Anxiety

Both students in Korea and Macau suffer from the fear of speaking English. Based on Manley's (2015) and this research, the self-perceived lower English proficiency and expectations of making no mistake could contribute to the reticence in the English classroom to a great extent. Students may want to speak up but are not confident of their English ability. To avoid being embarrassed by revealing any possible errors in their speech, they would rather stay silent in the class. As a result, they only give a response when teachers specifically nominate them. Another possible situation is that students are pondering over their wording or hesitating to give an answer to a question. In fact, teachers may feel uncomfortable with the awkward silence or be in need of catching up in their lesson plan, and thus hasten to prompt a verbal reaction from the students. In our opinion, instructors in the Asian context can focus more on the passive students by giving more hints to them and praising their achievement more frequently. The greater engagement in class should relieve speaking anxiety over time. Moreover, Manley (2015), paraphrasing the words of Lee and Ng (2010), stated that the feedback and responses from students could be different if a longer waiting time were allowed. This indicates that with an increase in the level of instructors' patience, students are likely to produce more meaningful responses actively.

Peer Influence

Learning with classmates is deemed to be effective in relieving student speaking anxiety. As is suggested by researchers (Cao & Philp, 2006; Cheng, 2000; Manley, 2015), using group work, pair work, and discussions is a practicable means to encourage students to speak up more actively. This view is in accordance with the reflection of the interviewees of this research. Participants from both high- and low-anxiety groups generally agree that they are more likely to shift the focus from accuracy to meaning and even fluency when they interact with peers. Some of the students even seek opportunities to speak outside the classroom with their peers, which can enable them to make English a useful tool in their daily life rather than required knowledge to be crammed for passing tests and high-stakes exams. Thus, putting

students into groups can reduce the pressure they get from the presence of teachers, and they can avoid the fear of making mistakes and thus focus more on building meaning in English conversations. Peer influence here appears positive for language learning. However, there could be a possibility that students compare themselves with peers and thereby struggle with thinking lowly of themselves. Such self-consciousness is quite natural and automatic to some extent, and teachers should allocate students into groups taking into consideration their capability and personality.

Teacher–Student Relationship

The authors take note of the importance of rapport between students and teachers in the course of learning English. Abstracting the results of other studies (Cao & Philp, 2006; Cheng, 2000; Manley, 2015) as well as the results of this research, students are found to be less anxious speaking with someone they are familiar with, namely their peers in most cases. Teachers can thus get to know more about their students and try to be friendly rather than establish a leader–subordinate relationship. The closer the relationship is between students and teachers, the less nervous students will feel about speaking English in front of teachers. Additionally, the experience and daily observations of the authors, which are in accordance with the opinions of the interview participants, is that the better the teacher–student relationship is, the more likely students will take teachers' instructions as friendly advice rather than criticism. In addition, paying more attention to wording and asking questions consistent with different students' level can give incentive to highly anxious students. Some of the interviewees shared that they feel reluctant to listen to teachers either because the topics are boring or the questions are difficult. When they were asked for some suggestions, they all agreed that they would be more engaged if the questions from teachers were made easier. For the same reason, teachers may try to be trendy, choose hot topics that their students are interested in for discussions and classroom tasks in order to encourage a sense of involvement. Moreover, smiling more and being humorous can draw the teachers and students closer as teachers appear to be more approachable in this light.

CONCLUSIONS

In order to investigate how foreign language classroom anxiety affects English learning of Macau's EFL secondary school graduates-to-be and college first-year learners, as well as the correlative influence that different constructs (in particular, Peer Influence) of FLCAS have on both groups, this study compared the latent constructs suggested by Horwitz et al. (1986) and Park (2012) with two hypothesized models proposed in this study. CFA has been conducted with the responses collected from 531 secondary respondents and 74 college respondents.

Three main conclusions can be drawn from the study: first, concerning the construct categories, the results of the study showed none of the four models with different categories of constructs provided acceptable suitability of indices. Despite this, when further descriptive and correlation analysis was carried out, it was found that each of the five constructs suggested in this study highly correlated with the overall FLCA as well as inter-correlated with the other four constructs in the scale. All values have been found to be statistically significant (see Table 3).

The second conclusion concerns the comparison of the level of FLCA between secondary and tertiary learners, and the comparison between genders. As indicated in the results, secondary learners exhibited a higher level of anxiety than the tertiary learners. However, gender difference was not found to be remarkable, as the mean scores of the two genders are very close (male: 3.1104 vs. female: 3.0807) and the two genders each exhibited a slightly higher level of anxiety for certain factors.

The final conclusion is about the construct that contributes to a higher level of FLCA. Through further analysis, it has been discovered that the influence of negative evaluation from teachers is not as dominant as other constructs. As indicated in Table 5, the mean scores of the construct Fear of Negative Evaluation, regardless of education levels, is the lowest among the five constructs, suggesting a relatively lower anxiety level. This result corresponds with the means of all 33 items: Two of the five items with the lowest mean scores belong to the category Fear of Negative Evaluation, suggesting that learners do not view teachers' comments as threatening and intimidating as other factors. In addition, as mentioned in the results, the influence of Test Anxiety

is less dominating than what Horwitz et al. (1986) suggested in their study. Similar to Fear of Negative Evaluation, two of the five lowest anxiety items belong to Test Anxiety. Contrastively, Peer Influence and Communication Apprehension occupy the highest proportion of the items with a high mean score, occupying two out of the five and five out of the ten highest anxiety items, respectively. In other words, while speech anxiety deserves more focus in English pedagogy, peer influence should also be taken into consideration as a source of English learning anxiety.

As for implications for future teaching, teachers may consider introducing more group work in their classes. Wording and levels of questions that cater to the capability of the students should be considered, and test coverage can be made more explicit so as to lessen unnecessary worry of students that arises from the perceived need to prepare intensely for assessments.

THE AUTHORS

Riko LEI, Ut Meng is a student at Beijing Foreign Studies University. She has great interest in English and has an insight into the difficulties of ESL/EFL students studying English. She excels in English, especially in speaking and writing. She came in first in English in Form 6. In 2017, she won the Best Speaker of Chinese-Medium Schools award in the 16th Macao-wide Speech Contest. In 2018, she became the first Macau student to win the Gold Award in the category of creative writing at the Genius Olympiad. Her research interests include cognitive linguistics, foreign language learning anxiety, and overcoming learning difficulties of English. Email: riko.research@gmail.com

Alan CHAN, Ka Lon is currently an instructor at the English Language Centre of the University of Macau. He taught English at a secondary school in Macau for over 14 years and puts great effort into the development of curriculum and new methods of instruction. He has also been a training teacher for many international STEAM competitions, leading students to set new Macau records. With his dedication to education, he was awarded the Distinguished Teacher Honour for the school year 2016–2017 by the local educational bureau. His research interests include EFL writing pedagogy, writing correction feedback, and language learning anxiety. Email: alanchan@um.edu.mo

REFERENCES

Aida, Y. (1994). Examination of Horwitz, Horwitz, and Cope's Construct of Foreign Language Anxiety: The case of students of Japanese. *Modern Language Journal*, *78*,155−168.

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588−606.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*, 238−246.

Browne, M. W., & Mels, G. (1990). *RAMONA user's guide*. Department of Psychology, Ohio State University.

Cao, Y., & Philp, J. (2006). Interactional context and willingness to communicate: A comparison of behavior in whole class, group and dyadic interaction. *System*, *34*(4), 480−493.

Cheng, X. (2000). Asian students' reticence revisited. *System*, *28*(3), 435−446.

Cheng, Y-S., Horwitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning*, *49*, 417−446.

Hogan, J. (2015, September 13). Koreans and studying English (part 1). *Hi Expat*. https://www.englishspectrum.com/koreans-and-studying-english-part-i/

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal*, *70*(2), 125−132.

Higher Education Bureau, Macau. (2019). *Summary of higher education data 2018*. https://www.dses.gov.mo/doc/2018/stat2018.pdf

Jeon, H-C., & Choi, H-S. (2006). *The economics of English* (CEO Information No. 178). Samsung Economic Research Institute.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). Guilford.

Landau, S., & Everitt, B. S. (2004). *A handbook of statistical analyses using SPSS*. Chapman & Hall.

Lee, W., & Ng, S. (2010). Reducing student reticence through teacher interaction strategy. *ELT Journal*, *64*(3), 302−313.

Liu. (2019, November 14). S. Korean students take all-important college entrance exam. *Xinhua*. http://www.xinhuanet.com/english/2019-11/14/c_138554571.htm

Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *Modern Language Journal*, *92*, 71−86.

Macau Special Administrative Region Education and Youth Affairs Bureau. (2019, December 3). *Macao-China: High-performing and fast-improving basic education system according to PISA 2018*. https://www.gov.mo/zh-hant/news/310069/

MacIntyre, P. D., & Gardner, R. C. (1989). Anxiety and second-language learning: Toward a theoretical clarification. *Language Learning*, *39*(2), 251−275. https://doi.org/10.1111/j.1467-1770.1989.tb00423.x

Mak, B. (2011). An exploration of speaking-in-class anxiety with Chinese ESL learners. *System*, *39*, 202−214.

Manley, K. (2015). *Comparative study of foreign language anxiety in Korean and Chinese students* [Master's thesis, St. Cloud State University]. The Repository, St. Cloud State University. https://repository.stcloudstate.edu/engl_etds

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher-order factor models and their invariance across groups. *Psychological Bulletin*, *97*, 562−582.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, *103*(3), 391−410.

Matsuda, S., & Gobel, P. (2004). Anxiety and predictors of performance in the foreign language classroom. *System*, *32*, 21−36.

McCroskey, J. C. (1978). Validity of the PRCA as an index of oral communication apprehension. *Communication Monographs*, *45*(3), 192−203.

McDonald, R. P., & Ho, M. R. (2002). Principles and practice in reporting structural equation analysis. *Psychological Methods*, *7*, 64−82.

Park, G-P. (2012). Investigation into the constructs of the FLCAS. *English Teaching*, *67*(2), 207−220.

Park, S-S. (2018, November 17). Check your English ability with "notorious" Korean college entrance exam. *The Korean Times*. https://koreatimes.co.kr/www/nation/2019/09/177_258803.html

PISA 2018 student performance of Korea (2019a). OECD. https://gpseducation.oecd.org/CountryProfile?primaryCountry=KOR&treshold=10&topic=PI

PISA 2018 student performance of Macau (2019b). OECD. https://gpseducation.oecd.org/CountryProfile?primaryCountry=MAC&treshold=10&topic=PI

Sarason, I. G. (1978). The Test Anxiety Scale: Concept and research. In C. D. Spielberger & I. G. Sarason (Eds.), *Series in clinical and community psychology* (pp. 193−216). Hemisphere.

Schumaker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. Lawrence Erlbaum.

Schumaker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling* (2nd ed.). Lawrence Erlbaum.

Shao, X. (2014). A study of Chinese college students' English reading anxiety. *American Journal of Educational Research*, *2*(5), 299−303.

South Korea Education. (n.d.). *Asia Society*. https://asiasociety.org/education/

south-korean-education

South Korea: Learning systems. (n.d.). *National Center on Education and the Economy*. http://ncee.org/what-we-do/center-on-international-education-benchmarking/top-performing-countries/south-korea-overview/south-korea-in structional-systems/

Steiger, J. H. (1989). *EZPATH: A supplementary module for SYSTAT and SYSGRAPH*. SYSTAT.

Tai, C. (2018, November 15). Why South Koreans are trapped in a lifetime of study. *SCMP*. https://www.scmp.com/week-asia/society/article/2173414/ schools-never-out-why-south-koreans-are-trapped-lifetime-study

Toth, Z. (2008). A foreign language anxiety scale for Hungarian learners of English. *WoPaLp*, *2*, 55−78.

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, *33*, 448−457.

Young, D. J. (1994). New directions in language anxiety research. In C. A. Klee (Ed.), *Faces in a crowd: The individual learner in multisection courses* (pp. 3−46). Heinle & Heinle.

Zhao, N. (2007). A study of high school students' English learning anxiety. *Asian EFL Journal*, *9*(3), 22−34.

APPENDIX

Foreign Language Classroom Anxiety Scale	
Scale: (1) Strongly Disagree (2) Disagree (3) Neural (4) Agree (5) Strongly Agree	
Chinese Translation	**English**
1 上英語課要說話的時候, 我總欠缺自信.	I never feel quite sure of myself when I am speaking in my foreign language class.
2[R] 上英語課時, 我不會擔心會犯錯。	I don't worry about making mistakes in language class.
3 在英語課堂上當老師叫到我時, 我會顫抖.	I tremble when I know that I'm going to be called on in language class.
4 在課堂上我如果不明白老師說的話, 我就會很緊張.	It frightens me when I don't understand what the teacher is saying in the foreign language.
5[R] 即使我要上更多的英語課時, 我一點也不煩惱.	It wouldn't bother me at all to take more foreign language classes.
6 上英語課時, 我總是在想課外的事情.	During language class, I find myself thinking about things that have nothing to do with the course.
7 我總是在想別的同學學英語學得比我好.	I keep thinking that the other students are better at languages than I am.
8[R] 我在進行英語課的一些測驗時, 通常感到輕鬆.	I am usually at ease during tests in my language class.
9 當我在英語課堂上, 沒有準備過的情況下說話, 我會感到慌張.	I start to panic when I have to speak without preparation in language class.
10 我會擔心我的英語課堂成績不合格.	I worry about the consequences of failing my foreign language class.
11[R] 我不明白為什麼有些人會因為上英語課而感到難過.	I don't understand why some people get so upset over foreign language classes.
12 上英語課時, 我會慌得把會的也忘了.	In language class, I can get so nervous I forget things I know.
13 我覺得上英語課時主動回答問題會令我感到尷尬.	It embarrasses me to volunteer answers in my language class.
14[R] 我與英語為母語者說話時不會覺得緊張.	It would not be nervous speaking the foreign language with native speakers.
15 如果有不明白老師改正我的地方, 我會覺得難過.	I get upset when I don't understand what the teacher is correcting.
16 即使我為英語課準備了很多, 我依然會覺得焦慮.	Even if I am well prepared for language class, I feel anxious about it.
17 我時常不太想上我的英語課.	I often feel like not going to my language class.
18[R] 我在英語課上說英語會覺得自信.	I feel confident when I speak in foreign language class.

19	我很害怕我的英語老師會改正我的每一個錯誤.	I am afraid that my language teacher is ready to correct every mistake I make.
20	在英語課堂上當老師叫到我時, 我心總是跳得很快.	I can feel my heart pounding when I'm going to be called on in language class.
21	我愈是為英語測驗溫習, 我愈是感到混亂.	The more I study for a language test, the more confused I get.
22^R	我不為要好好準備英語課而感到壓力.	I don't feel pressure to prepare very well for language class.
23	我總是感到其他學生說英語比我好.	I always feel that the other students speak the foreign language better than I do.
24	要在其他同學面前說英語, 我會感到很難為情.	I feel very self-conscious about speaking the foreign language in front of other students.
25	我覺得英語課的進程太快, 很擔心會跟不上.	Language class moves so quickly I worry about getting left behind.
26	比起其他課堂, 上英語課令我感到更焦慮.	I feel more tense and nervous in my language class than in my other classes.
27	我在英語課上說話時感到緊張與混亂.	I get nervous and confused when I am speaking in my language class.
28^R	每當我要上英語課時, 我總是感到胸有成竹且心情輕鬆.	When I'm on my way to language class, I feel very sure and relaxed.
29	當我不明白老師在英語課上說的所有的字的時候, 我會感到緊張.	I get nervous when I don't understand every word the language teacher says.
30	學英語要掌握這麼多語法, 會讓我擔心.	I feel overwhelmed by the number of rules you have to learn to speak a foreign language.
31	我在說英語時, 會因為擔心同學嘲笑而感到焦慮.	I am afraid that the other students will laugh at me when I speak the foreign language.
32^R	在英語為母語者在一起我感到很自在.	I would probably feel comfortable around native speakers of the foreign language.
33	當老師問我一個我無準備的問題時, 我會很緊張.	I get nervous when the language teacher asks questions which I haven't prepared in advance.

Note. Superscript ^{R} indicates reverse scored items.

Creating Level Tests for University EFL Courses at a Korean University

Tory S. Thorkelson

Hanyang University, Seoul, Korea

With the popularity of tests like the TOEIC, TOEFL and IELTS, and their proven records of successes in verifying the language abilities of non-native English speakers for specific purposes, it is probably natural that institutions in a variety of countries require them as entry or graduation requirements. However, there are situations where a high score on one of these tests may not be in and of itself adequate to prove that students have the necessary skills to succeed in their studies – specifically where the classes they are taking go well beyond the typical four skills of listening, speaking, reading, and writing – and so this paper will show how class-specific level tests were created and administered at a private university in Seoul between 2006 and 2014 with a discussion of the successes and failures encountered.

Keywords: TOEIC, IELTS, creating level tests, EFL, skills-based, university, Korea

INTRODUCTION

Wikipedia lists over 100 English proficiency tests ("List of Proficiency Tests," 2019), but most of these will probably be unfamiliar to the typical English instructor. Adding to the confusion, is the inclusion of terms like "proficiency" and "competency" in the test titles themselves, with "proficiency" appearing in eight of the listed test titles and "competency" in four. However, as Ingram and Wiley (1992) state "'Language proficiency'...refers to a learners ability to mobilize the language in order to carry out language tasks; the term 'communicative competence' is used both rigorously in ways that differ little from language proficiency and loosely to mean the ability to communicate" (p. 33). Additionally, Lowe (1988, as cited in Bachman, 1990), in referring

to the Interagency Language Roundtable (ILR), defined proficiency as follows:

> Proficiency equals achievement (ILR functions, content, accuracy) plus functional evidence of international strategies for creativity expressed in a single global rating of general language ability expressed over a wide range of functions and topics at any given...level. (p. 5)

Thus, there is confusion over what these terms actually mean when used for the purposes of assessing language ability in general.

Finally, terms like "assessment" and "attainment" are dispersed among tests that more clearly state that they are based on a recognized standard of language evaluation such as the Common European Framework of Reference for Languages (CEFR; Cambridge Assessment English, n.d.) or the American Council on the Teaching of Foreign Languages proficiency guidelines (ACTFL, n.d.). However, the appropriateness of using a commercial test as an entrance or graduation requirement for undergraduate students at Korean universities raises the issue of whether these tests are an acceptable assessment tool for overall English ability. When the author arrived at Hanyang University in 1998, a high TOEIC score was considered as an exemption for the Freshman English Conversation classes, and students were automatically assigned an A grade by the university as long as they submitted a valid TOEIC score before the end of their first term. Similarly designed policies are still enforced at numerous universities in Korea for incoming freshmen, international, and graduate students, including Gwangju Institute of Science and Technology, Korea University, and Hanyang University among others, with the intention of verifying language ability rather than rewarding students with earned course credits (GIST, n.d.; Hanyang University, n.d.; Study in Korea, 2016). While the regulations and data for the Hanyang University policy mentioned above are no longer available, In'nami and Koizumi (2017) discuss a similar situation in Japan:

> ...academic staff often create rules that administrative staff follow, academic staff are mainly responsible for issues that arise from the use of such rules. One may argue that awarding EFL credits is typically regarded as relatively low stakes, because credits for

foreign language courses are also necessary for graduation. Despite this relatively low-stakes situation, however, credit-awarding policy may be of fundamental importance because it could affect learners' long-term language development and, by extension, university education policies. If credit awarding is not appropriately executed, students who do not have sufficient English proficiency and need to improve may be inappropriately exempted from taking courses that are essential for their learning. Universities that give credits without firm justification may then run into difficulty in ensuring accountability in education quality to students and society. (p. 4)

Therefore, while administrative staff or university policies may see these widely recognized tests as valid and reliable for both entry and graduation, these practices may in fact be harming the language skills that university programs like those at many universities are designed to foster. Due to the above considerations, this paper will address the reasons why tests like the TOEIC, TOEFL, and IELTS may not be the best measuring tools for assessing university students before their university life begins and the process that took place during the creation, administration, and application of level testing for all courses offered within the skills-based program at Hanyang University's English Language and Literature Department (ELL) at the university's Seoul Campus (see Thorkelson, 2019).

Program Background

When the ELL program was getting off the ground in 2006, almost all four of the former Practical English Education Center (PEEC) faculty had been administering their own versions of entrance "tests" for our former department's classes informally to try to ensure that students enrolled in departmental classes were at the appropriate level(s) for these classes and to assess their skill levels upon entry to Hanyang University. As all Hanyang freshmen were required to take the first-year Practical English Conversation classes at that time, mixed-level classes grouped by majors were the norm rather than the exception. As Prodromou and Clandfield (2007) point out:

All classes are made up of mixed levels. As soon as you put two people together, you have a mixed-level situation, especially if

mixed levels are seen as more than a question of ability as demonstrated in tests of language proficiency. Mixed-level classes are...the result of the different learning styles of students, the pace at which they each learn, their level and kind of motivation, their personal interests, their background knowledge, and any social problems they may be facing. (p. 57)

The majority of the faculty felt that, with the creation of a new program, using standardized level tests for all classes and levels was preferable to depending on TOEIC scores like the university did. In research by Wilson (2001), it was found that both the TOEFL and TOEIC suffered from the same issue:

...overestimation of TOEFL LC mean for native-Korean speakers in the TOEFL testing context was deemed to parallel conceptually the...finding that the use of previously developed regression-based guidelines for inferring level of LPI [Language Proficiency Interview]-assessed EFL speaking proficiency from TOEIC scores resulted in systematic overestimation of LPI rating for native-Korean speakers in the TOEIC testing context. (p. 36)

This study, while the only one of its kind that the author was aware of before the ELL program was even conceived, caused doubts about the validity and reliability of using overall TOEIC scores as a stand-alone indicator of overall English proficiency for the ELL program's students. In addition, Sewell (2005) noted the following:

As an acceptable test mark, often in the 700 to 800 out of 990 range in Korea, is required before a job application will even be considered, there is a strong desire among many Koreans to achieve high TOEIC scores, a desire which is the basis for a large publishing and preparation course industry.... While the TOEIC purports by definition to be a test of communicative ability, such preparation books and courses do not generally teach English in any communicative or interactive way. Instead they focus on having students memorize structures and vocabulary items commonly found on different parts of the test, memorize grammar rules to dissect parts five and six, and learn where to find the answers to listening and reading passages without having to listen to or read the entire passage. Additionally, such books and courses teach a multitude of general test-taking

strategies and provide numerous practice tests. (p. 10)

These are issues that would probably have been noticeable to anyone who has been teaching in the Korean context for a few years, particularly when combined with the additional possible concerns related to the TOEIC mentioned by Sewell (2005) with regard to the TOEIC's validity as a speaking and writing measurement:

> While the TOEIC does not investigate a candidate's speaking or writing skills, ETS has also attempted to establish the concurrent validity of the test for these abilities. For speaking, ETS has focused on correlating the TOEIC to the US Foreign Service Institute's Language Proficiency Interview (LPI). Results from this have shown a correlation of between 0.71 and 0.83 (Woodford 1982, p. 14; Wilson 1993, p. 6), only slightly higher than Hughes' suggested minimum correlation of 0.70 for a relatively lows-takes test (1989, p. 24). However as noted in section 4.1 above, within a Korean sample the correlation was much lower at between 0.48 and 0.57 (Wilson, 2001). Other correlations between the TOEIC and speaking tests have produced some individual values as low as 0.49 (Hirai, 2002). These results mirror a general tenor in the independent literature to see the TOEIC as having only limited validity as a speaking test. With one exception, the TOEIC's claim of being a valid writing test seems to have had little attention since the TOEIC's initial validation study. That study (Woodford, 1982, p. 15) reported a correlation of 0.83 to an apparently in-house developed writing test, a figure which contrasts with Hirai's correlation of 0.66 for the TOEIC against the BULATS test of writing (Hirai, 2002). Hirai additionally strongly criticizes the original ETS writing test and suggests a TOEIC score is "practically meaningless as a measure of writing skill" (Hirai, 2002). (p. 14)

Further, with 3 or 4 sections of the same class offered at the same times and on the same days by different instructors, it was the ideal situation to attempt to minimize the adverse effects of both a commercial test like the TOEIC and those for mixed-level classes listed above by assigning students to the appropriate levels using a set of tests the faculty themselves designed and administered.

The question of whether to use a norm-referenced, criterion-referenced, or objective-referenced test was perhaps the biggest issue

faced by the faculty. According to Gronlund and Linn (1990), the differences between these test types are as described in Table 1.

TABLE 1. The Test Types

Test Type	Description
Norm-Referenced Test (NR)	Provides a measure of performance interpretable in terms of *an individual's relative standing in some known group*.
Criterion-Referenced Test (CR)	Provides a measure of performance interpretable in terms of a *clearly defined and delimited domain of learning tasks*.
Objective-Reference Test (OR)	Provides a measure of performance that is interpretable in terms of a *specific instructional objective (or objectives)*.

Note. From Gronlund and Linn (1990, p. 14).

At the beginning of the development of the level tests for our classes, it was only natural that the tests we were most familiar with influenced our decisions but − since the TOEIC test and IELTS test were the tests the faculty knew best and since they are norm-referenced and criterion-referenced, respectively, the tests used within our program evolved from criterion- and objective-referenced at the beginning to somewhere in between criterion- and norm-referenced as the faculty became more aware of the skills and abilities of the average incoming student over time. According to Vu et al. (2016), these two tests can be described as in Table 2.

Finally, as Frain (2009) notes, "Norm-referenced testing, which is common in Korea, can be characterized as being developed independently of any particular course of instruction" (p. 30). As a result of discussions among the faculty, it was decided to create in-house tests to evaluate and place our students in the respective classes, and the author, as head teacher at that time, was put in charge of the creation and administration of the level tests for all of the skills-based classes offered by the ELL Department. This arrangement continued from 2006 until the level tests were discontinued by the department in 2014. Accordingly, the rest of this paper will outline the evolutionary process of the level tests used for these courses (see Appendix A).

TABLE 2. Validity and Reliability of the TOEIC and IELTS Tests

Item	TOEIC	IELTS
Origin	Developed in 1979	Developed in 1989
Testing Organization(s)	ETS	Cambridge, British Council, IDP
Theory	**Theory of Language** *None. (Sewell, 2005, p. 13) *Fair, accurate and meaningful English-language assessment for the workplace specifically. (ETS)	**Communicative Language Competence:** *Language as a communication tool *The role of context of use *Interactiveness in language use
Assessment	**Norm-Referenced** *No direct provision of learners' language ability. *ETS's guidelines not used to interpret TOEIC scores in reference to the CEFR.	**Criterion-Referenced**
Validity and Reliability	MC Questions → High validity	High validity and much reliability
Common Practice	"Social moderation" *Two language skills test (L & R) still used. *This test should not be equated to the CEFR levels (in which four language skills are assessed) (Construct under representation)	Four language skills tested (objectively & subjectively)

Note. From Vu, Nguyen, & Nguyen, 2016, p. 17.

Issues in Level-Test Creation and Administration

The Students

Within the ELL Department, there are approximately 300 students enrolled in the undergraduate English Language and Literature degree program. They are required to take two classes taught in English by our faculty and can take another two to three optional classes. As our classes have been focusing on skills development rather than basic English fluency (see Appendix A) or the traditional four-skills structure of many ESL or EFL programs, the importance of a level test becomes essential to ensure that students are placed in the appropriate class or level for

their needs (for a comparison chart with a fairly accurate equivalency chart for the common frameworks and tests for evaluating English abilities, see EF, n.d.). The range of student abilities within the student population posed a clear challenge in this regard as abilities ranged from false beginners all the way to advanced levels. There were no truly proficient students in all skill areas, as many students spoke English quite fluently but could not write at the advanced level, for example, which meant that students could be in different classes and levels for different skills. This is perhaps one of the key reasons the tests were discontinued by the department, as many students did not want to be in different levels for different skills – or to be separated from their closest friends – but the faculty tried to enforce the results of the level tests as much as possible. Due to departmental and university policies, there are no official pre-requisites for any of the classes offered, so it was even more important that students enrolled in the ELL classes and levels took the classes they were assigned to.

The Level Tests

In 2006, the level tests were created for all writing and speaking classes. The writing tests were for Writing 1: Sentences; Writing 2: Paragraphs; and Writing 3: Essays. Students were given at least two topics in each case, and had 30 to 60 minutes to write a few sentences or a paragraph about one of the given topics. For the Writing 3 class, they were also given 2–3 topics to choose from and had up to 90 minutes to write a five-paragraph essay. However, most students finished the writing tests within 30 minutes for Writing 1 and 2, and 60 minutes for Writing 3. It was the faculty's intention to avoid the issue of time for these tests as mentioned by an interviewee in Teemant (2010):

> Time is sometimes a factor in my performance on an exam. Honestly, it takes me a little bit more time to read and actually understand clearly.... Most of the time, I know the answer, [but] I do something wrong because I did it really quick.... Sometimes I would like to have more time. (p. 8)

By 2011, when Writing 1 included sentences and paragraphs and Writing 2 was still essay-focused, more topics were added so that there were separate sheets for the list of topics and the double-sided writing sheet. Writing tests were scored out of a total of 21 points, initially

broken down into three categories worth seven points each. The categories were grammar/vocabulary, content, and format scored separately, but on later versions of the tests, the scores were simply out of 21 (see Appendix C for sample writing tests). For administrative purposes, the student names, email addresses and smart phone numbers were added to the writing test forms for cases where additional information or follow-up became necessary (see Appendices C and D to view how this information was collected for writing and speaking tests). The test criteria were loosely based on the IELTS public band descriptions since the author was most familiar with applying these (see British Council, n.d.-a, b). However, as only one of the program's faculty was trained to administer the IELTS test, this caused some issues with inter-rater reliability and validity as will be discussed in a later section.

The speaking test format was based on the author's experience with speaking tests given at Korean Air (KAL) between 1998 and 2006 and at YBM-Sisa from 1996 to 1998. These level tests were used to evaluate the English abilities of stewardesses, engineers, and flight crew and to assign incoming students at YBM into one of six proficiency levels. The ELL students were evaluated on their fluency, grammar/vocabulary, and ability to express opinions about the content related to the class and/or their experiences, which was a more consistent system than those at KAL or YBM. Speaking tests were for five minutes, and instructors chose random questions from a list. These questions started out as fairly generic conversation questions based on lists from http://iteslj.org/questions/ or https://esldiscussions.com/, but as courses became more skills-focused, level tests became more topic-specific (see Appendix D for sample speaking tests). Again, the IELTS speaking band descriptors (IELTS, n.d.) were used as a foundation for rating students out of a possible 21 points, with 7 points each for the three criteria mentioned above.

Initially, students were assigned a numeric grade out of 21 and a designation as Beginner (B), Intermediate (I), or Advanced (A), which matched up with the three classes/levels offered for all subjects taught. As there was an imbalance in abilities with smaller numbers of students in the Beginner and Advanced levels most terms, High Beginner (HB) and Low/High Intermediate (LI/HI) were also used to indicate students who could go into either Beginner or Intermediate levels (HB/LI) or be moved up into the Advanced class as necessary to balance out class

numbers (HI). With a small number of faculty, where there were any questions or issues about assigned levels, at least one other faculty member regraded the writing level tests in question and then the two instructors discussed their ratings to finalize the class and level before final lists were created and posted for students to access before classes were officially started in Week 2. For speaking tests, students might be asked to do another speaking test with another instructor as needed. As the add/drop period was normally at the beginning of Week 2, it was important to ensure that students were in the appropriate level before class lists were posted (2006−2009) or the administration informed them of their assigned class/level (2010−2014). This also caused some problems with the classes and program, and is discussed below.

Finally, as the instructors involved in the ELL program changed every two to three years on average, a set of level-test instructions was created by the head teacher to help new instructors become familiar with how the level tests worked (see Appendix B).

Advantages of the Level Tests

By creating and using in-house tests, the possible issues with the available commercial tests were avoided, including whether they were valid and reliable for the ELL program's purposes as mentioned above. For students, this meant that they did not have to pay for or take more tests specifically for the ELL program's classes. Secondly, the tests could be easily revised and updated each term for each class and new tests could be created as needed as the curriculum changed every two to four years. Thirdly, over time as the faculty's knowledge of individual student's abilities grew, the level-testing process became more norm-referenced than criterion- or objective-referenced since the faculty taught the same students in various classes over a three- to four-year period. This resulted in fewer errors being made in placing students in the appropriate classes and levels over time. Finally, a standard format with new questions each term meant that students were being assessed on the same criteria by all faculty, at least in theory, and that students were less likely to be able to prepare memorized answers beforehand for the writing and speaking tests. Teemant (2010) discusses student memorization for testing:

> ... [a] problem in testing for ESL students is the amount of memorization required for most tests. Memorizing content for tests

is complicated by the fact that they are memorizing in their second language. One student said, "It is hard [to] memorize in Portuguese, imagine in English." Other students described memorizing content-area information as a "double effort" or "harder" because it is "my second language." One student even suggested that her ability to memorize is different in her two languages: "In Korea, I think I really good at memorizing, and some how [sic] in U.S.A., I forget how to memorize." (p. 7)

Disadvantages of the Level Tests

As with any new program, there were issues to be overcome. The initial problems stemmed from students who were either unhappy with their assigned levels or wanted to be with their friends in other classes. As there were no official prerequisites, instructors had to include warnings in the class descriptions reminding students that they had to take a level test to be enrolled in the department's classes and that they could be assigned to classes other than the ones they preferred. Another concern was the perception by some of the other faculty that the level tests were a waste of time as they were done over two classes in Week 1. However, as many students used the add/drop period to try out various classes before they committed in Week 2 – or even skipped Week 1 entirely – this was both a weak justification for disparaging the level tests and a possibly valid concern, respectively. In these cases, faculty could request a writing sample for the affected students, and evaluate them using the same criteria as the level tests, but this was not recommended and was never offered openly to all students. It was a final option and rarely used. In attempting to match the tests' contents and styles to the classes the students were level-testing for, the tests included content and topics that were somewhat generic but also tried to mimic what students would experience in the classes themselves to some extent. Further, it was natural that students might not have been prepared for the test formats offered, which was an intentional challenge to make sure they were aware of what they were signing up for. Teemant (2010) found the following:

> ESL students' general testing preferences were first a preference for oral tests because that was what they were accustomed to in their home country. Second, students felt that their format preferences would depend on the subject matter of the course. In physics or

math, one student preferred multiple-choice tests because "you can take out [answers] that don't make sense,...so that helps you out." Another student felt that it is better to be tested "chapter by chapter" in a math class. For another student, multiple-choice formats were more appropriate for "technical" material such as chemistry elements, and essays would be more suitable for "theory" courses such as political science.

When asked about their specific test format preferences (e.g., true/false, multiple-choice, etc.), ESL students most often preferred a multiple-choice test format. Multiple-choice tests were the easiest for them to do well on. Essay exams and short-answer/completion test formats were considered the most difficult. When asked if these format preferences would be the same in both their native and second languages, these students responded either that all formats would be easier in their native languages or their preferences would be the same for both languages. (p. 9)

This would suggest that speaking tests in general could have been easier for many students than the writing tests – and was probably true at first when the speaking tests were more generic, but became less so as the tests evolved to include more challenging questions as well as more class-specific ones (see Appendices C and D for sample level tests).

One issue that was not anticipated was that an instructor who was with the program for the first five years was encouraging students to request that they be placed in her classes on the level tests and then – when the department moved to only requiring scores/levels instead of the actual tests – was discovered to be assigning students based on personal preferences rather than their scores on the level tests per se. It became so serious a concern that the department chair at the time requested and reviewed that entire instructor's accumulated level tests for all classes. The instructor in question left soon after, but the negative feelings their actions caused were almost definitely a contributing factor to the end of the level tests in 2014 as were the other issues mentioned in this section.

DISCUSSION

The process of creating and administering level tests for 300 or so students a term was a definite challenge. The lack of prerequisites for

departmental courses was one of the biggest issues since students often did not take courses in the order they were designed to be taken. This resulted in students who had not taken the Writing 1 classes trying to take Writing 2, 3, or 4 classes. Such students would have been assigned to lower-level classes and would have been unhappy with this for a number of reasons. This also required a "review" period in Week 2 of these courses to ensure that all the students knew what was covered in the Writing 1 class before beginning the Writing 2 material, for example. It got to be such a serious problem that students were being actively discouraged from taking Writing 4 (Essays and Research Papers) since they were almost certain to get a low grade or even fail if they had ignored the faculty's recommendations to take Writing 1 and 2 first. The speaking classes covered presentations, interviews, and discussion/debate, so this could have been less of an issue since the skills learned in the presentation classes might not have impacted their ability to handle job interviews, for example. However, the discussion class did require presentations so it was not true in that case.

Ultimately, as stated previously, it was a combination of student unhappiness and objections from the other faculty that resulted in the end of the official level testing conducted during Week 1 for the ELL Department's students. However, most of the faculty continued to use other methods to informally test students during the first few classes by having them introduce themselves or their partner(s) in speaking classes and writing some kind of homework assignment/needs assessment in the writing classes. Thus, the lessons learned from the level-testing experiences were not entirely wasted or forgotten.

CONCLUSIONS

In the end, despite all the issues and benefits discussed in this paper, the level-testing efforts by the ELL Department's faculty resulted in the betterment of the overall program and clear improvements in the skills and abilities of the students who took the writing and speaking classes between 2006 and 2014. The advantages outweighed the disadvantages, especially in later years as the tests evolved based on feedback from students and the other faculty to the point where − with a few exceptions − they allowed students with similar skills and levels to learn and thrive

in classes that were at the appropriate level(s) for them. It is unfortunate that the few students who were unhappy with the classes they ended up in and the negative perceptions of some of the faculty within the department lead to the level tests being gotten rid of entirely, but the lessons learned from administering and creating the individual classes from a criterion-referenced perspective gave the faculty a deeper understanding of the intentions and skills focuses of these classes as they were being developed and taught. This impacted the ELL Department's overall program in positive ways during its first few years and continues to do so up until the present time for the faculty who remain from 2006. The following from Frain (2009) sums up well how the level tests could have provided further benefits for students in the ELL program's classes:

> It is a custom in Korea to be constantly testing, as Koreans seem enamored with numerical rank and progress, a cultural trait that relies heavily on NR tests that force students to compete against each other so they can be better segregated. By adopting an alternative testing method, this endless cycle of testing can finally abate as students realize that they can measure their own strengths and progress in English. Hopefully, this student perception may lead towards self-assessment and reduce this demoralizing and unnecessary cycle of testing. (pp. 56–57)

THE AUTHOR

Tory S. Thorkelson (BA, BEd, MEd in TESL/TEFL, PhD in language studies / curriculum development) is a proud Canadian who has been an active KOTESOL member since 1998 and has presented at many local and international conferences. He is a past president of KOTESOL and the KOTESOL Seoul Chapter, and is an active KOTESOL Teacher Training member. Dr. Thorkelson is an associate professor for Hanyang University's English Language and Literature Program. He has co-authored research studies and textbooks, including the university-level textbook *World Class English*, with a team of fellow KOTESOL members. Currently, he is a regular contributor to *EFL Magazine*. Email: thorkor@hotmail.com

REFERENCES

ACTFL. (n.d.). *ACTFL proficiency guidelines 2012: English*. https://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012/english

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

British Council. (n.d.-a). IELTS TASK 1 Writing band descriptors (public version). https://takeielts.britishcouncil.org/sites/default/files/ielts_task_1_writing_band_descriptors.pdf

British Council. (n.d.-b). IELTS TASK 2 Writing band descriptors (public version). https://takeielts.britishcouncil.org/sites/default/files/ielts_task_2_writing_band_descriptors.pdf

Cambridge Assessment English. (n.d.). *International language standards*. https://www.cambridgeenglish.org/ exams-and-tests/cefr/

EF. (n.d.). *English test score converter*. English First. https://www.ef.com/wwen/english-tests/test-comparison/score-converter/

Frain, T. J. (2009). *A comparative study of Korean university students before and after a criterion referenced test* [Unpublished master's thesis]. University of Southern Queensland.

GIST. (n.d.). *Admissions overview*. Gwangju Institute of Science and Technology. https://www.gist.ac.kr/iadm/html/sub02/0201.html

Gronlund, N. E., & Linn, R. L. (1990). *Measurement and evaluation in teaching* (6th ed.). Macmillan.

Hanyang University. (n.d.). *Graduation: Foreign language examination*. http://gsis.hanyang.ac.kr/eng/sub/sub03_0502.php

Ingram, D. E., & Wylie E. (1992). Communicative competence and language proficiency testing. *Asian Studies Review*, *15*(3), 30−37. https://doi.org/10.1080/10357823.1992.9755377

IELTS. (n.d.). Speaking: Band descriptors (public version). https://www.ielts.org/-/media/pdfs/speaking-band-descriptors.ashx?la=en

In'nami, Y., & Koizumi, R. (2017). Using EIKEN, TOEFL, and TOEIC to award EFL course credits in Japanese universities. *Language Assessment Quarterly*, *14*(3), 274−293. https://doi.org/10.1080/15434303.2016.1262375

List of language proficiency tests. (2019, November 17). In *Wikipedia*. https://en.wikipedia.org/wiki/List_of_language_proficiency_tests

Prodromou, L., & Clandfield, L. (2007). *Dealing with difficulties: Solutions, strategies, and suggestions for successful teaching*. Delta.

Sewell, H. D. (2005). *The TOEIC: Reliability and validity within the Korean context* [Unpublished manuscript]. University of Birmingham. https://www.birmingham.ac.uk/documents/college-artslaw/cels/essays/testing/testing/sewelltesting.pdf

Study in Korea. (2016). TOPIK: Test of proficiency in Korean. http://www.

studyinkorea.go.kr/en/sub/college_info/room/allnew_topik.do

Teemant, A. (2010). ESL student perspectives on university classroom testing practices. *Journal of the Scholarship of Teaching and Learning*, *10*(3), 89−105.

Thorkelson, T. S. (2019). The journey from specialist to generalist: Creating a skills-based program for undergraduates at a private university in Seoul. *Korea TESOL Journal*, *15*(1), 137−166.

Vu, T. P. A., Nguyen, Q. T., & Nguyen, T. T. (2016, October). *Identifying the criteria for selecting learning outcomes assessment tools: The case of TOEIC and IELTS* [Paper presentation]. 4th British Council New Directions in English Language Assessment, Hanoi, Vietnam.

Wilson, K. M. (2001). *Overestimation of LPI ratings for native-Korean speakers in the TOEIC testing context: Search for explanation* (ETS Research Report RR-01-15).

APPENDIX A

Classes Offered: 2006–2014 (PEEC*/ELL Department)

Class Title	Term(s) Taught	Year(s) Taught
Introduction to Acting (S2)	Spring and Fall	2004*–2006, 2018–Present
Presentation Skills (S2)	Spring and Fall	2005*–Present
The Story of English (Global Englishes)	Fall	2005*–2006, (2017–Present)
English Speaking 1 (later Speaking 3): Interviews	Spring	2007–2010, 2011–2017
English Writing 2: Paragraphs	Spring	2007–2011
English Speaking 3: Discussion	Spring	2007–2012
English Writing 1: Sentences	Fall	2007–2011
English Writing 3: Essays	Fall	2007–2011
Writing 1: Sentences & Paragraphs	Fall	2011–2014
Writing 2: Critical Essay & Proposal	Spring	2011–2014
Writing 3: Business Correspondence & Project Essay	Spring	2011–2015
Speaking 2: Presentations & Debate	Fall	2011–2014
Project Essay Capstone Design 1 (New Writing 4)	Spring	2012–2015 (Graduate Paper cancelled in 2015)
Project Essay Capstone Design 2 (New Writing 4)	Fall	2014–Present (Same course as Design 1)

Note. From Thorkelson (2019). ELL course name or designation in brackets.

APPENDIX B

Level Testing Instructions for ELL Faculty

1.   Level tests are on your desk or in your mailbox (I made 20 copies of the Writing 1 level test and asked the TA to open the office before 9 a.m. on Monday).
2.   The writing tests are administered and graded by you (out of 21 points, as I recall), and please fill in the score next to the name on your class sheets OR have them include name, student number, and email address on the test itself if they are not listed so that the office has a way to contact them. I usually have the students fill in their information on the attendance sheet as well.
3.   We will have four assistants to help us with the Speaking 2 level test and two teachers in the offices and two in other classrooms. On Day 2, simply hold a shorter class and test anyone who was absent. Again, we want to get all level testing done by the end of the first week if at all possible. Preferences given to students in the English Language and Literature Department, but everyone must level-test to be in the class and anyone who does not will be eliminated from the class lists on the last day of the add/drop period presumably.
4.   Please give me your lists of student names/student numbers by Thursday at the latest, as I need to post them by Friday (main office and on our office doors as well as the classroom doors). Each teacher will get copies of the completed lists as well, so you can send the students to the appropriate class/level/teacher.
5.   The students are responsible for re-registering in the right class, and make that clear to them often. We want to minimize the number of lost souls as much as possible.

APPENDIX C

Writing Level Tests

Writing 1 Level Test: Sentences
Directions: Write a few sentences or a paragraph about **ONE** of the topics below. You will have 30 minutes to write. Good luck!

**Family and Friends. My hobbies.   My favorite club. My home.
University Tuition.   A food I hate. Smart Campus.   My major.**

_____
_____
_____ [15 lines originally provided.]

Name: _____   Student Number: _____
Cell: _____   Score: _____/21

Writing 2 Level Test: Critical Essay & Proposal

Topics Sheet
Directions: Write an essay about one of the topics from the list. You will have up to 60 minutes to write. Good luck!

1.   Some people think that a "gap year" between high school and university is a good idea; others think it is a waste of time and money. What do you think and why?
2.   Some think technology is the same as progress; others disagree. What do you think?
3.   Some people argue that smart technology is making us stupid; others disagree. What do you think?
4.   Happy people have been proven to be more likely to be satisfied and successful in their lives. What do you think?
5.   Korea's government is encouraging companies to hire more high school graduates. How do you think this will affect the employment of university grads?
6.   The humanities are dying, but many argue that they are necessary to create well-rounded citizens. What do you think?

Writing 3 Level Test: Critical Essay & Proposal

Topic Number: _____
Directions: Write an essay about one of the topics from the list. You will have up to 60 minutes to write. Good luck!

_____
_____
_____ [39 lines originally provided.]

Writing 4 Level Test: Capstone:

Directions: Write an essay about one of the topics below. You will have 60 minutes to write. Good luck!

Topic 1: Some feel that pessimists are more realistic than optimists. Others say that optimists are more positive and reap more rewards from life as a result. What do you think and why?

Topic 2: Today, smart technology is everywhere, but some people think that smart technology is actually making us stupid. What do you think?

Topic 3: Some people say that the humanities are a useless area of study. Others say that everyone should study some humanities to better understand society. What do you think and why?

_____
_____
_____ [30 lines originally provided.]

Name: _____   Student Number: _____
Cell: _____   Score: _____/21

APPENDIX D

Sample Speaking Tests

Speaking 1 (Discussion & Debate) Level Test

Teacher's Sheet
- Interviews should last about five minutes per student.
- Scoring is out of 21 points and rated for Fluency, Accuracy, Coherence and Cohesion, Grammar/vocabulary and pronunciation with advanced students able to make a well-structured argument for/against the topic as well as demonstrating a mastery of the language.

Warm-up (about 1 minute)
- Greeting.
- How was your summer vacation?
- Why are you interested in this class?
- What other classes do you plan to take this term?
- What do you know about this class? etc.

Topics (2–3 minutes)
- Tell me about your family/hobbies.
- Explain why your major is important to your life.
- Tell me why technology is good/bad for us.
- Explain why university clubs are important.
- Tell me why women should (or should not) join the Korean military.
- Tell me why tourists should come to Korea/your hometown.
- Tell me why education is important for life.
- Tell me why English is (is not) important for you.
- Tell me why stewardesses should (or should not) be able to wear pants on duty.
- Tell me why kimchee is such a healthy food.
- Tell me why the two Koreas should unify.
- Tell me why you think Hanyang University should be number 1 in Korea/the world.
- Tell me why living in a city is better than living in the country.
- Tell me why the FTA is a good/bad for Korea.
- Tell me about your favorite book/band/movie and why I should buy/see it.

- Tell me why you should be the next president of Korea.
- Tell me why computers/cell phones are essential to our lives.
- Tell me why debate/discussion/critical thinking are essential life skills.
- Other ideas......

Closing (about 1 minute)
- Level test results will be posted on our office doors by the end of August at the latest. Please check there or with the ELL main office to find out which class you will be in. There is only one Advanced Class.
- Remember you must re-register through your department on September 4.
- Thanks for coming and goodbye!

Speaking 2 (Presentations) Level Test

Teacher's Sheet
- Interviews should last about five minutes per student.
- Scoring is out of 21 points and rated for Fluency, Accuracy, Coherence, and Cohesion, with advanced students able to make a well-structured argument for/against the topic as well as demonstrating a mastery of the language.

Warm-up (about 1 minute):
- Greeting.
- How was your summer vacation?
- Why are you interested in this class?
- What other classes do you plan to take this term?
- What do you know about this class? etc.

Topics (2−3 minutes)
- Tell me about your family/hobbies.
- Explain how to make your favorite food.
- Tell me how to get from your home to this university.
- Explain why university clubs are important.
- Tell me why women should (or should not) join the Korean military.
- Tell me why tourists should come to Korea/your hometown.
- Tell me why education is important for life.
- Tell me why English is (is not) important for you.
- Tell me why you chose your major.

- Tell me why kimchee is such a healthy food.
- Tell me why the two Koreas should unify.
- Tell me why you like Hanyang University.
- Tell me why living in a city is better than living in the country.
- Tell me why the FTA is a good/bad idea for Korea.
- Tell me about your favorite book/band/movie.
- Tell me why you should be the next president of Korea.
- Tell me why computers/cell phones are essential to our lives.
- Other ideas......

Closing (about 1 minute)
- Level test results will be posted on our office doors by the end of August at the latest. Please check there or with the ELL main office to find out which class you will be in. There is only one Advanced Class (Tory's).
- Remember you must re-register through your department on September 4.
- Thanks for coming and goodbye!

Interview Class Level Test

Teacher's Sheet
- Interviews should last about five minutes per student.
- Scoring is out of 21 and rated for Fluency, Accuracy, Coherence and Cohesion with advanced students able to make a well structured argument for/against the topic as well as demonstrating a mastery of the language.

Warm-up (about 1 minute)
- Greeting.
- How was your summer vacation?
- Why are you interested in this class?
- What other classes do you plan to take this term?
- What do you know about this class?, etc.

Topics (2-3 minutes)
- Tell me about your family/hobbies.
- Explain how you plan to get your dream job.
- Tell me what you expect to learn in this class.
- Explain why interviewing is an important skill to master.
- Tell me who your favorite talk show host is and why.

- Tell me which you prefer and why: being a high-paid lawyer or a low-paid teacher.
- Tell me why education is important for life.
- Tell me why English is (is not) important for you.
- Tell me why you chose your major.
- Tell me why I should hire you to be my teaching assistant for this class.
- Tell me why the five-day workweek is a good thing for Korea.
- Tell me why you are a unique and talented person.
- Tell me what your strengths and weaknesses are.
- Briefly introduce yourself to me.
- Tell me why internships are a good way to get a job.
- Tell me why you should be the next president of Korea.
- Tell me why computers/cell phones are essential for the modern workplace.
- Tell me your plan for getting a job in the current recession.
- Do you think job sharing is a good idea? Why or why not?
- What is a reasonable monthly salary? Why?
- Should CEOs work their way up through the company or be talented people hired from outside? Why?
- Tell me why job sharing/internships/volunteer service for students are good ideas.

Closing (about 1 minute)
- Level-test results will be posted on our office doors by the end of this week at the latest. Please check there or with the ELL main office to find out which class you will be in. There is only one Advanced Class (Tory's).
- Remember you must re-register through your department.
- Thanks for coming and good bye!

Level Test Information Sheet

Speaking ___: _____

Name	Student No.	Major/Dept.	Email	Cell Phone	Score

[20 rows originally provided.]

Semantic-Based DDL for EFL Learners in the Classroom Based on Analysis of Adjective Use Between British and Japanese Students

Kunihiko Miura

The University of Shimane, Hamada, Japan

The aim of this study is to investigate adjective use by British elementary school and lower-secondary school students and Japanese lower-secondary school students for pedagogical use to adopt the data-driven learning (DDL) approach in the classroom. It is based on three written corpora of British elementary and secondary school students and Japanese lower-secondary school students. The two British students' corpora collected the written language from the BBC 500 Words competition, while the Japanese students' corpus collected the written language as a writing task in English classes. This research clarifies adjective use through different types of analyses: word list analysis, correspondence analysis, n-gram analysis, and semantic analysis. The analyses could identify particular usages of adjectives among these corpora. Moreover, this study endeavors to adopt the results for the DDL approach in the classroom to show a variety of activities concerning vocabulary use and collocations, and to learn from British students' expressions as a target language for Japanese EFL learners. Furthermore, this study also provides impressive suggestions for English teachers and EFL learners in other countries to experience a new type of DDL in the classroom.

Keywords: learner corpus, DDL, authentic materials

INTRODUCTION

With the growth of corpus linguistics, the number of reports regarding corpus-based teaching and learning has increased both overseas and in Japan. However, most of these reports, especially data-driven learning (DDL), are mainly in academic fields for university students.

The DDL approach in elementary and secondary school-level settings have not been thoroughly explored. In particular, most DDL approaches at the university level adopt mega corpora, such as the British National Corpus (BNC) or the Contemporary American English Corpus (COCA), as authentic materials or a target language. However, the levels of English in these corpora seem to be very difficult for EFL (English as a foreign language) learners to understand the content of the concordance lines. This study thus adopts written English of British elementary school and lower-secondary school students as a target language for Japanese EFL learners using the DDL approach in the classroom. By comparing British students' language use focused on adjectives with that of Japanese students, the DDL materials would be very effective in helping Japanese EFL learners fill in the gap between their language use and the target language, which vary in adjective use and collocations of "adjective + noun." Moreover, this study also offers numerous suggestions for similarly situated teachers and EFL learners in other countries.

LITERATURE REVIEW

In an increasingly globalized world, wherein English has become the world's lingua franca, there is a necessity for students to learn English. In Korea, instruction in English began in elementary schools in 1997. In Japan, English became compulsory for third- and fourth-grade students in 2002. In 2020, English will be implemented as an official subject in the fifth- and six-grade elementary curriculum. Teaching English as a foreign language (TESOL) in secondary schools has changed from a method of grammar-translation to a communicative language teaching (CLT) method. Task-based learning (TBL) has adopted pair work and group work as part of its methodology. In the wider context of English language teaching (ELT) and English as a foreign language (EFL) learning, Japanese English language education seems to be more similar to Korean English language education than to that of other Asian countries. Writing about early English language education in elementary schools, Park (1999) has noted that "English was adopted as one of the regular subjects in elementary education in Korea as of 1997, ahead of some of our neighboring Asian countries" (p. 6). Because of this early adoption of ELT, sharing Korea's experience may prove very valuable

for Japanese educators and educators in other Asian countries. On this point, Craig (1998) has demonstrated the value of Korea's experience with elementary school English language instruction and the importance of exchanging ideas and resources with other countries. Park and Kim (2014) have introduced the English Program in Korea (EPIK), which "came into operation in 1995 with the missions to improve the English-speaking abilities of students and teachers, to develop cultural exchange, and to reform ELT methodologies in Korea" (p. 51). Dailey (2009) has noted that the aim of adopting CLT and TBL into the curriculum is to enable Koreans to communicate in English but that TBL is challenging for Korean students because it is a Western method of teaching language. From this point of view, some problematic points of English instruction in Korea seem to be large class sizes, teacher-centered approaches such as grammar-translation, and the difficulty of adopting CLT and TBL in pair work or group work in the classroom. Given English education in Korea.

With the growth of corpus linguistics in educational fields, some researchers have attempted to adopt a corpus as a pedagogical experiment. In particular, previous studies overseas and in Japan have reported data-driven learning (DDL) in the classroom from a teacher-centered approach to a learner-centered approach. Concerning the effectiveness of DDL, Nesselhauf (2005) surveyed how German advanced learners of English use collocations and why they make mistakes. Kim and Chun (2008) examined the increasing vocabulary awareness/usability and autonomy of Korean college students. Rahimi and Momeni (2012) conducted an empirical study of teaching collocations for Iranian pre-university students, while Guan (2013) mentioned the effectiveness of DDL in a Chinese EFL class. Furthermore, Vyatkina (2017) examined German lexico-grammatical constructions (verb-preposition collocations) for North American college students at an intermediate level.

With the implementation of corpus-based approaches and DDL in Korean education fields, there are studies and reports that offer interesting insight into EFL in Japan and other Asian countries. Luo and Zhou (2017) provide us with an overview of previous studies of DDL on L2 writing from 2010 to 2016. Their work shows three primary streams of DDL studies on L2 writing: (a) the effects of DDL on writing to compare writing with corpora and dictionaries as resources, (b) the effect of using different types of corpora on writing, using general

corpora and specialized corpora to compare the native speakers' corpora with learner corpora, and (c) the effectiveness of using Google Search as a reference tool, which is more flexible and user-friendly than most concordancers. The study points out that most previous studies of DDL on writing use general corpora, such as the BNC and the COCA, and suggests adopting specialized corpora to prevent learners from spending too much time on different corpora. Using the corpus technique helps teachers to develop their students' English proficiency as a real language resource and facilitate the adoption of DDL in the classroom by providing teacher-training programs or in-service teacher training.

Luo and Zhou (2017) also discuss how previous DDL studies focused on advanced or upper-intermediate learners, with few studies focusing on lower-level learners. These previous studies report that lower-level learners benefit from DDL in terms of their positive attitude, but it is difficult for them to read concordance lines when using online corpora as general corpora. The report recommends the development of DDL for young learners as a means for evaluating lower-level learners. In addition, the study also points out the duration of previous studies on DDL. There were 12 studies with a duration of more than five weeks; the others were all short-term studies. Furthermore, this study shows an interesting tendency of DDL implementation. Most of the previous studies adopted the direct DDL approach, even though the indirect DDL approach might be more effective for lower-level learners to read, given the simplified concordance lines. Therefore, the corpus technique and a knowledge of developing specialized, small corpora are necessary for elementary- and secondary-school teachers to implement DDL at an appropriate level.

Yoon (2005) points out that there is little implementation of corpus-based studies in EFL classrooms. The study suggests that teachers use corpora to develop teaching materials or syllabus design; teachers could also use corpora for their studies. The study points out that the corpus-based approach would be valuable for EFL teachers, allowing them to access authentic data, which they can adopt to make their own teaching materials and give learners an opportunity to be exposed to authentic languages. On this point, corpus-based studies would solve the problems of English textbooks that seem to use language that is artificial and different from authentic expression.

Lee (2011) suggests integrating corpora into Korean secondary schools in four forms: pedagogically used specialized corpora, a Korean secondary learner corpus, an online database of corpus resources, and a

corpus workshop for teacher training. As one of the major problems involved in adopting the corpora into the classroom, the study points out that few teachers and learners have the required knowledge to use corpora. In order to solve the problem, it suggests creating local and small corpora such as textbook corpora, graded-reader corpora, and learner corpora. Lee's work also mentions the merit of adopting learner corpora into the classroom to compare learners' language with native speakers' language. From this, learners could increase their ability to speak like native speakers when they previously could not use real, authentic language. From this point of view, developing a native speaker corpus as a target language would be quite valuable not only for learning language use but also for acquainting students with different cultures of language use.

Kim (2019) points out that the indirect approach of using corpora at the elementary school level is less traditional than the teacher-centered direct approach. The study attempts to implement the DDL approach by considering the primary learners' level in three groups, including 18 six-grade students aged 12 years, and each English teacher in the Korean elementary schools. The study focuses on three prepositions, *in*, *at*, and *into*, and reports the effectiveness of DDL at primary school learners' level.

Blappert (1997) focuses on the usage of *going to*, *will*, and *plan to*, as these future planning phrases are considered difficult for many Korean students. The study attempted to use the COBUILD corpus and implement DDL to adopt "identify–classify–generalize" (p. 12) for a lower-intermediate conversation class at a university in Seoul, South Korea. The average scores of eight students increased from 60% at the pre-test stage to 87.5% at the post-test stage, demonstrating the effectiveness of DDL. This shows the benefits for low-level students in creating an awareness of the future planning usage through exploring authentic language.

Lee (2006) examines the usage of amplifier "-*ly* adverb" collocations to compare the writing of Korean university students taking two undergraduate-level English courses with that of English native speakers (a sub-corpus of the Louvain Corpus of Native English Essays [LOCNESS]). The study shows that Korean students' usage of "-*ly* adverb" collocations with adjectives is very limited when compared to that of English native speakers. Hong (2010) focuses on the usage of determiners for thirty advanced Korean learners (25 female and 5 male

students in the department of English language and literature at a university in Korea) and shows how DDL is effective in raising learners' consciousness of how language is used.

Carlstrom (2014) points out the difficulty of implementing the DDL in the classroom owing to a lack of resources, training teachers, and DDL materials, and the difficulty in training learners to use software. The study implemented the DDL approach for university students in the spring semester at Gachon University and suggested that paper-based DDL for lower-level learners consider the following four points: simplicity, focus, intuitiveness, and interactivity. The study showed the worksheet's effectiveness regarding DDL and the improvement in language production through the pre- and post-test.

Kim and Chun (2008) investigated the effectiveness of the DDL for 48 college students (37 female, 11 male), registered in a major course offered by the English education department at a university in Korea. The course was designed to enhance vocabulary awareness and usability, and to foster autonomy in learners by deepening their lexical knowledge and suggesting the benefits of adopting a small corpus. The corpus would develop learners' knowledge regarding collocations, raise their grammar consciousness, and foster learners' language learning autonomy. The study shows that the DDL approach is significantly effective in fostering learner awareness of vocabulary and grammar, and in enhancing learner confidence in learning language.

In Japan, Chujo et al. (2006) attempted to use a Japanese–English parallel corpus for freshmen in three Japanese universities. In addition, Chujo et al. (2012) conducted a three-year comparative case study in a beginner-level EFL class using computer-based and paper-based DDL. Nishigaki et al. (2015) implemented traditional and DDL instruction at two junior high schools. However, few studies have focused on adjective–noun collocations and their use with the DDL approach in the classroom.

Regarding research on the acquisition of adjective-noun collocations for children and young learners, the following studies are quite interesting and provide us with valuable results. In experiments with 36-month-olds and 24-month-olds, Mintz and Gleitman (2002) suggested that young word learners require access to the taxonomy of the object type so that the relevant property can be identified for novel adjectives. Nicoladis and Rhemtulla (2012) verified that children recognize abstract syntactic frames from at least two years of age and showed that three- and four-year-old children become more sensitive to the

semantic/syntactic roles of words relative to word order. Klibanoff and Waxman (2000) examined the acquisition of novel adjectives in English-speaking preschool-aged children three to four years old and focused on two aspects of adjectival interpretation. Also, Hall et al. (1993) conducted interesting experiments with children (two- and four-year-olds) on learning either a count noun or an adjective, and examined the effect of object-kind familiarity. By conducting studies focused on children and young learners, some research has zoned in on adjective use in other languages, such as Dutch adjectives *groot* ("big") and *klein* ("small") by two- to seven-year-old children (Tribushinina, 2013). For French-English preschool bilingual children, they suggest that cross-linguistic transfer might be better understood as an epiphenomenon of speech production. While there are few studies concerning adjective-noun combinations, Kochmar and Briscoe (2015) attempted error correction in content words in learner writing and reported that learners confuse the use of words that are similar in meaning or spelling, overuse words with general meaning, or select words based on their L1. Through overviewing previous studies, it is clear that very few DDL studies have focused on adjective-noun collocation, especially on a semantic analysis of adjectives.

There are several advantages of DDL for pedagogical purposes. Gilquin and Granger (2010) mentioned three advantages of DDL: (a) brings authenticity into the classroom for vocabulary expansion or heightened awareness of language patterns, (b) aids learners' need to correct their own interlanguage features by comparing native/expert writers or by consulting a learner corpus, and (c) provides an element of discovery, which arguably makes learning more motivating and fun. Concerning authenticity, Wills (2003) mentioned that two types of corpora, "pedagogic corpus" and "local learner corpus" (Sedlhofer, 2002), as being helpful. Moreover, regarding the goals of DDL, Gliquin and Granger (2010) mentioned that it is precisely one of the goals of DDL to develop a more autonomous learning style. Based on these previous studies regarding acquisition of adjective and adjective-noun collocations and given the effectiveness of DDL and its advantages and goals, this research provides quite interesting results for five- to nine-year-old British students' (5−9 BSs) and 10- to 13-year-old British students' (10−13 BSs) language use and for 13- to 15-year-old Japanese lower-secondary students (13−15 JSs) concerning adjective and adjective-noun collocations. Furthermore, this study, based on semantic

analysis of adjective and adjective-noun collocations, suggests several different types of DDL approaches for the classroom and provides several examples of paper-based DDL materials for classroom use.

RESEARCH DESIGN

This study proposes two elements for applying the DDL approach to the classroom. One is to clarify the language use of British students, comparing whether or not the characteristic use of adjectives can be identified in the 5−9 BSs and the 10−13 BSs. The other is to investigate the adjective use of British students as compared with Japanese students. If there were differences in adjective use, these differences in language use would become effective authentic DDL material for Japanese lower-secondary school students. Furthermore, conducting a semantic analysis of adjective use by British students will make it possible to provide an opportunity for Japanese EFL learners to learn various types of adjectives that are unfamiliar for them via the DDL approach in the classroom.

Research Questions

In this study, the following three research questions were investigated.

> RQ1. Can any quantitative differences in the adjective use of students be observed among 5−9 BSs, 10−13 BSs, and Japanese students?
> RQ2. Can any qualitative differences in the adjective use of students be observed among 5−9 BSs, 10−13 BSs, and Japanese students?
> RQ3. Can the results of qualitative analyses be used as authentic materials with a variety of DDL approaches in the classroom?

Data Collection and Building Corpora

This study collected three written corpora: that of 5−9 BSs, 10−13 BSs, and 13−15 JSs. The British student corpora were collected from the written data of the BBC's 500 Words story-writing competition in the UK, which are original stories of British students chosen as the best 50

stories from 2014 to 2017. Each text was uniformly collected in relation to the number of words. The Japanese student corpus was collected from written data given as a written task to students aged 13–15 at a Japanese national secondary school The written task was to write an email to a pen pal in another country each year from 2006 to 2008. These three written samples were annotated with the CLAWS 7 tagger developed by Lancaster University in the UK. The annotation of the parts of speech (POSs) was considered an effective method for extracting adjectives and adjective collocations to examine the language use in detail.

Before focusing on adjective use in these three corpora, overviewing each corpus is necessary for quantitative analysis. Table 1 shows the size of each corpus.

TABLE 1. Corpus Size of the BBC_5-9, BBC_10-13, and LC_13-15

Corpus	Tokens	Types	TTR
BBC_5-9	46136	7249	0.15
BBC_10-13	46912	8318	0.17
LC_13-15	33308	2347	0.07

Note. TTR = type-to-token ratio.

The type-to-token ratio (TTR) for each corpus is as follows: BBC_5-9 (0.15), BBC_10-13 (0.17), and LC_13-15 (0.07). This result shows that the TTR (0.17) of the 10–13 BSs is the highest (0.17), second is the TTR (0.15) of the 5–9 BSs, and the TTR (0.07) of the 13–15 JSs is the lowest. It is clear that the language used by the 10–13 BSs is much more varied than that of the 5–9 BSs and that the variety of their language use is much greater than that of the 13–15 JSs. This result suggests that the language used by British students can be used as a model language for Japanese EFL learners as their target language because the difficulty of the language used by the British students is not significantly higher than that of the Japanese students.

METHOD

A word list of each corpus was created for quantitative analysis. It provided an overview of each corpus size and compared the 20 most frequent adjectives in each word list to grasp the characteristics of use.

Correspondence analysis also clarified characteristic adjectives in each corpus. Moreover, n-gram analysis as qualitative analysis showed the 20 most frequent collocation patterns of "adjective + noun" and classified some of the categories from the "adjective + noun" collocations, which would make it possible to discover what students described. Furthermore, a semantic analysis would reveal what conception students intended and elucidate the differences in comparing these corpora. By adopting the results, this study demonstrates a number of DDL approaches for learning adjective use with regards to vocabulary, collocations, and effective expressions for Japanese EFL learners, based on n-gram analysis and semantic analysis.

RESULTS AND DISCUSSION

This study adopts the following various analyses of the data in order to investigate each of the three research questions.

Research Question 1

Research Question 1 is "Can any quantitative differences in the adjective use of students be observed among the 5−9 BSs, the 10−13 BSs, and the Japanese students?"

By counting the number of adjectives in each corpus, Table 2 shows the number of general adjectives and adjective-types compared to the number of words in each corpus.

TABLE 2. Total Number of Adjectives and Number of Adjective-Types in Each Corpus.

Corpus	Total Number of Adjectives	Number of Adjective-Types
BBC_5-9	3227 (6.99%)	1039 (2.25%)
BBC_10-13	3261 (6.95%)	1290 (2.74%)
LC_13-15	1877 (5.64%)	243 (0.07%)

Note. The percentages indicate the number of adjectives/adjective-types to total words in each corpus.

Table 2 shows quite interesting results concerning the total instances of adjective use in each corpus. It can be seen that the 5−9 BSs use adjectives the most among these three corpora. In particular, it is notable that the instances of adjective use among the 10−13 BSs is much higher than that among the 5−9 BSs. The instances of adjective type use among the 13−15 JSs is the least among the corpora. The results show that older British students use a much greater variety of adjectives than younger British students do and that the adjective use of Japanese EFL learners is very limited in comparison with younger native English speakers.

Table 3 shows the total number of comparative and superlative adjectives and the number of adjective-types in each corpus.

TABLE 3. Total Number of Comparative Adjectives and Superlative Adjectives and Number of These Types

Corpus	Total No. of Comparative Adjectives	No. of Comparative Adjective-Types
BBC_5-9	70(0.15%)	35(0.07%)
BBC_10-13	62(0.13%)	34(0.07%)
LC_13-15	54(0.16%)	11(0.03%)
Corpus	**Total No. of Superlative Adjectives**	**No. of Superlative Adjective-Types**
BBC_5-9	55(0.12%)	27(0.05%)
BBC_10-13	63(0.13%)	26(0.05%)
LC_13-15	62(0.19%)	10(0.03%)

Focusing on comparative adjectives, Table 3 reveals that the number of comparative adjectives used among the 13−15 JSs is greater than that of either British student group. This overuse of comparative adjectives by Japanese EFL learners seems to be influenced by the second-year English textbook for lower-secondary schools in Japan, which contains a target sentence incorporating comparative adjective use. Concerning the number of comparative adjective-types used by British students, the types are almost the same for the 5−9 BSs and the 10−13 BSs. Regarding superlative adjective use, the total number of superlatives used by the 13−15 JSs is more than that used by both British ages. However, the number of superlative adjective-types of Japanese EFL learners is less than that of the British students. The results show that Japanese EFL learners overuse a limited number of superlative adjectives compared to native English speakers (NESs).

Research Question 2

Research Question 2 is "Can any qualitative differences in the adjective use of students be observed among the 5–9 BSs, the 10–13 BSs, and the Japanese students?"

Regarding qualitative analyses, this study adopted the following analytical procedures: (a) word list analysis of the 10 most frequent adjectives in each corpus, (b) correspondence analysis to identify distinct adjectives in each corpus, (c) bigram analysis of these distinct adjectives in each corpus, and (d) semantic analysis to understand a characteristic of adjective use from the viewpoint of semantic analysis. As the first step, to deepen our understanding regarding adjective use in each corpus, an analysis of word lists of adjectives gave a general picture of these three corpora. Table 4 shows the 20 most frequently used adjectives in each corpus.

TABLE 4. Frequency Lists of the 20 Most Frequently Used Adjectives in Three Corpora

	BBC_5-9	Freq.	R.F.	BBC_10-13	Freq.	R.F.	LC_13-15	Freq.	R.F.
1	little	73	0.16	old	54	0.12	old	343	1.03
2	big	70	0.15	new	44	0.09	Japanese	159	0.48
3	old	63	0.14	small	40	0.09	junior	154	0.46
4	good	40	0.09	black	37	0.08	good	103	0.31
5	full	39	0.08	little	30	0.06	nice	71	0.21
6	blue	33	0.07	happy	30	0.06	favorite	53	0.16
7	red	33	0.07	long	28	0.06	very	44	0.13
8	new	33	0.07	beautiful	26	0.06	English	39	0.12
9	great	32	0.07	perfect	24	0.05	best	37	0.11
10	cold	32	0.07	young	22	0.05	fine	35	0.11
11	other	31	0.07	huge	22	0.05	interesting	33	0.10
12	white	28	0.06	white	22	0.05	dear	32	0.10
13	black	27	0.06	dark	22	0.05	beautiful	27	0.08
14	beautiful	26	0.06	other	22	0.05	big	23	0.07
15	pink	24	0.05	ready	22	0.05	happy	22	0.07
16	special	22	0.05	blue	20	0.04	exciting	19	0.06
17	green	21	0.05	great	19	0.04	fun	17	0.05
18	bad	21	0.05	bright	19	0.04	difficult	17	0.05
19	small	20	0.04	big	18	0.04	cool	16	0.05
20	happy	20	0.04	red	18	0.04	hard	15	0.05

Note. Freq. = frequency; R.F. = relative frequency.

By comparing the 20 most frequently used adjectives of the BBC_5-9 with the BBC_10-13, it is found that there are 13 adjectives in common: *little*, *big*, *old*, *blue*, *red*, *new*, *great*, *other*, *white*, *black*, *beautiful*, *small*, and *happy*. Distinct adjectives used in only one of the three word lists are as follows: 6 adjectives in the BBC_5-9 (*full*, *cold*, *pink*, *special*, *green*,

bad), 6 adjectives in the BBC_10-13 (*perfect*, *young*, *huge*, *dark*, *ready*, *bright*). Also, Japanese EFL learners use the following nine adjectives only in their corpus (*nice*, *favorite*, *very*, *English*, *best*, *fine*, *interesting*, *dear*, *hard*).

Correspondence analysis is effective in identifying more clearly distinct adjective types in each corpus. Figure 1 shows the result of correspondence analysis of adjectives in each corpus. The correspondence analysis shows the distinct adjective types in each corpus clearly as follows: Nine adjectives in BBC_5-9 (*scared*, *full*, *little*, *quick*, *green*, *red*, *busy*, *cold*, *blue*), 11 adjectives in BBC_10-13 (*golden*, *quiet*, *late*, *young*, *black*, *strange*, *grey*, *huge*, *brown*, *bright*, *worst*), and 10 adjectives in LC_13-15 (*fun*, *nice*, *cool*, *Japanese*, *English*, *junior*, *old*, *favorite*, *best*, *interesting*). The 5–9 BSs used adjectives concerning colors (*green*, *red*, *blue*), their feelings (*scared*, *busy*) and amounts of things (*full*, *little*). The 10–13 BSs mention colors (*golden*, *black*, *grey*, *brown*), the shapes of things (*huge*, *sharp*), and degrees of luminance (*bright*, *dark*). The 13–15 JSs mainly talk about their feelings (*fun*, *nicer*, *cool*, *interesting*) and their favorite things (*favorite*). This correspondence analysis specified distinct adjectives in each corpus.

FIGURE 1. Correspondence Analysis of 100 Most Frequently Used Adjectives in Each Corpus

Next, this study attempted to identify distinct adjective types using semantic-based correspondence analysis, to grasp the trends in adjective use in each corpus from a different angle. Figure 2 shows noticeable results concerning the distribution of semantic tags by Wmatrix analysis of each corpus to clearly show the varied semantic features of adjectives. The aim of the semantic-based correspondence analysis was to gain useful information about the differences in adjective use among the 5–9 BSs, the 10–13 BSs, and the 13–15 JSs. The semantic tag-based procedure contained two steps: (a) extracting all adjectives based on a POS (part of speech) tag using CLAWS 7 from these three corpora that were of three types: general adjectives, general comparative adjectives, and general superlative adjectives and (b) making word frequency lists of adjectives and semantic tags using Wmatrix for semantic-tag based correspondence analysis. By executing this procedure, it was found that the number of types of semantic tags used in relation to the adjectives among these three corpora is 65.

FIGURE 2. Semantic Tag-Based Correspondence Analysis of All Adjectives in Three Corpora
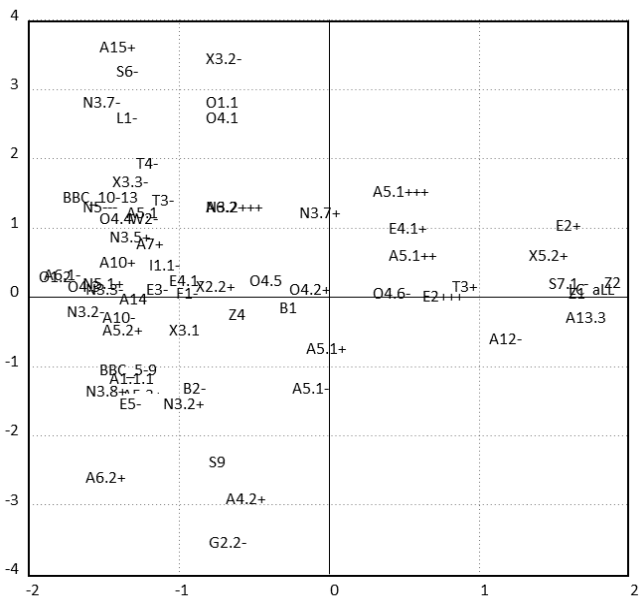
The results in Figure 2 provide invaluable input concerning the semantics of adjectives and clearly show the position of distribution concerning each semantic tag among these three corpora. There were ten distinct semantic tags for the 5−9 BSs as follows: A1.1.1 (general and abstract term: e.g., *busy*), N3.8+ (speed [fast]: e.g., *quick*), E5 (bravery and fear: e.g., *angry*), N3.2+ (size: e.g., *big*), B2- (disease: e.g., *cold*), X3.1 (sensory [taste]: e.g., *sweet*), A5.2+ (evaluation [true/false]: e.g., *true*), A5.3+ (evaluation [accurate]: e.g., *right*), O4.3 (color and color pattern: e.g., *black*, *blue*,...) and A10- (closed [hiding/hidden]: e.g., *secret*).

The ten distinct semantic tags for the 10−13 BSs were the following: N5 (quantities: e.g., *empty*), X3.3 (sensory [touch]: e.g., *quiet*), T4- (time [early/late]: e.g., *late*), O4.4 (shape: e.g., *sharp*), A5.1 (evaluation [bad]: e.g., *worst*), T3 (time [old, new and young: age]: e.g., *new*, *young*), N3.5 (measurement [weight]: e.g., *heavy*), N3.2 (measurement [size]: e.g., *huge*), A6.2- (comparing [unusual]: e.g., *strange*), and L1.1- (dead: e.g., *dead*).

For the 13−15 JSs, there were the following distinct semantic tags: Z1 (personal name: e.g., *English*), S7.1- (no power: e.g., *junior*), A13.3 (degree [boosters]: e.g., very), E2+ (like: e.g., *dear*, *favorite*), A12- (difficult: e.g., *difficult*), T3+ (time [old, new and young: age]: e.g., *old*), and O4.6 (temperature: e.g., *cool*).

As qualitative analysis, this research conducted a bigram analysis to focus on distinct adjectives by correspondence analysis and classify the collocation patterns in each corpus. For the 5−9 BSs, mainly two types of collocations of adjectives concerning *little* and adjectives of colors were observed. The bigrams for *little* for the 5−9 BSs are as follows: *little Grimlick*, *little girl*, *little bit*, *little sister*, *little ghost*, *come little*, *little cry*, *little sisters*, *little dragon*, and *little boy*. The collocation patterns are "little + personal pronoun," "verb + little," "little + verb," and the expression *little bit*. With regard to colors such as *green*, *red*, and *blue*, the following collocation patterns were observed: *green grass*, *red bus*, *flossy red*, *red paint*, *bright red*, *red light*, *blue ribbon*, *clear blue*, *blue monkey*, and *blue eyes*. When mentioning colors, the 5−9 BSs used a variety of nouns after an adjective related to color.

For the 10−13 BSs, the following adjective use regarding colors was observed: *world grey*, *the grey*, *old brown*, *brown eyes*, and *darkest part*. Furthermore, the 10−13 BSs used the following adjective collocations: *all quiet* and *too late*. On the other hand, the 13−15 JSs used typical

collocation patterns regarding *fun, nice*, *favorite*, and *best* as follows: *lot fun*, *it's fun*, *nice day*, *nice city*, *nice places*, *my favorite*, *favorite class*, *favorite singer*, *favorite player*, *favorite subject*, *favorite food*, *favorite sports*, *songs best*, *best friend*, *sport best*, *song best*, and *English best*.

Also, with regard to 13–15 JSs' use of adjectives such as *Japanese* and *English*, a variety of nouns followed these adjectives: *Japanese girl*, *Japanese boy*, *I'm Japanese*, *Japanese spoken*, *speak Japanese*, *Japanese food*, *Japanese drums*, *Japanese used*, *studied Japanese*, *know Japanese*, *spoken Japanese*, *Japanese yesterday*, *Japanese class*, *study Japanese*, and *Japanese country*. Regarding *English*, the following collocation patterns were observed: *speak English*, *English little*, *study English*, *can English*, *studied English*, *English used*, *English song*, *is English*, *English yesterday*, *English night*, *English class*, *because English*, *read English*, *English country*, *can't English*, *English best*, *can English*, and *so English*.

From bigram analysis, it was found that the writing topics affected these collocation patterns; the British students were given free writing topics for their original stories. Therefore, the British students wrote their creative stories and described their original characters, and they could express persons, animals, and places using a wide variety of adjectives. In particular, using a variety of adjectives concerning colors is quite interesting, and these variations of "adjective + noun" patterns would be useful for use in DDL in the classroom for Japanese EFL learners.
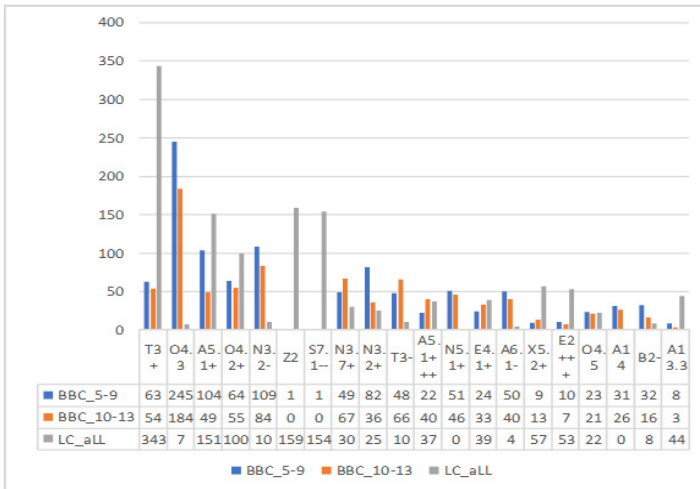
On the other hand, the collocation patterns of the 13–15 JSs seemed to be affected by the Japanese textbooks of English because most of their collocation patterns are part of the main target grammar structures in each grade of lower-secondary school. The study results show that British students used a wide variety of adjectives in their original stories, while the limited number of adjectives used by the Japanese students were overused. This was affected by the main grammar structures appearing in their English textbooks. As authentic materials, adopting the younger British students' written corpora would be suitable for primary Japanese EFL learners in elementary school and lower-secondary school, and for EFL learners in other countries. The information obtained from this study concerning adjective frequency lists, distinct adjective uses by classified correspondence analysis, and semantic-tag correspondence analysis would be suitable for making authentic materials for vocabulary teaching and grammar teaching through collaborative learning using the DDL approach in the classroom.

Research Question 3

Research Question 3 is "Can the results of qualitative analyses determine whether authentic materials (corpus data) are useful for a DDL approach in the classroom?"

This study suggests a new perspective on vocabulary and collocation learning of adjectives: adoption of the results of adjective frequency lists based on semantic-tag analysis. This study classified 65 different types of semantic tags for adjectives among the three corpora. The results make it possible to adopt a data-driven learning (DDL) approach in the classroom. The results of this study related to adjectives in the three corpora can be adopted for classroom use to serve two purposes: (a) the adoption of the adjective frequency list for focusing on high-frequency adjectives used by the 5–9 and 10–13 BSs as model language for Japanese EFL learners to learn varied adjective uses and the "adjective + noun" collocation and (b) the adoption of the results of the semantic-tag analysis of adjectives for recognizing and learning a variety adjectives based on different types of semantic categories used by British students that cannot be observed in the adjective use of Japanese EFL learners. The 20 most frequent semantic tags for adjectives among the three corpora in this study appear in Figure 3.

FIGURE 3. The 20 Most Frequently Used Semantic Tags of Adjectives Among the Three Corpora

	T3+	O4.3	A5.1+	O4.2+	N3.2-	Z2	S7.1--	N3.7+	N3.2+	T3-	A5.1+++	N5.1+	E4.1+	A6.1-	X5.2+	E2++	O4.5	A1.4	B2-	A1.3.3
BBC_5-9	63	245	104	64	109	1	1	49	82	48	22	51	24	50	9	10	23	31	32	8
BBC_10-13	54	184	49	55	84	0	0	67	36	66	40	46	33	40	13	7	21	26	16	3
LC_aLL	343	7	151	100	10	159	154	30	25	10	37	0	39	4	57	53	22	0	8	44

Focusing on notable semantic tag uses among 5−9 BSs and 10−13 BSs such as the BBC_5-9 and the BBC_10-13, it would seem reasonable to adopt them as authentic materials for DDL in the classroom. From Figure 3, the following semantic tags of adjectives are distinct in British students and 9/11 most frequent semantic tag sets are the following: O4.3 (color), N3.2 (size: small), A5.1+ (evaluation: good/bad), N3.2 (size: big), O4.2+ (judgement of appearance), T3+ (time: old), N5.1+ (entire: maximum), A6.1- (comparing: different), and T3- (time: new and young). The following two semantic tags appear in only one of the two British student corpora: N3.7 (short and narrow) in the 5−9 BSs corpus and A5.1+++ (evaluation: best, perfect) in the 10−13 BSs corpus. Table 5 shows the result of the above semantic tag analysis divided into 11 semantic categories and examples of adjectives corresponding to the category of each semantic tag.

TABLE **5. Recommended List of Adjectives Categorized by Semantic Tags in British Students' Written Data for DDL in the Classroom**

Semantic Tag	No. of Types	Adjectives
O4.3	9	black, blue, white, red, bright, yellow, golden, orange
N3.2+	3	little, small, tiny
A5.1	5	good, great, fine, super, magnificent
N3.2-	2	big, large
O4.2+	4	nice, beautiful, amazing, lovely
T3+	1	old
N5.1	2	full, whole
A6.1	1	other
N3.7+	5	long, tall, deep, high, wide
T3-	2	new, young
A5+++	2	best, perfect

In reference to Table 5 above, these adjectives may be used with different types of DDL approaches to collaborative learning: (a) guessing the keyword in context by reading the concordance lines on the right and left of the keyword and (b) discovering the use of distinct adjectives and nouns used in "adjective + noun" collocations by adopting the results of semantic tag-based analysis.

The following examples show how to make authentic materials considering the two different concepts for using the DDL approach in the classroom.

Data-Driven Learning Type 1

Type 1 of the DDL approach is to guess the keyword in a given context by reading the concordance lines on the right and left of the missing keyword.

The aim of this type of task is basic to DDL: discovery learning that focuses on the following six focusing and learning points: vocabulary, POS, collocations, grammar, content, and native English speakers' (NESs) expressions. The following three tasks are examples of DDL Type 1, and Figure 4 is an example of part of a DDL worksheet.

Task A: Guess the keyword in context by reading the concordance lines to the right and left of the missing keyword and share your idea(s) with your group members.

Task B: Read the concordance lines to the left of the missing keyword and report your discoveries concerning the usage of vocabulary, collocations, and grammar.

Task C: Read the concordance lines to the right of the missing keyword and report your discoveries concerning the usage of vocabulary, collocations, and grammar.

The DDL Type 1 approach looks simple at first glance, but it can be said to be an open-ended task for learners to focus on some points concerning vocabulary, POSs, collocations, grammar, content, and NES expressions, and to allow them to discover some features of these language uses. The worksheet example is a concordance of the keyword *red*. Japanese EFL learners may discover some distinct NES language use on the left and right of the keyword *red*.

FIGURE 4. DDL Worksheet Sorted in Alphabetical Order of the Column to the Right of the Missing Keyword

sun began to rise,	and shimmering ;_; red for
enly the seatbelt light flashed	and the pilot announced that
as ever he went as	as a beetroot with anger !
,_, the berries were still	as rubies ._. Just as
they were ,_, the glowing	berries ._. He filled his
1 tree branch with its tiny	berries and prickly leaves ,_,
a blue shell that had	blobs that looked like fireballs
hopped on ._. His rosy	breast warmed her heart and
the street ._. As the	bus arrived ,_, he nimbly
hurried off the large ,_,	bus into the enormous crowd
was the London Eye and	buses stacked at the traffic
over a nice cup of	bush tea ,_, stated Mma
crimson ,_, illuminated and blinking	button ,_, its label read
caught up and noticed a	button marked stoppe ._. Fleazle
Swept Lands and Seas of	by Dulcie Cooper ,_, aged 9
boot of our trusty ,_,	car ._. All the same ,
precisely in front of a	carpet ._. Lucky was led
a lavish building with plush	carpet and gilt furniture ._.
the bright sunshine ._. My	coat gleamed in the sun
large ,_, wearing a bright	coat with gold buttons and
1 trees ,_, colonies of menacing	crabs and turtles hiding in

Data-Driven Learning Type 2

The DDL Type 2 approach focuses on "adjective + noun" collocations by extraction by the result of semantic tag-based analysis and finds the use of distinct adjectives and nouns.

The aim of this DDL Type 2 approach is to focus on the variety of adjectives and nouns and classify them into categories by semantic tag analysis to learn as not only a "word" but also part of a two-word "collocation." The following three tasks are examples of the DDL Type 2 approach, and Figure 5 shows an example of the DDL collocation worksheet.

Task A: Guess the keyword in context by reading the concordance lines on the right and left of the keyword and share your idea(s) with your group members.

Task B: Focus on "adjective + noun" collocations in the context of

the concordance lines. What type of adjectives are used from concordance A to concordance D? Share your discoveries with your group members.

Task C: Focus on "adjective + noun" collocations in the context of the concordance lines. What type of nouns are used from concordance A to concordance D? Share your discoveries with your group members.

FIGURE 5. DDL Worksheet of "Adjective + Noun" Collocation by Semantic Tag Analysis of Adjectives

	Concoradance A	
1	as no time to waste . My first	idea was to make ice-shoes so t
2	some more . And thats when an	thing happens : she gets up out
3	to make his living by breeding	horses from the stallion and fr
4	hildren and the animals had an	time , everyone was getting on
5	Charlotte very carefully . The	thing was that if you asked a q
6	treet , all the people saw the	green carpet of grass coming fr
7	should he do ? Then he had an	idea . If he could make prosthe
	Concordance B	
1	! Long ago his wife had had a	baby called May , she was now a
2	Come on George , wakey wakey a	girl comes today . But then the
3	Dog By Amabel Smith , aged 9 A	morning , the birds are singing
4	f you were sat in my place . A	butterfly , a shiny sun , Take
5	ion tents and there was even a	baking tent where everyone brou
6	y sprinkled salt all over five	cakes . Another old lady , with
7	a tweed skirt , wellies and a	diamond tiara that glittered so
8	realised that it was the most	horse he had ever seen . He cou
9	where I could see out in to a	garden , where the wind gently
10	ffed my edging and sprayed the	glass case and wiped it so that
11	p the stage . She turns into a	spider . She is a lovely patien
12	dy floss . On her back she had	wings , she was a faerie ! Hell
13	ved a witch and her uncommonly	daughter , Daisy . Daisy was th
14	ange limousine pulled up and a	lady hopped out in a fancy blac
15	he exclaimed , " What is this	sword doing at the bottom of th
16	ds , telling the sunflower how	it had once looked . Suddenly s
	Concordance C	
1	were too late ... Petal was a	lady and never let anyone down
2	and every year it would get a	glossy lick of bright red paint
3	on the shelf . Goodness . What	manners ! Its almost as if Ive
4	put it in her bag with all the	white teeth . But then she woul
5	a beautiful spider . She is a	patient pink footed tarantula .
6	to his prawns . Oh no ! In his	blue bowl was disgusting , slop
	Concordance D	
1	in town , and it was a really	town ! It wasnt called Nicetown
2	icetown for nothing . But even	towns can have problems . The t
3	is full of friendly worms and	food like leaves and sausages .
4	nice peice of pork and what a	bottom for ham ! * But no . Now
5	programme about space and that	Professor Brian Cox was on it .
6	ve knight like me ! Oh , and a	flock of extremely woolly sheep

The worksheet is based on the semantic-tag analysis of "O4.2+" as defined by "judgement of appearance." The four adjectives adopted for this worksheet to learn "adjective + noun" collocations are as follows: *nice*, *beautiful*, *amazing*, and *lovely*. The aim of the focus on the semantic tag "O4.2" and choosing these four adjectives is, for the adjectives *nice* and *beautiful*, to give Japanese EFL learners a chance to learn about the variety of use of nouns in an "adjective + noun" collocation. Moreover, regarding the two adjectives *amazing* and *lovely*, as Japanese EFL learners have rarely used these two adjectives in their language practice, this DDL Type 2 approach creates a good opportunity to learn the collocation of these adjectives.

As discussed above, the use of the results of qualitative analyses to adopt different types of DDL approaches would be very efficient for Japanese EFL learners to acquire a wider variety of English vocabulary and collocations. Also, the DDL approach would give Japanese EFL learners the opportunity to discover rules of grammar and focus on parts of speech in context, as well as the variety in NESs' expressions through DDL as corroborative and discovery learning in the classroom.

CONCLUSIONS

This study compared the use of adjectives of 5–9 BSs and10–13 BSs with that of 13–15 JSs and whether there was observable distinct language use. The language studied was written English in three corpora. The two British students' data were collected from the BBC's 500 Words writing competitions from 2014 through 2017 and the Japanese EFL learners' data was collected from a written in-class task in English class at the end of each year at a lower-secondary school between 2006 and 2008.

This study aimed to conduct quantitative and qualitative analysis of three language samples – two of young British NESs and one of Japanese EFL learners – and explore the possibility of adopting the results for DDL in the EFL classroom. The results of quantitative analyses show that the type–token ratio for the 10–13 BSs was a little higher than that of the 5–9 BSs and least for the 13–15 JSs. The results of qualitative analyses indicate that the number of adjectives used by British students is much more than that used by Japanese EFL students.

The qualitative analyses revealed that the most frequent adjective use among British students is almost the same, but the 10−13 BSs use more different types of adjectives than the 5−9 BSs. The results suggest that adjective use by Japanese students is very limited compared to both samples of the British students' language use. In the qualitative analyses, this study also uses semantic-tag analysis concerning adjectives. In consequence, it could be successful to identify distinct semantic categories in each language use among these three different types of language.

Furthermore, bi-gram analysis identified distinct "adjective + noun" collocations in the British students' language use. The results of quantitative and qualitative analyses were valuable in dealing with the possibility of applying DDL to EFL English classes. This study suggests two types of DDL approaches: (a) DDL Type 1, which is based on an adjective frequency word list as an open-ended DDL approach for learning vocabulary, collocations, parts of speech, grammar, and NES expressions, and (b) DDL Type 2, which is based on adjective semantic-tag analysis to make it possible for Japanese EFL learners to learn different adjectives in the same semantic tag category at the same time and discover a variety of different types of "adjective + noun" collocations. Moreover, it also gives Japanese EFL learners the opportunity to notice a variety of noun uses and increase their language use of "adjective + noun" collocations.

Having reviewed previous studies of DDL overseas and in Japan, there seems to be little DDL research focusing on the primary-learner level or suggesting authentic materials that are suitable for primary-level Japanese EFL learners. Accordingly, this study is valuable and novel in the field of comparative language study, developing practical DDL approaches based on analyses of adjective use concerning 5- to 9-year-old and 10- to 13-year-old British students' written language as a target language for Japanese EFL learners and primary and lower-secondary school EFL learners in other countries.

Previous studies in Korea regarding corpus-based and DDL approaches provide us with some interesting and useful information to solve the difficulties associated with increasing future availability of DDL in secondary schools. The small amount of research and DDL implementation at the secondary school level makes it more difficult for teachers to learn the DDL method, especially since most researchers point to the necessity of training teachers in the use of corpora at

universities or at workshops by a teacher who has experience in their pedagogical uses.

Regarding the educational environment, especially computer facilities and the amount of time required to use general corpora, some researchers suggest that we create paper-based DDL. The goal of this would be to mitigate learners' lack of computer skills and their difficulty in reading concordance lines, as well as make it unnecessary to teach these computer skills. Some researchers suggest that we create a small, specialized corpus suited to the leaners' level. This would be used to create textbook corpora, graded-reader corpora, and learner corpora. Having a teacher create these corpora would require significant preparation time. Therefore, teachers' web networks for dealing with the DDL approach would be valuable and beneficial; this would allow teachers to create specialized corpora and share their implementations in daily English classes in secondary schools. Lee (2011) suggests that an online community be built for teachers to use corpora in the classroom: "It may be necessary to build an online community for teachers who wish to use corpora in their classroom in which they can share their experiences in implementing and using corpora in their classroom" (p. 174).

Given the previous studies on DDL and the difficulties in implementing DDL with young learners, this study suggests a semantic-based DDL approach and provides authentic written language material produced by a group of British students between the ages of 5 and 9, and another group between the ages of 10 and 13. Furthermore, this study provides a sematic-based DDL approach to teaching adjectives to widen young learners' knowledge of synonyms. The corpus size of this study was very small and was created over three years based on the BBC 500 Words competitions between 2014 and 2017. Therefore, future studies might continue collecting data to create a much larger, specialized corpus, which is necessary for developing a sufficient amount of authentic material for semantic-based DDL in the future. For future studies, we plan to develop and provide semantic-based DDL material for teachers and learners in elementary and secondary schools. Sharing and exchanging DDL implementation with English teachers in other countries would be quite valuable for both teachers and learners.

THE AUTHOR

Kunihiko Miura is a professor of the Faculty of Policy Studies and a teacher on the EFL teacher-training program at the University of Shimane in Hamada, Japan. He received his MA in TESOL from the University of Birmingham and his PhD in English linguistics from Daito Bunka University, Tokyo. Email: k-miura@ u-shimane.ac.jp

REFERENCES

Blappert, T. (1997). Data-driven learning: Theory and classroom implementation. In Korea TESOL (Eds.), *Proceedings of the 1997 Korea TESOL Conference* (pp. 11−18). Korea TESOL. https://koreatesol.org/sites/ default/files/pdf_publications/KOTESOL-Proceeds1997web.pdf

Carlstrom, B. (2014). Data-driven made easy. In M. Pinto & D. Shaffer (Eds.), *KOTESOL Proceedings 2013* (pp. 95−102). Korea TESOL. https://www. academia.edu/10291111/Data-Driven_Learning_Made_Easy

Chujo, K., Anthony, L., Oghigian, K., & Uchibori, A. (2012). Paper-based, computer-based, and combined data-driven learning using a web-based concordance. *Language Education in Asia*, *3*(2), 132−145.

Chujo, K., Utiyama, M., & Miura, S. (2006). Using a Japanese parallel corpus for teaching English vocabulary to beginning-level students. *English Corpus Studies*, *13*, 153−172. https://doi.org/10.5746/LEiA/12/V3/I2/A02/ Chujo_Anthony_Oghigian_Uchibori

Craig, B. A. (1998). Extending English instruction to primary schools: Lessons from East Asian EFL contexts. In Y. R. Heldi, H. Ardi, & A. Arianto (Eds.), *The Proceedings of the Seventh International Symposium on English Teaching* (Vol. 1, pp. 373−384). Atlantis Press.

Dailey, A. (2009). *Implementing task-based language teaching in Korean classrooms*. University of Birmingham. https://www.birmingham.ac.uk/ Documents/college-artslaw/cels/essays/languageteaching/DailyLTMImpleme ntingTask-BasedLanguageTeachinginKorean.pdf

Gilquin, G., & Granger, S. (2010). How can data-driven learning be used in language teaching? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 359−370). Routledge.

Guan, X. (2013). A study on the application of data-driven learning in vocabulary teaching and learning in China's EFL class. *Journal of Language Teaching and Research*, *4*(1), 105−112. https://doi.org/10.4304/jltr.4.1.105-112

Hall, D. G., Waxman, S. R., & Hurwitz, W. M. (1993). How two- and four-year-old children interpret adjectives and count nouns. *Child Development*, *64*(6), 1651−1664. https://doi.org/10.2307/1131461

Hong, S. C. (2010). EFL learners' consciousness-raising through a corpus-based approach. *English Teaching*, *65*(1), 57–86. https://doi.org/10.15858/engtea.65.1.201003.57

Park, J. K. (1999). Teaching English: Korean culture and context. *English Teaching*, *54*(3), 3–21.

Park, J. K., & Kim, M. K. (2014). Teaching and learning of EIL in Korean culture and context. In R. Marlina & R. A. Giri (Eds.), *The pedagogy of English as an international language* (pp. 47–61). Springer.

Kim, H. (2019). The perception of teachers and learners towards an exploratory corpus-based grammar instruction in a Korean EFL primary school context. *Primary English Education*, *25*(1), 123–152. https://doi.org/10.25231/pee.2019.25.1.123

Kim, H., & Chun, S. (2008). Fostering lexis awareness and autonomy by corpus-based data-driven learning. *English Teaching*, *63*(2), 213–235. https://doi.org/10.15858/engtea.63.2.200806.213

Klibanoff, R. S., & Waxman, S. R. (2000). Basic level object categories support the acquisition of novel adjectives: Evidence from preschool-aged children. *Child Development*, *71*(3), 649–659. https://doi.org/10.1111/1467-8624.00173

Kochmar, E., & Briscoe, T. (2015). Using learner data to improve error correction in adjective–noun combinations. In *Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 233–242). Association for Computational Linguistics. https://doi.org/10.3115/v1/W15-0627

Lee, S. (2006). A corpus-based analysis of Korean EFL learner's use of amplifier collocations. *English Teaching*, *61*(1), 3–17.

Lee, S. (2011). Challenges of using corpora in language teaching and learning: Implications for secondary education. *Linguistic Research*, *28*(1), 159–178. https://doi.org/10.17250/khisli.28.1.201104.009

Luo, Q., & Zhou, J. (2017). Data-driven learning in second language writing class: A survey of empirical studies. *International Journal of Emerging Technologies in Learning*, *12*(3), 182–196. https://doi.org/10.3991/ijet.v12i03.6523

Mintz, T. H., & Gleitman, L. R. (2002). Adjectives really do modify nouns: The incremental and restricted nature of early adjective acquisition. *Cognition*, *84*(3), 267–293. https://doi.org/10.1016/S0010-0277(02)00047-1

Nesselhauf, N. (2005). *Collocations in a learner corpus* (Vol. 14). John Benjamins. https://doi.org/10.1075/scl.14

Nicoladis, E., & Rhemtulla, M. (2012). Children's acquisition of word order depends on syntactic/semantic role: Evidence from adjective-noun order. *First Language*, *32*(4), 479–493. https://doi.org/10.1177/0142723711422629

Rahimi, M., & Momeni, G. (2012). The effect of teaching collocations on English language proficiency. *Procedia – Social and Behavioral Sciences*, *31*, 37–

42. https://doi.org/10.1016/j.sbspro.2011.12.013

Tribushinina, E. (2013). Adjective semantics, world knowledge and visual context: Comprehension of size terms by 2- to 7-year-old Dutch-speaking children. *Journal of Psycholinguistic Research*, *42*(3), 205−225. https://doi.org/10.1007/s10936-012-9217-3

Vyatkina, N. (2016). Data-driven learning of collocations: Learner performance, proficiency, and perceptions. *Language Learning Technology*, *20*, 159−179.

Yoon, H. (2005). Corpus-based studies in L2 teaching: Implications for an EFL context. *English Teaching*, *60*(2), 285−305.

Foreign Language Learners' Reading Attitudes in Print and Digital Settings

Rosalyn G. Mirasol & Eunjin Kim
University of Santo Tomas, Manila, Philippines

The study investigated the academic and recreational reading attitudes of Korean students in print and digital settings. There is a strong relationship between attitudes and reading. A good attitude towards reading contributes to a high level of academic achievement. As more and more Koreans study in the Philippines, it is important to determine their reading attitudes in digital settings. The findings indicate that the participants were somewhat positive with academic reading and recreational reading in digital settings. However, they were neutral in their attitudes toward reading in the print setting. It was also found that female participants were more inclined to read in the digital setting while male participants prefer reading recreationally in the digital setting. The females had more positive attitudes towards reading, both in the print and digital settings. It is important for educators to consider digital reading platforms so that learners can better connect with the text.

Keywords: reading attitudes, academic reading, recreational reading, print setting, digital setting

INTRODUCTION

The world is filled with a vast amount of information. This allows people to acquire necessary education not only to further and satisfy their pursuit of knowledge but, most importantly, to be able to survive the challenges of life. With the advent of technology and cyberspace, all types of information can easily be accessed by almost anybody. Technology and cyberspace have created a new world and led the people into an information-oriented society. With this massive amount of information that is available to anyone, it is a challenge for learners to

be more critical with the kinds of information available on the internet. Reading is one of the important skills to help learners achieve a high level of critical literacy. No one can deny the fact that reading enables one to improve his/her ability to think critically and become a wise user of this information technology. Reading broadens the "experience view" and helps one to have high ideology (Lee, 2009).

Reading is essential to the success of a student in school and in life. One who is not able to read will not be able to succeed (U.S. Department of Education, 2005). Reading is one of the most important parts of the learning process and is seen as a channel for life-long learning and intellectual growth. Reading transfers experiences to the individual so that the reader may expand one's horizon as well as identify, extend, and intensify his/her interests and gain a deeper understanding of the world (Green, 2002). Reading is an important component in learning a language, and it is an essential tool for lifelong learning (Mokatsi, 2005; Pandian, 1997). Reading skills lie at the heart of formal education; it is difficult to achieve things without the ability to read fluently with good comprehension (Senturk, 2015). It was once thought that reading was interpreting visual information of any given code or system (Lone, 2011). But now there are different views and definitions of reading. Frey and Fisher (2006) and Smith and Robinson (1980) define reading as a process to understand a writer's message. The first description of reading and its processes can be traced back to Thorndike in 1971. According to Thorndike (1971), reading is a process of reasoning. Reading is one of the macro skills and is made up of various processes; therefore, reading is affected by a multitude of factors.

Research has shown that there is a strong correlation between reading attitudes and high reading proficiency and reading comprehension (Bokhorst-Heng & Pereira, 2008; McKenna et al., 2012; Schooten et al., 2004). Adolescents' reading proficiency and comprehension are generally affected by their attitudes towards reading. The assessed reading attitude can be used to predict future reading behavior and capacity for achievement (Jang, 2006). Thus, negative attitudes towards reading tend to produce poor reading habits and a low level of comprehension, while a positive and pro-active attitude towards reading disposes adolescents towards a good reading habit with a high level of reading comprehension. But with the proliferation of technological gadgets and games easily accessed by anybody, reading has become the lowest priority among adolescents today. It is these

gadgets and games that have shifted adolescents' attitudes drastically.

According to a study on internet usage conducted by Pew Internet and American Life Project, more than a million people have participated in literary activities through the internet (Lenhart et al., 2001). Ninety-three percent (93%) of American adolescents are internet users (Lenhart et al., 2007), while 21.5% of Korean adolescents are using their gadgets to read (Korea Culture and Tourism Institute, 2012). The findings of this survey imply that there is a need to investigate the reading attitudes of adolescents in terms of digital setting. Hence, it is very important for there to be a systematic and valid way of understanding and analyzing their reading attitudes.

Adolescents' quality of reading is generally affected by their attitudes towards it. Attitude towards reading influences reading comprehension and reading difficulties. Improving adolescents' proficiency and reading comprehension requires an understanding of their attitudes towards reading. McKenna et al. (2012) cited several reasons why one should use a broader lens to determine attitude in conducting inquiry into the reading attitudes of adolescents. First, "the measurement of attitudes holds the potential to contribute to our understanding of reading comprehension and reading difficulties" (pp. 283–284). Second, the study of reading attitudes broadens our understanding of adolescents' reading identities. And, finally, a reliable and valid assessment of their attitude can provide teachers with the tools they need to address the affective goals.

Several studies have seen a dramatic decrease in reading attitudes among elementary students from Grade 2 to Grade 6 (Anderson et al., 1985; Barnett & Irwin, 1994; McKenna, Kear, et al., 1995; McKenna, Stratton, et al., 1995). The same findings were found among Korean elementary learners (Choi, 2010; Yoon & Kim, 2008). Therefore, it should be of great interest among educators to explore the reading attitudes of adolescent learners in terms of purpose and setting.

LITERATURE REVIEW

Youth and Digital Reading

Our era is described as the information or digital age. This means that information is readily accessible at one's fingertips almost anywhere

on the globe. More and more people are using the computer and other electronic tools as a source of knowledge ranging from simple questions to academic discourses. In fact, the digital age also includes the digitization of recreation, so that with the use of handy gadgets, one can actually play games while traveling or waiting or simply in between work or study.

"Global literacy has grown substantially in the last two centuries" (Roser & Ortiz-Ospina, 2013, section 3). This means that a lot of people across the globe can read and write. But what is alarming is that despite this growth, there has been a decline in the proportion of children who read daily from childhood to the tween and teen years. According to Calkins (2012), if American students were assessed, only 15% of the population would perform at a level at par with standards. This shows that the remaining 85% is failing. The National Assessment of Education Progress said that less than one third of America's fourth-grade students are able to read at or above the proficient level of achievement (Martinez et al., 2008). One study documented a drop of 48% in 6- to 8-year-olds and 24% in 15- to 17-year-olds who are daily readers (Common Sense Media, 2014). In the same study, it was found that reading for pleasure drops off dramatically as children get older, and the rates among all children − especially teens − have fallen precipitously in recent years. "Frequency of reading books for fun is significantly lower for learners aged 12−17 than for children age 6−11; frequency of reading books for school is also lower for learners aged 12−17 than for learners ages 6−11" (Scholastic, 2013, p. 5). Recurrence of perusing books for entertainment only is fundamentally lower for students aged 12−17 than for youngsters age 6−11; recurrence of perusing books for school is likewise lower for students ages 12−17 than for students ages 6−11 (Scholastic, 2013). One reason is that reading competes with many different activities as children grow older (Scholastic, 2015).

Despite this being the digital information age, people are still using print materials (Liu, 2003, 2006). There are various reasons for this tendency to make use of print materials. The first reason is research: Much older and age-old information are still in printed form (Manguel, 1997; Tveit & Mangen, 2014; van der Weel, 2011). Another reason for the preference of reading in print rather than in digital is the phenomenon called "haptic dissonance" (Gerlach & Buxmann, 2011). Gerlach and Buxmann describe *haptic dissonance* as a feeling of expectation in the experience of reading. This expectation is the element

of experience that something is missing. It denotes the level of feeling connected to reading, which is generally regarded as occurring in print reading. Readers report that this experience does not happen in digital reading (Gerlach & Buxmann, 2011; Pattuelli & Rabina, 2010; Rose, 2011; Scarry, 2001). Print reading is considered more intense and personal, as it provides the tactile experience of reading, for example, leafing through the pages and jotting down notes. Finally, another reason is psychological: People used to reading in print tend to prefer reading printed materials.

Research indicates, however, that young readers prefer alternative reading materials such as periodicals, comic books, and websites (Gordon & Lu, 2008). However, reading in print from print books continues to dominate classrooms, Wolk (2010) argued that it would remain at the forefront of reading, but more and more, young readers have a more positive attitude towards reading from an e-reader (Allen, 2013; Larson, 2010).

Reading Motivation

Motivation plays an important part in the reading process. Motivation refers to the reason or reasons for doing things. When readers are not clear about their motivation, then reading is mechanical. On the contrary, when they have purpose, the engagement seems to be more serious and personal. Engagement in reading is directly related to reader motivation (Schraw & Bruninig, 1999). According to Beers (1998), readers lacking motivation feel that they are not reading at all. They consider reading as a mere function to comply with certain social roles. It can also be said that reading is taken as one of the skills they have to master at the word level.

People have different attitudes towards reading. According to Spear-Swerling (2004), struggling readers give up on reading tasks easily. They do not engage in text reading, whether print or digital, because of a lack of motivation to do so (Schraw & Bruninig, 1999). There is therefore a very strong correlation between motivation and struggle in reading.

Because of a lack of motivation, struggling readers do not really make an effort to devise strategies and approaches in their reading. Reading is done mechanically under close supervision. Without supervision, struggling readers may finally abandon reading and do other

things instead. Unmotivated readers have negative feelings about readers, and they do not identify themselves as readers. They typically see reading as functional and approach reading as a set of skills at the word level. Struggling readers often lack necessary strategies to learn effectively with text.

The impact of the influence of motivation towards reading is the "Matthew Effect." According to Stanovich (1986), the "Matthew Effect" causes good readers to get better while poor readers become more limited in their ability. Good readers tend to be actively involved in selecting their electives, activities in school, and other programs they find interesting and useful. On the other hand, struggling readers simply accept what is offered to them and take for granted its use, purpose, and other values for consideration. The "Matthew Effect" indicates that motivation is at the heart of attitude towards reading. A student with higher motivation tends to be engaging, active, and participatory in determining school subjects, organizations, and other aspects of student life. In contrast, struggling readers live according to the flow of the current; hence, they go wherever the current leads them.

The home environment is a factor in motivation to read. Harmer (2004) emphasized that "the attitudes of parents and other siblings will be crucial because their approval and encouragement in reading will affect the students' motivation and interests to read" (p. 99). Therefore, both the students' reading attitudes and reading motivation are aspects that students develop within the family setting (Johari et al., 2013). It is also said that children who see their parents reading every day, whether for academics or for pleasure, will be more likely to be avid readers in the future. One can also say that if a child's home has a culture of reading, there is a greater probability that the child will not abandon reading when he or she grows up (Davidovitch et al., 2016).

Another source of motivation comes from school. But the motivation that comes from the school is not more positive than that from the home. The notion that reading is imposed upon students in school does not sit well. At school, students usually read not because they want to, but because they have to (Khairuddin, 2013). According to Tunde-Awe (2014), students' reading and learning are limited to what is specifically required to achieve their limited objectives, such as success in examinations and job procurement. Adolescents do not like reading materials and tasks imposed on them; they would like to be given choices when it comes to what and when to read (Guthrie et al., as cited

in Mante-Estacio, 2012). Both attitude and motivation, therefore, play a big role in reading.

Fox and Alexander (2009) assert that motivation also impacts the metacognitive knowledge capacity of the individual. This means that a person well-motivated towards reading has better ways to employ reading strategies and deeper reading comprehension. The opposite is true for the struggling readers and readers lacking motivation. This is further affirmed by McKenna and Kear (1990), who claim that there is a reciprocal relation between reading quantity and reading achievement. The more an individual reads, the more their reading achievement increases. Reading achievement connotes deeper comprehension and finding meaning in reading. For this reason, an increase in reading achievement motivates an individual to increase the amount of reading. In a way, McKenna and Kear present a cycle of reading where motivation is highlighted at the center. With motivation, one reads willingly; this joyful encounter with the text leads to meaningful experiences, and this leads to deeper comprehension and an increased amount of reading.

Second Language Reading Attitudes

It has been indicated earlier that there is a strong link between attitude toward reading and achievement. This link also applies to the second language reading attitude. Studies have shown that a good attitude towards reading greatly affects the reading achievement and performance of second language learners (Kamhi-Stein, 2003; Yamashita 2004, 2007, 2013). The question, however, is whether their attitude towards reading in their native language is the same as their attitude in reading in the second language, and whether the second language actually reshapes attitude in so far as culturally the second language is in a different context and provides a different experience.

Day and Bamford (1998) argued that a positive attitude towards second language reading is influenced by their existing attitude manifested in the reading activities practiced at home. Ro and Chen (2014) replicated the study conducted by Day and Bamford (1998) among Korean adolescents and found that having a positive attitude towards reading in their home language is carried on as a habit and attitude in dealing with second language reading. Moreover, the study showed that experience in the target-language culture and years of

previous English study did not show significance statistically with the first research. This suggests that attitude towards reading is a carryover from the home language, and not really something that will develop over time as students engage in the second language. Finally, poor reading habits and lack of time are considered as major factors affecting reading attitudes among Korean adolescents (Ro & Chen, 2014).

Factors Influencing the Attitude Towards Reading

Family has a strong influence in adolescents' attitude towards reading, whether in digital or in print or both settings. It is normally assumed that the behavior of the individual is reflective of the traits and characteristics of the family. This assumption also applies to their attitude towards reading.

Studies have shown that there are two primary ways a family can have an impact on children's reading development. First, there is the likelihood of a genetic link (DeFries et al., 1997). Genetic contributions towards reading development are best studied via genetically sensitive research design such as twin or adoption studies (Nation, 2006). Second, family history can influence the children's reading skills through the provision or lack of certain environmental experiences. Family provisions refer to the situations and conditions the family creates for their children. For example, parents who are readers will most likely influence their children to appreciate and get into reading at a young age. In a home where books and other reading materials, whether in print or in digital forms, are part of the children's everyday activity, or a family activity, this will generate a good reading habit among the children. This is termed *home literacy environment*, where the availability of books and time has been linked to the children's development of reading skills (Molfese et al., 2003).

While there are many other factors in the family such as socio-economic level, educational attainment of parents, and other environmental factors affecting attitude towards reading, it is the amount of exposure the children have in the culture of reading that influences the children's reading habits and abilities (Leslie & Allen, 1999; Samuelsson & Lundberg, 2003). Within the context of reading, the way in which children appraise their reading capabilities is expected to influence motivational aspects. This includes interest in reading and reading persistence, which influences the children's reading achievement

(Chapman & Tunmer, 1995; Jacobs et al., 2002).

THE STUDY: METHOD

Aims

This study determined the reading attitudes of selected Korean adolescents. Specifically, the following research questions were addressed:

RQ1. Is there a significant difference between reading attitudes and gender?

RQ2. Is there a significant difference between reading academically in a print setting and reading academically in a digital setting?

RQ3. Is there a significant difference between reading recreationally in a print setting and reading recreationally in a digital setting?

Participants

The respondents of the study were selected Korean adolescents enrolled in nine comprehensive educational institutions in Manila. Among the 75 respondents, 43 were females (57%) and 32 were males (43%). In terms of age distribution, 63% of the respondents were 18 years old; 26% were 17 years old; 7% were 16 years old, and 4% were15 years old.

Materials: The Survey Instrument

The McKenna et al.'s (2012) adolescent reading attitudes survey was used to describe the reading attitudes of the respondents in print and digital settings. The instrument was a 6-point Likert scale, which requires respondents to rate a series of activities and events from 1 to 6 on the basis of perceived desirability. A 6-point scale was selected because adolescents are typically capable of discriminating among this many discrete bits of information (e.g., Case & Khanna, 1981; Chi, 1978; Chi

& Klahr, 1975; Nitko & Brookhart, 2010). In addition, an even number of points avoids a neutral midpoint and does not permit raters to select a neutral middle option as a means of masking their judgments (Nunnally & Bernstein, 1994; see also Henk et al., 2011; McKenna, Kear, et al., 1995). Every item was rated using a Likert-type scale (1 = *very bad*; 2 = *bad*; 3 = *somewhat bad*; 4 = *somewhat good*; 5 = *good*; 6 = *very good*). Intermediate positions were undefined.

Procedures

The instrument was made available in printed form and through an online survey. The printed form was personally distributed to the respondents. Before administering the survey, the authors explained the purpose of the study and that the responses would be treated with utmost confidentiality. For the online responses, the researchers used Google Docs.

Korean students for the online survey were pre-selected; hence, orientation was done through distance communication using email. The researchers contacted them individually either through email, calls, or through KakaoTalk, an electronic messaging app popular among Koreans. In the same way, the researchers explained to the respondents the aims and objectives of the research. The respondents were told that they could withdraw at any time if they wished to. While the conversation was going on, the target participant was already looking at the Google Docs online survey to ensure that the target participant understood the instructions. (They were also assured of the confidentiality of their responses.) The participants were assured that their responses would be treated with utmost confidentiality. After the e-conversation, the respondents were asked to fill out the survey questionnaire online.

RESULTS

Reading Attitudes and Gender

An independent-samples *t*-test was conducted to compare the attitudes of the respondents toward reading printed materials versus

digital reading materials. It was found that there was a significant difference in attitudes towards digital reading materials between the male (M = 38.42, SD = 9.58) and female respondents (M = 43.93, SD = 9.056); $t(73)$ = -2.542, p = 0.013. This implies that women were more inclined to read text in a digital format than in print form. On the other hand, the male respondents in the study seemed to prefer digital text when reading as a form of recreation (see Tables 1 and 2).

TABLE 1. Reading Attitudes and Gender

Group	Gender	N	Mean	Std. Deviation	Std. Error Mean
AvePrint	Male	32	37.4018	2.19497	.38802
	Female	43	38.3889	2.94788	.44955
AveDigital	Male	32	38.4219	9.58166	1.69381
	Female	43	43.9302	9.05577	1.38099

TABLE 2. Reading Attitudes and Gender: *t*-Test Results

| Treat-ment | Variances | Levene's Test for Equality of Variances | | *t*-Test for Equality of Means | | | | | | |
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference Lower	Upper
AvePrint	Equal variances assumed	.860	.357	-1.593	73	.116	-.98706	.61970	-2.22212	.24801
	Equal variances not assumed			-1.662	72.998	.101	-.98706	.59385	-2.17059	.19647
AveDigit	Equal variances assumed	.280	.598	-2.542	73	.013	-5.50836	2.16719	-9.82756	-1.18915
	Equal variances not assumed			-2.520	64.783	.014	-5.50836	2.18544	-9.87326	-1.14346

Reading Academically in Print Settings and in Digital Settings

The results reveal that more respondents preferred reading academic text digitally. However, the mean differences between the male and female responses were not significant at a .05 level of significance. On the other hand, the mean difference between the male and female responses in their attitudes toward reading academic texts in print was

significant at $p = .049$. This implies that female attitudes toward reading texts in either print or digital form are more positive than their male counterparts (see Tables 3 and 4).

TABLE 3. Reading Academically in Print Settings and in Digital Settings

Group	Gender	N	Mean	Std. Deviation	Std. Error Mean
AveDig	Male	32	39.1250	9.14665	1.61692
	Female	43	40.9767	7.76895	1.18475
Aveprnt	Male	32	35.4688	10.06105	1.77856
	Female	42	49.6977	8.22173	1.25380

TABLE 4. Reading Academically in Print Settings and in Digital Settings: *t*-Test Results

Treat-ment	Variances	Levene's Test for Equality of Variances		*t*-Test for Equality of Means						95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference		Lower	Upper
RecPrint	Equal variances assumed	.000	.986	-.946	73	.347	-1.85174	1.95683		-5.75171	-2.04822
	Equal variances not assumed	1.045		-.924	60.377	.359	-1.85174	2.00451		-5.86084	-2.15736
RecDig	Equal variances assumed	.037	.310	-2.002	73	.049	-4.22892	2.11253		-8.43919	-.01866
	Equal variances not assumed			-1.943	58.757	.057	-4.22892	2.17607		-8.58361	.12577

Reading Recreationally in Print Settings and in Digital Settings

Based on the independent *t*-test, males and females showed a significant difference in attitude ($p > .007$) when asked to rate/describe their feelings toward recreational text in a print setting. More female respondents ($M = 48.16$) preferred recreational reading of text in print than having to read it digitally ($M = 38.67$). Furthermore, the female participants had a more positive attitude towards reading text in print as a form of recreation as opposed to the male respondents of the study (see Tables 5 and 6).

TABLE 5. Recreational Reading in Print and Digital Settings

Group	Gender	N	Mean	Std. Deviation	Std. Error Mean
AveDig	Male	32	41.3750	10.24459	1.81100
	Female	43	48.1628	10.78122	1.64412
Aveprnt	Male	32	36.6563	7.68161	1.35793
	Female	42	38.6744	7.24313	1.10457

TABLE 6. Recreational Reading in Print Settings and Digital Settings: *t*-Test Results

Treat-ment	Levene's Test for Equality of Variances			*t*-Test for Equality of Means						
	Variances	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
RecPrint	Equal variances assumed	.984	.325	-2.754	73	.007	-6.78779	2.46461	-11.69975	-1.87583
	Equal variances not assumed			-2.775	68.709	.007	-6.78779	2.44599	-11.66778	-1.90780
RecDig	Equal variances assumed	.037	.848	-1.163	73	.249	-2.01817	1.73523	-5.47647	1.44013
	Equal variances not assumed			-1.153	64.691	.253	-2.01817	1.75044	-5.51435	1.47802

DISCUSSION

Based on the findings, it is interesting to note that there was a significant difference in attitudes towards digital reading materials between the male and female respondents. The result implies that females are more inclined to read texts in digital form than in print form. However, the results found that the male respondents of the study preferred digital text when reading as a form of recreation. Difference between gender motivations has often been found in other studies. There has been speculation that the cultural influences of non-literary activities (e.g., sports, music) on males and the perceived feminine quality of reading have brought about this disparity and that girls score higher than

boys on tests of reading ability (Loveless, 2015). Abdorahimzadeh (2014) emphasized that differences in gender regarding topic interest do have a psychological reality. Females tend to read literature that cater to feminine topics while males also showed a surprising interest in female topics. In any case, reading motivation with gender as a factor usually relies on the interest held by the two genders and how it affects their focus and attention on a given text.

Looking into the results of the study, it is evident that non-cognitive assessment in reading performance (e.g., reading attitudes) is an important point to consider by all teachers across levels. Although this study focused on adolescent learners, it is necessary to consider the reading attitudes and motivation of learners of all ages before considering any instructional plan. Teachers should also take into consideration their strategies being employed in the classroom. This is because the pleasure of learning in areas of personal interest and curiosity is hard to foster in traditional reading classes, where the teacher selects the same texts for every student (Yamashita, 2013). Extensive reading, on the other hand, is an approach to reading pedagogy that encourages students to engage in a large amount of reading. Day and Bramford (2002) stated that the approach operates best under these circumstances: (a) when a variety of reading materials are available, (b) when the students are given the chance to choose what they want to read, (c) where the students are allowed to read unassisted, (d) where the teacher serves only as a model, and (e) where reading in itself is seen as a reward. It was determined that the use of extensive reading had a positive effect on both attitude and motivation. The study then takes the stance that extensive reading exerts a significant effect on the aspects of reading attitude that may foster intrinsic motivation (i.e., positive attitude and intellectual satisfaction) than on those that may relate to extrinsic motivation (i.e., higher grades and career benefits). Furthermore, educators should highly consider the digital environment to be an integral part of the learning platforms in the classroom. As learners have already gained access to cyberspace, it is an opportune time to look at the benefits of digital learning, especially for learners with high positive attitudes in learning digitally or recreationally.

CONCLUSIONS

It is evident from this study that adolescents today have more positive attitudes towards digital reading than print reading. The results of the study revealed that the participants held positive attitudes in digital settings, both in academic reading and recreational reading. However, they were neutral/indifferent in their attitudes toward reading in the print setting, both academically and recreationally. It was also found that female participants were more inclined to read text in a digital setting than in a print setting, while male participants preferred reading recreationally in a digital setting. The female participants had more positive attitudes towards reading, both in print and digital settings. The gender difference was substantial in measuring the attitudes of the students. The gender factor was consistent with the international findings that males are less likely to deem recreational reading enjoyable, leading to the conclusion that "most of the gender gap can be explained by boys being less engaged, and less engaged students show lower performance" (Merga, 2014, p. 168). Finally, the findings of the study challenge ESL pedagogy to seriously take into consideration the learning platforms of the students. Educators should consider learning platforms as part of the learning environment of foreign language learners. In assessing students' attitude towards reading, appropriate and contextual design that addresses gender differences could be an important factor in designing ESL curriculum for foreign students.

THE AUTHORS

Rosalyn G. Mirasol is an associate professor in the Department of English, Faculty of Arts and Letters of the University of Santo Tomas, Manila, in the Philippines. She is also a research associate at the Research Center for Social Sciences and Education at the same university. She received her Doctor of Philosophy in Reading at the Philippine Normal University in Manila, Philippines. Email: rgmirasol@ust.edu.ph

Eunji Kim is a graduate of English language studies of the Faculty of Arts and Letters of the University of Santo Tomas in Manila, Philippines. Email: eungikim94@naver.com

REFERENCES

Abdorahimzahed, S. (2014). Gender differences and EFL reading comprehension: Revisiting topic interest and test performance. *System*, *42*(1), 70−80. https://doi.org/10.1016/j.system.2013.11.008

Allen, D. D. (2013). *Attitude toward digital and print-based reading: A survey for elementary students* [Doctoral dissertation, University of South Florida]. Scholar Commons. https://scholarcommons.usf.edu/etd/4858/

Anderson, M. A., Tollefson, N. A., & Gilbert, E. C. (1985). Giftedness and reading: A cross-sectional view of differences in reading attitudes and behaviors. *Gifted Child Quarterly*, *29*(4), 186−189. https://doi.org/10.1177/001698628502900411.

Barnett, J. E., & Irwin, L. (1994). The effects of classroom activities on elementary students' reading attitudes. *Reading Improvement*, *31*, 113−120.

Beers, K. (1998). Choosing not to read: Understanding why some middle schoolers just say no. In K. Beers (Ed.), *Into focus: Understanding and creating middle school readers* (pp. 37−64). Christopher-Gordon.

Bokhorst-Heng, W., & Pereira, D. (2008). *Non-at-risk adolescents' attitudes towards reading in a Singapore secondary school* (pp. 285−301). Blackwell.

Calkins, L. (2012). *Pathways to common core: Accelerating achievement*. Heinemann.

Case, R., & Khanna, F. (1981). The missing links: Stages in children's progression from sensorimotor to logical thought. *New Directions for Child and Adolescent Development*, *1981*(12), 21−32. https://doi.org/10.1002/cd.23219811204

Chapman, J., & Tunmer, W. (1995). Development of young children's reading self-concepts: An examination of emerging subcomponents and their relationship with reading achievement. *Journal of Educational Psychology*, *87*, 154−167.

Chi, M. T. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73−96). Erlbaum.

Chi, M. T., & Klahr, D. (1975). Span and rate of apprehension in children and adults. *Journal of Experimental Child Psychology*, *19*(3), 434−439.

Choi, S. (2010). The correlation analysis between reading motivation and attitude toward reading of middle school students. *Journal of Reading Research*, *23*, 345−380.

Common Sense Media. (2014, May 7). *Children, teens, and reading: A Common Sense Media research brief*. https://www.commonsensemedia.org/research/children-teens-and-reading

Davidovitch, N., Yavich, R., & Druckman, E. (2016). Don't throw out paper and pens yet: On the reading habit of students. *Journal of International Education Research*, *12*(4), 129−144. https://doi.org/10.19030/jier.v12i4.9799

Day, R. R., & Bamford, J. (1998). *Extensive reading in the second language classroom*. Cambridge University Press.

DeFries, J. C., Alarcon, M., & Olson, R. K. (1997). Genetic aetiologies of reading and spelling deficits: Developmental differences. In C. Hulme & M. Snowling (Eds.), *Dyslexia: Biology cognition and intervention* (pp. 20–37). Whurr.

Fox, E., & Alexander, P. A. (2009). Text comprehension: A retrospective, perspective, and prospective. In S. E. Israel & G. G. Duffy (Eds.), *Handbook of research on reading comprehension* (pp. 227–239). Routledge.

Frey, N., & Fisher, D. (2006). *Language arts workshop: Purposeful reading and writing instruction*. Merrill Prentice-Hall.

Gerlach, J., & Buxmann, P. (2011). Investigating the acceptance of electronic books: The impact of haptic dissonance on innovation adoption. *In ECIS 2011 Proceedings*, *141*. AIS eLibrary. http://aisel.aisnet.org/ecis2011/141

Gordon, C. A., & Lu, Y. L. (2008). "I hate to read, or do I?" Low achievers and their reading. *School Library Media Research*, *11*, 1–15. http://www.ala.org/aasl/pubs/slr/vol11

Green, P. (2002). Teachers intervention in children's reading. *Journal of Child Hood Education*, *46*(3), 147–149.

Harmer, J. (2004). *The practice of English language teaching*. Longman.

Henk, W. A., McKenna, M. C., & Conraddi, K. (2011). Developing affective instrumentation for use in literacy research. In N. K. Ducke & M. H. Mallette (Eds.), *Literacy research methodologies* (2nd ed., pp. 242–269). Guilford.

Jacobs, J., Lanza, S., Osgood, D., Eccles, J., & Wigfield, A. (2002). Changes in children's self-competence and values: Gender and domain differences across grades one through twelve. *Child Development*, *73*, 509–527.

Jang, S. J. (2006). Research on the effects of team teaching upon two secondary school teachers. *Educational Research*, *48*, 177–194.

Johari, A., Tom, A., Morni, A., & Sahari, S. (2013). Students' reading practices and environments. *Indonesian Journal of Applied Linguistics*, *3*(1), 17–28.

Kamhi-Stein, L. D. (2003). Reading in two languages: How attitudes toward home language and beliefs about reading affect the behaviors of "underprepared" second language college readers. *TESOL Quarterly*, *37*, 35–71. https://doi.org/10.2307/3588465

Khairudin, Z. (2013). A study of students' reading interests in a second language. *International Education Studies*, *6*(11), 160–170. https://doi.org/10.5539/ies.v6n11p160

Korea Culture and Tourism Institute. (2012). *Survey of Korean status of reading*. Ministry of Culture, Sports and Tourism. http://www.korea.net/AboutUs/Ministry-of-Culture-Sports-and-Tourism

Larson, L. C. (2010). Digital readers: The next chapter in e-book reading and response. *The Reading Teacher*, *64*(1), 15–22.

Lee, S. (2009). *A study on self-concept formation and reading attitude change through reading instruction of subject orientation* [Unpublished manuscript]. Chonnam National University.

Lenhart, A., Madden, M., MacGill, A. R., & Smith, A. (2007). *Teens and social media*. Pew Internet and American Life Project. http://www.pewinternet. org/PPF/r/230/report_display.asp

Lenhart, A., Simon, M., & Graziano, M. (2001). *The internet and education: Findings of the Pew Internet and American Life Project*. Pew Charitable Trusts.

Leslie, L., & Allen, L. (1999). Factors that predict success in early literacy intervention projects. *Reading Research Quarterly*, *39*, 72−93.

Liu, Z. (2003). Trends in transforming scholarly communication and their implications. *Information Processing and Management*, *39*(6), 889−898.

Liu, Z. (2006). Print vs. electronic resources: A study of user perceptions, preferences, and use. *Information Processing and Management*, *42*(2), 583−592.

Lone, F. A. (2011). *Reading habits of rural and urban college students in the 21st century*. University of Nebraska − Lincoln. http://digitalcommons. unl.edu/libphilprac/586

Loveless, N. S. (2015). Towards a manifesto on research-creation. *RACAR: Revue d'art canadienne / Canadian Art Review*, *40*(1), 52−54.

Manguel, A. (1997). *A history of reading*. Flamingo.

Mante-Estacio, J. (2012). Dimensions of reading motivation among Filipino bilinguals. *TESOL Journal*, *7*, 10−29.

Martinez, R., Aricak, O., & Jewell, J. (2008). Influence of reading attitude on reading achievement: A test of the temporal-interaction model. *Psychology in the Schools*, *45*(10), 1010−1023.

McKenna, M., & Kear, D. (1990). Measuring attitude toward reading: A new tool for teachers. *The Reading Teacher*, *43*, 626−639.

McKenna, M., Simkin., S., Conradi, K., & Lawrence, C. (2008). *Assessment for reading instruction* (2nd ed.). Guilford Press.

McKenna, M. C., Conradi, K., Lawrence, C., Jang, B. G., & Meyer, P. J. (2012). Reading attitudes of middle school students: Results of a U.S. survey. *Reading Research Quarterly*, *47*(3), 284−286.

McKenna, M. C., Kear, D. J., & Ellsworth, R. A. (1995). Children's attitudes toward reading: A national survey. *Reading Research Quarterly*, *30*(4), 934−956. https://doi.org/10.2307/748205

McKenna, M. C., Stratton, B. D., Grinder, M. C., & Jenkins, S. J. (1995). Differential effects of whole language and traditional instruction on reading attitudes. *Journal of Reading Behavior*, *27*, 19−44.

Merga, M. K. (2014). Peer group and friend influences on the social acceptability of adolescent book reading. *Journal of Adolescent and Adult Literacy*, *57*(6) 472−482.

Mokatsi, R. (2005). *Sharing resources: How library networks can help reach education goals*. Book Aid International. https://pdfs.semanticscholar.org/c5ad/3692f8bfc620b24a54a3c25b5e56a6898b41.pdf

Molfese, V. J., Modglin, A. A., & Molfese, D. L. (2003). Newborn and preschool predictors of second grade reading scores: An evaluation of categorical and continuous scores. *Journal of Learning Disabilities*, *34*, 545−554.

Nation, K. (2006). Reading and genetics: An introduction. *Journal of Research in Reading*, *29*, 1−10.

Nitko, A. J., & Brookhart, S. M. (2010). *Educational assessment of students* (6th ed.). Prentice Hall.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Organisation for Economic Cooperation and Development (OECD). (2010). *PISA 2009 results: Executive summary*. https://www.oecd.org/pisa/pisaproducts/46619703.pdf

Pandian, A. (1997). *Reading in Malaysia*. Universiti Kebangsaan Malaysia.

Pattuelli, M. C., & Rabina, D. (2010). Forms, effect, function: LIS students' attitudes towards portable e-book readers. *ASLIB Proceedings: New Information Perspectives*, *62*(3), 228−244.

Reynolds, W. M., & Miller, G. E. (2003). Current perspectives in educational psychology. In W. M. Reynolds & G. E. Miller (Eds.), *Handbook of psychology: Educational psychology* (Vol. 7, pp. 3−20). Wiley.

Ro, E., & Chen, C. (2014). Pleasure reading behavior and attitude of non-academic ESL students: A replication study. *Reading in a Foreign Language*, *26*, 49−72.

Rose, E. (2011). The phenomenology of on-screen reading: University students' lived experience of digitised text. *British Journal of Educational Technology*, *42*(3), 515−526. https://doi.org/10.1111/j.1467-8535.2009.01043.x

Roser, M., & Ortiz-Ospina, E. (2013). *Literacy*. Our World in Data. https://ourworldindata.org/literacy/

Samuelsson, S., & Lundberg, I. (2003). The impact of environmental factors on components of reading and dyslexia. *Annals of Dyslexia*, *53*, 210−212.

Scarry, E. (2001). *Dreaming by the book*. Princeton University Press.

Schooten, E., Glopper, K. D., & Stoel, R. D. (2004). Development of attitude toward reading adolescent literature and literary reading behavior. *Poetics*, *32*(5), 343−386.

Scholastic. (2013). *Kids and family reading report* (4th ed.). http://mediaroom.scholastic.com/files/kfrr2013-wappendix.pdf

Scholastic. (2015). *Scholastic 2015 Kids and family reading report*. https://www.scholastic.com/parents/books-and-reading/reading-resources/language-and-literacy-milestones/scholastic-2015-kids-and-family-reading-report.html

Schraw, B., & Bruning, R. (1999). How implicit models of reading affect motivation to read and reading engagement. *Scientific Studies of Reading*, *3*(3), 281–302.

Senturk, B. (2015). EFL Turkish university students' attitudes and motivation towards reading in English. *Procedia – Social and Behavioral Sciences*, *199*, 704–712. https://doi.org/10.1016/j.sbspro.2015.07.601

Smith, N., & Robinson, H. (1980). *Reading instruction for today's children*. Prentice Hall.

Spear-Swerling, L. (2004). A road map for understanding reading disability and other reading problems: Origins, prevention, and intervention. In R. Ruddel & N. Unrau, (Eds.), *Theoretical models and processes of reading* (5th ed., pp. 517–573). International Reading Association.

Stanovich, K. (1986). Matthew effect in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, *21*, 306–407.

Thorndike, R. L. (1971). *Educational measurement* (2nd ed.). American Council on Education.

Tveit, Å. K., & Mangen, A. (2014). A joker in the class: Teenage readers' attitudes and preferences to reading on different devices. *Library and Information Science Research*, *36*(3), 179–184.

U.S. Department of Education. (2005). *Helping your child become a reader*. Educational Publications Center. https://www2.ed.gov/parents/academic/help/reader/index.html

van der Weel, A. (2011). *Changing our textual minds: Towards a digital order of knowledge*. Manchester University Press.

Wolk, S. (2010). What should students read? *The Phi Delta Kappan*, *91*, 8–16.

Yamashita, J. (2004). Reading attitudes in L1 and L2, and their influence on L2 extensive reading. *Reading in a Foreign Language*, *16*, 1–19.

Yamashita, J. (2007). The relationship of reading attitudes between L1 and L2: An investigation of adult EFL learners in Japan. *TESOL Quarterly*, *41*, 81–105. https://doi.org/10.1002/j.1545-7249.2007.tb00041.x

Yamashita, J. (2013). Effects of extensive reading on reading attitudes in a foreign language. *Reading in a Foreign Language*, *25*, 248–263.

Yoon, J., & Kim, Y. R. (2008). The developmental trends in middle schoolers' attitude toward reading. *Journal of Korean Language and Literature Education*, *43*, 159–184.

Brief Reports

Constructionist Approaches for Korean EFL Teaching and Learning: From Learner Errors Towards Construction-Based Instruction

Yunjung Nam
Georgia State University, Atlanta, Georgia, United States

This brief report aims to explore the potential pedagogical benefits of Construction Grammar (CxG) approaches in the Korean EFL context, particularly for analyzing learner errors or classroom instructions. Constructions are defined as form-meaning pairings (Goldberg, 1995), ranging from idioms, abstract phrasal patterns, syntactic construction, as well as morphemes and words. The paper first presents a summary of constructs and concepts of CxG in the field of second language acquisition. It is followed by the review of recent studies conducted on the CxG approaches in EFL contexts. The paper then focuses on the Korean EFL context specifically, and explains how CxG can help EFL Korean teachers better understand and address learner errors. A specific type of construction, resultative verb-argument construction, will be discussed to illustrate the potentials of CxG approaches for teaching and learning. Specifically, with CxG approaches, learners do not need to engage in rote learning of arbitrary grammatical rules and exceptions, which has not been helpful in actual proficiency improvement. This paper discusses the implications and provides future suggestions.

Keywords: construction-based instruction, Korean EFL, learner errors

INTRODUCTION

Construction Grammar (henceforth, CxG) is an umbrella term for various linguistic approaches that explore the concept of "constructions" (Ellis, 2013; Goldberg, 2013). CxG grew out of cognitive semantics, including Charles Fillmore's frame semantics and George Lakoff's experientially based approach to language (Goldberg, 1995). The concept

of a "construction" can be traced back to the Saussurean notion of the linguistic sign as arbitrary form–meaning pairings (Hoffmann & Trousdale, 2013). However, the current version of form–meaning pairings of CxG is an extended notion of Saussurean sign in that it covers idioms, abstract phrasal patterns, and syntactic constructions, as well as morphemes and words (see Table 1). These different types of constructions are considered on a lexicon–syntax continuum by Construction Grammarians.

TABLE 1. Constructions at Varying Levels of Complexity and Abstraction

Construction	Meanings or *Examples*
Word	*banana, avocado*
Idiom (filled)	*give the Devil his due*
Idiom (partially filled)	*jog* (someone's) *memory*
Verb-Argument Constructions	
Ditransitive: Subj V Obj1 Obj 2	X causes Y to receive Z: *Pat kicked me a ball*.
Resultative: Subj V Obj Resultative Phrase	X causes Y to become Z: *Pat kicked the door open*.

Note. Adapted from Goldberg (1995, 2006).

The notion of interdependence of lexis and grammar has been introduced to EFL teachers. However, there is still a tendency, it seems, to teach grammar and vocabulary separately in most EFL classrooms (Salem, 2007). The dichotomy of grammar and vocabulary still forms the basic structure of English instruction in Korea. In most cases, grammar rules are presented to learners in grammar class with complex grammatical terminology, and students are required to apply the rules in exercises, using vocabulary they have memorized separately from the class or as homework. This tradition of English classroom instruction has not been helpful for the learners, so EFL classroom teachers have been seeking a better way to help learners improve English proficiency.

In this paper, I will provide a brief summary of constructs and concepts in L2 construction learning. It will be followed by a review of several recent studies on applied CxG in second language acquisition and EFL settings. Finally, this paper will discuss what pedagogical implications CxG approaches or constructionist approaches would provide for Korean EFL teachers.

CONSTRUCTION GRAMMAR AND SECOND LANGUAGE ACQUISITION

L2 learners share the same goal with L1 learners: to understand how a language works based on their linguistics experiences. For both L1 and L2 acquisition, there are major influential factors in acquiring constructions. First, frequency plays an important role in facilitating learning and entrenchment of constructions, which highlights the significant role of input in language acquisition. To be more specific, high token frequency is more likely to promote learning of irregular forms and idioms, while high type frequency strengthens representational schema of morphological or syntactic patterns for further use with new items (Ellis, 2013). However, frequency effect cannot be simply stated as "the more frequent, the faster they learn" because frequency effect might vary depending on distribution patterns and state of the learner's language development (Ellis & Wulff, 2015). Second, contingency of form–function mapping is also important since it has been found to be the catalyst in associative learning (Ellis, 2006b, 2013). If certain form–function pairings show high reliability, learners learn those constructions more easily. Third, prototypicality of meaning, as the most representative instance of a category, is a factor influencing construction acquisition. According to Ellis and Ferreira-Junior (2009), learners were found to learn the more prototypical and generic constructions first, in the process of learning VACs (verb–argument constructions). Finally, salience is a factor related to learned attention for associative learning. Salience is defined as intensity of subjective experience of stimuli (Ellis, 2006b). The more salient a construction, the more attention it receives from learners.

In the process of construction acquisition, L2 learners cannot avoid L1 interference, blocking, and overshadowing. These factors are important to be considered in L2 learning because they pose challenges to L2 learners in the process of construction acquisition (Ellis, 2006a, 2006b). Even frequent cues are overshadowed by the learner's L1 experiences, so the learner does not successfully learn from the input. Nevertheless, L2 acquisition undergoes reconstruction with focused instruction (Ellis, 2013), which helps learners explicitly and consciously process target constructions. Once they consolidate the form–function pairs with explicit instruction, implicit learning plays its role of updating

the statistical information about frequency and contingency of form–function mapping (Ellis & Wulff, 2015).

Therefore, for L2 acquisition of linguistic constructions, both explicit and implicit learning have a great impact on the development of the learner's language system. While L2 learners form generalizations and abstractions based on continuously given input, there are some unlearnable aspects of L2 through implicit learning. Thus, explicit instruction is also necessary for form–function mapping. This warrants consideration of construction grammar in EFL classrooms where explicit instruction and input enhancement should be carefully planned to help learners.

CONSTRUCTION GRAMMAR IN EFL TEACHING AND LEARNING

There have been attempts to apply the concept of CxG to L2 learning and teaching (Holme, 2010; Littlemore, 2009; Tyler, 2012). Both Littlemore (2009) and Tyler (2012) discussed the usefulness of CxG in L2 learning and teaching in the framework of cognitive linguistics. Littlemore (2009) presented a general overview of why and how CxG is potentially useful in second language teaching and learning. Tyler (2012) provided experimental evidence for the superiority of CxG-based instruction over traditional instruction. In Tyler et al.'s (2011) study, the Vietnamese EFL learner participants receiving CxG-based intervention demonstrated significant greater improvement, compared to those with traditional instruction. Holme (2010) also presented a pedagogical experiment conducted in Hong Kong, which found that the CxG approach led to more uptake and improved accuracy in writing.

More recent studies have discussed CxG from pedagogical perspectives specifically in EFL formal educational settings (Herbst, 2016; Torres-Martínez, 2014, 2015, 2017). Herbst (2016) recently suggested seven principles for pedagogical construction grammar for teaching and learning of foreign languages. He presented examples from widely used textbooks in the German EFL context to demonstrate weaknesses of teaching materials in terms of grammatical terminology and its description. Thus, he argued that CxG would help make informed decisions about how much terminology and which grammatical terms

should be used to better help learners. Torres-Martínez (2014, 2015, 2017) combined insights from corpus linguistics and CxG in teaching constructions such as hedging strings (e.g., *kind of*, *sort of*), and multiword verbs (e.g., *pull over*). He suggested that multiword verbs (MWVs) should be considered and learned as chunks, in relation to VACs. Pedagogical tasks for noticing and categorizing the associations between MWV-VACs are presented as those tasks help learners make generalizations of form–meaning mapping, just as how one-word verbs are learned. Both researchers presented the potential benefits of using CxG approach in EFL classrooms.

Along with the trend, application of CxG to Korean EFL context has currently been receiving attention from researchers (Kim et al., 2013; Rah & Kim, 2018; Sung & Yang, 2016; Sung, 2018). Kim et al. (2013) replicated the sentence sorting task (Bencini & Goldberg, 2000) to explore Korean EFL learners' development pattern of constructional knowledge with three different proficiency groups. The results indicated that Korean EFL learners' constructional knowledge incrementally developed as their proficiency increased. Two most recent studies (Sung & Yang, 2016; Rah & Kim, 2018) explored effects of the construction-based approach on teaching resultative constructions, one of the most challenging constructions to Korean EFL learners. Sung and Yang (2016) found that Korean secondary students who received construction-centered instruction had greater improvement in translation tasks than those who received form-centered instruction. They also suggested the facilitative role of light verbs (e.g., *make*) and positive network effects of teaching marked construction in resultative construction learning. Rah and Kim (2018) conducted a similar study with college students in Korea, focusing on network effect. Their findings indicated that construction-centered instructions were more effective than form-centered instructions in improving participants' performance. Also, participants were found to perform better when they were given a target construction in relation to other constructions (e.g., *network*) than when presented with a single target construction separately. Sung (2018) investigated the effects of CxG-based instruction on a younger population group, Korean middle school students. It was found that CxG-based instruction helped the middle school students improve correct production of verb–participle constructions and even learn complex structures including figurative and marked target constructions.

The studies mentioned suggest the effectiveness of construction-

centered instruction in teaching constructions to EFL leaners. This presents a new possibility of utilizing the construction-based approach to facilitate Korean EFL learners' development of English constructions. In line with this new possibility, the next section will focus on how the CxG approach can help Korean EFL classroom teachers.

CONSTRUCTIONIST APPROACHES FOR KOREAN EFL TEACHING AND LEARNING

In this section, I will illustrate how CxG can help teachers better understand and address learner errors, using one type of construction as an example. Then, I will explain why utilizing construction-based instruction would be beneficial for Korean EFL teachers.

Understanding and Addressing Learner Errors from the Korean EFL Context

Errors are unavoidable in foreign language learning. However, learner errors are important since they may signal lack of mastery in specific learning targets. They function as signposts for what is missing or what should be addressed in learner language. Based on the students' error rate or error patterns, the teacher can plan for more effective future lessons accordingly and develop more effective teaching and learning materials. In this section, one useful type of student error will be discussed from the perspective of CxG approaches.

One important type of construction is argument structure constructions, pairings of form and function for basic clauses (Goldberg, 2006). Some representative argument structure constructions of English include intransitive motion, caused motion, ditransitive, and resultative (Goldberg, 2006, p. 73). Among those argument structure constructions, it is quite well known that the resultative construction (*X causes Y to become Z state*) is the most difficult for Korean EFL learners.

The difficulty might be attributable to the differences between Korean and English. The Korean resultative construction is a paring of the form Subj-Obj-Resultative Phrase-V and the meaning *X causes Y to become Z state* (Sung & Yang, 2016). In this form, the Korean

resultative phrase is marked by an adjunct particle *-key*, similar to *-ly* of English and the prototype verb is *mandul-da* (E. *make*). This contributed to L1 transfer, which was observed in producing resultative constructions in Kim's study (2016). Korean learners demonstrated the tendency to use adjuncts to produce resultative constructions, which led to errors, like **comb her hair smoothly* or *comb her hair to make it smooth* when the intended target answer was *comb her hair smooth*. Thus, without signal-like words with *-ly* or the verb *make*, it might be difficult for them to perceive a given statement as a resultative construction unless they are explicitly taught with construction-based instruction.

The difficulty could be further elaborated by considering possible responses to an English-to-Korean translation task. If Korean learners were asked to translate one of the least prototypical examples for a resultative construction, *She kissed him unconscious* (Goldberg, 2006), a couple of possible interpretations would include:

1) She gave him an unconscious kiss.
2) She unconsciously kissed him.

The possible comprehension errors above suggest that Korean learners might use the strategy of word-based translation rather than construction-based processing. However, the actual meaning of the given sentence is *She caused him to become unconscious by kissing*.

In most instructional settings in Korea, the resultative constructions are taught with form-centered instruction. Most learners may learn the structure S-V-O-OC, often with mentioning of complex transitive verbs (i.e., *find*, *make*; Greenbaum & Quirk, 1990). The verb *kiss*, from the task above, does not belong to the category, so it would be extremely difficult – if not impossible – to comprehend the intended meaning within this framework. EFL teachers probably have no choice but to tell their students that it is an exception to the rule. However, with the construction-based approach, it would be easier to provide students with more meaningful and systematic explanations by directing their attention to construction-level rather than word-level meaning.

Towards Construction-Based Instruction

The previous section discussed why the CxG approach would be a better option for teachers when understanding learner errors and

providing meaningful and systematic explanations to students regarding the errors. Based on the discussion above, it is clear that the construction-based approach would be superior to the traditional approach in teaching and learning.

However, traditional form-centered instruction is still prevalent in Korean EFL classrooms. Teachers still rely heavily on traditional grammatical categories, such as object, complement, and so on. Of course, form-centered instruction would be more helpful for EFL learners than no instruction at all, but the major problem of this approach is that it does not adequately account for the behavior of words in context (Littlemore, 2009). In classrooms, arbitrary grammar rules are presented to students as something they should memorize. The meaningful form–function mapping does not have due attention. This might be the reason why many Korean EFL learners have a difficult time understanding and producing constructions like resultative constructions.

Since CxG does not downplay either form or meaning, teachers can provide "meaningful" accounts of certain "grammatical" phenomena at the same time. Teachers can explain why certain words work better with a certain construction in a more meaningful way and how seemingly similar constructions are different (Littlemore, 2009). Instead of presenting arbitrary grammar rules to students, teachers can provide more effectively learnable and meaningful explanations in class using the CxG approach.

IMPLICATIONS, LIMITATIONS, AND FUTURE RESEARCH

In applying the CxG approach in the Korean EFL classroom, the following topics are some of the aspects that should be carefully considered and examined in terms of pedagogical implications. Limitations and future research directions will also be discussed.

L1 Transfer

There might be some negative transfer from L1 constructions to L2 constructions, which hinder L2 construction learning. For example, to express transfer of possession, the Korean language uses case markers while word order does not determine argument roles in ditransitive

constructions. However, in English, Korean L1 learners need to pay attention to the new linguistics cues, word order. Thus, it is common to fine errors like *Peter gave a book Karen*. (Year & Gordon, 2009). Teaching activities should help leaners to unlearn, or de-entrench, L1-based form–function parings (Della Putta, 2016) since it would be difficult for learners to get rid of the L1 influence by themselves. Korean EFL teachers should develop activities and learning materials carefully considering this possible L1 transfer.

Regarding L1 transfer among Korean EFL learners, most recent studies are still limited to argument structure constructions such as ditransitive or resultative constructions. Though they are the most challenging learning targets for Korean learners, other constructions could also be explored to provide a more comprehensive account for construction learning. To examine Korean learner's knowledge of English constructions and suggest more effective pedagogical strategies, researchers could utilize Korean learner corpus data, conduct gap-filling experiments (Römer et al., 2014), or conduct a longitudinal usage-based study (Eskildsen, 2012) in Korean EFL classrooms.

Usage-Based Approach and Input

Goldberg (2006) suggested that construction learning could be enhanced by prototypical exemplars with high frequency (i.e., input skewed to prototypical exemplars). In the usage-based approach, frequency, contingency, and prototypicality of input are considered as important. For those three factors mentioned so far, Römer et al.'s (2014) study found a similarity between L1 and L2 construction learning: Both were affected by type and token frequency, contingency of form–function pairing, and prototypicality. However, unlike L1 acquisition, usage or frequency itself is not sufficient for L2 acquisition. In the EFL context, a large and representative input is necessary for learners to build a rational model and strengthen the association of form–meaning. Furthermore, the facilitative effect of skewed input in SLA has been inconclusive with contrasting study findings (Year & Gordon, 2009). Thus, teachers should be aware of the fact that the role of prototypical exemplars as input in L2 classroom instruction might be influenced by various other factors. Future research could examine how construction-based input and different factors interact with each other in a usage-based framework.

Network of Constructions and Use of Metalanguage

Constructions are related in a hierarchical network (e.g., subject–predicate construction). Rah and Kim (2018) suggested that presenting the target construction (e.g., resultative construction) in a network can lead to better scores among Korean EFL learners. While it would be a good idea to present the network and relationship between/among constructions, use of metalanguage to explain constructions might add unexpected complications. Instead of using metalanguage, visualization of the prototype verbs with clausal constructions could be an option, as described by Tyler (2012). However, Herbst (2016) suggested that grammatical terminology should be limited to a minimum number of useful ones, where the use of grammatical terminology is conducive to understanding different means of expressing similar meanings in different languages. Thus, for researchers and classroom teachers in the Korean EFL context, how much and which terminology would be helpful still remains to be explored and decided.

CONCLUSIONS

In conclusion, this paper attempted to explore the potential pedagogical benefits of CxG in the Korean EFL context for analyzing learner errors or classroom instructions. With CxG, learners do not need to involve themselves in rote learning of arbitrary grammatical rules and exceptions, which has not been helpful in actual proficiency improvement. It can provide teachers with directions for developing materials and lesson planning to make their instruction more effective for language learning. However, there are some prerequisites for successful application of construction-based instruction to Korean EFL classrooms. National curriculum standards and guidelines should reflect construction-based instruction, with insights from corpus linguistics and the usage-based approach. Textbooks also should be carefully analyzed and revamped if necessary. More importantly, good quality training for pre-service and in-service teachers should be administered so that teachers are equipped with theoretical foundations and necessary pedagogical strategies for the classroom.

THE AUTHOR

Yunjung Nam is a PhD student in the Department of Applied Linguistics and ESL at Georgia State University in Atlanta, Georgia, U.S.A. Her research interests include writing assessment, language assessment literacy, classroom assessment, and teacher development. She used to be an EFL high school teacher in South Korea. Email: ynam3@gsu.edu

REFERENCES

Bencini, G. M. L., & Goldberg, A. E. (2000). The contribution of argument structure constructions to sentence meaning. *Journal of Memory and Language*, *43*(4), 640–651. https://doi.org/10.1006/jmla.2000.2757

Della Putta, P. (2016). Do we also need to unlearn constructions? The case of constructional negative transfer from Spanish to Italian and its pedagogical implications. In G. Gaëtanelle & S. De Knop, (Eds.), *Applied construction grammar* (pp. 21–51). De Gruyter Mouton.

Ellis, N. C. (2006a). Language acquisition as rational contingency learning. *Applied Linguistics*, *27*(1), 1–24. https://doi.org/10.1093/applin/ami038

Ellis, N. C. (2006b). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, *27*(2), 164–194. https://doi.org/10.1093/applin/aml015

Ellis, N. C. (2013). Construction grammar and second language acquisition. In T. Hoffman & G. Trousdale (Eds.), *The Oxford handbook of construction grammar* (pp. 365–378). Oxford University Press.

Ellis, N. C., & Ferreira-Junior, F. (2009). Construction learning as a function of frequency, frequency distribution, and function. *The Modern Language Journal*, *93*(3), 370–385. https://doi.org/10.1111/j.1540-4781.2009.00896.x

Ellis, N. C., O'Donnell, M. B., & Römer, U. (2014). Second language processing of verb–argument constructions is sensitive to form, function, frequency, contingency, and prototypicality. *Linguistic Approaches to Bilingualism*, *4*(4), 405–431. https://doi.org/10.1075/lab.4.4.01ell

Ellis, N. C., & Wulff, S. (2015). Usage-based approaches in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in Second language acquisition: An introduction* (2nd ed., pp. 75–93). Routledge.

Eskildsen, S. W. (2012). L2 Negation constructions at work. *Language Learning*, *62*(2), 335–372. https://doi.org/10.1111/j.1467-9922.2012.00698.x

Herbst, T. (2016). Foreign language learning is construction learning: Principles of pedagogical construction grammar. In G. Gaëtanelle & S. De Knop, (Eds.), *Applied construction grammar* (pp. 21–51). De Gruyter Mouton.

Hoffmann, T., & Trousdale, G. (2013). Construction grammar: Introduction. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford handbook of construction grammar* (pp. 1−12). Oxford University Press.

Holme, R. (2010). A construction grammar for the classroom. *IRAL: International Review of Applied Linguistics in Language Teaching*, *48*(4), 355−377.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford University Press.

Goldberg, A. E. (2013). Constructionist approaches. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford handbook of construction grammar* (pp. 15−31). Oxford University Press.

Greenbaum, S., & Quirk, R. (1990). *A student's grammar of the English language*. Longman.

Kim, S. (2016). *L1 influence on L2 learning of English resultative constructions: The syntactic and semantic structure of Korean students' interlanguage* [Unpublished doctoral dissertation]. Seoul National University.

Kim, H., Choi, H., & Yang, H. (2013). Developmental patterns of Korean EFL learners' English argument structure constructions. *Social and Behavioral Sciences*, *97*, 397−404. https://doi.org/10.1016/j.sbspro.2013.10.251

Littlemore, J. (2009). *Applying cognitive linguistics to second language learning and teaching*. Palgrave Macmillan.

Rah, Y., & Kim, H. (2018). Construction-based approach to teaching the English resultative construction to Korean EFL learners. *System*, *72* (Supplement C), 1−12. https://doi.org/10.1016/j.system.2017.09.027

Römer, U., O'Donnell, M. B., & Ellis, N. C. (2014). Second language learner knowledge of verb−argument constructions: Effects of language transfer and typology. *The Modern Language Journal*, *98*(4), 952−975. https://doi.org/10.1111/modl.12149

Salem, I. (2007). The lexico-grammatical continuum viewed through student error. *ELT Journal*, *61*(3), 211−219. http://ezproxy.gsu.edu/login?url= https://search-proquest-com.ezproxy.gsu.edu/docview/85665515?accountid=11226

Sung, M., & Yang, H. (2016). Effects of construction-centered instruction on Korean students' learning of English transitive resultative constructions. In S. De Knop & G. Gilquin (Eds.), *Applied construction grammar* (pp. 89−114). De Gruyter.

Sung, M. (2018). *Effects of construction-grammar-based instruction on the learning of English verb-particle constructions by Korean middle school students* [Unpublished doctoral dissertation]. Seoul National University.

Torres-Martínez, S. (2014). Introducing conversational grammar in EFL: A case for hedging strings: Bringing insights from corpus linguistics and construction grammar into the English language classroom. *English Today*,

30(2), 24−32. https://doi.org/10.1017/S0266078414000091

Torres-Martínez, S. (2015). A constructionist approach to the teaching of phrasal verbs: Dispelling the verb+particle myth in multiword verb instruction. *English Today*, *31*(3), 46−58. https://doi.org/10.1017/S0266078415000255

Torres-Martínez, S. (2017). Working out multiword verbs within an applied cognitive construction grammar framework. *European Journal of Applied Linguistics*, *5*(1), 55−86. https://doi.org/10.1515/eujal-2016-0003

Tyler, A. (2012). *Cognitive linguistics and second language learning*. Routledge.

Tyler, A., Ho, V., & Mueller, C. (2011). Applying cognitive linguistics to instructed L2 learning: The English modal verbs. *AILA Review*, *23*, 30−49. https://doi.org/10.1075/aila.23.03tyl

Year, J., & Gordon, P. (2009). Korean speakers' acquisition of the English ditransitive construction: The role of verb prototype, input distribution, and frequency. *The Modern Language Journal*, *93*(3), 399−417. https://doi.org/10.1111/j.1540-4781.2009.00898.x

Misinformation and Stereotype Perpetuation: A Recount of My Experiences with Textbooks Used in the Korean Private Language Academy Arena

Travis Frank
Banana Kids English Academy, Seoul, Korea

INTRODUCTION

As teachers, it is very important for us to not take things as a given or at face value. Rather, we are required to go further than being merely acceptors of the way things are, and to be active scrutinizers and agents of change in the systems and frameworks that we find ourselves placed in.

Though I have known this to be an irrefutable part of our responsibilities as South African teachers residing and working in our home country, the need to uphold the aforementioned ethos of activism and to be a proponent of change for the better is something that I have felt most prominently over the past four years, in which I have been based in the Republic of Korea as an English teacher. My main motivation in this regard has stemmed from an analysis of the issues of racism, averseness to otherness, the emboldening of cultural appropriation and stereotypes, and the oppression of women that is to be found in the textbooks that are commonly used and endorsed in the Korean private language academy arena. Whilst the books cannot be criticized on the basis of grammar and technical accuracy in terms of hard and cold proofreading and editing, this in itself should not render the materials as being suitable for pedagogical purposes − particularly in the modern ESL classroom, which acts as an interface or bridge between cultural, ethnic, and linguistic differences.

My acute awareness of the issues pertaining to the books in question began only in my current occupation as a matter of evaluating the materials I have on hand at the moment to teach as well as those that

I have used in my previous occupations or roles undertaken at different language academies in South Korea. The commonplaces that are to be found in the textbook series that are used is too compelling to overlook, suggesting an underlying trend of narrow-mindedness and misinformed approaches on the parts of authors and publishers. I have taught kindergarten, elementary, and middle school students in the private language academy arena, and the books that have been used fall short of the needs of diversity, equality, and accurate subculture representation and sensitivity in many respects.

PRESCHOOL TEXTBOOK ANALYSIS

The kindergarten books that I have used focused more on storytelling as a means of imagination development and vocabulary acquaintance; reading fluency is also an important objective. The first of two series that I would like to bring under the spotlight here for analysis purposes is the Oxford Reading Tree (ORT, n.d.) series. The main problem that I have with this series of books is that it does not accurately represent modern British society. The lack of representation of minorities is alarming, as is the lack of representation of the multiple ways in which people from different cultures can interact with one another. For instance, I am yet to see intermarriage as being incorporated as a viable option that can even be presented to students as a possibility, let alone as an actual reality in modern society. This concept is central to the educational experience of children when they deal with Western teachers who are either dating, engaged to, or actually married to Koreans. This was the case at one of the schools that I worked at, where I saw my British supervisor go through all of the above-mentioned motions, and in the course of the entire year that I worked there, he never made an appearance with her at the school! The fact that the textbooks we used didn't embrace the concept could have, in some inadvertent way, acted as dissuasion to them both from making such a public appearance, which is further aided by the fact that such discourse didn't permeate the classroom environment. Drawing all of this back full swing to the *Wonders* (n.d.) textbook series, intercultural friendships and relations of all sort were very poorly explored, but, in the Korean educational and sociological setting, such exploration could have served

as the initial catalyst in the forming of a sense of otherness that young minds feel and that cannot be broken or circumnavigated.

The other problem that I had with the syllabus taught at one of my former schools is the fact that the Disney Books (n.d.) that were taught focused their attention on Western – and predominantly white – characters, largely outweighing the focus that could have been placed on books that had Asians and other people of color as their main protagonists. A healthy balance could have gone much further than the skewed tendency to over-focus on one ideal and marginalize the rest. So, even though we did read the likes of *Mulan*, *Aladdin,* and *Pocahontas*, the students had already been conditioned to see the characters as being one-dimensional in terms of race, ethnicity, language, and culture when it comes to formulating the idea of what constitutes the ideal literary and visual embodiment of a cartoon character. The mere suggestion that these "givens" that shaped, for instance, how Cinderella was typically depicted, seemed preposterous to students when I suggested that we try to create our own African, Asiatic, Pacific-Islander, and other ethnic versions of the character in question. For me, it felt very difficult to try to bring students to the awareness that the imagination of the reader should remain sovereign when it comes to shaping subjective consciousness and interpretation of a literary text, which is a process that is naturally imbued with a great deal of prejudice and personal upbringing factors – especially at that age. Yet, how might those things be shaped through words, images, and real tangible examples of people's lives as represented to them by their own personal teachers? We have a big part to play in selecting the right texts (or creating our own), and these texts should embody our lives and the influence that our students and loved ones, on an intercultural basis, have had on our lives.

As a result of misrepresentations of race in all its complexities in the books that children are exposed to (a factor that is compounded by the homogeneity of the society unto which they belong and, in many cases, their lack of interaction with foreigners), it is difficult to teach students about what *race* actually means. Further, classifying racial groups and explaining the origins of terminologies that may be considered either to be offensive or acceptable in different parts of the world also becomes quite a stifled endeavor in the face of narrow racial demarcation parameters. In most instances, the only three racial groups that the students I teach can identify are white, black, and Asian. This presents a bit of a problem when trying to introduce the concepts of Latino,

Pacific Islander, mixed race, mulatto, Hispanic, and a host of other terms that would shed more light on just how intricate the topic is (and more pertinently, how pivotal it is to use the correct term when describing or addressing somebody).

This oversimplification of the concept of race stems from an ESL downfall akin to the manner in which language as a whole is increasingly watered down so that something easy to remember and easy to pronounce is used to describe intricate concepts. This is evidenced daily in our interactions with students, in which the question "How are you?" is usually met with the overplayed and clichéd "I'm fine" being offered as a sole response. The thesaurus provides us with a plethora of synonyms that could alternatively be offered in lieu of *fine*, including *all right*, *good*, *content*, *satisfied*, *pleased*, *doing well*, etc. Further, the need to foster good relations and not enter into what would otherwise be deemed as "delicate discourse" is something that is quite uncommon in Korean culture, where an attitude of non-addressing of touchy issues is normally viewed as more preferable than addressing matters head on to shed more light on a problem. So, how is it possible to incorporate the topic of race into our textbooks (or at least into the discussions that we have in our classrooms to unpack the textbooks unto which lessons are usually anchored) in a meaningful, informative, and non-offensive and non-hostile manner?

Perhaps the answer can be found at the elementary level; if the topic is too sensitive to discuss with kindergarten students, who are usually bound not to understand the issues at stake in any event, perhaps efforts should be directed towards bringing students at the elementary level to a understanding of the topic. This is primarily because students start to question things for themselves more readily and more independently at this age. "Why am I black and she is white?" "Why do men do that?" "Why do women do this?" "How is language gender-based − and should it be to begin with?" All these types of questions − and more − come to the fore at the elementary stage. However, though these queries are wonderful, I find that, all too often, the textbooks that the students are exposed to are not addressing and delimiting these sociological phenomena by assigning labels to them and actually allowing them to enter into academic discourse to begin with.

One of the main examples of this omission of treatment of social phenomena that I am currently encountering is with the ORT series and the gaps that exist between the diversity of modern British society and

the lack of representation of such dynamics in the books. Ethnicity is not explained to students, even though different races appear in the books (with the white family consisting of Mom, Dad, Kipper, Chip, and Biff as the central protagonists; Wilf, Wilma, and their parents as the black family; Anneena and Nadim as the Pakistani-British children; and Lee, Lin, and Grandpa Chen as the Chinese immigrants). Yet, I only know this because I extrapolated such information from my own pre-existing general knowledge, which, in itself, lends itself to a certain amount of unfounded assumptions (e.g., the Chinese immigrants may indeed be Taiwanese, as intended by the unstated intentions of the authors).

As for the black children in the ORT stories, no mention is made of their ethnic origin, leaving it up in the air as to whether or not they happen to hail from African or Caribbean descent. Interestingly enough, assuming that the characters of color happen to be Chinese, Pakistani, and for the sake of argument, let's say Jamaican, it would really suffice to make mention of the fact that all of these countries happen to be former British colonies! Once they begin to understand how these power dynamics work, students will begin to get a better and richer awareness of the political, linguistic, economic, social, and postcolonial factors that underpin the development of modern British society, and even more importantly, racial relations across the globe. I feel that Korean students can be very sympathetic unto this plight given their strained and troubled relations with Japan on a postcolonial level.

As for the characters that are to be found in ORT, there are a few problems that are to be noted here. Too little of the spotlight is given to the Pakistani and Chinese students, which by extension is a bit of an insult, on an indirect level, to the Korean students that the system we teach under purports to serve. If anything, it emboldens the stereotype that Asians are by nature silent, not opinionated and submissive, rather just being there on scene to be seen but not heard. Inadvertently, this also creates an overemphasis on the narrow fissure between black and white as the focus shifts between the black and white families who do the majority of the speaking and focalizing. The majority of the stories are about the lives of the white family, and an overemphasis is placed on the events that take place in this household as opposed the dynamics of other households.

Another problem is that the only teacher in the school that all these children attend is a white female. I know that, like South Africa due to colonial influences, Britain's primary school grades have one teacher

assigned to a class to cover all the subjects, but this is not my complaint; what doesn't rest well with me is that they could have chosen to incorporate teachers of color, but they didn't. Perhaps the storyline could have had some of the children in different grades with a different teacher – there are endless possibilities if one wants to embrace and showcase diversity. Yet, it seems that diversity and intercultural celebration is not at the forefront of the authors' and publishers' minds in the development of this series. As concerns teaching about the changing nature and dynamics of the family unit – a foundational springboard for education at the kindergarten level – there are many realities that are well overlooked. Why are none of the children mixed race? Why does interracial marriage not have a place in the literature? Further, there is also a concern about the manner in which Wilf (an abbreviation of "Wilfred," I would imagine) and Wilma were named much in the same manner as Asian parents stereotypically assign English names to their children that sound very similar (e.g., Karl and Karen, or Lenny and Lena). Though this, in itself, is not a problem, and it probably does serve to bolster a great sense of solidarity and identity amongst Korean parents wanting to connect with the text and its merit. The problem lies in the fact that only the two black characters are used as experimental figures to apply this naming system, whilst characters from other cultures, races, and ethnicities are not subjected to the same naming system. Equality is, quite paradoxically, of equal import as diversity in some respects, and it begs the question as to what criteria was applied to select the black characters as the only eligible ones for such purposes.

The final major problem that I have with the ORT series pertains to the manner in which insufficient attention is paid to showcasing the plight of the British career women, as only Mrs. May is shown as being economically active – and she is either on the verge of approaching retirement or already over the hill and still teaching! Wilf and Wilma's mom and Chip, Biff, and Kipper's mom – both of whom must at least be twenty years the junior of Mrs. May – however, are shown as being housewives (only by extrapolating the fact that they never go to work and, though the fathers in the stories are also not shown as having jobs, the patriarchal aspect of the British job market that is still prevalent leads one to assume that only the male figures are the breadwinners). Though this may not be the authors' intentions (the series involves a joint authorship), they didn't do a stellar job of spelling out that the mothers have jobs and professions – which is a big problem because the

children look around the ages of ten to twelve years old, and their mothers don't need to be at home to care for them on a full-time basis.

ELEMENTARY SCHOOL TEXTBOOK ANALYSIS

As we move on to the elementary school aspect of the equation, new book series come into the picture. Specifically, I would like to make an informed analysis of the books that I have encountered whilst teaching from Lyceum Learning's The Fun Club series. The main problem in this series is the fissure that it creates on a visual level between different racial representations either in illustrations or photographs. On the topic of photographs, I am yet to teach one book from the series where there is a person of color appearing in the form of a photograph! The only photographed people are white people, and this is very alarming. Characters of color are only represented as illustrations, and it is such a subtle way of dehumanizing minorities that many students don't see at first how people of color are reduced in terms of personality and identity so as not to be deemed even eligible to appear in photographs. The series also makes many stereotypical violations in terms of education levels, occupations, gender roles, and the like. One of the most poignant depictions of such violations comes from the book *The Fun Club Goes to the Post Office* (Shaffer, 2009). In this book, The Fun Club consists of non-white students − one Asian girl, one Hispanic boy, and one black boy − being taken on a field trip to the post office by their white teacher, Ms. Dimple. The employee at the post office is a black lady, and the disparity is very sharp between the post office employee and Ms. Dimple, emboldening the tensions of class, social status, and race in stereotypical and prejudicial parameters. What is more alarming, however, is to see Korean students reaffirm the merit of such arguments in most instances, arguing that white teachers are better, more intelligent, and so on, and that black people are more adept at athletic pursuits or manual labor! These books are the antithesis of the embracing of humanity and diversity, and they can only serve to impede teachers who are trying to change misguided and pre-existing mindsets that have been warped by culture and mass media. The problems that are encountered at the kindergarten level concerning the lack of representation of interracial friendships, marriages, business association,

academic relations, and a host of other normal social bonds are also found in this series. The patriarchal aspect of the male figures being leaders and decision-makers is still a concurrent theme that runs through all of the series' books, and it is something that overshadows the accomplishments of single career women in the stories, such as Ms. Dimple.

MIDDLE SCHOOL TEXTBOOK ANALYSIS

The final part of this recount will deal with the middle school syllabus that I was teaching for a time before it became too monotonous and pointless to teach. The book series in question is entitled *Side by Side Plus* (Molinsky & Bliss, 2015). I have dealt with two of the books of the series to date – Books 2 and 3 – and they have not made any headway at all in more closely approximating representing the intricacies of what it means to live "side by side" in all its infinite possibilities and intricacies. Economic disparities between different race groups in the series become more compounded by representations of the income and class divergences between U.S. citizens and immigrants, not taking account of the fact that many immigrants to the United States have gone on to become very wealthy, affluent, and influential. Intermarriage is not shown at all, and this is so very unimaginable to not depict when it comes to the United States, which has one of the highest rates of interracial marriages in the world!

What was also surprising for me to see is how homosexual couples are not represented nor given a voice at all. The students I teach at this level are regularly asking me questions about transgender realities and the prevalence of homosexuality or lesbianism in South Africa, and I always do my best to address the queries in a factual yet professional manner that nips any raunchiness or lewdness in the bud (students sometimes really don't know where to draw the line when it comes to being invasive and prying with their line of questioning for the better part). Why these textbooks cannot do the same is beyond me, save for the fact that I know Korean parents will not, in most cases, readily approve of such discourse taking place in the classrooms to which their children belong. This is fueled by the fact that the state of the Republic of Korea is not willing to embrace homosexuality or transgender

identities as orthodox and acceptable ways of life. This is pandering to the power that parents have in determining the educational experience that their children undergo – without even consulting native English teachers about how syllabi should be tailored to accurately inform students of the nature of the Western societies their children are bound to come into contact with in some way, shape, or form.

To return our attention to the textbooks at hand (as the Korean private language education system and its inner politics fall slightly outside the ambit of this paper's scope), there are a few more aspects with which I was quite displeased regarding the illustrations incorporated. The manner in which minorities are represented as homogeneous in terms of behavior, marital status, and not forming the focus of conversation snippets, but rather only being the characters who answer questions posed to them, really doesn't sit well with me. It is an element that I have found to be common amongst many of the textbooks that I teach from as well. The way in which all the African-American characters have Afros – and only in one color, namely black – is so removed from what modern African-Americanism is all about! No weaves, no dyed or straightened hair, no outlandish clothing, no subcultures, and no means of personal expression: Everyone just gets painted with the same brush. This also applies to the white characters in the books – men don't have long hair, no tattoos or piercings are to be found, no woman has short hair or is bald, etc. – all the characters look very conservative and preppy.

My greatest concern with this textbook series, however, is the manner in which it illustrates elitism and individuality as an exclusively Caucasian or white American societal attribute that is virtually inaccessible to anyone else. Hence, it only displays rich white people relaxing in cafes and engaging in haute couture activities such as visiting art galleries, never showing people of color engaging in such activities. It is also worth noting that the series doesn't make any mention of political, religious, ethnic, socioeconomic, legal, and diaspora-related factors that have given rise to America being called the "melting-pot" nation. This is a serious problem, given the fact that most Korean students who I have spoken to aspire to relocate to America one day either for the purposes of studying, working, or settling down with a partner and starting a family (if they had to choose any other country besides Korea, that is).

Hence, the begging question is "What is the education system doing

to prepare such students adequately for what awaits them across the ocean? Wouldn't it be better if cultures and societies were displayed and depicted more accurately by the people who were actually born in those countries or who had at the very least lived there for a lengthy spell?" What surprises me even more is either U.S. citizens or U.S. residents wrote these books (a joint authorship applies here as well, which is also just as shocking as how the ORT series was written by two British citizens!). So, is it also a case of making the books bland and utterly removed from the true state of affairs concerning modern American society and the places that immigrants and minorities hold in it? What is the point of marketing textbooks that try to reflect the extent to which foreign cultures mirror and embody the ideals of Korean culture? What about diversity and appreciating difference?

CONCLUSIONS: BOLD STEPS FORWARD

The way forward needs a few bold steps to be taken by us as native English teachers as well as by Korean teachers of the English language whom we work with and need insight from in certain instances. Collaborative efforts need to be made to develop in-house academic textbook publications that are more reflective of the diversity of different societies, subcultures, religions, histories, languages, rituals, social interactions, and the like. The more we keep students in a stereotypical and misguided frame of mind, by shackling them to outdated and socially irrelevant or misinformed textbooks, the more they won't be able to see reality for what it truly is – let alone change the mindsets that they approach the classroom with. It will also be beneficial to incorporate colloquialisms, dialects, and a range of other local and regional factors to show how varied and organic English as a dynamic international language is, to show that it is very susceptible to and reflective of local and inter-linguistic influences as opposed to being a predictable and standardized plug-and-play language that is spoken in the same manner the world over. We can learn a lot from the shortcomings of these textbooks – and we have the power to change things for the better!

THE AUTHOR

Travis Frank is a creative writer, and his interests include memoirs, biographies, diaries, letters, poetry, short stories and novellas, and literature as a bridge in cultural communication. The examination of the content of the text in terms of stereotypes and cultural appropriations also falls within the ambit of his research. He received a Master of Arts in TESOL from Woosong University. Email: travis.frank14@gmail.com

REFERENCES

Disney Books. (n.d.). *Disney Publishing Worldwide*. https://books.disney.com/

Molinsky, S. J., & Bliss, B. (2015). *Side by Side Plus*, Books 1–4. Pearson Education ESL.

Oxford Reading Tree. (n.d.). Oxford University Press. https://global.oup.com/education/content/primary/series/oxford-reading-tree/?region=international

Shaffer, J. (2009). *The Fun Club goes to the post office* (The Fun Club Series: Leveled Literacy Intervention). Heinemann.

Wonders Reading Program. (n.d.). *Wonders*: A K-6 literacy curriculum. McGraw Hill. https://www.mheducation.com/prek-12/program/microsites/MKTSP-BGA07M0/wonders.html

Book Review

Classroom Writing Assessment and Feedback in L2 School Contexts

Reviewed by YunDeok Choi

INTRODUCTION

Writing effectively and competently in English is one of the most desired academic abilities by second language (L2) students. Nevertheless, it is a long-standing issue among many stakeholders, particularly writing teachers. Of particular concern is how to assess student writing so that the assessment can lead to an improvement in their learning of writing in schools. *Classroom Writing Assessment and Feedback in L2 School Contexts*, written by Icy Lee, addresses the topic situated in pre-college contexts. Lee, an L2 writing teacher-educator and researcher, wrote this compact volume with the purpose of examining how classroom writing assessment and provision of feedback can be best practiced as a catalyst for primary and secondary school ESL/EFL students' learning of writing. From the Introduction, the author outlines her hope that the book will be a great resource for three groups of readers: (a) writing teachers needing practical advice and tips for assessing writing and providing feedback to young L2 learners, (b) teacher educators training L2 school practitioners in efficient classroom writing assessment and feedback procedures, and (c) L2 writing researchers wishing to identify directions for future studies on classroom assessment and feedback in the domain of L2 writing.

SUMMARY

The book consists of ten chapters, and, after the introductory chapter, the following nine chapters are well organized according to four interconnected themes: L2 classroom writing assessment (Chapters 2 to 4), feedback in L2 writing (Chapters 5 to 7), use of portfolios and technology in support of feedback as new writing assessment tools in L2 classroom contexts (Chapters 8 and 9), and L2 classroom writing teachers' assessment and feedback literacy (Chapter 10).

Lee opens Chapter 1 by outlining the main purposes of the book and establishes the background for the remaining chapters. Setting the scene, the author makes a clear distinction between three approaches to assessment: assessment for learning (AfL), assessment as learning (AaL), and assessment of learning (AoL). She clearly states that AfL/AaL, an integrated part of instruction intended to enhance learning, is the focus of the book. AaL, which is a sub-part of AfL, places heavy emphasis on the learner's role in classroom assessment and learning. As opposed to AfL/AaL, AoL reflects traditional views on assessment as a measurement tool used for assigning grades and serving an accountability function. In addition, feedback is conceptualized in accordance to three stages of learning: where I am going (i.e., "feed up"), how I am going to get there ("feed back"), and where I am going next ("feed forward"; Hattie & Timperley, 2007). The three approaches to assessment and conceptualization of feedback, aligning with the sequential learning stages, are continuously reiterated throughout the entire book. The opening chapter ends with an overview of subsequent chapters.

Building on Chapter 1, the next three chapters center on the learning-oriented approaches to classroom writing assessment (AfL/AaL) compared with AoL. Chapter 2 scrutinizes main goals and theoretical bases for AfL/AaL in the classroom and discusses guiding principles for healthy assessment practices in a writing classroom. In Chapter 3, Lee clarifies the concepts of AfL and reviews key research findings relevant to AfL in L1 and L2 writing. She also considers issues along with four main factors (teacher, student, school, and system) that impact sound practices of AfL in school contexts. She concludes the chapter with a summary of instructional fundamentals underlying effective implementation of AfL in L2 writing classrooms. In Chapter 4, the theoretical background and strategies of AaL for fostering young students

as active learners in L2 writing classrooms are further expounded. The remaining part of the chapter is devoted to a review of previous studies on AaL in writing, thereby uncovering under-explored areas that are worthy of investigation in future studies.

Chapters 5 through 8 carry on consistently with the theme of feedback and stages of learning in AfL/AaL-oriented writing classrooms. In Chapter 5, Lee frames a discussion of feedback in L2 writing by going into theories on the topic (e.g., second language acquisition, composition, and writing theories), its role as a mediating learning experience that facilitates AfL/AaL, and characteristics of good feedback. Grounded in activity theory, a subordinate of sociocultural theory, the discussion incorporates proposals for innovations needed for effective feedback practices in the L2 writing classroom. The author concludes the chapter with a brief overview of common types of feedback (teacher, peer, and technology-enhanced) further discussed in detail in Chapters 6, 7, and 9, respectively. Expanding upon the general overview in Chapter 5, key research findings on teacher feedback in L2 writing are reviewed in terms of focus (e.g., language errors, content) and modes (written corrective feedback, written commentary, and oral feedback) in Chapter 6. The author addresses not only a gap between research-driven best practices for teachers in providing written feedback and actual practices in L2 classrooms, but also factors causing the observed mismatch (e.g., examination culture). Lee emphasizes the importance of understanding the teachers' practice of feedback provision in contexts, and concludes the chapter with eight governing principles for effective implementation of teacher feedback (e.g., being selective and individualized) in L2 school contexts. Moving on to Chapter 7, the author explains the role of peer feedback in terms of supporting theories (e.g., process writing, collaborative learning), and reviews important research findings and pedagogical implications for effective writing classroom assessment practice in L2 pre-university contexts. The author wraps up the chapter with a batch of useful strategies classroom writing teachers can employ to incorporate peer-feedback activities as a vital learning strategy to promote schoolchildren's L2 writing.

In Chapters 8 and 9, Lee turns the conversation to portfolio-based assessment and the use of technology for classroom assessment and feedback in L2 writing. Lee opens Chapter 8 with a rationale for potential uses of portfolio assessment. Portfolio assessment, predominantly used in postsecondary-level contexts for L2 school

students, is reviewed according to characteristics, types, and possible uses for both AfL/AaL and AoL purposes in the writing process in relation to the three stages of learning: "before writing: where I am going," "during writing: how I am going," and "after writing: where to next," and different levels of feedback (e.g., feedback regarding a task and process). Based on research findings, the author assesses portfolio writing as educational and assessment tools in writing classrooms, as well as calling attention to challenges of the evaluation method in L2 school contexts. She ends the chapter by calling for future research to expand the limited territory of this research area: portfolio assessments in the pre-college writing classroom. Chapter 9 introduces technology-enhanced tasks (digital storytelling, blog-based writing, and collaborative writing on wikis) apposite for young L2 students and reviews previous research that lends support for the use of these tasks. After a survey of the tasks, automated writing evaluation, screencast feedback, and web-based resources (e.g., Microsoft Word tools, concordancing) are discussed as technology-assisted teacher-, self-, and peer-evaluation tools with possible benefits and pitfalls of each of the tools when used in elementary and secondary school contexts. This penultimate chapter closes with an illustration of Writing ePlatform, designed for Grades 4 to 9 students in Hong Kong, which illustrates possible uses of technology as pedagogical means for enhancing AfL/AaL in the writing classroom.

The book rounds off with Chapter 10 on school teachers' assessment literacy, in general, and L2 classroom teachers' assessment and feedback literacy for writing, in particular. Lee canvasses assessment literacy (including feedback literacy) required for L2 school teachers to effectively implement classroom writing assessment. She also reviews the literature on assessment literacy among pre- and in-service school teachers and significant factors contributing to the development of L2 teachers' assessment literacy for classroom writing assessment. The author ends the chapter, stressing the importance of developing classroom writing teachers' assessment literacy through appropriate and sufficient training to advance L2 students' learning of writing.

EVALUATION

Overall, this text is a valuable addition to the area of classroom writing instruction and assessment in light of the fact that it addresses classroom writing assessment in conjunction with feedback and focuses on AfL/AaL, especially in less explored contexts (L2 primary to secondary schools). As the author points out, this is an under-researched subject that needs more attention in the field of L2 teaching and learning, as most studies on the subject have been conducted in university contexts. In this regard, this monograph is likely to appeal to both researchers and practitioners who are engaged in classroom writing instruction, especially in L2 pre-college contexts. It provides researchers with fundamental knowledge about small-scale contextualized assessment and feedback in the L2 writing classroom in terms of theoretical underpinnings and research findings. In addition, various practical guidelines and tips, as well as plenty of resources (e.g., sample teacher and peer feedback forms, links to examples of technology-enhanced tasks completed by schoolchildren, peer feedback rating scales, URLs of web-based concordancing) are provided in Chapters 2 through 9. They are expected to be of great benefit to school writing teachers.

Despite due merits, one consideration is left to be desired – a broader scope of assessment approaches. The book focuses primarily on AfL/AaL, without enough consideration of the means for obtaining synergy between summative and formative assessments. In many classrooms, where examination-oriented learning and teaching are dominantly practiced, due to realistic constraints like national educational policies, it is not feasible, nor reasonable, to carry AfL/AaL forward, separating them from AoL. This unrealistic approach might sound obscure or irrelevant to some teachers who instruct under such constraints. Rather, it would make much more sense for these teachers to find ways to keep a balance between the two approaches. In fact, some researchers suggest the potential benefits of using both summative and formative approaches in L2 classroom assessment (e.g., Carless, 2008; Weigle, 2002). Notwithstanding this noted issue, this book will definitely serve as a useful reference for its intended audience and prospective readers.

THE REVIEWER

YunDeok Choi holds a PhD in applied linguistics and technology from Iowa State University of Science and Technology. Her specializations include ESL/EFL writing and speaking performance assessment as well as test development and validation. She has been teaching Language Testing and Program Evaluation, Research Methods, and Introduction to Statistics at Sungkyunkwan University and the International Graduate School of English in Seoul, Korea. Email: yundeokchoi@gmail.com

REFERENCES

Carless, D. (2008). Developing productive synergies between formative and summative assessment processes. In M-F. Hui, & D. L. Grossman (Eds.), *Improving teacher education through action research* (pp. 9–23). Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112.

Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.

Korea TESOL Journal
General Information for Contributors

As an academic journal in the field of English language teaching (ELT), the *Korea TESOL Journal* welcomes the submission of manuscripts that meet the general criteria of significance and scientific excellence. Submissions should be of practical import, dealing with aspects of the Korean ELT context or directly applicable to it. As a journal that is dedicated to the nurturing of research among ELT practitioners, the Journal also welcomes quality submissions from the early-career researcher.

The *Korea TESOL Journal* invites submissions in three categories:

1. Full-Length Articles. Contributors are strongly encouraged to submit manuscripts of 5,000 to 8,000 words in length, including references, tables, etc.

2. Brief Reports. The Journal also invites short reports (approximately 2,500 words). These manuscripts may present preliminary findings, focus on some aspect of a larger study, or summarize research done in the pursuit of advanced studies.

3. Reviews. The Journal invites succinct, evaluative reviews of scholarly or professional books, or instructional-support resources (such as computer software, video or audio material, and tests). Reviews should provide a descriptive and evaluative summary and a brief discussion of the significance of the work in the context of current theory and practice. Submissions should generally be no longer than 1,500 words.

Manuscripts are accepted for peer review with the understanding that the same work has not been submitted elsewhere (i.e., not pending review or currently under review) and has not been previously published, online or in print.

Manuscripts should follow APA style guidelines (*Publication Manual of the American Psychological Association*, 7th ed.), especially for in-text citations, references, tables, and figures. Submissions should be made with tables, figures, and other graphics included in the manuscript text (and upon request, as separate files). All figures should be created in black and white, and graphs must display distinctive shades or patterning for readability. Manuscripts should be submitted as MS Word (DOC or DOCx) files.

The *Korea TESOL Journal* accepts submissions for two issues annually, on a rolling submissions basis.

Inquiries and manuscripts: journal@koreatesol.org

For more information on submissions to the *Korea TESOL Journal*, including paper submission deadlines, evaluation criteria, and formatting requirements, visit:

https://koreatesol.org/content/call-papers-korea-tesol-journal

LOOKING FOR FLEXIBLE STUDY OPTIONS?

The University of Birmingham's Department of English Language and Applied Linguistics offers flexible personal development opportunities for professionals wishing to develop their skills and expertise. Our distance learning Masters programmes are delivered part-time over 30 months, to fit around your existing commitments.

APPLIED LINGUISTICS MA

This programme is for professionals wishing to further their personal development, and those who are interested in learning more about possible applications of language research. You will study topics including corpus linguistics, sociolinguistics, lexis, functional grammar, spoken and written discourse, and multi-modal communication.

TEACHING ENGLISH TO SPEAKERS OF OTHER LANGUAGES (TESOL) MA

This programme is for practising teachers of English as a second or foreign language who wish to develop their knowledge of classroom methodology and materials design. You will study topics such as language teaching methodology, second language acquisition, syllabus and materials, pedagogic grammar, lexis, and teaching young learners.

KEY FACTS

- Start in February, April, July, October or December
- Study entirely online
- All modules assessed by written assignment
- Pay per module

Find out more and apply:
www.birmingham.ac.uk/elal-dl