ISSN 2411-7390

jle

JOURNAL OF LANGUAGE & EDUCATION

Volume 6 Issue 1, 2020

HIGHER SCHOOL OF ECONOMICS
NATIONAL RESEARCH UNIVERSITY

Contents

Raitskaya, L., & Tikhonova, E. (2020). Pressure to Publish Internationally: Scholarly Writing Coming to the Fore. *Journal of Language and Education, 6*(1), 4-7. https://doi.org/10.17323/jle.2020.10631

Pressure to Publish Internationally: Scholarly Writing Coming to the Fore

Lilia Raitskaya[1], Elena Tikhonova[2,3]

[1] Moscow State Institute of International Relations (MGIMO University)
[2] National Research University Higher School of Economics
[3] Peoples' Friendship University of Russia (RUDN University)

Correspondence concerning this article should be addressed to Elena Tikhonova, National Research University Higher School of Economics, 11 Pokrovsky Bulvar, Pokrovka Complex, room S313, Moscow, Russian Federation, 109028. E-mail: etihonova@hse.ru

JLE editors touch upon the trends and challenges arising out of the changing landscape of scholarly communication as well as two sets of major problems non-Anglophone researchers face in publishing their research in international English-language journals. Firstly, if not desk rejected, they encounter continuous revisions of their submissions to such journals. Secondly, English as lingua franca of international scholarly communication may lead to some disengagement of national scholarly elites who essentially publish in English and to a wider national scientific community decoupled from English and limited to their native language communication. Given the challenges, the editorial offers a refined and widened JLE scope regarding language- and education-related issues of scholarly written communication.

Keywords: scholarly publishing, academic english, scholarly communication, writing for publication, international journal, english as an additional language (EAL), intelligibility

JLE is doing its best to sensibly react to new challenges brought about by language- and education-related issues. With much focus on academic requirements for publishing research internationally, academics, faculty members, PhD students, and fully-fledged researchers have to seek ways to produce high- or average-quality articles and get them published. Ceteris paribus, quality implies intelligibility of scholarly text. Though some researchers find it nearly impossible 'to objectively judge if a manuscript is intelligible or not' (Flowerdew & Ho Wang, 2016), international journals have to stick to their publishing standards, sometimes obscure and illusive. Depending on their command of English, non-Anglophone researchers meet with various hurdles in submitting research to international journals. In many contexts, such journals often stand for European and North American. In addition, they are substantially English language journals, with the best of them being overwhelmingly Anglophone. On the whole, English is indisputably considered as the dominant language for scholarly communication across the globe. In a sense, dominance may imply both functions of lingua franca and internationalisation-induced and unwelcome usage of the dominant language in numerous science settings.

Researchers in non-English speaking countries have to essentially write and publish in English. Their command of English and obstacles it may pose (primarily efforts to improve their manuscripts via literacy or scholarly editorial services and their academic writing skills) lead to longer cycles of research publishing. Their submissions tend to undergo continuous negotiation between journal editors, reviewers and authors. The non-Anglophone academia at large raises a hot issue of language policy, monolingualism and multilingualism in national science and scholarly communication within national states. There is a rising concern that internationalisation of science via publications in international English-language journals confronts with national language communication in wider national scholarly communities. On the one hand, internationalisation is narrowed to the research at national levels conducted by only a small share of non-Anglophone scholars who may be classified as expert English users. On the other hand, national research published in international English-language journals is often inaccessible for national scholars at large.

Expert language users account for a low percentage of non-Anglophone researchers. They have fewer barriers to international publishing. But their expertise and career trajectories associate with another set of challenges.

Harbord highlights that the thrust toward disengagement of the English-speaking (L2) national academic elite from their national science community (Harbord, 2018) is based on factors discouraging local language publishing. Why publish in English? Research in English often means reaching a larger and more competent professional audience and getting career benefits for non-Anglophone academics. In contrast, local language publishing often implies periphery of science. Thus, English has become the symbol of global high-profile science.

Cheung sets a number of prerequisites for a successful publication in peer-reviewed journals in English. They cover 'gaining entry into a particular discourse community … and making good use of situated knowledge' (Cheung, 2010). The criteria of the discourse community encompass common goals, participatory mechanisms, information exchange, community-specific genres, highly specialized terminology, and a high-level expertise (Swales, 1990). Entering a new discourse community, i.e. the global English-speaking scholarly writing world, requires lots of effort and expertise. Researchers as novices have to remain at the margins for some time or even permanently as opposed to the core of global science (Cheung, 2010). Only after having familiarized themselves with the conventions of the discourse community, non-Anglophone academics may get a chance to enter it, given the fact that the appropriate English discourse is a kind of entrance ticket but not a warranty to be published or cited internationally.

Scholars with English as an Additional Language (EAL) 'have to overcome considerable difficulties in order to publish their research in international journals' (Flowerdew & Ho Wang, 2016). Flowerdew & Ho Wang summed up the key non-standard ('non-canonical') features of texts by EAL authors who are not expert users of English. The irregularities cause minor and major language revisions of EAL authors' submissions to international journals. Surface-level, or minor corrections include determiners, singular/ plural forms, punctuation, spelling errors, verb forms, confusions between parts of speech. The major revisions entail word/ morpheme level, additions (more frequently at the lexical and grammatical levels of clause/ clause complex and prepositional groups), deletions (clause/ clause complex level, prepositional groups, nouns, conjunctions, adverbs), rearrangements – replacements and substitutions to alter the meaning (prepositional groups, adverbs, nominal groups, verbal groups, clauses) (Flowerdew & Ho Wang, 2016).

EAL researchers have to constantly improve their command of English and academic writing skills, seek clarity of their scholarly texts, better text organisation through logic cohesion, well-grounded thesis development, and better general readability. Most of future academics and scientists studied under a writing-enriched curriculum at university. While working with journal editors and reviewers, some of them claimed that 'they had learned a lot from the editor about revising their manuscripts in the process' (Flowerdew & Ho Wang, 2016). Nevertheless, non-English speaking scientists across the globe are often forced to turn to intermediaries who help their manuscripts get better shaped and more intelligible. Two types of so-called 'literacy brokers' are outlined as (1) academic brokers, focusing on the content of articles; (2) language brokers, more concerned with language-related problems (Flowerdew & Ho Wang, 2016).

From time to time Anglophone editors and reviewers revisit the statement that 'non-standard English used by EAL writers should be accepted in international refereed journal articles' in case their language is intelligible (Jenkins, 2013). But if Q1 Scopus-indexed Anglophone journals are considered, only non-Anglophone expert users of English or those supported by scholarly editorial and proofreading services get their articles published there. There is a gaping difference in the English discourse between Anglophone and non-native English-language international journals. Though the latter may occasionally enter Q1 and Q2 in the Scopus database, when they approach expert level in their English, the bulk of Q1 and Q2 journals are either Anglophone or close to the discourse standards of English language native speakers.

Anyway, English scholarly text intelligibility lays foundation for high-level international publication. Linguists, editors, reviewers, and educators may contribute much to improve understanding of the phenomenon and ways to foster researchers' skills in writing for publishing in English.

Scholarly writing is not only about writing, but also about rhetoric schemas in various languages, thesis development and logic. Each language determines the ways people think. Writing in English is successful when an author constructs sentences, paragraphs and the whole text due to the conventions of the English language, barring native tongue interference. If otherwise, perfect grammar and proper discourse may fail to save the

writing. Wrong or foreign rhetoric moves and steps in an article coupled with non-English text organization and thesis development may distort the understanding and perception of the text (Swales, 1990). L2 and EAL researchers have to go a long and painstaking way to become efficient scholarly communicator and expert level users of English to succeed internationally.

Given the trends and new challenges above, below you will find a letter from the JLE Editors communicating the re-shaped journal scope to the Readers.

> *Dear JLE Readers,*
>
> *JLE editors are inspired to announce that the Journal of Language and Education has decided to refocus its scope adding studies on scholarly communication to the field of languages and education. This move is part of a strategic relaunch of the journal to provide greater attention to sharing and publishing research findings so that they could get available to a wider academic community. As you see from the above, the landscape of scholarly communication is changing under the influence of technology, internationalisation, global and national policies. These challenges demand new considerations. We are willing to provide an international source of peer-reviewed information on scholarly communication at large, with language- and education-related issues in closeup. It will cover such niches and issues as academic writing, writing for publishing, science editing, canonical patterns and non-standard irregularities of scholarly and academic texts, English as an Additional Language (EAL), English as lingua franca and multilingualism, literacy brokerage, journal writing conventions, structural and thesis-driven aspects of scholarly writing, research readability, self-awareness of writing confidence, publishing norms, fostering academic writing skills, writing-enriched university curricula, moves and steps theory and practice in scholarly writing, rhetoric schemas.*
>
> *Moreover, after several months of planning and re-thinking our strategy, we decided to expand the Education section coverage. The journal will publish original articles on changing universities and their missions, education reframing, innovative models of teaching and progressive learning technologies (flipped class, mixed learning, deep active learning, etc.), virtual education and MOOCs, gamification and game-based curricula, redefining of quality of tertiary education and international rankings, mobility and autonomy in higher education. The list is certain to be regularly updated as the educational settings are prone to fast and constant change.*
>
> *Looking forward to new inspiring submissions from our authors!*
>
> *Best Regards,*
> *JLE Editors*

Conflict of interest

The authors declare that they have no conflict of interest.

References

Cheung, Y.L. (2010). First publication in refereed English journals: Difficulties, coping strategies, and recommendations for student training. *ScienceDirect System*, *38*, 134-141. https://doi.org/ 10.1016/j.system.2009.12.012

Edmunds, M.W. & Waldrop, J. (2018). What is Scholarly Writing? *The Journal for Nurse Practitioners*, *14*(8), A11-A12. https://doi.org/10.1016/j.nurpra.2018.08.012

Flowerdew, J. & Ho Wang, S. (2016). Author's editor revisions to manuscripts published in international journals. *Journal of Second Language Writing*, *32*, 39-52. https://doi.org/10.1016/j.jslw.2016.03.004

Harbord, J. (2018). Language policy and the disengagement of the international academic elite. Science Editing, 5(1), 32-38. https://doi.org/10.6087/kcse.115

Jenkins, J. (2013). *English as a lingua franca in the international university: The politics of academic English language policy*. Routledge.

Swales, J. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.

Zarkov, D. (2019). What we do and what we don't: Paradoxes of academic writing for publishing. *European Journal of Women's Studies, 26*(4), 357-359. https://doi.org/10.1177/1350506819870632

Antonova, M.B. (2020). The Container Image Schema as the Conceptual Basis of English Adjectives' Semantics. *Journal of Language and Education, 6*(1), 8-17. https://doi.org/10.17323/jle.2020.9751

The Container Image Schema as the Conceptual Basis of English Adjectives' Semantics

Marina B. Antonova
National Research University Higher School of Economics

Correspondence concerning this article should be addressed to Marina B. Antonova, School of Foreign Languages / Department of English for Economic and Mathematical Studies, National Research University Higher School of Economics, 34 Tallinskaya Ulitsa, rom 420, Moscow, 123458, Russian Federation.
E-mail: mbantonova@hse.ru

This paper focuses on the cognitive foundation of the semantics of English adjectives that denote mental and moral characteristics of human beings. Research into these adjectives seems a challenging task because they denote abstract qualities that cannot be perceived through vision, hearing, or touch; and here a question arises: How are abstract qualities interpreted in English encoded through adjectives? To answer it, this study follows the idea of two-level semantics, i.e. word semantics is treated as a two-level phenomenon that comprises the semantic (external) level and the conceptual (deep) one. This study is the first to address adjectival semantics from this perspective. Here a novel approach to revealing the cognitive foundation of adjectives is introduced: given that adjectives originated from old syncretic items and a word cognitive model forms at the moment of word creation, cognitive models underlying adjectives' semantics are unearthed via analysis of their etymological data. Our contribution is two-fold. First, the approach has revealed that the image schema CONTAINER guides semantics of an array of various adjectives independent of their morphemic structure or date of origin. The examples demonstrate that abstract human qualities are interpreted via the following container features: boundary, container substance, size, hardness/softness of a container shell, etc. The semantics of affixed or compound adjectives appear to stem from the integration of concepts represented by an affix and a root or two roots, respectively. Second, the findings show that the value given to every container feature appears to predetermine the evaluation conveyed by an adjective. Container features tend to possess ambivalent value, realizing the positive or negative one due to the interaction with a frame in which the CONTAINER is incorporated, therefore the same polysemantic adjective may develop both positive and negative meanings. To reveal the whole inventory of cognitive models that govern adjectival semantics in English, further research needs to be conducted.

Keywords: conceptual basis, image schema, semantics, frame, evaluation meaning, adjective, English

Introduction

One of the most revolutionary ideas introduced by cognitive linguistics concerns the distinction between the semantic, or language, and the conceptual levels. Initially, this idea was developed within the framework of the Two-level Semantics theory pioneered by Bierwisch (Bierwisch, 1983; Bierwisch & Lang, 1987; Bierwisch & Bosch, 1995). The theory argues "the conceptual nature of the meanings conveyed by natural language expressions" (Lang & Maienborn, 2011, p. 709). In other words, the semantics of a language item is treated as a complex phenomenon: the meanings that a language item conveys rest on a certain conceptual basis that accounts for them. The conceptual basis is structured by cognitive models (Lakoff, 1987; Evans, 2006; Jackendoff, 2011; Lang & Maienborn, 2011; Cruse, 2017; Boldyrev, 2016).

Being part and parcel of culture, every language operates with its own systems of concepts and cognitive models to regulate the construction of associated meanings. Since "the connection between words and the outside world is mediated through concepts" (Cruse 2017, p. 241), to unearth the conceptual bases of various language

items and categories seems an important task. To the best of the author's knowledge, however, few studies have been conducted in this field so far; in particular, conceptual bases were modelled for English idioms (Zykova, 2018), and euphemisms belonging to such Romano-Germanic languages as Italian, Spanish, French, English, and German (Porohnickaya, 2014).

As for adjectives, and English ones are no exception, they have not yet been investigated in this respect. However, adjectival conceptual basis modelling seems to be a challenging aspect of linguistic research, bearing in mind the ambiguous character of the adjectival category, which still remains the most debatable among linguists. By way of illustration, English adjectives have been investigated from various perspectives: structural linguistics, generative linguistics, and cognitive semantics. Researchers state different, even polar opinions on the adjectival status, from classifying qualitative and relative adjectives as two separate parts of speech to including adjectives into the verb or adverb categories (Chafe, 1970; Givon, 1970; Lakoff, 1970; Dixon, 1982; Baker, 2003). Moreover, adjectives were characterized as "a notorious swing category" (Givon, 2014, p. 13-14). Cross-linguistically, the category referred to as "adjectives" varies more significantly than the noun or verb ones. In some languages, and in English, the adjectival category constitutes a large, open class of words, while "in a number of other languages, such as Hausa, there is a small closed class of words… In yet other languages, such as Chinese, it is claimed that adjectives do not exist" (Croft, 2003, p. 183). Recent years have seen similar discussions; in particular, the distinction between adjectives and determiners has been justified for Serbian (LaTerza, 2015), a double semantic value of the Gothic adjective *sama* has been stated (Ratkus, 2018), and a mixed category combining morphemic and syntactic peculiarities of nouns and adjectives has been singled out in the Tungus language (Nikolaeva, 2008). The ambiguity of adjectives leads to difficulty in their identification in texts, proper use, and proper adjective ordering (Abd Rahim, 2013; Sosnina, 2015; Berg, 2019; Trotzke & Wittenberg, 2019). Although the dominant opinion is that adjectives are not a universal language category, researchers agree that "every language seems to be equipped with adjectival concepts" (Pustet, 2006, p. 61).

Adjectives constitute the bulk of English vocabulary, playing an essential role in expressing our emotions, feelings, and attitudes, which has compelled computational linguistics to thoroughly investigate this language category. Researchers have argued that all adjectival peculiarities are substantiated by their nature (Cristani et al., 2018). This view seems quite reasonable since difficulties in determining the categorial features of adjectives may be linked to their syncretic origin, which also makes it important to gain insight into the adjectival conceptual basis.

Initially, quality was not encoded via adjectives; at the pre-logical reasoning stage qualities were comprehended not individually but inseparably from their referents. Properties of referents comprised unique combinations that allow the recognition of certain referents. Old words could function as denominations of both an object (referent) and its qualities (Kacnel'son, 2001, p. 413; Zhirmunskij, 1976, p. 210). The assumption that the Proto-Indo-European language did not differentiate between the noun and the adjective is substantiated by modern compound words composed by two nouns (noun+noun), with the first constituent playing the role of an attribute (Potebnya, 1968, p. 62; Klimov, 1977, p. 215; Gamkrelidze, 1984, p. 279). The distinction between the entity (object) and its properties (qualities) resulted in the formation of a new lexico-grammatical category – the adjective. This consecutiveness of language category formation can be evidenced by the fact that children acquire the names of objects before the names of their qualities (Hiramatsu et al., 2010). Moreover, there exists a certain consecutiveness of adjective acquisition by children (Weicker & Schulz, 2019).

Alongside the theory of the noun origin of adjectives (Jespersen 1951; Zhirmunskij, 1976; Potebnya, 1926; Kacnel'son, 2001; Kubryakova, 2008, and others), other points of view exist on this issue: the halophrastic approach considers phrases, not nouns, as a primary language item (e.g. Piaget 1926; Schuchardt 1928; Gans 1981).

According to his typological research, Dixon has determined the nominative area of adjectives, subdividing them into core and peripheral semantic types. The core semantic types of adjectives cover DIMENSION, AGE, VALUE, COLOUR, while the peripheral ones denote PHYSICAL PROPERTY, HUMAN PROPENSITY, and SPEED (Dixon, 2004, p. 3-5).

The present study aims at modelling the conceptual basis of English adjectives denoting such human propensities as mental and moral ones.

The adjectival categories of mental and moral characteristics have been chosen for the following reasons: first, they nominate essential human characteristics that began to be recognized before the Old English period; second, these adjectives denote abstract qualities that cannot be perceived through the senses of sight, hearing, smell, or touch but manifest themselves indirectly in human behavior, reasoning, and activity. Maturana and Varela (1992) stated that "we have language, there is no limit to what we can describe, imagine, and relate" (p. 212), but finding out how these abstract qualities are expected to be verbalized seems a challenging task. Third, given these adjectives convey evaluation – positive or negative, they reflect both the culture's evaluation standards and a speaker's interpretation of the performed quality. The contribution of this study is, therefore, twofold: the cognitive basis of English adjectives denoting abstract human beings' qualities is modelled and the influence of the cognitive basis on the evaluation meaning of the adjectives is considered.

Method

Theoretical Background

The following sections demonstrate the method and procedure for modelling the conceptual basis of the adjectives and dwell upon one of the major ways that mental and moral characteristics of human beings appear to be interpreted in English, namely via the CONTAINER image schema. According to Johnson, who introduced this term, image schema implies "a recurring dynamic pattern of our perceptual interactions and motor programs that gives coherence and structure to our experience" (Johnson, 1987, p. xiv). It should be mentioned that CONTAINER belongs to the major image schemas (Hampe, 2005, p. 2). The theory of image schemas is widely utilized in various cognitive studies (e.g. Alessandroni, 2017; Dancygier, 2017; Senkbeil, 2017; Martinez et al, 2018; Ioannou 2018; Vernillo, 2019), and it has proved to be of great application potential, even inspiring new research in the area of human-computer interactions (Hurtienne, 2017).

Being anthropocentric in nature, our conceptual systems ultimately depends on humans' physical and cultural experiences (Lakoff, 1987, p. xiv). Evans argues that "our embodiment is directly responsible for" the ways we perceive and interpret the world and the understanding of space boundaries emerges from human beings' "ongoing embodied experience" (Evans, 2006, pp. 177, 185). Since our physical bodies are "bounded and set off from the rest of the world by the surface of our skins," they can be treated as *containers* with "an in-out orientation" that is projected "onto other physical objects that are bounded by surfaces" or in the case they are not, "we impose boundaries marking off territory". In fact, "there are few human instincts more basic than territoriality" (Lakoff & Johnson, 2003, p. 30). In fact, CONTAINER reflects certain spatial relations, but space awareness and awareness of oneself within space belong to the oldest and most fundamental ideas (Meletinskij, 1976; Ivanov, 1992; Podosinov, 1999). Kubryakova reasoned that the concept of container is ubiquitous both from ontological and philosophical perspectives (Kubryakova, 1999, p. 7).

The CONTAINER image schema can be extended to a large number of abstract concepts (Lakoff, 1987, p. 272), in particular, to that of the mind: "Mind is a container" (Lakoff & Johnson, 2003, p. 149). In turn, ideas and knowledge are interpreted as discrete material objects that can be placed into a container, thus presenting a container substance: "Ideas are objects" (Lakoff & Johnson, 2003, p. 31). In accordance with this, the CONTAINER image schema, although it should be treated as an integral whole, includes the following structural elements: INTERIOR/CONTAINER SUBSTANCE, BOUNDARY, and EXTERIOR CONTAINER (Lakoff, 1987, p. 272).

Materials

The adjectives for the current research have been gathered from The Original Roget's Thesaurus of English words and phrases (1987). Adjectives denoting mental characteristics of human beings belong to the sections "Mental" and "Intellect". As for moral qualities, adjectives encoding them may be found in the following sections: dueness, ethics, friendship, innocence, justice, kindness, love, modesty, pity, probity, sociability, virtue, arrogance, cunning, deception, enmity, improbity, injustice, jealousy, hatred, pride, selfishness, undueness, unsociability, vice, etc. The number of adjectives denoting mental and moral characteristics total 409 and 426 respectively.

Remarkably, words of negative evaluation predominate over those of positive meaning: mental characteristics – 233 vs 176 adjectives, moral characteristics – 267 vs 159 adjectives. In terms of derivation, the adjectives

under study may include only one stem, therefore possessing one semantic centre, or two stems in the case of compounds, therefore comprising two semantic centres. Analysis of such structurally different adjectives can provide reliable and valid conclusions regarding the conceptual basis of these English adjectives' semantics.

Procedure

Given adjectives' ontogenesis, namely their stemming from a syncretic word, and the fact that the conceptual basis of word semantics forms at the moment of word creation, an analysis of adjectival etymological data could become the starting point of their conceptual basis modelling. To demonstrate the process of revealing the conceptual basis of adjectival semantics, let us consider the following example. Etymological data attest that the adjective *accomplished* XV c.[1] is derived from the verb *accomplish*, which originates from the Old French *acompliss-* originating from the verb *acomplir*, from the Vulgar Latin **accomplere*, from *complere* 'to fill', formed by the prefix *com-* and the root *plere* 'to fill', which goes back to the Proto-Indo-European root **pele-1* 'to fill' (OED). The action of filling implies the mandatory presence of some *container* that is filled; consequently, in this case, the acquisition of knowledge or skills is interpreted as filling a container – the mind, thus bringing out the characteristics of the inner part of a container, or its *interior*.

Similar associations gave rise to the adjectives *proficient*, *qualified*, and *affectionate*, which came into English in the XVI c., originating from the Proto-Indo-European[2] root **dh-/*dhē-* 'to set, put'[3] (OED; Vaan, 2008, p. 198-199; Watkins, 2000). Insofar as the adjective *affectionate* is concerned, the heart is viewed as a container, while love and affection are associated with objects put inside.

Quite a number of adjectives characterizing human beings' mental activity have stemmed from the PIE root **gno-* 'to know': *keen* Old English*, acquainted* XV c., *canny* XVI c., *knowing* XVII c*., knowledgeable* XIX c. (OED). If a person knows something, it means he/she has cumulated some thoughts, knowledge, or skills, and hence some thoughts, knowledge, or skills are perceived as objects that have been placed in the mind – a container.

Adjectives derived with the help of negative affixes form their semantics due to the integration of the concepts represented by the root and the affix. Meanings antonymous to the discussed above result from the integration of the concept CONTAINER represented by the roots and the concept ABSENCE represented by the affixes un-, -less: *unknowing* XIV c*., unacquainted* XVI c., *unqualified* XVI c., *unaccomplished* XVIII c., *unaffectionate* XIX c., *affectless* XX c., and *unknowledgeable* XX c. This conceptual integration creates the image of an empty container, in terms of which the deficiency in knowledge and skills or fondness and tenderness is described. Another example here is the adjective *ignorant* XIV c., which originates from the Latin *ignorantem* (nominative *ignorans*), from *ignorare*, from *in-* 'not' + Old Latin *gnarus* 'aware, acquainted with', from the Proto-Latin suffixed form **gno-ro-*, from PIE **gno-* (OED), which again forms an image of an empty container.

Roots of some adjectives go back to the PIE **dheu-1* 'dust, vapor, smoke' or to the PIE roots meaning 'thick', for example: *dull* XIII c., *doltish* XVI c., *dozy* XX c. originate from PIE **dheu-1*; *thick* XVI c. originates from PIE **tegu-* 'thick'; and *dense* XIX c. originates from PIE **dens-* 'thick' (OED; Vaan, 2008, p. 167). The logic behind these associations may be substantiated by people's experience that evidences that every container holds its own specific substance, which secures containers' functions. So obstruction of the intellectual process may be seen to be caused by an improper extremely dense container substance.

As for compound adjectives, their semantics also stems from integrated concepts but in this case is represented by two constituents of a compound. The head element of compounds may frequently convey the idea of a container in quite an explicit way, for example, *dull-headed* XVI c. *airheaded* 1970, *fat-brained* XVI c., *hard-hearted* XIII c., *stone-hearted* XVI c. and so forth.

Further, implementation of the proposed method in modelling the conceptual basis of the adjectives in question will be demonstrated and other container characteristics involved in its formation will be described.

[1] The date after the word indicates the time of its first known use in the discussed meaning as attested in The Oxford English Dictionary (1989).
[2] henceforth – PIE.
[3] In light of space limitations, from this point on the etymological data are given in a reduced form.

Analysis of the Adjectives Denoting Mental and Moral Human beings' Characteristics

Given that states are conceptualized as containers (Lakoff & Johnson, 2003, p. 31), meditating and being engrossed in thought can also be considered as filling container, while the absence of thought or forgetfulness are interpreted as empty containers: *mindless* Old English*, mindful* XIV c., *unmindful* XIV c., *mental (inf. 'insane')* XX c. go back to PIE **men-* 'to think'; *thoughtful XIII c., thoughtless* XVI c., *thinking* XVII c. originate from PIE **tong-* 'to think' (OED); and *rational* XIV c., *irrational* XIV c. go back to PIE *re- 'to think' (OED; Vaan, 2008, p. 519-520).

It is worth emphasizing that Wierzbicka labels the concepts THINK and KNOW as mental predicates belonging to basic human concepts, or semantic primitives, which are verbalized in various ways in different languages (Wierzbicka, 1996, p. 35).

Although CONTAINER acts as an integral image (image schema), in some cases different conceptual constituents of the image might come into focus. Some human qualities are nominated via the association with the container BOUNDARY. Thus, for instance, an area where people dwell, whether it is a house, town, or state, is enclosed with real physical boundaries and thus it can be referred to as a container. Dwellers in any particular area must obtain the necessary knowledge regarding its laws, standards of behavior, and traditions, which enables them to act under various circumstances properly: *politic* XV c. (archaic) 'prudent, sagacious' originates from PIE **tpolh-* 'citadel; enclosed space, often on high ground, hilltop' (OED).

Societies impose certain limits for speaking about one's own merits in public and a person knowing these limits is estimated positively: *modest* XVI c. originates from PIE **med-1* 'to measure' (OED; Vaan, 2008, p. 384; Watkins, 2000; Pokorny, 1959, p. 705-706). Conversely, a person going beyond 'the boundary' of ethical norms is characterized negatively: *immodest* XVI c.

A person capable of understanding other people and showing compassion with them is seen as having easily penetrated the container shell, or having gone through boundaries: *ruthful* XIII c. (archaic) 'feeling sorrow or pity' originates from PIE **kreue-2* 'to push, strike'; *feeling* XVII c. originates from PIE *pal- 'to touch, shake, strike softly' (OED). On the contrary, a hard shell makes an impervious barrier for external events and people's emotions: *ruthless* XIV c.*, unfeeling* XVI c.; *callous* XVII c. originates from PIE *kal-1 'hard' (Pokorny, 1959, p. 523-524); *hard* Old English 'showing no kindness or sympathy', *hardened* XIV c. originates from PIE *kar-/ker- 'hard'; *remorseless* XVI c. goes back to PIE **mer-2* 'to harm' (OED).

Ideas, notions, and problems that are to be understood may be interpreted as certain EXTERIOR containers penetrated by the mind. In this case, the mind becomes an instrument that beats or cuts in order to finally destroy the container shell and to cross its boundaries, for example: *studious* XIV c. originates from PIE **(s)teu-1* 'to stick, knock, beat' (OED; Pokorny, 1959, p. 1032). In his commentary on the PIE root **(s)kel-1* 'to cut', Watkins points out that human cognitive activity is compared to this quite concrete, physical operation (Watkins, 2000). Likewise, for the English language, Lakoff and Johnson (2003) single out the conceptual metaphor "Ideas are cutting instruments" (p. 49). The semantics of a large number of adjectives has formed on the basis of this interpretation of mind, for example: *sharp-witted* XVI c.*, razor-sharp* XIX c.; *shrewd-brained* XVII c.*, shrewd-headed* XIX c.*, shrewd-pated* XVII c. – *shrewd* originates from PIE **skreu-* 'to cut, cutting tool' (OED). Frequently, the first meaning of an adjectives is not 'clever' but 'sharp', which is the case with the following words: *sharp* Old English 'clever' goes back to PIE **sker-1* 'to cut' (OED); *smart* XIV c. 'clever' (inf.) originates from PIE *smerd- 'pain' – "from the notion of 'cutting' wit" (OED); *incisive* XIX c. 'intelligently analytical, clear-thinking' originates from PIE **(s)kel-1* 'to cut' (Watkins, 2000); *acute* XVI c.*, cute* XIX c. 'clever' originate from PIE **ak-* 'be sharp, pierce' (OED; Vaan, 2008, p. 23). Conversely, an unsatisfactory performance of the mind is associated with a blunt, useless instrument: *nescient* XVII c. goes back to PIE **ne* 'not' and **skei-* 'to cut, split, incise' (OED; Vaan, 2008, p. 545).

The typicality of interpreting the mind as a cutting instrument can be justified by the back semantic derivation, for instance, the aforementioned adjective *keen*, originates from PIE **gno-* 'to know'. When this adjective first appeared in Old English, it encoded an abstract quality of intellect activity –'clever, wise, expert, skillful', but later, in the XIII c., the word commenced to denote a perceivable quality of an object – 'having a sharp edge or point' (OED; Barnhart, 1988).

The degree of understanding something can be represented as the depth of penetration into a container: *deep* XIII c. 'of penetrating mind, having the power to enter far into a subject' goes back to PIE **dheub-* 'deep, hollow' (OED; CODEE).

It is apparent that other cognitive models should participate in the interpretation of mental and moral qualities and these models, as well as their productivity, are still to be revealed. This is proved by the results obtained by Allan (2008) who investigated adjectives denoting human intellectual qualities in terms of identifying source domains for conceptual metaphors and metonymies for cleverness and stupidity (36 source domains were discovered).

Results and Discussion

So far, we have shown that the conceptual basis of an adjective can be unearthed via addressing its etymological data. It should be mentioned that an old root denotes an action, feature, or thing in the most general manner and can correlate with diverse frames, whose perspective affects the evaluation conveyed by adjectives. To flesh this out, let us bring in some examples. If the content of the container implies some knowledge and a cutting instrument implies the mind, the quality 'sharp' is considered positive and therefore the adjective develops a positive meaning: *cute* 'clever' goes back to PIE **ak-* 'be sharp, pierce' (OED; Vaan, 2008, p. 23). In cases where a container refers to another person's mind and penetration occurs to deceive the person, the quality 'sharp' acquires a negative evaluation: *cute* 'cunning'. The phenomenon of polar evaluation meanings of one and the same word is referred to as enantiosemy. It is by no means a rare case in the language system, and additional examples may be cited. Similarly, the adjective *nice* possesses several opposite evaluation meanings: besides 'foolish, stupid' XIII c. and 'kind, good-natured, thoughtful' XIX c. belonging to the scope of this study, the word evolved the meanings of 'timid', 'fussy', 'delicate', 'precise, careful', and 'agreeable, delightful' (some of them have become obsolete and remain only in idioms and set expressions). Regarding *nice*, Weekley (1967) states, "The sense development has been extraordinary, even for an adjective" (p. 983). All these meanings arose due to the adjective's correlation with different frames, i.e. different situations where the adjective is applicable. The word *nice* originates from PIE **ne-* 'not' and **skei-* 'to cut, split' and literally means 'not to cut'. If a blunt instrument targets an object, it fails to cut it, which metaphorically indicates the situation when the mind, not sharp enough, fails to pierce into a new domain of knowledge; hence the negative meaning 'foolish, stupid' emerges. If a dull instrument directs a person, it cannot injure anyone, which metaphorically indicates a situation when someone's behavior or words do not hurt people's feelings; hence the positive meaning 'kind, good-natured, thoughtful' forms.

The concept of CONTAINER and its conceptual constituents can be represented explicitly through the components of a compound adjective. The first element of an adjective codes such peculiarities of a container as size, depth, substance, and indestructibility, to name but a few. The head element of a compound, i.e. the second one, denotes 'a container': *-head, -skull, -pat, -brain, -mind, -thought, -think, -wit, -heart*. Note that *-thought* and *-think* imply states, which are seen as containers, while *brain, mind* and *wit* are interpreted as "the seat of intelligence" (OED), i.e. the container including intelligence.

It is common knowledge that a larger container can hold more, metaphorically – more kindness, sympathy or intelligence, while small containers imply the presence of small quantities, for example: *long-headed* XVIII c., *big-hearted* XIX c., *highbrow* XIX c., *high-browed* XX c.; *shallow-hearted* XVI c., *shallow-brained* XVI c.*, shallow-headed* XVII c., *shallow-witted* XVII c.*, shallow-minded* XIX c., *shallow-thoughted* XIX c.*, small-minded* XIX c., *pinheaded XX c., lowbrow* XX c. etc. The examples demonstrate that compound adjectives denoting intellectual qualities may consistently incorporate the same first element (eg. *shallow-*) and different head elements, which virtually are transnominations of mind (eg. *-minded, -witted, -thoughted, -brained, -headed*).

However, a large size appears to be ambiguous from the evaluation point of view: sometimes adjectives denoting high intellectual abilities acquire pejorative connotation like in *egg-headed* XX c. and *pointy-headed* XX c. Insofar as moral qualities are concerned, a large container may imply an excessive 'scope' of self-conceit, vanity, or arrogance: *big-headed* XX c.*, swell-headed* XIX c.*, swollen-headed* XX c.

As mentioned above, a container may be characterized in terms of the substance it holds; the substance deficiency and its wrongness explain the malfunction of the mind, for instance: *empty-headed* XVII c.*, empty-pated* XIX c.*, empty-skulled* XIX c., *dull-witted* XIV c., *dull-headed* XVI c. (for the etymology of *dull,* see above); *air-headed* XX c.*; muddy-minded* XVII c., *muddy-brained* XVII c.*, muddy-brained* XVII c.*; addle-pated* XVIII c.*, addle-headed* XVIII c.*, addlebrained* XIX c.*; wooly-minded* XX c.*, meat-headed* XX c., *fat-brained* XVI c.*, fat-witted* XVI c.*, fat-headed* XVIII c., and so forth.

A substance may also be characterized in terms of its weight and an unduly heavy or light substance indicates its inefficiency: *thick-witted* XVII c.*, thick-skulled* XVII c., *heavy-headed* XVII c.*, light-minded* XVI c.

The softness of a container's substance may convey opposite evaluations depending on whether it metaphorically represents cognitive abilities or attitudes to other people. In the former case, adjectives convey the negative evaluation of a mind's functionality: *soft-minded* XVI c.*, soft-brained* XVII c., *soft-pated* XVII c.*,* or *soft-headed* XVII c. Regarding moral traits, a gentle or soft heart can be easily touched by other people's misery, suffering, or sorrow, thus producing a positive evaluation: *tender-hearted* XV c., *soft-hearted* XVI c.

A solid container of stone, wood, bone, or iron lacks the interior space to hold anything; hence it fails to meet people's needs. Such a container metaphorically implies the absence of intelligence or kindness: *hard-hearted* XIII c., *stone-hearted* XVI c.*, blockheaded* XVI c., *ironhearted* XVII c.*, timber-headed* XVII c.*, wooden-headed* XIX c., *bone-headed* XX c., etc. The root *block* originates from PIE **bhelg-* 'a thick plank, beam' (OED). Interestingly, all compound adjectives based on the association with the quality of a container shell convey a negative evaluation, however, most of them initially lacked connotations and the negative figurative meaning applied to human beings developed much later, which is proved by the diachronic information recorded in dictionaries. For example, the primary, literal meaning of the adjective *thickheaded* XIX c. is "having a thick head", which is realized in biological terms like *thick-headed mullet*, *thick-headed shrike* (OED). The skull thickness of pachycephalosaurus motivated the coinage of its name "thick-headed lizard". Metaphorical meaning has also enriched the semantics of the following adjectives: *pachydermatous* XIX c., *thick-skulled* XVII c.*, thick-skinned* XVII c.*, case-hardened* XVIII c. At present, a person extremely sensitive to criticism or insults is disapprovingly referred to as *thin-skinned* (the primary meaning is 'having a thin skin').

Apart from that, mental and moral characteristics may be interpreted via temperature. It is common knowledge that human beings possess warm blood, which gave rise to reasoning metaphorically that warmth accounts for sympathetic pity and concern for the sufferings or misfortunes of others, while the absence of compassion and love is associated with cold: *warm-hearted* XVI c.*, cold-blooded* XVI c., *cold-hearted* XVII c.

The interpretation of abstract human qualities through the CONTAINER image schema is one of the oldest; the first adjectives whose semantics was formed by this cognitive model date back to the Old English period (*sharp* 'clever', *keen* 'clever', *hard* 'showing no kindness or sympathy'); and still, the CONTAINER image schema participates in producing adjectival semantics in Modern English (*unknowledgeable* 1969, *bubble-headed* 1950s, *airheaded* 1972, *pointy-headed* 1972, etc.). It is noteworthy that the Old English adjectives are structurally simple, i.e. they include one root, and compound two-root adjectives are attested from the Middle English period.

The CONTAINER image schema underlying adjectival semantics may find reflection in phraseology and may influence the formation of the contexts where human intellectual or moral qualities are discussed. Here it is necessary to recall that Potebnya believed the closest etymological meaning of a word to help get a better understanding of its usage (Potebnya, 1926). For instance, the idiom "dull as ditchwater" directly refers to the substance of a container. The following contexts carry references to exterior containers – "Travellers are to be mindful of their *surroundings*" or to container substance – "What I said was very *thoughtless*" (Merriam-Webster). So, to determine how the conceptual basis of adjectives affects their use in context seems a fruitful direction for further research.

Conclusion

This paper focused on modelling the conceptual basis of the semantics of English adjectives denoting human beings' abstract mental and moral qualities. To infer the conceptual basis, it was proposed that we consider

etymological data about the adjectives in question to find out how these abstract qualities were interpreted at the moment of their nomination via a language sign. One of the ways to interpret mental and moral qualities is via the CONTAINER image schema, which seems to account for semantics and the evaluation meaning of the adjectives.

This study has shown that the CONTAINER image schema significantly contributes to this conceptual basis, accounting for 20% of the adjectives that were investigated. As mentioned above, this schema arises from the so-called embodied experience, which is important for interpreting such abstract human beings' qualities as mental and moral ones. The semantics of a word may be guided by any element of the CONTAINER image schema – INTERIOR, BOUNDARY, or EXTERIOR CONTAINER. In particular, a human's possession of knowledge and morality, or deficiency in them, may be associated with the fullness or emptiness of the container, the ability/inability to cross the container boundary, or an appropriate/inappropriate container substance.

In addition, the research has revealed that adjectival semantics may form on the basis of two integrated concepts, represented either by a root and an affix (in the case of derived adjectives) or by two roots (in the case of compound adjectives). Conceptual integration helps us encode various container characteristics.

The obtained findings also demonstrate that it is this image schema that predetermines the evaluation meaning of the adjectives, as evaluation meaning directly correlates with the particular characteristic of a container element. For example, the small size or hard shell of a container, or a thick or light container substance correlate with negative container value. Interestingly, a container characteristic may bear ambivalent value, for example, "soft", "hard", "large", or "cutting" can be estimated positively or negatively depending on the frame that incorporates the concept of CONTAINER (e.g. *soft-minded, soft-hearted, big-headed*). This enables polysemantic adjectives to comprise opposite evaluation meanings (e.g. *nice*, *cute)*. This study has also shown that the CONTAINER image schema commenced to produce adjectival semantics as long ago as in the Old English period; and it can underlie semantics of structurally diverse words – primary, derived, or compound ones.

It goes without saying that cognitive models are not material things; their knowledge and exploitation are intuitive. However, it is this linguistic intuition that enables language speakers to coin innovations, enhance an adjective's semantics, and express an evaluation of people's qualities. Despite this, cognitive models underlying adjectival semantics do not exist as a separate language phenomenon, as they reflect ways that abstract human qualities are interpreted in English by correlating with conceptual metaphors, in particular, "Minds are containers" and "Ideas are objects". In some cases, this connection becomes apparent and is recognized by language speakers: *She has a razor-sharp mind; He is an incisive critic; He is sharp-witted, He is hard-hearted,* etc.

It could thus be assumed that the CONTAINER image schema affects adjectival semantics in different European languages. In order to compare its role in the interpretation of abstract human qualities among different languages, further research needs to be conducted.

References

Abd Rahim, N. (2013). Adjective identification in television advertisements. In *Procedia – Social and Behavioral Sciences*, *103*, 86-94. https://doi.org/10.1016/j.sbspro.2013.10.311.

Alessandroni, N., & Rodrigues, C. (2017). Is container a natural and embodied image schema? A developmental, pragmatic, and cultural proposal. *Human Development*, *60*(4), 144-178. https://doi.org/10.1159/000478841.

Allan, K. (2008). *Metaphor and Metonymy: A Diachronic Approach.* Wiley-Blackwell.

Baker, M. (2003). *Lexical categories: Verbs, Nouns and Adjectives*. Cambridge University Press.

Barnhart, R.K., & Steinmetz, S. (Eds.). (1988). *The Barnhart Dictionary of Etymology*. Wilson.

Berg, T. (2019). Adjective phrases with doubly modified heads: How lexical information influences word order and constituent structure. *English Language and Linguistics*, *23*(2), 341-361. https://doi.org/10.1017/S1360674317000430.

Bierwisch, M. (1983). *Semantische und konzeptuelle Repräsentation lexikalischer Einheiten* [Semantics and conceptual representation of lexical items]. In R. Růžička & W. Motsch (Eds.), *Untersuchungen zur Semantik*

(pp. 61-99). Akademie Verlag.
Bierwisch, M., & Lang, E. (1987). *Grammatische und konzeptuelle Aspekte von Dimensionsadjectiven* [Grammatical and conceptual aspects of dimensional adjectives] (Vol. 26/27, pp. 1-28). Akademie-Verlag.
Bierwisch, M., & Bosch, P. (1995). *Semantic and Conceptual Knowledge*. IBM Deutschland.
Boldyrev, N.N. (2016). Cognitive schemas of linguistic interpretation. *Voprosy kognitivnoj lingvistiki*, *4*(49), 10-20. http://doi.org/10.20916/1812-3228-2016-4-10-20.
Chafe, W. L. (1970). *Meaning and the structure of language*. University of Chicago Press.
CODEE. (2003). *The Concise Oxford Dictionary of English Etymology*. Oxford University Press.
Cristani, M., Chitó, I., Tomazzoli, C., & Zorzi, M. (2018). A sample algorithm for the lexical classification of comparable adjectives. *Procedia Computer Science*, *126*, 626-635. https://doi.org/10.1016/j.procs.2018.07.297.
Croft, W. (2003). *Typology and universals*. Cambridge University Press.
Cruse, D. (2017). The lexicon. In M. Aronoff & J. Rees-Miller (Eds.), *The Handbook of Linguistics* (pp. 235-255). Wiley.
Dancygier, B., & Vandelanotte, L. (2017). Image-schematic scaffolding in textual and visual artefacts. *Journal of Pragmatics*, *122,* 91-106. https://doi.org/10.1016/j.pragma.2017.07013.
Dixon, R. M. W. (1982). *Where have all the adjectives gone?* Mouton Publishers.
Dixon, R. M. W. (2004). Adjective classes in typological perspective. In R. M. W. Dixon & A. Aikhenvald (Eds.), *Adjective classes. A cross-linguistic typology* (pp. 1-49). Oxford University Press.
Evans, V., & Green M. (2006). *Cognitive linguistics: An Introduction*. Edinburgh University Press.
Gamkrelidze, T. V., & Ivanov, V. V. (1984). *Indoevropejskij yazyk i indoevropejcy* [The Indo-European language and Indo-Europeans] (Vol. 2). Tbilisi University.
Gans, E. (1981). *The Origin of Language: A Formal Theory of Representation*. University of California Press.
Givon, T. (1970). Notes on the semantic structure of English adjectives. *Language*, *46*(4), 816-837.
Givon T. (2014). *On understanding grammar*. Academic Press.
Hampe, B., & Grady, J. E. (2005). *From perception to meaning: Image schemas in cognitive linguistics*. Mouton de Gruyter.
Hiramatsu K., Rulf K. E., & Epstein S. (2010). When knowledge causes failure: Children's extension of novel adjectives and the interpretation of one. In *Lingua*, *120*(5), 1209-1218. https://doi.org/10.1017/S1366728917000232.
Hurtienne, J. (2017). How cognitive linguistics inspires HCI: Image schemas and image-schematic metaphors. In *International Journal of Human-Computer Interaction*, *33*(1), 1-20. https://doi.org/10.1080/10447318.2016.1232227.
Ioannou, G. (2018). Constructions and image-schema preservation. A historical-comparative analysis of PAY in Greek and English. *Lingua, 206*, 85-111. https://doi.org/10.1016/j.lingua.2018.02.001
Ivanov, V. V. (1992). Levyj i pravyj [Left and right]. In S.A. Tokarev (Ed.), *Mify narodov mira. Encyclopeadia* [Myths of the world's nations] (Vol. 2, pp. 43-44). Soviet Encyclopeadia.
Jackendoff, R. (2011). Conceptual Semantics. In C. Maienborn, K. von Heusinger & P. Portner (Eds.), *Semantics. An International Handbook of Natural Language Meaning* (Vol. 1, pp. 688-709). Mouton de Gruyter.
Jespersen, O. (1951). *The philosophy of grammar*. Allen & Unwin.
Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. University of Chicago Press.
Kacnel'son, S. D. (2001). *Kategorii yazyka i myshleniya: Iz nauchnogo naslediya* [Language categories and cognition]. Yazyki Slavyanskoj kul'tury.
Klimov, G. A. (1977). *Tipologiya yazykov aktivnogo stroya* [The typology of active-stative languages]. Nauka.
Kubryakova, E. S. (1999). Semantika v kognitivnoj lingvistike (o koncepte kontejnera i formah ego ob"ektivacii v yazyke) [Semantics from cognitive linguistics perspective (on the concept 'container' and its language representation)]. *Izvestiya RAN. Seriya literatura i yazyk, 58*(5-6), 3-12.
Kubryakova, E. S. (2008). *Chasti rechi v onomasiologicheskom osveshchenii* [Parts of speech from the onomaseological point of view]. LKI.
Lakoff G. (1970). Pronominalization, negation, and the analysis of adverbs. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English transformational grammar* (pp. 145-165). Wiltham.
Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. University of Chicago Press.
Lakoff, G., & Johnson M. (2003). *Metaphors we live by*. The University of Chicago Press.
Lang, E., & Maienborn, C. (2011). Two-level Semantics: Semantic Form and Conceptual Structure. In C. Maienborn, K. von Heusinger & P. Portner (Eds.), *Semantics. An International Handbook of Natural Language*

Meaning (Vol. 1, pp. 709-740). Mouton de Gruyter.
LaTerza, I. (2015). Adjectives and determiners. *Lingua*, *168*, 85-103. http://dx.doi.org/10.1016/j.lingua.2015.09.007.
Martinez, I., Español, S., & Pérez, D. (2018). The interactive origin and the aesthetic modelling of image-schemas and primary metaphors. *Integrative Psychological and Behavioral Science*, *52*(4), 646-671. https://doi.org/10.1007/s12124-018-9432-z.
Maturana, H.R., & Varela, F.J. (1992). *The tree of knowledge: The biological roots of human understanding*. Shambala Publications.
Meletinskij, E. M. (1976). *Poetika mifa* [The myth poetics]. Nauka.
Merriam-Webster (n.d.). Merriam-Webster dictionary online. An Encyclopeadia Britannica Company. http://www.merriam-webster.com/
Nokolaeva I. (2008). Between nouns and adjectives: A constructional view. *Lingua*, *118*(7), 969-996. https://doi.org/10.1016/j.lingua.2006.07.003.
OED. (1989). The Oxford English Dictionary (2nd ed.). The Clarendon Press.
Piaget, J. (1926). *The Language and Thought of the Child*. Routledge & Kegan.
Podosinov, A. V. (1999). *Orientaciya po stranam sveta v arhaicheskih kul'turah* E*vrazii* [Cardinal direction orientation in archaic cultures]. Yazyki Russkoj kul'tury.
Pokorny, J. (1959). *Indogermanisches Etymologisches Wörterbuch* [Indo-Germanic Etymology Dictionary]. http://www.proto-indo-european.ru/dic-pokorny/iew-p.htm
Porohnickaya, L. V. (2014). *Konceptual'nye osnovaniya ehvfemii v yazyke (na materiale anglijskogo, nemeckogo, francuzskogo, ispanskogo i ital'yanskogo yazykov)* [The conceptual basis of euphemia in language: Analysis of English, German, French, Spanish and Italian] [Doctoral dissertation]. Moscow State Linguistic University.
Potebnya, A.A. (1968). *Iz zapisok po russkoj grammatike* [Some notes on Russian grammar] (Vol. 3). Prosveshchenie.
Potebnya, A.A. (1926). *Thought and Language*. State Publishing House.
Pustet R. (2006). Adjectives. In K. Brown (Ed.), *Encyclopedia of Language & Linguistics* (2nd ed., pp. 60-63). https://doi.org/10.1016/B0-08-044854-2/00282-0.
Ratkus, A. (2018). This is not the same: The ambiguity of a Gothic adjective. *Folia Linguistica Historica*, *52*(539-2), 475-494. https://doi.org/10.1515/flih-2018-0017.
Roget, P. M. (1987). *The Original Roget's Thesaurus of English Words and Phrases*. Longman.
Schuchardt-Brevier, H. (1928). *Ein Vademecum der allgemeinen Sprachwissenschaft* [A vademecum of general linguistics] (2nd ed.). Niemeyer.
Senkbeil, K. (2017). Image schemas across modes and across cultures: Communicating horror in Philip Roth's Nemesis and Ridley Scott's Alien. *Language and Literature, 26*(4), 323-339. https://doi.org/10.1177/0963947017739741.
Sosnina, A. (2015). Semantic relations of the adjective 'empty' in Modern English language. *Procedia – Social and Behavioral Sciences*, *200,* 531-536. https://doi.org/10.1016/j.sbspro.2015.08.017.
Trotzke, A., & Wittenberg, E. (2019). Long-standing issues in adjective order and corpus evidence for a multifactorial approach. *Linguistics*, 57 (2), 273-282. https://orcid.org/0000-0002-3943-4282.
Vaan, de M. (2008). *Etymological Dictionary of Latin and the Other Italic Languages*. Brill.
Vernillo, P. (2019). The role of the image schema in the analysis of the semantic variation of action verbs. Data from IMAGACT. *CEUR Workshop Proceedings*, 2347. http://doi.org/urn:nbn:de:0074-2347-1.
Watkins, C. (2000). *The American Heritage Dictionary of Indo-European Roots*. HMH. http://www.proto-indo-european.ru/dic-watkins/ier-p.htm
Weekley, E. (1967). *An Etymological Dictionary of Modern English* (Vol. 2). Dover Publications.
Weicker, M., & Schulz, P. (2019). Red train, big train, broken train: Semantic aspects of adjectives in child language. *Language Acquisition and Language Disorders*, *63*, 203-221. https://doi.org/10.1075/lald.63.11wei.
Wierzbicka, A. (1996). *Semantics: Primes and Universals*. Oxford University Press.
Zhirmunskij, V.M. (1976*). Obshchee i germanskoe yazykoznaniev* [General and Germanic linguistics]. Nauka.
Zykova I.V. (2018). Synesthesia in the process of phraseologism-formation: A new approach. *Yearbook of Phraseology*, *9*(1), 85-110. http://doi.org/10.1515/phras-2018-0006.

Bahari, A. (2020). Use of Nonlinear Dynamic Motivational Strategies to Manage L2 Academic Entitlement and Psychological Reactance. *Journal of Language and Education, 6*(1), 18-38. https://doi.org/10.17323/jle.2020.10099

Use of Nonlinear Dynamic Motivational Strategies to Manage L2 Academic Entitlement and Psychological Reactance

Akbar Bahari
University of Qom

Correspondence concerning this article should be addressed to Akbar Bahari, University of Qom, Ghadir Blvd., 3716146611, Qom, Qom, Iran. E-mail: bahariakbar2020@gmail.com

Expanding the dynamicity and nonlinearity of L2 motivation introduced by Bahari (2019a) based on the complex dynamics systems theory has served as the theoretical framework to introduce and contextualize nonlinear dynamic motivational strategies (NDMSs). The present study used the NDMSs to manage L2 academic entitlement and psychological reactance as two obstructive factors in the L2 learning-teaching process. For conceptualization purposes, a mixed methods approach was conducted among teachers as well as learners to examine the effectiveness of proposed strategies as a pedagogical tool to manage and minimize these obstructive factors in academic contexts. The observed effectiveness of the NDMSs at managing and minimizing the analyzed obstructive factors along with replacing teacher-centered and test-oriented L2 classrooms with a learner-friendly motivating L2 classroom has significant pedagogical and theoretical implications. The major finding of the study following a rigorous methodological triangulation of the data that was collected confirms the effectiveness of the NDMSs as an L2 teaching strategy to cater to the diversity of individual differences for the purpose of improving teacher-learner interactions. Drawing on the results, it can be safely concluded that the NDMSs as the independent variable of the study showed significant impact on managing and minimizing academic entitlement and psychological reactance.

Keywords: academic entitlement (AE), psychological reactance (PR), nonlinear dynamic motivational strategies (NDMSs), complex dynamics systems theory (CDST)

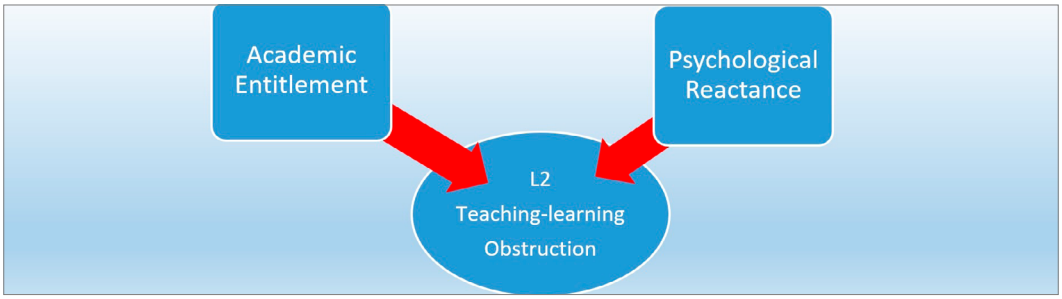
Introduction

When a second language (L2) learner believes that his/her lack of achievement is the result of a teacher's decision, he/she feels entitled to the achievement (Major, 1994) and expresses oppositional behavior via anger mechanisms. Similarly, in situations where individual autonomy or freedom is restrained, oppositional behavior is the common behavior (Brehm, 1996) which leads to resistance, incivility, and dissent as different aspects of reactance. The same oppositional behavior occurs when nonlinear dynamic L2 motivational factors are restrained (Bahari, 2018). Given the obstructive nature of academic entitlement (AE) and psychological reactance (PR) in the L2 teaching-learning context, the present study proposed the use of nonlinear dynamic motivational strategies (NDMSs) as a valid tool for managing and minimizing the obstructive influence of these factors in L2 teaching-learning.

Literature Review

Academic entitlement (AE) is a shift in values of education that undermines the face of education by offering achievement without any effort or skill (Morrow, 1994) or expressing anger over a low grade (Chowning & Campbell 2009; Ciani, Summers, & Easter, 2008; Greenberger, Lessard, Chen, & Farruggia, 2008). AE can trigger failure in learning contexts as an obstructive factor in teacher-learner relationships. While searching for the origins of AE among academically entitled students, several studies (Campbell, Bonacci, Shelton, Exline, & Bushman, 2004; Foster, Campbell, & Twenge, 2003; Hoover 2007; Lombardi, 2007; Twenge 2006) have addressed learner-related concepts (e.g. motivation, narcissism, and (inflated) self-esteem). Under reactance theory (RT), displaying oppositional behavior is a common response in human behavior (Brehm, 1996) however, most of the

Figure 1
L2 Teaching-Learning Obstructive Factors



studies on learning have applied psychological reactance apropos of the learner under labels such as uncivil behavior (Achacoso, 2002; Chowning & Campbell, 2009; Ciani, et al., 2008; Greenberger, et al., 2008; Kopp & Finney, 2013; Lippmann, Bulanda, & Wagenaar, 2009; Nutt, 2013) or an uncivil classroom (Bjorklund & Rehling, 2009; Clark & Springer, 2007; Cortina, Magley, Williams, & Langhout, 2001; Feldmann, 2001). However, this is mostly done without considering the role for the teacher in stirring up uncivil behavior, which is clearly against the basic principle of RT. The basic principle of PR is that oppositional behavior is a common response in human behavior (Brehm, 1996) that is applied to situations where individual autonomy or freedom is restrained by some mechanisms. Therefore, when learner autonomy is restrained, psychological reactance on the part of the learner is a common behavior. The L2 teaching-learning context, when it is the same as other learning contexts, might be the environment where such oppositional behaviors originate. The present study suggests the use of the NDMSs to avoid such an obstructive environment. The question then arises, "Do NDMSs have the potential to manage AE and PR in L2 teaching-learning contexts?" Imagine a teacher taking demotivating measures instead of catering to the motivational needs of L2 learners, or restraining learner autonomy via test-score manipulation. Now the next question arises, "Can this teacher manage reactance, prevent incivility, minimize resistance, and manage dissent while taking anti-motivational measures in L2 teaching-learning contexts?" Restricting learners' pedagogical preferences by ignoring their motivational factors is an example of restricting freedom/autonomy in L2 teaching-learning contexts, which increases the chances of reactance (Jost, Banaji, & Nosek, 2004). Accordingly, adopting test-oriented instruction, threatening policies, reactance-inducing statements and feedback (Bahari, 2019a) obstruct the learning process to a large extent. Given the fact that restrictive measures are met with backlash (Laurin, Kay, Proudfoot, & Fitzsimons, 2013) they need to be avoided in keeping with the internalized concepts of self and identity (Ushioda & Dörnyei, 2017) in L2 motivation.

Based on the reported ineffectiveness of linear patterns in explaining and predicting the relationships observed in SLA data, (Bahari, 2019a, 2020a; de Bot & Larsen Freeman, 2011; de Bot, Lowie, & Verspoor, 2007; Dörnyei, 2014; Fusella, 2013; Hiver, 2015; Kikuchi, 2015; Larsen-Freeman & Cameron, 2008; MacIntyre & Legatto, 2011), the present study tried to incorporate the principles of complex dynamics systems theory (CDST) as a solution to this deficiency. To this end, the NDMSs arranged based on the CDST were used to manage and minimize AE and PR. The CDST considers the components of the system (either with two variables or with innumerable variables) in a global way and confirms that a nonlinear process is at work within the system and in interactions between every internal and external stimuli (Henry, Dörnyei, & Davydenko, 2015; Jiang & Dewaele, 2015).

Given the complete interconnectedness of variables in the language learning context, the proposed NDMSs could potentially influence many other variables in the L2 learning context (e.g. self-efficacy, anxiety, etc.); however, in the present study they are expected to be applied to manage and minimize AE and PR. In other words, given the dynamic trajectory of interacting variables in learning contexts and the ever-changing interactions between them, the study focused on examining the effectiveness of the NDMSs for managing and minimizing two variables out of innumerable dynamic variables present in the language learning context. Therefore, it is impossible to adopt all of the relevant features of a complex system such as the L2 learning system in a single study with respect to feasibility and other limitations. This aspect was rarely approached in previous studies, which approached L2 motivation apropos of strategies (Griffiths, 2013; Oxford, 2017; Quoidbach, Mikolajczak, & Gross, 2015; Schunk & Zimmerman, 2012) or as a static factor (Moskovsky, Racheva, Assulaimani, & Harkins, 2016) or a learner-context interaction subject (Thompson & Erdil-Moody, 2016) or introducing its influential factors (Lyubomirsky & Layous, 2013; Rusk & Waters, 2015; Sheldon, Boehm, & Lyubomirsky, 2013).

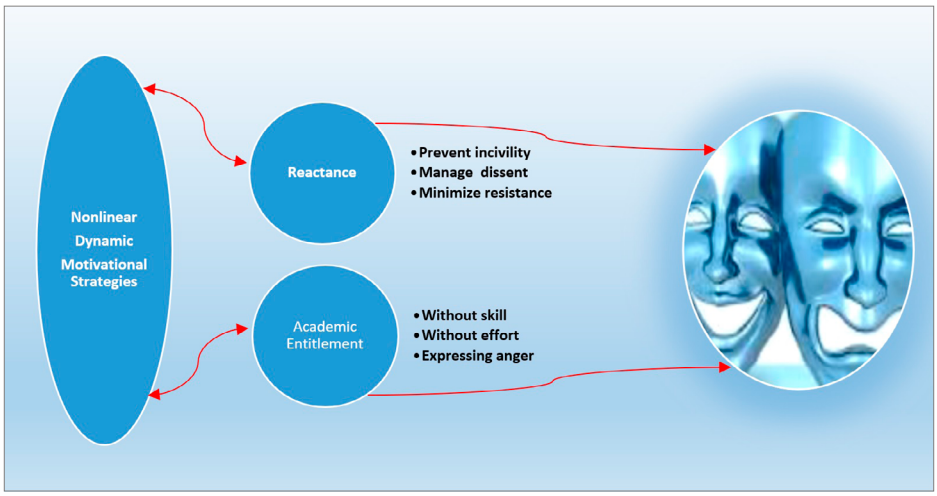
Drawing on CDST and the reported effectiveness of the NDMSs at attenuating the frequency of resorting to oppositional behavior on the part of the learners (Bahari, 2019b), the NDMSs were adopted as a multifaceted pedagogical tool to measure the rate of oppositional behavior among the participants in keeping with the psycho-socio-cultural aspects of L2 learner motivation (Bahari, 2018). To this end, the dynamic psychological, social, and cultural features of L2 motivation were embedded in statements reflecting a dynamically oriented taxonomic structure (Hiver & Al-Hoorie, 2016) in keeping with complex systems (de Bot, Lowie, & Verspoor 2007; Dörnyei, MacIntyre, & Henry, 2015; Dörnyei, Ibrahim, & Muir, 2016; Hiver & Al-Hoorie, 2016; Larsen-Freeman, 2015b; Larsen-Freeman & Cameron, 2008; Serafini, 2017; Thompson, 2017). With the constant changes and adaptation in L2 motivation under CDST, the proposed strategies are meant to intervene (to minimize and manage AE and PR) in interactional synchrony (i.e. the active interacting process) via creative communicative behavior. The main point here is the applicability of CDST to L2 motivation in terms of meeting the nonlinearity and dynamicity of L2 motivation at individual levels (Bahari, 2019a, 2020a) and not its applicability in terms of cognitive processing where information processing is considered a linear process. This is the point where it renders CDST inapplicable.

Description of the Study

The study was conducted to utilize the motivational surges at the individual level (Bahari, 2019b) and unlock the potential behind their nonlinearity and heterogeneity via the NDMSs. The proposed strategies are intended to integrate all of the components of the motivational superstructure from the directed motivational currents described as goal-oriented surges to other intense motivational experiences (Dörnyei, Henry, & Muir, 2016) concerning nonlinear-dynamic L2 motivation. This process starts at the individual level by identifying the motivational factors in members of the learning group and moves onto the group level where identified motivational factors are examined for compatibility. Following that they are nonlinearly integrated and dynamically adopted by the teacher. The NDMSs are applied at three stages: the pre-motivational stage, the motivational stage, and the post-motivational stage. The first stage consists of steps starting with potential motivation diagnosis and ending with nonlinear integration. Drawing on nonlinearity and dynamicity, even the proposed hierarchy allows a dynamic order, which means there is no need to complete all the steps in a linear process and the order can change dynamically and nonlinearly based on what motivates the individual learner at that moment, without trying to impose what motivates one learner compared to another or the whole learner group. To diagnose the motivational disposition of L2 learners, the majority of studies suggest a focus on learners' Motivational Self System (Csizer & Lukacs, 2010; Dörnyei, 2009, 2010; Lamb, 2012; Thompson & Erdil-Moody, 2016) or its variations such as intrinsic or extrinsic motivation, etc. (Csiz´er & Magid, 2014; Dörnyei & Chan, 2013; You, Dörnyei, & Csiz´er, 2016). However, the missing point in these studies is the lack of addressing the nonlinearity and dynamicity of L2 motivation via a comprehensive approach. To bridge this gap, we need to identify tedious and boring experiences from the past as well as the attractive and enjoyable experiences of the learners (Dörnyei, Ibrahim, & Muir, 2015), apropos of dynamics of motivation (Waninge, Dörnyei, & de Bot, 2014) at the nonlinear dynamic individual level instead of the group level (Bahari, 2019a). The present study hopes to approach this problem through the rigorous integration of the elements of psychological reactance (Brehm, 1966; Brehm & Brehm, 1981) into the NDMSs via continual cognitive-motivational functioning (Dörnyei, 2010; Larsen–Freeman, 2015a) to manage and minimize oppositional behaviors on the part of the learners. To this end, it is critical to make sure that previously identified motivational factors are dynamically and nonlinearly compatible in terms of motivational intensity, motivational imagery, and motivational behavior (You & Chan, 2015) with respect to gender differences (Henry & Cliffordson, 2013; You, Dörnyei, & Csiz´er, 2016) and can act together to unlock the potential behind nonlinear dynamic motivation. The dynamic compatibility of the NDMSs should not be confused with 'cohesive groups' which has been used in the literature. While the former is an attempt to find out the rate of compatibility among motivational strategies in order to sort and categorize them under multiple dynamic motivational strategies, the latter is an attempt to group the learners under a single group which is labeled as a 'cohesive group' regardless of the nonlinear and dynamic nature of the motivational factor in each and every member of the learning group.

Given that restricting behavioral options can lead to a preference for the restricted action (Laurin et al., 2013) and that ignoring motivation or demotivation can negatively influence L2 teaching-learning (Chang, 2010; Kikuchi, 2009; Kim, 2009; Oxford, 2017; Quoidbach, Mikolajczak, & Gross, 2015; Trang & Baldauf, 2007), the present study explored the connections of these L2-obstructive factors versus L2-facilitating NDMSs to find a solution to change the opposition-inducing teaching-learning context into a motivation-inducing context.

Figure 2
NDMSs to Manage Psychological Reactance and Academic Entitlement



To this end, the NDMSs were proposed as a facilitator to deal with the challenge of AE and PR in L2 teaching-learning contexts. A mixed methods approach was applied to examine the relationship between the NDMSs and L2 learner-teacher attitudes towards incivility prevention, resistance minimizing, and dissent management with a focus on three aspects: learner-teacher anxiety, frustration, and self-doubt. To familiarize participants with the NDMSs proposed by (Bahari, 2018) and its application in L2 teaching-learning contexts, several workshops (5 sessions for professors' group and 10 sessions for students' group) were conducted before administering questionnaires to provide practical experience for the participants before responding to the NDMSs-oriented questionnaires. Different strands of data collection were used to answer the following questions:

RQ1: How effective are NDMSs as a valid tool for managing and minimizing PR in L2 teaching-learning contexts at three levels of incivility prevention, dissent management, and resistance minimizing?

RQ2: How effective are NDMSs as a valid tool for managing and minimizing AE in L2 teaching-learning contexts?

RQ3: What patterns can be observed between teacher-learner responses about the effectiveness of NDMSs for managing AE and PR in L2 teaching-learning contexts?

RQ4: What relationship can be seen between teachers' and students' responses about the effectiveness of NDMSs (arranged on psycho-socio-cultural aspects of L2 motivation) as a tool to minimize/ manage AE and PR?

Method

Setting and Participants

147 participants (36 professors in English Language Teaching Methodology and 111 M.A. students in ELT Methodology) were drawn from three branches of Azad University in Tehran, Iran. To facilitate QUAL-QUAN analysis, the participants were divided into two groups: professors (female=33% and male=67%) and students (male=41% and female=59%). Professors' ages ranged 32-55 and Students' ages ranged 25-42. Permission to participate in the research was obtained from the participants. The size of the population made it impossible to use random sampling for the purpose of generalizability. The study adopted intact group design to sample the participants and conduct the study.

Data Sources

The required data in response to the first three research questions were gathered by administering author-made questionnaires in English to the L2 teachers and learners (see Appendices A, B). Using the same statements to teacher-learner participants to elicit their attitudes on the effectiveness of the NDMSs on AE and PR

management, the questionnaires were distributed among teachers as well as learners to collect their opinions. For the fourth research question an author-made semi-structured interview (see Appendix C) was rigorously prepared and administered among 50% of the participants (face-to-face/online).

NDMSs - PR Questionnaire

The NDMSs - PR Questionnaire was used to collect the required data in response to the first research question to measure the oppositional behavior. An author-developed survey with 45 items was used to examine the effectiveness of the proposed strategies at three levels of incivility prevention, resistance minimizing, and dissent management from three perspectives: frustration, self-doubt, and learner-teacher anxiety (see Appendix A). The items are rated along a 6-step Likert continuum (e.g., 1 = *strongly agree* to 6 = *strongly disagree*). This was done in keeping with the studies (e.g. Chomeya, 2010) reporting higher discrimination and reliability values of the Likert's 6-point scale versus the Likert's 5-point scale. The questionnaire took approximately 40–45 minutes to complete. Subjects were asked to respond to the items and they were encouraged to ask for explanation if they did not understand any of the items (face-to-face/online). The first fifteen items on the questionnaire assess learner attitudes towards incivility prevention. These are termed incivility prevention *(a=.68),* the belief that learner incivility can be prevented by adopting some strategies with respect to frustration, self-doubt, and learner-teacher anxiety (e.g., "I think friendship strategy can prevent learner incivility and reduce learner-teacher anxiety). The second fifteen items on the questionnaire assess learners' attitudes towards resistance minimizing. These are termed resistance minimizing *(a=.77),* the belief that resistance can be minimized by adopting some strategies with respect to frustration, self-doubt, and learner-teacher anxiety (e.g. "I feel less resistance and frustration when a controversial subject is delivered unbiasedly"). The third fifteen items on the questionnaire assess learners' attitudes towards dissent management. These are termed dissent management *(a=.70),* the belief that one can manage dissent by adopting some strategies with respect to frustration, self-doubt, and learner-teacher anxiety (e.g. "I believe that meeting students' dynamic motivational factors by the teacher can reduce the level of dissent and create a friendly environment with less anxiety and self-doubt"). The internal consistency of the 45 subscales measured in line with Wigfield and Guthrie (1995; Table 1) ranged from .68 to .77 at three levels of resistance minimizing, incivility prevention, and dissent management.

Table 1
Reliabilities of the NDMSs – Psychological Reactance Scale

Subscales	N of Items	Reliability
Incivility Prevention	15	.68
Resistance Minimizing	15	.77
Dissent Management	15	.70

NDMSs - AE Questionnaire

To collect the required data for the second research question, a researcher-made questionnaire with 16 items was prepared and distributed among learners and teachers (see Appendix B) to elicit their opinions concerning L2 AE management via the NDMSs. The author believes that teachers and learners have a mutual role for the creation and rise of the obstructive factors within the L2 teaching environment; therefore we need to collect their opinions on the same statements to avoid making a unidirectional decision. The internal consistency of the items that was measured in line with Wigfield and Guthrie (1995) ranged between .70 and .77.

Interview to collect data on the effectiveness of NDMSs for managing and minimizing AE and PR

To answer the fourth research question, an interview was administered among teachers as well as learners to find out what relationships can be observed between teachers' and learners' responses about the effectiveness of the NDMSs as a tool for minimizing/managing AE and PR (see Appendix C). NDMSs-AE and PR Self-Interview is a 5-part survey (see Appendix C) developed by the author to examine the efficiency of the NDMSs at three levels of preventing, minimizing, and managing obstructive factors. The first part elicits the interviewees' (i.e. teacher/learner) personal experiences (e.g., Have you experienced/witnessed psychological reactance and academic entitlement in terms of incivility, resistance, and dissent?). The elicited responses were interpreted

and coded as 1= positive experience, 2=negative experience, 3=no experience, which were termed *experience (a=.74)*. The second part elicits the interviewees' attitudes concerning the effectiveness of the NDMSs on psychological reactance and academic entitlement with regard to previous experiences by asking questions (e.g., How influential are 'nonlinear dynamic motivation-oriented strategies' in L2 teaching-learning by telling about your own experiences?). The elicited responses were interpreted and coded as 1= influential, 2= uninfluential, 3=undecided. The third part elicits the interviewees' responses regarding the need for psychological reactance and academic entitlement self-management in L2 teaching-learning by asking questions (e.g., How necessary is psychological reactance management in L2 teaching-learning?). The elicited responses were interpreted and coded as 1=necessary, 2=not necessary, and 3=undecided. These are termed *need (a=.72)*. The final construct elicits responses concerning the effectiveness of the NDMSs that can facilitate psychological reactance and academic entitlement self-management in L2 teaching-learning (e.g. eliminating test-oriented classes, providing novel activities, and improving learner achievement) by asking question (e.g., Do you think that eliminating test-oriented classes can facilitate preventing incivility, minimizing resistance, and managing dissent in face-to-face/online L2 teaching-learning?). The elicited responses were interpreted and coded as 1= Yes, 2= No, and 3= Undecided. To determine the internal consistency, the subscales (LL^1, LL^2, and LSL) were measured in line with Wigfield and Guthrie (1997) alphas and the results ranged from .71 to .75 (Table 2).

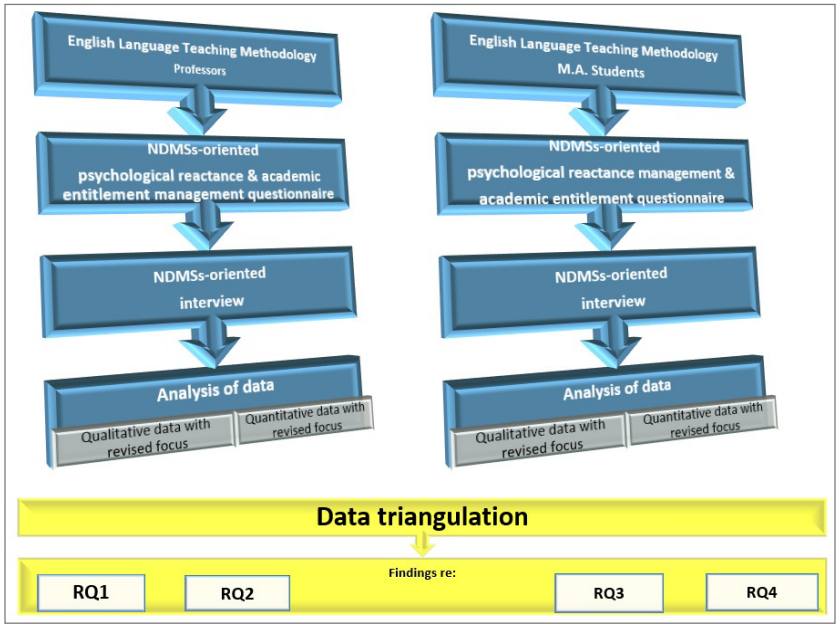
Table 2
Reliabilities for the NDM-oriented reactance management interview Subscales

Subscale	N of Items	Reliability
Experience	2	.74
Effectiveness	3	.71
Need	2	.72
Strategy	3	.75

Data Analysis

Thematic analysis of QUAL-QUAN data was conducted via mixing the data for the purpose of meta-inferences (Figure 3) in line with (Tashakkori & Teddlie, 2003). To ensure the correctness of the results, an iterative analysis was done prior to the selection of QUAL-QUAN analyses. Quantitative analyses were used to examine the effectiveness of the NDMSs for managing and minimizing AE and PR. Given the unpaired and categorical nature of the collected data, to test the effectiveness of the NDMSs for managing or minimizing AE and PR, the Pearson Chi Square Test was used.

Figure 3
Visual Representation of the Study Design



The transcribed interviews were scrutinized for codable NDMSs-oriented statements. In keeping with Urdan and Mestas (2006), the presence/absence of references (explicit/implicit) to obstructive factors in elicited responses was determined as the criterion for data coding. Sorting and categorizing elicited data in keeping with Saldaña (2013) provided a picture of the orientation of the collected responses following the use of sub-coding techniques and preparing a list of codes (see Appendix C). To resolve the discrepancies and ensure inter-rater reliability, three expert EFL researchers were invited and the results of the analysis confirmed the inter-rater agreement at 78% per interview on average.

Results

The results of analyzing participants' responses to the NDMSs-PR questionnaire are displayed at three levels of civility prevention, resistance minimizing, and dissent management. The results showed that the majority of the participants (100% of students and 48% of professors) held a positive view about the effectiveness of the NDMSs for preventing, minimizing, and managing psychological reactance. Drawing on the obtained results, the study confirms the effectiveness of the NDMSs and suggests that future studies should delve into the potential behind the NDMSs by contextualizing it in other contexts. The mean of the observed standard deviations M=1.185 shows (see Table 3) that there were no polarized responses (except for negative responses elicited from male professors) and the majority of the participants (51.7%) strongly believe in the effectiveness of the NDMSs.

Table 3
Descriptive Statistics on NDMSs & PR (Incivility Prevention)

NDMSs & PR (Incivility Prevention)					
		Frequency	**Percent**	**Valid Percent**	**Cumulative Percent**
Valid	strongly agree	76	51.7	51.7	51.7
	agree	40	27.2	27.2	78.9
	slightly agree	11	7.5	7.5	86.4
	slightly disagree	12	8.2	8.2	94.6
	disagree	8	5.4	5.4	100.0
	Total	**147**	**100.0**	**100.0**	

The observed frequencies concerning the participants' responses on the effectiveness of the NDMSs are displayed in the frequency table of the Chi-Square Test (see Table 4). The majority of the respondents (127 out of 147) held a positive opinion concerning the effectiveness of the proposed strategies to prevent incivility in L2 learning environments. The upper range of the Table (i.e. 6) shows no cases of negative opinions, while the expected number was 42.0. This can be interpreted as the lack of a strong negative opinion on the part of the participants (particularly teachers who mostly expressed negative opinions). In other words, there might be some other reasons for opting disagree/slightly disagree beyond the effectiveness of the proposed strategies (e.g. they might think that other factors should be included, etc.).

Table 4
Chi-Square Test for NDMSs & PR (Incivility Prevention)

Frequencies				
	NDMSs & PR (Incivility Prevention)			
	Category	**Observed N**	**Expected N**	**Residual**
1	strongly agree	76	7.0	69.0
2	agree	40	14.0	26.0
3	slightly agree	11	21.0	-10.0
4	slightly disagree	12	28.0	-16.0
5	disagree	8	35.0	-27.0
6	Strongly disagree	0	42.0	-42.0
Total		**147**		

The test statistics (see Table 5) compare the expected and observed values. In this case, the discrepancy was very large and statistically significant. (Asymp. Sig. = .000). The results of the Chi-Square test concerning the observed frequency of responses between participants (i.e. teachers and learners) and a significant interaction was found X^2 (5) = p < .05. Therefore, the majority of students as well as the majority of female teachers (except for 62% of male teachers) had a positive opinion about the effectiveness of the NDMSs as a valid tool to prevent incivility among L2 learners.

Table 5
Chi-Square Test Statistics (Incivility Prevention)

Test Statistics			
			NDMSs & PR (Incivility Prevention)
Chi-Square			805.162[a]
df			5
Asymp. Sig.			.000
Monte Carlo Sig.	Sig.		.000[b]
	99% Confidence Interval	Lower Bound	.000
		Upper Bound	.000

In response to part of the first research question, a Chi-Square test was run to provide a picture of the observed frequencies concerning the participants' responses regarding the effectiveness of the NDMSs for the dissent management variable (see Table 4). The majority of the respondents (126 out of 147) held a positive opinion concerning the effectiveness of the proposed strategies for dissent management in L2 learning environments. A look at the upper range of the Table (i.e. 6) reveals that the observed negative cases was zero, while the expected number was 42.0. This reflects the lack of a strong negative opinion on the part of the participants (even male teachers who mostly expressed negative opinions but did not select the strongly negative choice).

Table 6
Chi-Square Test (Dissent Management)

Frequencies				
	NDMSs & PR (Dissent Management)			
	Category	**Observed N**	**Expected N**	**Residual**
1	strongly agree	89	7.0	82.0
2	agree	31	14.0	17.0
3	slightly agree	6	21.0	-15.0
4	slightly disagree	12	28.0	-16.0
5	disagree	9	35.0	-26.0
6		0	42.0	-42.0
Total		**147**		

The Chi-Square test statistics (see Table 7) compared the expected and observed values. In this case the discrepancy was very large and statistically significant. (Asymp. Sig. = .000). The Chi-Square calculated the frequency of responses between participants (i.e. teachers and learners), and a significant interaction was found X^2 (5) = p < .05. Therefore, the majority of students as well as the majority of female teacher (except for male teachers) had positive opinions on the effectiveness of the NDMSs as a valid tool to manage dissent among L2 learners.

Table 7
Chi-Square Test Statistics

Test Statistics			
			NDMSs & PR (Dissent Management)
Chi-Square			1062.386[a]
df			5
Asymp. Sig.			.000
Monte Carlo Sig.	Sig.		.000[b]
	99% Confidence Interval	Lower Bound	.000
		Upper Bound	.000

In response to the other part of the first research question, a Chi-Square Test was run (see Table 8). The goal was to evaluate whether the NDMSs are effective tools to minimize resistance. Based on the results, 63 out of 147 strongly agreed, 52 out of 147 agreed, and 11 out of 147 participants held a positive view on the effectiveness of the proposed strategies for minimizing resistance. As with the previous analyses of the PR levels (incivility prevention and dissent management), we observe that there was no strongly disagree choice observed in the elicited responses from the participants, which reflects their tendency towards the positive opinions.

Table 8
Chi-Square Test Frequencies (Resistance Minimizing)

Frequencies				
	NDMSs & PR (Resistance Minimizing)			
	Category	**Observed N**	**Expected N**	**Residual**
1	strongly agree	63	7.0	56.0
2	agree	52	14.0	38.0
3	slightly agree	11	21.0	-10.0
4	slightly disagree	12	28.0	-16.0
5	disagree	9	35.0	-26.0
6		0	42.0	-42.0
Total		**147**		

Table 8 shows the Chi-Square analysis that calculated the frequency of responses between participants (i.e. teachers and learners), and a significant interaction was found X^2 (5) = p < .001. From the test statistics output table we can observe the Chi Squared statistic, $x2 = 626.362$, degrees of freedom 5, corresponding to p < 0.001. Therefore, we can conclude with 99.9% confidence that there is very strong evidence of the association between adopting the NDMSs and minimizing resistance.

Table 9
Test Statistics

Test Statistics			
			NDMSs & PR (Resistance Minimizing)
Chi-Square			626.362[a]
df			5
Asymp. Sig.			.000
Monte Carlo Sig.	Sig.		.000[b]
	99% Confidence Interval	Lower Bound	.000
		Upper Bound	.000

The results of the above analyses confirm the effectiveness of the NDMSs for managing and minimizing PR at three forms of incivility, resistance, and dissent. Accordingly, both groups believed that the NDMSs have the potential to manage these obstructive factors in classroom contexts. The negative responses by 52% of the professors (42% of the female professors and 62% of the male professors) can be interpreted as their traditionally-established preferences for teacher-centered L2 teaching. This also reflects their orientation towards monologic

instruction instead of a dialogic one, which ends up in a non-learner-friendly context with a high risk of causing AE and PR in L2 teaching-learning.

Running a Chi-square test in response to the second research question revealed the following results. The results of the Chi-Square test frequencies showed a significant number of responses were positive (ranging from strongly agree =79, agree =51, to slightly agree =8), with only 9 out of 147 negative opinions on the effectiveness of the NDMSs for managing and minimizing AE among L2 learners (Table 10).

Table 10
Chi-Square Test Frequencies

Frequencies				
	NDMSs & Academic Entitlement			
	Category	**Observed N**	**Expected N**	**Residual**
1	strongly agree	79	7.0	72.0
2	agree	51	14.0	37.0
3	slightly agree	8	21.0	-13.0
4	slightly disagree	6	28.0	-22.0
5	disagree	3	35.0	-32.0
6		0	42.0	-42.0
Total		**147**		

The results of the Chi-Square Test (see Table 11) calculating the frequency of responses elicited from the participants (i.e. teachers and learners), showed a significant interaction between the NDMSs and AE X^2 (5) = p < .001. We can observe from the test statistics output table the Chi Squared statistic, $x2$ = 934.948, degrees of freedom 5, corresponding to p < 0.001. Therefore, based on the obtained results the present study confirms with 99.9% confidence that there is very strong evidence of the association between adopting the NDMSs and minimizing resistance among the L2 learners in the study.

Table 11
Chi-Square Test Statistics on NDMSs & Academic Entitlement

Test Statistics			
			NDMSs & Academic Entitlement
Chi-Square			934.948[a]
df			5
Asymp. Sig.			.000
Monte Carlo Sig.	Sig.		.000[b]
	99% Confidence Interval	Lower Bound	.000
		Upper Bound	.000

In an answer to the third question, searching for patterns between teacher-learner responses about the effectiveness of the NDMSs for managing AE and PR, the results of the analyses revealed that there was a statistically significant correlation between different forms of psychological reactance (incivility prevention, resistance minimizing, and dissent management) with Sig. (2-tailed) < 0.05. A closer look at our results reveals that the strongest correlation was between incivility prevention and resistance management, where r = .936. It was based on N = 147 students and teachers and its 2-tailed significance, p=0.000, which means there is a 0.000 probability of finding this sample correlation or a larger one if the actual population correlation is zero. The results of the analysis also revealed that there was no correlation between PR and AE in the elicited data. Its strongest correlation was between the NDMSs & AE and incivility prevention where the correlation was .039 but p=.635 which was not statistically different from zero. That is there was a .039 chance of finding it if the population correlation was zero.

Table 12
Correlations

Correlations					
		NDMSs & PR (Incivility Prevention)	**NDMSs & PR (Resistance Minimizing)**	**NDMSs & PR (Dissent Management)**	**NDMSs & AE**
NDMSs & PR (Incivility Prevention)	Pearson Correlation	1	.936	.874	.039
	Sig. (2-tailed)		.000	.000	.635
	N	147	147	147	147
NDMSs & PR (Resistance Minimizing)	Pearson Correlation	.936	1	.918	.081
	Sig. (2-tailed)	.000		.000	.331
	N	147	147	147	147
NDMSs & PR (Dissent Management)	Pearson Correlation	.874	.918	1	-.049
	Sig. (2-tailed)	.000	.000		.558
	N	147	147	147	147
NDMSs & AE	Pearson Correlation	.039	.081	-.049	1
	Sig. (2-tailed)	.635	.331	.558	
	N	147	147	147	147

To answer the fourth research question, the elicited responses from half of the respondents (n=74) were converted into three types of answers (i.e. yes, no, undecided) then the data were analyzed in terms of frequency of each answer type (see Table 13) with respect to position (teacher/student). Both students and learners provided revealing data in terms of the necessity of providing a learner-friendly environment where a variety of motivational factors from psychological to cultural and social vantage points are catered for. Some of the interviewees also referred to the concept of identity preservation, which was not precisely the focus of the study but might be addressed in future studies. They believed that one of the outcomes of meeting nonlinear dynamic motivation at individual level (which is the focus of the present study) is that those students with modest backgrounds or from a different race will have an opportunity to introduce/discuss and reinforce their personal nonlinear dynamic motivational factors. The interviewees also described the NDMSs as effective tools to manage AE and PR in L2 learning because they can create motivational individual learners who have identified their nonlinear dynamic motivational factors (by themselves or with their teachers) and have recognized the large range of dynamicity and nonlinearity from one classmate to another.

Table 13
Interview Frequencies

Interview Frequencies				
Count				
		Position		**Total**
		Student	Teacher	
Interview	Yes	12	56	**68**
	Undecided	3	0	**3**
	No	3	0	**3**
Total		**18**	**56**	**74**

Discussion

The AE and the PR, in the learning context, negatively affect the learning process and it is critical to elaborate on this problem. The NDMSs were suggested as the solution to this problem, but the goal was to elicit students' opinions as well as teachers' opinions on their effectiveness to manage or minimize AE and PR. Drawing on psychological reactance theory, the required data were collected from both sides of the oppositional behavior (teachers and students) concerning the effectiveness of the proposed strategies. The quantitative results of

the study revealed that 85.71% of the participants held a positive view on the effectiveness of the proposed strategies. The data triangulation of the collected data from the questionnaire and the interview confirmed the effectiveness of the proposed strategies along with a number of suggestions for further studies in the field with the same focus. Individual differences are common and holding double standards is also common among some teachers. Therefore, it is highly critical to take precautionary measures and prevent oppositional behavior via the NDMSs in a learner-friendly environment. As one of the interviewees argued, how could a teacher with double standards fairly evaluate language proficiency in a test-oriented classroom with some students holding opposite standards (different religion, ethnicity, political party, etc.), or how could a demotivating L2 teacher using discriminatory statements, feedback, responses, etc. create an environment free from AE and PR? Given the probability of such discriminatory situations where there is a high chance of stirring psychological reactance and academic entitlement, some questions arise: Is it fair to blame only learner(s) for feeling entitled to achievement/psychological reactance?, Are teachers sufficiently well-informed and well-equipped to prevent academic entitlement/psychological reactance?

The findings of the study are in line with the reported relationship between AE and motivation (Graham, & Weiner, 1996; Lerner, 1987) and PR and motivation (Jost, Banaji, & Nosek, 2004). Given the obtained results, it seems imperative to conduct further studies in terms of demotivating L2 teaching environments. Ruling classrooms with an iron fist is an explicit violation of motivational factors and leads to a variety of obstructive factors. Additionally, with regards to similar characteristics of AE and PR with L2 motivation in terms of nonlinearity and dynamicity, further studies are necessary to clarify their connections. Accordingly, it is critical to investigate and determine the potential behind nonlinearity and dynamicity of language learning and language learners (Bahari, 2018; Buehl & Beck, 2015; Fives & Gill, 2015; Fives, Lacatena, & Gerard, 2015) along with learning-related findings of studies on psychological reactance (Chartrand, Dalton, & Fitzsimons, 2007; Dillard & Shen, 2005; Eagly, Mladinic, & Otto, 1994; Kim, Levine, & Allen, 2013; Quick & Considine, 2008; Quick & Stephenson, 2007; Rains, 2013) towards a comprehensive PR-free model of learning.

The misguided theoretical trend in L2 motivation studies is mainly under the influence of L2 self systems introduced by Dörnyei (2009) and the studies that tried to expand it (e.g. Liu & Thompson, 2018; Thompson, 2017) at the cost of confusing anti-ought-to-self with a form of PR (see Bahari, 2020a). The current theoretical trends fail to incorporate the dynamicity and nonlinearity of L2 motivation, which varies from one learner to another. Some peer-reviewed journals such as *System* and *Modern Language Journal,* without conducting rigorous review, publish articles contextualizing or expanding this deficient theoretical system and serve to mix up PR and its forms with anti-ought-to-self for readers. Psychological reactance is a psychological state of mind that might cause a range of oppositional expressions (e.g. incivility, dissent, and resistance). However, anti-ought-to-self is a type of self, which serves to negatively motivate the L2 learner to continue the learning process. Briefly, anti-ought-to-self negatively motivates the learner to do something positive (i.e. learning) while PR negatively motivates the learner not to do something positive (i.e. learning). While the former potentially facilitates learning, the latter practically obstructs learning.

Pedagogical Implications

The most salient pedagogical implication of the study is about test-oriented classes and their negative effect on managing and minimizing AE and PR. Such classes increase the chances of the emergence of AE and PR in the three forms of incivility, dissent, and resistance along with learner-teacher self-doubt, anxiety, and frustration. Such classes not only provide some teachers with a manipulative tool (i.e. test score manipulation) to threaten or oppress learner autonomy but also help them dodge their responsibility to prepare novel and creative activities compatible with learners' dynamic and nonlinear motivation. Some of the teacher-participants also implicitly confirmed the existence of such reactance-inducing conditions in L2 teaching-learning contexts. Therefore, some pedagogical reformations are needed to address these anti-learner features of test-oriented classes that affect L2 teaching-learning environments. The second implication is that reinforcing nonlinear dynamic motivation has the potential to prevent incivility, minimize resistance, and manage dissent along with catering to the motivational needs of the L2 learner group at the same time. The NDMSs-oriented pedagogy ensures learner-friendly environments where L2 learner motivation is catered to at individual level (Bahari, 2020b). In keeping with the dynamicity and nonlinearity of learner's motivation, the third implication of the study is the need to foster the collaborative meaning-making process through dialogic discourse instead of traditionally established monologic discourse in L2 teaching-learning. This is to integrate the dynamic and nonlinear features

of L2 motivation at the individual level with that of the group level during interactional synchrony. While the former discourse type permits argumentative virtues, the latter one fosters teacher-centered teaching beliefs.

Conclusion

Based on participants' beliefs, this study confirms the effectiveness of the NDMSs as a valid tool for minimizing and managing psychological reactance and academic entitlement in L2 learning-teaching contexts. Accordingly, it can safely be concluded that restraining learner's freedom of voice, preferences, and options within a demotivating, undemocratic L2 classroom causes oppositional behaviors, which need to be avoided on the part of the teacher. The study confirms a positive opinion among teachers as well as learners on the applicability of the NDMSs as a minimizing and managing tool for academic entitlement and psychological reactance. The study suggests new pedagogical reforms in terms of teachers' belief systems about teaching practices with a focus on the nonlinearity and dynamicity of motivation, which differs from learner to learner. Given the thin literature apropos of the nonlinearity and dynamicity of L2 motivation, future studies are suggested to retest the applicability and contextualization of the NDMSs in different learning contexts and at different language proficiency levels. The main point of the study is to benefit from the diversity of L2 motivation in learners to create an every-learner-motivated classroom.

References

Achacoso, M. V. (2002). *"What do you mean my grade is not an A?" An investigation of academic entitlement, causal attributions, and self-regulation in college students* [Unpublished doctoral dissertation]. The University of Texas at Austin.

Bahari, A. (2018). Nonlinear dynamic motivation-oriented telecollaborative model of language learning via formulaic sequences to foster learner autonomy. *The Journal of Teaching English with Technology, 18*(3), 65–81. http://www.tewtjournal.org

Bahari, A. (2019a). FonF practice model from theory to practice: CALL via focus on form approach and nonlinear dynamic motivation to develop listening and speaking proficiency. *Computers & Education*, *130*(3), 40-58. http://doi.org/10.1016/j.compedu.2018.11.009

Bahari, A. (2019b). The Impact of applying FonF practice model on developing L2 listening and speaking with a focus on intentional and incidental vocabulary acquisition in CALL context. *Revista de Lingüística y Lenguas Aplicadas, 14*, 1-15. http://doi.org/10.4995/rlyla.2019.10785

Bahari, A. (2020a). Mainstream theoretical trends and future directions of L2 motivation studies in classroom and CALL contexts. *CALL-EJ, 21*(1), 1–28.

Bahari, A. (2020b). Computer-Assisted Language Proficiency Assessment Tools and Strategies. *Open Learning: The Journal of Open, Distance and e-Learning, 35*(2), 34–59. http://doi.org/10.1080/02680513.2020.1726738

Bjorklund, W. L., & Rehling, D. L. (2009). Student perceptions of classroom incivility. *College Teaching, 58*(1), 15–18. https://doi.org/10.1080/87567550903252801

de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition 10*(1), 7–21. https://doi.org/10.1017/S1366728906002732

de Bot, K., & Larsen Freeman, L. (2011). Researching second language development from a dynamic systems theory perspective. In M. H. Verspoor, K. de Bot & W. Lowie (Eds.), *A dynamic approach to second language development: Methods and techniques* (pp. 5–24). John Benjamins.

Brehm, J. W. (1966). *A theory of psychological reactance.* Academic Press.

Brehm, S., & Brehm, J. W. (1981). *Psychological reactance: A theory of freedom and control.* Academic Press.

Buehl, M. M., & Beck, J. S. (2015). The relationship between teachers' beliefs and teachers' practices. In H. Fives & M. Gregoire Hill (Eds.), *International handbook of research on teachers' beliefs* (pp. 66–84). Routledge.

Campbell, W. K., Bonacci, A. M., Shelton, J., Exline, J. J., & Bushman, B. J. (2004). Psychological entitlement: Interpersonal consequences and validation of a self-report measure. *Journal of Personality Assessment, 83*(1), 29-45. https://doi.org/10.1207/s15327752jpa8301_04

Chartrand, T. L., Dalton, A. N., & Fitzsimons, G. J. (2007). Nonconscious relationship reactance: When significant others prime opposing goals. *Journal of Experimental Social Psychology*, *43*, 719–726. https://doi.org/10.1016/j.

jesp.2006.08.003
Chomeya, R. (2010). Quality of Psychology Test between Likert Scale 5 and 6 Points. *Journal of Social Sciences, 6*(3), 399–403.
Chowning, K., & Campbell, N. J. (2009). Development and validation of a measure of academic entitlement: Individual differences in students' externalized responsibility and entitled expectations. *Journal of Educational Psychology, 101*(4), 982–997. https://doi.org/10.1037/a0016351
Ciani, K. D., Summers, J. J., & Easter, M. A. (2008). Gender differences in academic entitlement among college students. *The Journal of Genetic Psychology, 169*(4), 332–344. https://doi.org/10.3200/GNTP.169.4.332-344
Clark, C. M., & Springer, P. J. (2007). Thoughts on incivility: Student and faculty perceptions of uncivil behavior in nursing education. *Nursing Education Perspectives, 28*(2), 93–97. https://doi.org/10.12691/ajnr-4-3-2.
Cortina, L. M., Magley, V. J., Williams, J. H., & Langhout, R. D. (2001). Incivility at the workplace: Incidence and impact. *Journal of Occupational Health Psychology, 6,* 64-80. https://doi.org/10.1037/1076-8998.6.1.64
Csizer, K., & Luk_acs, G. (2010). The comparative analysis of motivation, attitudes and selves: The case of English and German in Hungary. *System, 38*(1), 1–13. https://doi.org/10.1016/j.system.2009.12.001
Csizer, K., & Magid, M. (Eds.). (2014). *The impact of self-concept on language learning*. Multilingual Matters.
Dillard, J. P., & Shen, L. (2005). On the nature of reactance and its role in persuasive health communication. *Communication Monographs*, *72*, 144-168. http://doi.org/10.1080/03637750500111815
Dörnyei, Z. (2009). *The L2 motivational self-system.* In Z. Dörnyei & E. Ushioda (Eds.), Motivation, language identity and the L2 self (pp. 9-42). Multilingual Matters. http://doi.org/10.14746/ssllt.2018.8.4.2
Dörnyei, Z. (2010). The relationship between language aptitude and language learning motivation: Individual differences from a dynamic systems perspective. In E.Macaro (Ed.), *Continuum companion to second language acquisition* (pp. 247–267). Continuum. http://doi.org/10.5040/9781474212397.ch-009
Dörnyei, Z., & Chan, L. (2013). Motivation and vision: An analysis of future L2 self-images, sensory styles, and imagery capacity across two target languages. *Language Learning*, *63*, 437–462. http://doi.org/10.1111/lang.12005
Dörnyei, Z., Henry, A., & Muir, C. (2016). *Motivational currents in language learning: Frameworks for focused interventions*. Routledge.
Dörnyei, Z., Ibrahim, Z., & Muir, C. (2015). Directed Motivational Currents: Regulating complex dynamic systems through motivational surges. In Z. Dörnyei, Z., MacIntyre, P., & Henry, A. (Eds.), *Motivational dynamics in language learning.* Multilingual Matters.
Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.
Eagly, A. H., Mladinic, A., & Otto, S. (1994). Cognitive and affective bases of attitudes toward social groups and social policies. *Journal of Experimental Social Psychology*, *30*, 113-137. http://doi.org/10.1006/jesp.1994.1006
Feldmann, L. J. (2001). Classroom civility is another of our instructor responsibilities. *College Teaching, 49,* 137–140. https://doi.org/10.1080/87567555.2001.10844595
Fives, H., & Gill, M. G. (Eds.). (2015). International handbook of research on teachers' beliefs. Routledge.
Fives, H., Lacatena, N., & Gerard, L. (2015). Teachers' beliefs about teaching (and learning). In H. Fives & M. G. Gill (Eds.), International handbook of research on teachers' beliefs (pp. 249–265). Routledge.
Foster, J. D., Campbell, W. K., & Twenge, J. M. (2003). Individual differences in narcissism: Inflated self views across the lifespan and around the world. *Journal of Research in Personality, 37*(6), 469-486. https://doi.org/10.1016/S0092-6566(03)00026-6
Fusella, P. V. (2013). Dynamic systems theory in cognitive science: Major elements, applications, and debates surrounding a revolutionary meta-theory. Dynamic Psychology, *1,* 258-307. http://dynapsyc.org/2013/Fusella.pdf
Graham, S., & Weiner, B. (1996). Theories and principles of motivation. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 63–84). Macmillan.
Greenberger, E., Lessard, J., Chen, C., & Farruggia, S. P. (2008). Self- entitled college students: Contributions of personality, parenting, and motivational factors. *Journal of Youth and Adolescence, 37*(10), 1193–1204. https://doi.org/10.1007/s10964-008-9284-9
Griffiths, C. (2013). *The strategy factor in successful language learning*. Multilingual Matters.
Henry, A., Davydenko, S., & Dörnyei, Z. (2015). The anatomy of directed motivational currents: Exploring intense and enduring periods of L2 motivation. *Modern Language Journal*, *99*, 329- 345. https://doi.org/10.1111/modl.12214
Hiver, P. (2015). Attractor states. In Z. Dörnyei, P. MacIntyre, & A. Henry (Eds.), Motivational dynamics in language learning (pp. 20–28). Multilingual Matters. https://doi.org/10.13140/RG.2.1.2501.8722
Hiver, P., & Al-Hoorie, A. H. (2016). A dynamic ensemble for second language research: Putting Complexity

Theory into practice. *Modern Language Journal*, *100*, 741–756. https://doi.org/10.1111/modl.12347
Hoover, E. (2007). Here's looking at you, kid: Study says many students are narcissists. *Chronicle of Higher Education, 53*(27). https://eric.ed.gov/?id=EJ757228
Jiang, Y., & Dewaele, J. M. (2015). What lies bubbling beneath the surface? A longitudinal perspective on fluctuations of ideal and Ought-to L2 self among Chinese learners of English. *International Review of Applied Linguistics in Language Teaching*, *53*, 331–354. https://doi.org/10.1515/iral-2015-0015
Jost, J. T., Banaji, M. R., & Nosek, B. A. (2004). A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo. *Political Psychology*, *25*, 881–919. https://doi.org/10.1111/j.1467-9221.2004.00402.x
Kikuchi, K. (2015). Demotivation in second language acquisition. Multilingual Matters.
Kim, K. J. (2009). Demotivating factors in secondary English education. *English Teaching, 64*(4), 249–267. https://doi.org/10.15858/engtea.64.4.200912.249
Kim, S., Levine, T., & Allen, M. (2013). Comparing separate process and intertwined models for reactance. *Communication Studies*, *64*, 273-295. http://doi.org/10.1080/10510974.2012.755639
Lamb, M. (2012). A self-system perspective on young adolescents' motivation to learn English in urban and rural settings. *Language Learning, 62*(4), 997–1023. https://doi.org/10.1111/j.1467-9922.2012.00719.x
Larsen–Freeman, D. (2015a). Ten 'lessons' from Complex Dynamic Systems Theory: What is on offer. In Z. Dörnyei, P. D. MacIntyre & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 11–19). Multilingual Matters. https://doi.org/10.21832/9781783092574-004
Larsen-Freeman, D. (2015b). Complexity theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (2nd ed., pp. 227-244). Routledge
Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford University Press.
Laurin, K., Kay, A. C., Proudfoot, D., & Fitzsimons, G, J. (2013). Response to restrictive policies: Reconciling system justification and psychological reactance. *Organizational Behavior and Human Decision Processes, 122,* 152–162. https://doi.org/10.1016/j.obhdp.2013.06.004
Lerner, M. J. (1987). Integrating societal and psychological rules of entitlement: The basic task of each social actor and fundamental problem for the social sciences. *Social Justice Research, 1,* 107–125. https://doi.org/10.1007/BF01049386
Liu, Y., & Thompson, A.S. (2018). Language learning motivation in China: An exploration of the L2MSS and psychological reactance. *System, 72*, 37–48. https://doi.org/10.1016/j.system.2017.09.025
Lombardi, J. V. (2007). The academic success entitlement. *Inside Higher Education*. http://www.insidehighered.com/
Lyubomirsky, S., & Layous, K. (2013). How do simple positive activities increase well-being? *Current Directions in Psychological Science, 22*, 57–62. http://doi.org/10.1177/0963721412469809
MacIntyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing and idiodynamic method to capture rapidly changing affect. *Applied Linguistics, 32*(2), 149–171. https://doi.org/10.1093/applin/amq037
Major, B. (1994). From social inequality to personal entitlement: The role of social comparisons, legitimacy appraisals, and group memberships. *Advances in Experimental Social Psychology, 26*, 293–355. https://doi.org/10.1016/S0065-2601(08)60156-2
Miron, A. M., & Brehm, J. W. (2006). Reactance theory-40 years later. *Social Psychology*, *37*, 9–18. https://doi.org/10.1024/0044-3514.37.1.9
Morrow, W. (1994). Entitlement and achievement in education. *Studies in Philosophy and Education, 13*(1), 33–47. https://doi.org/10.1007/BF01074084
Moskovsky, C., Racheva, S., Assulaimani, T., & Harkins, J. (2016). The L2 Motivational Self System and L2 Achievement: A study of Saudi EFL learners. *The Modern Language Journal, 100*(3), 641–654. https://doi.org/10.1111/modl.12340
Oxford, R. (2017). *Teaching and researching language learning strategies: Self-regulation in context* (2nd ed.). Routledge.
Quick, B. L., & Considine, J. R. (2008). Examining the use of forceful language when designing exercise persuasive messages for adults: A test of conceptualizing reactance arousal as a two-step process. *Health Communication*, *23*, 483-491. http://doi.org/10.1080/10410230802342150
Quick, B. L., & Stephenson, M. T. (2007). Further evidence that psychological reactance can be modeled as a combination of anger and negative cognitions. *Communication Research*, *34*, 255–276. http://doi.org/10.1177/0093650207300427
Quoidbach, J., Mikolajczak, M., & Gross, J. J. (2015). Positive interventions: An emotion regulation perspective.

Psychological Bulletin, 141, 655–693. http://doi.org/10.1037/a0038648
Rains, S. A. (2013). The nature of psychological reactance revisited: A meta-analytic review. *Human Communication Research*, *39*, 47-73. http://doi.org/10.1111/j.1468-2958.2012.01443.x
Rusk, R. D., & Waters, L. (2015). A psycho-social system approach to well-being: Empirically deriving the five domains of positive functioning. *The Journal of Positive Psychology, 10*, 141–152. http://doi.org/10.1080/17439760.2014.920409
Schunk, D. H., & Zimmerman, B. J. (2012). *Motivation and self-regulated learning: Theory, research, and applications*. Routledge.
Serafini, E. J. (2017). Exploring the dynamic long-term interaction between cognitive and psychosocial resources in adult second language development at varying proficiency. *The Modern Language Journal*, *101*(2), 369-390. http://doi.org/10.1111/modl.12400
Sheldon, K., Boehm, J., & Lyubomirsky, S. (2013). Variety is the spice of happiness: The hedonic adaptation prevention (HAP) model. In S. A. David, I. Boniwell, & A. C. Ayers (Eds.), *Oxford handbook of happiness* (pp. 901–914). Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199557257.013.0067
Tashakkori, A. & Teddlie, C. (2003). *Handbook of mixed methods in social & behavioral research*. Sage.
Thompson, A. S., (2017). Don't tell me what to do! The anti-ought-to self and language learning motivation. *System, 67,* 38–49. http://doi.org/10.1016/j.system.2017.04.004
Thompson, A. S., & Erdil-Moody, Z. (2016). Operationalizing multilingualism: Language learning motivation in Turkey. *International Journal of Bilingual Education and Bilingualism, 19*(3), 314–331. http://dx.doi.org/10.1080/13670050.2014.985631
Twenge, J. M. (2006). *Generation me: Why today's young Americans are more confident, assertive, entitled - And more miserable than ever before*. Free Press
Ushioda, E., & Dörnyei, Z. (2017). Beyond Global English: Motivation to learn languages in a multicultural world: Introduction to the special issue. *The Modern Language Journal, 101*(3), 451–454. https://doi.org/10.1111/modl.12407
You, C., & Chan, L. (2015). The dynamics of L2 imagery in future motivational self-guides. In Z. Dörnyei, P. D. MacIntyre & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 405–430). Multilingual Matters. https://doi.org/10.1057/9781137457134_3
You, C., Dörnyei, Z. & Csiz´er, K. (2016). Motivation, vision, and gender: A survey of learners of English in China, *language learning,* 66(1), 94–123. http://doi.org/10.1111/lang.12140
Waninge, F., Dörnyei, Z., & de Bot, K. (2014). Motivational dynamics in language learning: Change, stability, and context. *Modern Language Journal*, *98*, 704–723. https://doi.org/10.1111/modl.12118
Wigfield, A., & Guthrie, J. T. (1995). *Dimensions of children's motivations for reading*: *An initial study* (Research Rep. No. 34). National Reading Research Center.
Wigfield, A., & Guthrie, J.T. (1997). Relations of children's motivation for reading to the amount and breadth of their reading. *Journal of Educational Psychology, 89*(3), 420–432. https://doi.org/10.1037/0022-0663.89.3.420

Appendix A

This questionnaire asks L2 learner/teacher to rate their opinions concerning the effectiveness of the NDMSs to manage PR at three levels of incivility prevention, dissent management, and resistance minimizing

NDMSs - PR Questionnaire

Psychological Reactance Levels	Focus of statements	Statements	Strongly agree	Agree	Slightly agree	Slightly disagree	Disagree	Strongly disagree
Incivility Prevention	Anxiety	I think applying NDMSs along with learner-centered strategies can prevent learner incivility and reduce learner-teacher anxiety						
	Anxiety	I believe that NDMSs have the potential to reduce the negative impact of test-oriented classes which act against learner motivation and increase learner anxiety and learner incivility						
	Anxiety	I believe that NDMSs have the potential to minimize reactance-inducing statements used by some teachers which act against learner motivation and increase anxiety and incivility						
	Anxiety	I believe that NDMSs have the potential to reduce reactance-inducing statements by some motivating strategies to reduce learner anxiety and incivility						
	Anxiety	I believe that NDMSs have the potential to replace test-score manipulation with a learn-friendly context which minimizes learner anxiety and incivility						
	Frustration	I believe that NDMSs have the potential to improve learners' low achievement which has demotivating effects and causes frustration and incivility among learners						
	Frustration	I think lack of novel activities in L2 teaching causes frustration among learners but I believe that NDMSs have the potential to minimize frustration						
	Frustration	I believe that restrictive classes act against NDMSs and increase learner frustration and learner incivility						
	Frustration	I think test-score manipulation by teachers act against NDMSs and causes frustration and incivility						
	Frustration	I think meeting learners' motivational factors in classroom/online L2 teaching can reduce frustration and incivility						
	self-doubt	I think lack of novel activities in L2 teaching causes self-doubt among teachers which leads to incivility						
	self-doubt	I believe that restrictive classes act against NDMSs and increase learner self-doubt and learner incivility						
	self-doubt	I think test-score manipulation by teachers act against NDMSs and causes self-doubt and incivility						
	self-doubt	I think lack of objective criteria and subjective assessment increases self-doubt among learners and might lead to incivility over a low grade						
	self-doubt	I think meeting learners' motivational factors in classroom/online L2 teaching can reduce self-doubt and incivility						

Psychological Reactance Levels	**Focus of statements**	**Statements**	Strongly agree	Agree	Slightly agree	Slightly disagree	Disagree	Strongly disagree
Dissent Management	Anxiety	I believe that catering for learners' dynamic motivational factors via NDMSs can reduce the level of dissent and create a friendly environment with less anxiety and self-doubt						
	Anxiety	I believe that restrictive classes act against NDMSs and increase learner anxiety and learner dissent						
	Anxiety	I think test-score manipulation by teachers act against NDMSs and causes anxiety and dissent among learners						
	Anxiety	I think lack of objective criteria and subjective assessment increases anxiety among learners and might lead to dissentive behavior over a low grade						
	Anxiety	I think meeting learners' motivational factors in classroom/online L2 teaching can reduce anxiety and dissent						
	Frustration	I think low achievement has demotivating effects and causes frustration and dissent among learners therefore improving learner achievement can facilitate dissent management						
	Frustration	I think lack of novel activities in L2 teaching causes frustration among learners which leads to learner dissent						
	Frustration	I believe that threatening activities act against NDMSs and increase learner frustration and learner dissent						
	Frustration	I think test-score manipulation by teachers act against NDMSs and causes frustration and dissent among learners						
	Frustration	I think lack of objective criteria or ignoring them in subjective assessment of assignments has demotivating effects and increases frustration among learners which might lead to dissentive behaviors over a low grade						
	self-doubt	I believe that catering for learners' dynamic motivational factors via NDMSs can reduce the level of dissent and create a friendly environment with less anxiety and self-doubt						
	self-doubt	I think lack of novel activities in L2 teaching has demotivating effects and causes self-doubt among teachers which leads to teacher dissent						
	self-doubt	I believe that test-oriented classes act against NDMSs and increase learner self-doubt and learner dissent						
	self-doubt	I think test-score manipulation by teachers act against NDMSs and causes self-doubt and dissent among learners						
	self-doubt	I think lack of objective criteria or ignoring them in subjective assessment of assignments has demotivating effects and increases self-doubt among learners which might lead to dissentive behaviors over a low grade						

Psychological Reactance Levels	Focus of statements	Statements	Strongly agree	Agree	Slightly agree	Slightly disagree	Disagree	Strongly disagree
Resistance Minimizing	Anxiety	I believe that restrictive classes act against NDMSs and increase learner anxiety and learner resistance						
	Anxiety	I think test-score manipulation by teachers act against NDMSs and causes anxiety and resistance						
	Anxiety	I think lack of objective criteria and subjective assessment has demotivating effects and increases anxiety among learners which might lead to resistance over a low grade						
	Anxiety	I think lack of objective criteria or ignoring them in subjective assessment of assignments has demotivating effects and increases anxiety among learners which might lead to resistance over a low grade						
	Frustration	I feel less resistance and frustration when a controversial subject is delivered unbiasedly						
	Frustration	I think low achievement causes frustration and resistance among learners therefore improving learner achievement can minimize learner frustration and resistance						
	Frustration	I think threatening activities in L2 teaching has demotivating effects and causes frustration among learners which leads to learner resistance						
	Frustration	I believe that test-oriented classes increase learner frustration and learner resistance						
	Frustration	I think test-score manipulation by teachers causes frustration and resistance among learners						
	self-doubt	I think lack of objective criteria and subjective assessment increases self-doubt among learners which might lead to resistance over a low grade						
	self-doubt	I think test-score manipulation by teachers has demotivating effects and causes self-doubt and resistance among learners						
	self-doubt	I believe that restrictive classes increase learner self-doubt and learner resistance						
	self-doubt	I think lack of novel activities in L2 teaching causes self-doubt among teachers which leads to learner resistance						
	self-doubt	I think threatening activities in L2 teaching has demotivating effects and causes self-doubt among teachers which leads to teacher resistance						

Appendix B

NDMSs - AE Questionnaire

This questionnaire asks L2 learners/teachers to rate their opinions concerning the effectiveness of NDMSs to manage L2 academic entitlement

No.	Statement	Strongly agree	Agree	Partly agree	Slightly disagree	Disagree	Strongly disagree
1	There is no place for dictatorial teacher in a democratic learner-friendly L2 teaching-learning context						
2	Ruling L2 teaching-learning contexts with an iron fist encourages academic entitlement						
3	L2 learners are customers and teachers are responsible for customer services						
4	L2 teachers should adopt NDMSs to improve learners' low self-esteem						
5	Threatening policies are adopted by L2 teachers as pretexts to hide their lack of activity on providing novel and motivating materials to the class which meet nonlinear dynamic motivation of the learner group						
6	L2 teachers should adopt NDMSs to improve learners' low achievement						
7	Ruling class by double standards causes demotivation and academic entitlement						
8	L2 teacher should stop emphasizing on a single shortcoming and ignoring learner's motivational factors as a penalty for that single shortcoming in a vindictive manner						
9	Academic entitlement is an excuse to blame students who demand their violated right from a domineering teacher						
10	Academic entitlement is an excuse to blame students who demand their violated right from a domineering educational system						
12	Academic entitlement is an excuse to blame students who demand their violated right from a domineering test-oriented class						
13	Teachers blame learners for being entitled as a way of dodging their responsibilities						
14	Ignoring learners' various test-taking styles lead to academic entitlement						
15	L2 teachers' inflexibility in terms of assignments lead to academic entitlement						
16	L2 teachers' monologic teaching approach leads to academic entitlement						

Appendix C

NDMSs-AE and PR Self-Interview

Strategy	Theme	Prompt	Coding responses
Psychological Strategies	Explaining the effectiveness of psychological strategies on psychological reactance and academic entitlement self-management	Are psychological strategies influential in managing psychological reactance and academic entitlement?	Elicited responses are interpreted and coded as 1= Yes 2= Undecided 3= No
Cultural Strategies	Explaining the effectiveness of cultural strategies on psychological reactance and academic entitlement self-management	* Are cultural strategies influential to manage psychological reactance and academic entitlement in L2 teaching –learning contexts?	Elicited responses are interpreted and coded as 1= Yes 2= Undecided 3= No
		*Are psychological reactance and academic entitlement important obstructive factors in L2 teaching-learning based on your own experiences?	Elicited responses are interpreted and coded as 1= Yes 2= Undecided 3= No
		* Are you under the effectiveness of psychological reactance and academic entitlement in learning-teaching English?	Elicited responses are interpreted and coded as 1= Yes 2= Undecided 3= No
Social Strategies	Explaining the effectiveness social strategies on psychological reactance and academic entitlement self-management	* Are social strategies in L2 teaching –learning contexts influential?	Elicited responses are interpreted and coded as 1= Yes 2= Undecided 3= No
		* Are social strategies in managing psychological reactance and academic entitlement influential?	Elicited responses are interpreted and coded as 1= Yes 2= Undecided 3= No
Psycho-socio-cultural Strategies	Multiple strategies to manage psychological reactance and academic entitlement	*Do you think learner-centered strategies along with psycho-socio-cultural strategies can facilitate psychological reactance and academic entitlement self-management? How?	Elicited responses are interpreted and coded as 1= Yes 2= Undecided 3= No
		*Do you think that psycho-socio-cultural can facilitate preventing incivility, minimizing resistance, and managing dissent in L2 teaching-learning?	Elicited responses are interpreted and coded as 1= Yes 2= Undecided 3= No
		*Do you think that psycho-socio-cultural have the potential to minimize the negative impact of obstructive factors?	Elicited responses are interpreted and coded as 1= Yes 2= Undecided 3= No
		*Do you think that managing obstructive factors by psycho-socio-cultural strategies can develop L2 teaching-learning?	Elicited responses are interpreted and coded as 1= Yes 2= Undecided 3= No
Part 5	Finally Thank you for your time. Do you have any questions that you would like to ask?		

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 1, 2020

Fang, F., Chen, R., & Elyas, T. (2020). An Investigation of the Relationship between Global Perspective and Willingness to Communicate in English in a Chinese University Context. Journal of Language and Education, 6(1), 39-54. https://doi.org/10.17323/jle.2020.10175

An Investigation of the Relationship between Global Perspective and Willingness to Communicate in English in a Chinese University Context

Fan Fang[1,2], Runting Chen[1], Tariq Elyas[2]

[1] Shantou University
[2] King Abdulaziz University

Correspondence concerning this article should be addressed to Fan Fang, Department of Foreign Languages & Literature, College of Liberal Arts, Shantou University, 243 Daxue Rd, Jinping District, Shantou, Guangdong Province, China, 515063. E-mail: ffang@stu.edu.cn

In an era of increasing global connectivity, acquiring a global perspective (GP) and being able to communicate in English are important for both personal and professional development. People with a GP are expected to be more willing to engage in intercultural activities, and more willing to communicate in English. Although previous studies have focused on having a GP and willingness to communicate (WTC) in English, few have investigated whether there is a correlation between these two variables. This study, conducted at a Chinese university, aimed to determine the relationship between GP and WTC in English. Data were collected from students via a questionnaire (n = 114) and interviews (n = 7) at the end of the autumn semester in 2018. The findings demonstrate that the students' GP and their WTC in English through the English curriculum did not progress to a significant extent, and their GP and WTC in English were correlated. Furthermore, three themes that affected the interviewees' WTC were extracted from the interviews: global cognition, self-actualization, and intercultural experience. These findings suggest that it would be beneficial to learn English by developing a GP, and that educators can enhance students' WTC by helping them cultivate that perspective.

Keywords: english as a global language, english as a medium of instruction, english learning, english language teaching, global perspective, willingness to communicate in english

Introduction

Today, as the world is becoming increasingly interconnected, higher education has begun to pay attention to students' global learning and development (American Council on Education, 1995; Association of American Colleges & Universities, 2007). Therefore, in addition to intellectual development, university students also need to gain a better sense of themselves and their relationships with others and the world. In other words, they need to cultivate a global perspective (GP) to develop as a global citizen, not only in language learning but also in general education (Baker & Fang, 2019; Fang & Baker, 2018; Goren & Yemini, 2017;). As English has evolved as a global language in recent decades (Crystal, 2012), it has become more important to acquire a GP through the use of English. People from different lingua-cultures often use English as a common language to communicate if they do not know each other's native language (Seidlhofer, 2011). Moreover, English is the dominant language at many international universities, where it is often used as the medium of instruction (EMI) (Gundsambuu, 2019; Fang, 2018; Fenton-Smith, Humphreys & Walkinshaw, 2017). In this sense, English is a tool for people to connect with the world and further their personal development.

The key to building connections with the world is engaging in communication with people from different cultural backgrounds. Thus, a willingness to communicate (WTC) in a second language (L2), defined as 'a readiness to enter into discourse at a particular time with a specific person' (MacIntyre, Dörnyei, Clément, & Noels, 1998, p. 547), is essential for personal development because it enhances a person's ability to communicate with the world as a global citizen. However, while previous studies have addressed the theoretical foundation of GP and WTC, they have not sufficiently discussed the correlation between these two aspects. The present study, based

on the context of a Chinese university, aims to determine the relationship between students' GP and WTC in English learning and identify any correlation between these two variables. The findings of this study could serve as a reference for educators to think about students' development and to reflect on their teaching. This study could also help bridge the gap between students' ability to learn English in school and their personal and professional development in their future career or educational endeavors that could make their English learning more sustainable.

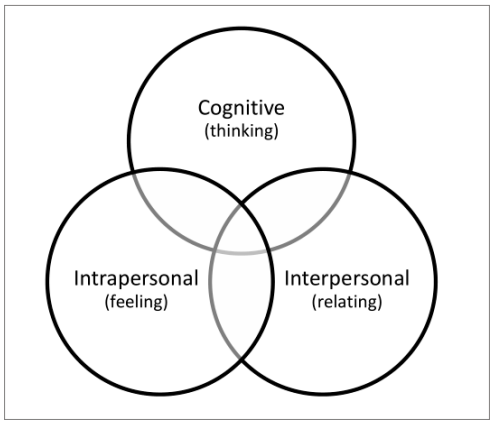
Literature Review

Global Perspective

The meaning of GP varies in different research areas. The present study adopts the explanation of Braskamp, Braskamp, and Engberg (2014), who illustrate that GP is a holistic perspective on human development in the global world. According to their explanation, GP is grounded in two theoretical perspectives on personal development: *intercultural maturity* and *intercultural communication*. The first is based on the work, *In Over our Heads*, by Kegan (1994), who proposed that people need to rely on their feelings and relationships with others and their own thinking in their life journey. Kegan (1994) also put forward three domains of human development: cognitive, intrapersonal, and interpersonal, which were later adapted by King and Magolda (2005) to *intercultural maturity*, describing the social-cultural development of college students. As the second perspective, *intercultural communication*, posited by intercultural communication scholars, demonstrates an individual's intercultural competence and sensitivity (Bennett & Bennett, 2004; Chen & Starosta, 1996). These scholars identified three other domains of personal development: cognitive, affective, and behavioral, which, respectively, are similar to thinking, feeling, and relating and believed that through intercultural communication, the skills associated with these three domains will be strengthened.

In the present study, the perspectives on *intercultural communication* and *intercultural maturity* were integrated into a GP with Kegan's (1994) three domains. Figure 1 depicts the model of GP that encompasses three interconnecting circles. According to Braskamp et al. (2014), the cognitive (thinking) circle is about knowledge and epistemology; this means that one needs to take complexity and multiple cultural perspectives into consideration when viewing knowledge acquisition. The second circle, the intrapersonal (feeling) circle, stresses self-identity, awareness of personal values, and discovery of the purpose of life. The interpersonal circle includes communicating with people from different cultural backgrounds, relating to others, and accepting their differences.

Figure 1
The GP Model

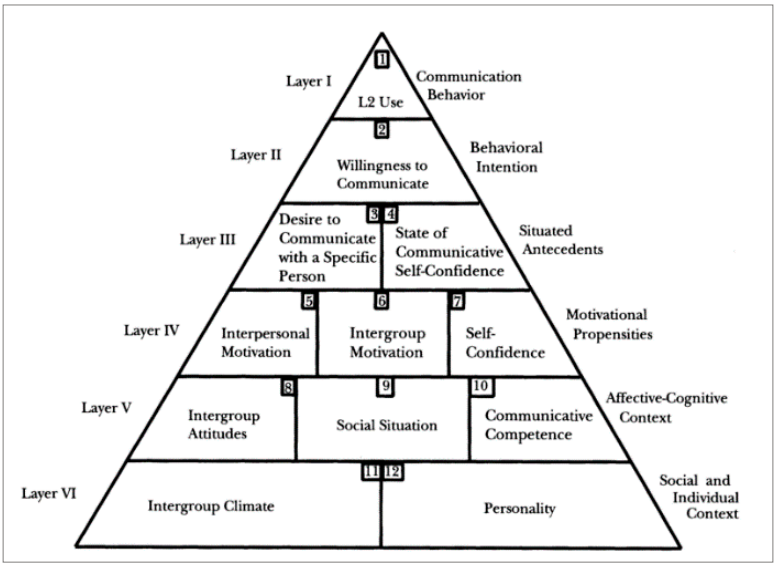


L2 WTC and Other Related Concepts

Originating from the concept of WTC in the first language (McCroskey, 1992; McCroskey & Richmond, 1987), L2 WTC was initially defined as a stable, trait-like predisposition that was influenced by communication

competency and communication anxiety (MacIntyre, 1994). L2 WTC was found to be affected by the interactions among various factors. For example, MacIntyre and Charos (1996) developed a model where L2 WTC was an integrative result of perceived L2 competence, motivation, L2 anxiety, integrativeness, and attitudes towards the learning situation. Later, MacIntyre et al. (1998) developed this model into a pyramid L2 WTC model (Figure 2) consisting of six layers, from the bottom layer that contained intergroup climate and personality, to the top layers containing L2 use. In this model, L2 WTC is not only a trait-like property, it also fluctuates according to different situations (see also Zhang, Beckmann & Beckmann, 2019).

Figure 2
The Pyramid L2 WTC Model



In the models mentioned above, integrativeness (intergroup motivation, intergroup attitudes, and intergroup climate) is one of the factors that influences L2 WTC. Integrativeness is associated with a learner's level of motivation for L2 learning. It refers to the inclination to approach an L2 group; thus, it will affect the L2 WTC (Gardner, 1985; MacIntyre & Charos, 1996; Peng, 2014). Learning English in English-speaking countries provides learners with many opportunities to approach an L2 group. However, in the Asian context, where L2 English learners may lack access to an L2 group, the replacement of integrativeness was considered, and the notion of international posture was posited by Yashima (2002). Consisting of an interest in foreign or international affairs, the willingness to go overseas to live or work, the readiness to interact with intercultural partners, and an open attitude towards different cultures, Yashima used equation modelling to show that international posture helps increase L2 WTC. The idea of international posture links learners to a larger world, which is similar to the intercultural domains of GP. The present study expands the concept of international posture to GP, and it attempts to determine if there is congruity between the educational outcome and English learning. It also aims to identify the ways in which students are more willing to engage in English communication.

To investigate the relationship between GP and WTC, this study adopted a mixed-methods approach, including disseminating questionnaires and conducting interviews. By doing so, it provides an overview at the group level as well as depth at an individual level. Specifically, this study is based on the following research questions:

Q1: What is the difference between GP and L2 WTC among students with different levels of English proficiency?
Q2: To what extent, if any, is L2 WTC correlated with GP?

Method

Research Context

This study was conducted at a university located in southeast China at the end of the autumn semester in 2018. At this university, the English Language Center (ELC) is in charge of teaching English to all university students.

The ELC has recruited teachers from both China and abroad, so the university is regarded as an emergent English as a lingua franca (ELF) community[1]. The goal of the ELC is to provide better teaching, facilitate an English learning environment, and enhance students' communication and critical thinking skills using creative and lively teaching activities. Every new student at this university must take an English placement test to determine their language proficiency level. Then, they can choose English classes based on their English proficiency level, and can also choose teachers using the school's registration system. Thus, students with different majors will take English classes together. At a minimum, students in all schools, except for the art school, have to pass ELC Level Four[2] courses to achieve a higher-intermediate proficiency level in English upon graduation.

Participants

The participants in this study were all university students taking the ELC courses. For the questionnaire, the respondents consisted of 114 students from four different English classes: level one (29 students, 25.4%), level two (30 students, 26.3%), level three (26 students, 22.8%), and level four (29 students, 25.4%). Of the 114 students, 64 (56.1%) were males and 48 (42.1%) were females (two did not mention their gender in the questionnaire). The students were all non-English majors from the following schools: Liberal Arts, Law, Business, Science, Engineering, Journalism, and Art. In all four classes, the teachers were Chinese. Seven students participated in the interviews because they provided contact information and showed interests in participating in further research. Of those, five were selected from the students who had completed the questionnaire and two were English majors with study abroad experience who were contacted directly by the researchers. Three of the seven interviewed students were male and four were female.

Instruments

Questionnaire

The questionnaire, containing two parts, a GP scale and an L2 WTC scale, was administered to the participating students at the end of the autumn semester in 2018. Because the students had different language proficiency levels, the questionnaire was translated from English into Chinese. The questionnaire took around 10 to 15 minutes to complete, and it was carried out during the class break time of the English classes.

Students' GP was measured with the GP scale adapted from the Global Perspective Inventory (GPI) (2017), which assesses GP from cognitive, intrapersonal, and interpersonal dimensions. Each dimension is divided into two aspects; thus, the GPI consists of six scales. The questionnaire used in this study selected two of the six items from each scale based on the students' familiarity, so 12 items were adopted and rated using a 5-point Likert scale[3] (see Appendix 1 for the items).

For L2 WTC, items from Mystkowska-Wiertelak and Mirostaw (2017), concerning L2 WTC inside and outside the classroom, were used as the measures. Eight items were used: four L2 WTC inside the classroom items and four L2 WTC outside the classroom items according to students' experience and level of understanding. As some items on the original scale were not familiar to the students, the items were adapted to facilitate understanding. For example, the item, 'I am willing to use computer-mediated communication to address an acquaintance of mine', was changed to 'I am willing to communicate with a foreigner (teacher, friend, etc.) in English on the internet'. These items were piloted with a group of students and found to be valid. The L2 WTC items were rated on a 6-point Likert scale[4] (see Appendix 1 for the items).

Interviews

This study used semi-structured interviews because this method enables researchers to collect similar data from different interviewees while facilitating the flexibility to ask follow-up questions and discover unexpected areas.

[1] ELF is defined as 'any use of English among speakers of different first languages for whom English is the communicative medium of choice, and often the only option' (Seidlhofer, 2011, p. 7).

[2] After completing the ELC courses and passing Level Four, students are expected to achieve English proficiency at an IELTS 6-6.5 (CEFR B2) level.

[3] The five-point scale is measured as: 1 strongly disagree, 2 disagree, 3 neither agree nor disagree, 4 agree, or 5 strongly agree (see Research Institute for Studies in Education, 2017).

[4] The six-point scale is measured as: 1 definitely not willing, 2 probably not willing, and 3 perhaps not willing 4 perhaps willing, 5 probably willing, or 6 definitely willing (see Mystkowska-Wiertelak & Pawlak, 2017)

The interviews were conducted after disseminating the questionnaire. The first round of interviews involved five students who completed the questionnaire, and it focused on the same themes (see the interview prompts in Appendix 2) related to their WTC in English and GP. In addition to these questions, follow-up questions were also asked based on the interviewees' answers and personal experiences. The second round of interviews was conducted with two senior English majors who had a study abroad experience. In addition to the questions listed above, this interview also focused on their study abroad experiences. In conducting the interviews, this study tried to determine the interaction between L2 WTC in English and GP. All the interviews were audio-recorded and then transcribed by the researchers with the permission of the participants.

Results

Questionnaire results

Statistical Package for the Social Sciences (SPSS) Version 16.0 was adopted to analyze the questionnaire data. The coefficient reliability of the GP and L2 WTC scales was measured, and their Cronbach α was .71 and .92, respectively, which are acceptable (Hair et al., 2010). The GP and L2 WTC of every participant were assessed by calculating the mean scores of all the items in each scale. Next, based on the two previously mentioned research questions, a One-Way Analysis of Variance (ANOVA) and Pearson Correlation Analysis were run.

The first research question is about the difference between GP and L2 WTC among students with different English proficiency levels. Because this requires examining the differences among more than two groups, and the result of the One-sample Kolmogorov-Smirnov test (Table 1) met the distribution assumption (Asymp. Sig. (2-tailed) > .05), ANOVA was chosen. Table 2 shows the means of the GP and L2 WTC scores of the students with different English proficiency levels. The mean of GP ranged from 3.17 to 3.47 for the four English proficiency levels, and L2 WTC ranged from 4.33 to 4.73. The results do not change much based on the students' English proficiency levels. Table 3 describes the result of the One-Way ANOVA among the students with different English proficiency levels. The statistics indicate that there was no significant difference in GP and L2 WTC among students from all four English proficiency levels.

Table 1
Results of the One-sample Kolmogorov-Smirnov Test

	GP MEAN	**L2 WTC MEAN**
Asymp.Sig. (2-tailed)	.155	.243

Table 2
Means of GP and L2 WTC of Students with Different English Proficiency Levels

		N	**Mean**
GP	level 1	28	3.47
	level 2	29	3.40
	level 3	25	3.17
	level 4	28	3.43
L2 WTC	level 1	28	4.67
	level 2	30	4.73
	level 3	26	4.33
	level 4	29	4.55

Table 3
Results of One-Way ANOVA among Students with Different English Proficiency Levels

		df	F	sig
GP	Between Groups	3	2.458	.067
	Within Groups	106		
	Total	109		
L2 WTC	Between Groups	3	.836	.477
	Within Groups	109		
	Total	112		

The second research question aimed to determine if L2 WTC is correlated with GP. Figure 3 shows the simple scatter plot between GP and L2 WTC, which indicates that there may be a weak to moderate positive correlation between the two variables (Field, 2005). After obtaining an overview from the scatter plot, a Pearson product moment correlation analysis was run in SPSS to examine the degree of the correlation between GP and L2 WTC. As seen in Table 4, in this study, L2 WTC was significantly correlated with GP ($r = .47$, $p < .05$). The effect size of the correlation between these two variables, according to Field (2005), reaches the medium effect size ($r = 0.3$).

Figure 3
Scatter Plot between GP and L2 WTC

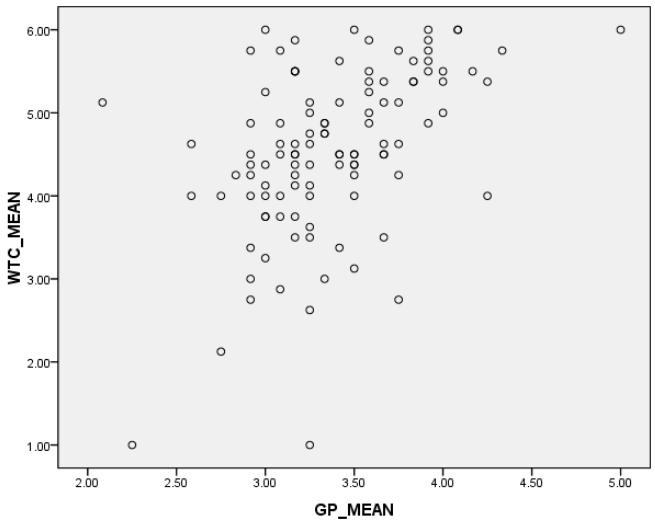


Table 4
Results of the Pearson Correlation Analysis of GP and L2 WTC

	Mean	SD	GP	L2 WTC
GP	3.37	.45	1	
L2 WTC	4.57	1.00	.47**	1

Interview Findings

After collecting the questionnaire data, five questionnaire participants (denoted as Participant A, Participant B, Participant C, Participant D, and Participant E) were invited for a first-round interview based on their GP and WTC scores obtained from the questionnaire. Because participants F and Participant G did not complete the questionnaire but had study abroad experience, they were scheduled for the second-round interview, the focus of which was different from the first round. Table 5 shows the GP and WTC scores from the questionnaires of the interviewees, excluding Participant F and Participant G. The audio recordings of the interviews were transcribed in Chinese and stored in NVivo 11. This study then used latent content analysis (Dörnyei, 2007) to identify essential themes that are related to the research questions and emergent issues. Thus, it adopted both

top-down coding based on the research questions and bottom-up coding based on the interviewees' answers (Miles et al., 2014). The coding scheme was developed with 26 codes under five themes: global cognition, self-actualization, intercultural experience, WTC in English, and English learning. A student who had taken a Research Methodology course was invited to read the coding scheme for one of the interviews, and an acceptable level of agreement was reached (86%). The following section outlines the relationship between GP and L2 WTC in English summarized from the responses from the seven interviewees.

Table 5
Interviewees' Scores of GP and WTC in English

Interviewee	GP Score (full score: 60)	WTC Score (full score: 48)
Participant A	33	17
Participant B	56	46
Participant C	40	24
Participant D	40	44
Participant E	49	47

Global Cognition and WTC in English

The theme of global cognition corresponds to the cognition domain of GP, which is about thinking in a complex and multicultural way. The interview participants were asked if they were interested in international news and cultures of different countries. None of the participants had the habit of following international news, but some of them, occasionally paid attention to international breaking news. In particular, Participant C followed some news stories in WeChat because she thought it would be useful for her further study and future career:

Extract 1:

> I follow several public accounts, such as thepaper.cn and huanqiu.com. Why did I pay attention to this news at the beginning? Because I wanted to take the civil servant exam and a postgraduate entrance exam at that time. I could pay attention to some public accounts. I think foreign news is okay.

Participant F expressed that she did not pay much attention to international affairs but mentioned that she would keep an eye on news related to China instead:

Extract 2:

> I don't really follow international news. Instead, I will focus on the topic of China, and it will also involve some foreign views. Take China's democracy as an example; I want to know what the point of view of foreigners is.

When asked about their interest in foreign cultures, most of the interviewees, such as Participant E, did not pay much attention to other cultures:

Extract 3:

> I pay some attention, but not much attention. I do not pay much attention to these cultures. Sometimes, when I watch videos or programs, I find their concepts are different from ours. I think it's very interesting, and I go to watch them.

Participant G expressed her curiosity about foreign cultures:

Extract 4:

> I'm interested in other people's cultures…Maybe it's because of curiosity, because the country itself has diversity, and then you listen to them talk about their culture, there will be a lot of different stories.

In terms of WTC in English, Participant C's questionnaire score and interview responses showed that she did not have a strong need to communicate in English:

Extract 5:

> At present, I do not have much desire to communicate with foreigners. Let's just say, I'm open to it… I was really nervous in the middle of the semester, because I felt that the students in front of me spoke very well, and this semester was very special. Just one of my classmates and I were sophomores in our class, and the others were freshmen. Besides, because they reached level two when they first entered the school, I felt they were really cool. And then it felt a little bit off. I seldom spoke English in class.

Similarly, Participant F did not demonstrate much WTC in English, except for classes where English is needed. However, Participant E and Participant G both expressed their WTC in English and with foreigners:

Extract 6:

Interviewer: Do you take the initiative to find opportunities to chat with others in English?
Participant G: Yes, especially in English classes or with people who can speak English. Not with someone who does not speak English.
Interviewer: With native speakers?
Participant G: Not necessarily. As far as I know they can speak English. If I'm sure other people do not speak English, I don't.

Extract 7:

Interviewer: Do you like to talk with others in English?
Participant E: Yeah, I like to talk to my roommates in English when I am free.
Interviewer: Is it a deliberate exercise, or is it just for fun?
Participant E: Both. For example, after watching a TV show, I also want to learn to speak. As usual, I speak English, and then I can improve my spoken English. Indeed, speaking English is also a necessary skill.
Interviewer: What other reasons do you have for learning English besides this ability?
Participant E: The most obvious reason is that it's cool. If you speak in English, people think this person is awesome.

These excerpts demonstrate that most of the interviewees are not interested in global affairs and foreign cultures. Even if they follow international news, as Participant C did, they did so for practical reasons, such as preparing for exams. For the participants, things that are happening in other countries are not pertinent, and compared their immediate or vital interests, this type of news and information about other cultures is not necessary in their life.

Self-Acutalization and WTC in English

In the present study, the concept of self-actualization refers to a person's expectation of himself/herself and how he/she fulfills his/her goals, especially in English and for life-long planning. Thus, self-actualization can also refer to the intrapersonal domain in GP. To determine one's self-actualization, the participants in this study were asked the following questions: 1) What is the role of English in your life now and in the future and 2) what is your goal in learning English?

For Participant A, English does not matter much for his future goals, and he is not interested in it. As an engineering student who is going to apply for a job after earning his bachelor's degree, the College English Test Band 6 certificate was sufficient for him, and he would probably stop learning English in the future. He mentioned: 'I'm not interesting in speaking English because I don't have interest in it.' As English is a compulsory subject for all the university students, he is simply not motivated to learn English. Participant D was not clear about what role English would play in her future:

Extract 8:

Interviewer: Do you have a plan for your future career? What would English be used for in your future?
Participant D: Because I do not know whether I will go abroad in the future, I do not have plans in this regard.
Interviewer: In addition to an overseas exchange program, do you think English will be helpful for your future work? Or will you not use English in the future?
Participant D: In terms of employment, it depends on what companies I will work for. Because some enterprises may need to use English to communicate.
Interviewer: Do you intend to join international corporations?
Participant D: It's not clear yet.

In contrast, participant B was clearer about the role of English in his future:

Extract 9:

> Certainly, I will keep learning English. After all, it doesn't matter what you want to do in the future. I want to study psychology in the future. For example, in the field of psychology, foreign studies, foreign experiments, and conclusions are much better than those in China. It is difficult to translate the latest foreign achievements into Chinese. If you are engaged in this field, and you want to get the latest conclusions, information, and theories, then you must have a good ability to read English literature. I think that is a necessary ability for you to work in any industry in the future and broaden your horizon.

Participant B did not indicate his WTC in English in the interview, but his scores in the L2 WTC questionnaire were reasonably high. However, he expressed that he would seldom use English in China and he did not like to learn English simply for passing exams. While participant C did not clearly state her L2 WTC, she has joined some English organizations where she needed to speak English:

Extract 10:

Interviewer: Do you like to communicate with others in English?
Participant C: I have joined several organizations where I need to communicate in English.
Interviewer: What organizations?
Participant C: The Case Study Group and English Lounge, but now I have to quit.
Interviewer: How did it feel in English Lounge?
Participant C: I think it is listening that helps me a lot.
Interviewer: What about speaking?
Participant C: I think it creates an environment for people to speak, but there are about 20 people in one activity, so not everyone has the chance to talk. We mainly lead others to speak.

In terms of WTC in English, participant A is not interested in anything related to English. However, if communicating in English is needed, he would comply:

Extract 11:

Interviewer: Have you ever tried to communicate with others in English?
Participant A: It is required in class. Sometimes, I meet some foreigners outside of class and need to communicate with them.
Interviewer: Have you ever had such an experience?
Participant A: Yes.
Interviewer: Are you able to communicate with them?
Participant A: I've basically tried it twice and I think I understood their English.
Interviewer: Do you think it feels good to talk to people in English?
Participant A: No, it's frightening, really. Because many people cannot speak English, at least I can. One time, no one could speak English to a foreigner, because many of them were middle-aged. Then, they grabbed me and said you should be young enough to speak English, so I said a few words.

In conclusion, Participant A and Participant B were both clear about their future use of English, although Participant A considered English to be useful in his future life while Participant B considered it to be useless, except for passing exams. Participant C has not made a decision about what she will do in the future. The complexity of self-actualization and WTC was revealed from the participants' comments.

Intercultural Experience and WTC in English

Intercultural experience is about meeting with people from other countries or cultural backgrounds; it is related to the interpersonal perspective in GP because it involves person-to-person communication (Jackson, 2020). Because most of the participants were freshmen and sophomores, their only chance to communicate in English was in their English class. They seldom talked with foreign teachers or students because they were afraid of being embarrassed if they could not understand what is being said.

Extract 12:

Interviewer: Are there any foreign teachers in your specialized courses?
Participant B: Yes, but I'm afraid to go. I'm afraid I will not understand.

Extract 13:

Interviewer: Are there any foreign teachers in your specialized courses?
Participant C: At present, I have had three semesters of English classes, and two of them were taught by Chinese teachers. I think Chinese teachers are a little better. For example, if there are some important questions that some people may not understand in English, I can say the important parts again in Chinese. I think this is very good, and it may be better for me. I also chose the courses taught by a Chinese teacher next semester.

Extract 14:

Interviewer: Have you ever met some foreign teachers and foreign students when you entered the university and then tried to communicate with them in English?
Participant E: Yes, but I seldom take the initiative. When I go to some English events, I meet foreign teachers and foreign students. I don't usually communicate with them.

Unlike the other interview participants, participants F and G are seniors who have both had several study abroad experiences. Participant F has taken part in two short projects in Nepal and Germany, respectively. According to her interview responses, these experiences enabled her to see what other countries are actually like, and she also thought that English would be beneficial if she chose to work abroad in the future. However, she does not have much interest in communicating with people from other countries because of her shyness and her fear of misunderstanding what is being said. Participant G joined an exchange program that enabled her to study at an American university for one semester. In the interview, she shared her interest in talking with people from different countries because she liked to learn about other people's cultures. She made several friends in America who were from different countries, and their conversations usually went well because her English did not impede communication, and she was able to handle the situation.

Discussion

In comparison with international posture (Yashima, 2002), GP includes the intercultural communication perspective as well as cognition and the interpersonal perspective. If people think globally, feel clear about their identity in the world, and hope to engage in intercultural communication, it is necessary for them to communicate in English (Pawlak & Mystkowska-Wiertelak, 2015; Peng, 2014, 2019). EMI education might be proved helpful and effective for those students' English learning journey. Therefore, the present study suspected that there is a link between GP and L2 WTC.

From the questionnaire results, the correlation coefficient between GP and L2 WTC in English indicates that, in the context of one Chinese university, a student with a high GP may be more likely to communicate in English. Inversely, a student who has high L2 WTC in English could possibly have a broader GP than those with lower L2 WTC in English. When English serves as a lingua franca to communicate with the world rather than a knowledge-based skill, GP and L2 WTC in English are probably correlated. The One-Way ANOVA analysis showed that there was no significant difference in GP and WTC among students with different English proficiency levels. In this study, both the mean scores of GP and L2 WTC in English were above the medium score (GP: 3.17–3.47 on a 5-point scale, L2 WTC: 4.33–4.73 on a 6-point scale). This also means that students may not make enough progress in GP and L2 WTC, which should be an important educational outcome of learning English. From level 1 to level 4, the students may still not be aware of their English-learning goals, and they may acquire knowledge about the English language but still not have a high level of WTC in English with others.

From the interview excerpts related to *global cognition* and *WTC in English*, it was found that all of the interviewees except Participant G, did not have strong global cognition, but their L2 WTC varied because their interest in and fear of talking in an L2 may have been impacted by other variables, such as a sense of being cool (Participant E) and a lack of confidence (Participant C). However, in Participant G's case, her interest in foreign cultures prompted her to communicate with foreigners in English, which indicates that a global cognition may enhance one's L2 WTC.

From the interview excerpts related to *self-actualization* and *WTC in English,* Participant B's case indicates that if people know why they need to learn English and why English is so important, they may have a greater WTC in English because the recognition of English will enhance their WTC. Participant A's case reveals the opposite; if one does not recognize the importance of the role that English plays in personal and professional development, it is difficult to embrace English, which impacts their WTC. However, although Participant C did not articulate her purpose for learning English, she joined English clubs where she can cultivate her ability to communicate in that language. This may be due to her personal interests and the environment in which she is learning English.

The interview excerpts related to *intercultural experience* and *WTC in English*, revealed that language can be either a barrier or a bridge to intercultural communication (Fang & Baker, 2018). For the participants who consider their English to be insufficient, English is a barrier to engaging in intercultural communication, such as communicating with foreign teachers or students. Even though some of them have a high WTC in English, their willingness does not overcome their fear of being embarrassed. However, for Participant G, who overcame those worries, language is a bridge that connects her to the world, helping her communicate with people from different countries. In turn, this intercultural experience may also enhance WTC and improve communication skills (Fang & Baker, 2018).

The results and findings of the questionnaire and the interviews suggest that there is room for students and educators to improve the English learning process, and students' WTC can be enhanced by cultivating their GP (Fang & Baker, 2018; Baker & Fang, 2019; Yashima, MacIntyre & Ikeda, 2018). On the one hand, students need to treat English as a communication tool from a sociocultural perspective, rather than as a subject and an exam they have to take. If their learning in English is limited to an exam-oriented and knowledge-based perspective, they may easily feel less motivated to learn English, and they might not realize the fun and importance of communicating, which is more essential for their English progress (Peng, 2014). On the other hand, educators could attempt to stimulate students' GP in class by discussing international affairs and foreign cultures, inviting students to talk about their goal of learning English, creating a communicative environment, and so on (see e.g., Peng, 2012, 2014; Yashima, MacIntyre & Ikeda, 2018). However, we need to realize that WTC and GP are both dynamic concepts, as students' 'momentary psychological reactions to contextual factors both facilitated and constrained their participation in the discussions, showing interesting situational dynamics' (Yashima et al., 2018, p. 132). Further understanding of the relationship between WTC and GP will help both teachers and students explore the concept of global citizenship education (Goren & Yemini, 2017; Fang & Baker, 2018).

Conclusion

Before making any conclusions, we will recognize some limitations of this study. Currently, the theoretical foundation of the relationship between GP and WTC is insufficient. Therefore, further research is needed to support the causality between these two variables. To determine the cause-effect relationship, the number of questionnaire and interview participants needs to be enlarged in the future. In particular, more interview data can be helpful for better understanding the correlation between GP and WTC. Moreover, because this study only involved classes taught by Chinese teachers, additional studies should include classes taught by international teachers.

In a non-English-speaking country such as China, being able to communicate in English is essential for a person who has a GP. Likewise, with a GP, a learner may be more willing to communicate in English than those who are unclear about their life goals and are not interested in engaging in global issues in the future. The present study found that GP and L2 WTC in English are two correlated variables, and there is room for students and educators to improve these two aspects in their learning and teaching. If students could make sense of the role of English in their life, they could practice their English in a more effective way than is possible if they only follow their teachers' instruction. If educators could use English classes and activities to activate the students' WTC in English and help them cultivate a global perspective, they might contribute to students' life-long development and facilitate a sustainable way for them to learn English.

Funding

This work was supported by the "The 13th Five-Year Plan Education and Science Special Research Projects on Philosophy and Social Science of Higher Education (2019)", Department of Education of Guangdong Province – "Effectiveness of the Implementation of English as a Medium in Higher Education of Guangdong Province" [Grant No.: 2019GXJK194].

Conflict of Interest

The authors declare that they have no conflict of interest.

References

American Council on Education. (1995). *Educating Americans for a world in flux: Ten ground rules for internationalizing higher education*. https://files.eric.ed.gov/fulltext/ED388165.pdf

Association of American Colleges and Universities (2007). Global learning for the new global century: Executive summary with findings from employer survey. https://www.aacu.org/sites/default/files/files/LEAP/GlobalCentury_final.pdf

Baker, W., & Fang, F. (2019). *From English language learners to intercultural citizens: Chinese student sojourners' development of intercultural citizenship in ELT and EMI programmes*. British Council.

Bennett, J. M. & Bennett, M.J. (2004). Developing intercultural sensitivity: An integrative approach to global and domestic diversity. In D. Landis, J. M. Bennett & M. J. Bennett (Eds.), *Handbook of intercultural training* (3rd ed., pp. 147–165). Sage.

Braskamp, L. A., Braskamp, D. C., & Engberg, M. E. (2014). *Global Perspective Inventory (GPI): Its purpose construction, potential uses, and psychometric characteristics*. Global Perspective Institute.

Braskamp, L. A., Braskamp, D. C., & Merrill, K. C. (2009). Assessing progress in global learning and development of students with education abroad experiences. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *18*, 101-118.

Chen, G. M. & Startosta, W. (1996). Intercultural communication competence: A synthesis. *Communication*

Yearbook, 19, 353–383. https://doi.org/10.1080/23808985.1996.11678935
Crystal, D. (2012). *English as a Global Language* (2nd edn.). Cambridge University Press.
Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.
Fang, F. (2018). Review of English as a medium of instruction in Chinese universities today: Current trends and future directions. *English Today*, *34*(1), 32-37. https://doi.org/10.1017/S0266078417000360
Fang, F., & Baker, W. (2018). 'A more inclusive mind towards the world': English language teaching and study abroad in China from Intercultural citizenship and English as a lingua franca perspectives. *Language Teaching Research*, *22*(5), 608-624. https://doi.org/10.1177/1362168817718574
Fenton-Smith, B., Humphreys, P. & Walkinshaw, I. (Eds.). (2017). *English medium instruction in higher education in Asia-Pacific*. Springer.
Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). SAGE Publications.
Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. Edward Arnold.
Goren, H., & Yemini, M. (2017). Global citizenship education redefined – A systematic review of empirical studies on global citizenship education. *International Journal of Educational Research*, *82*, 170-183. https://doi.org/10.1016/j.ijer.2017.02.004
Gundsambuu, S. (2019). Internationalization and English as a medium of instruction in Mongolian higher education: A new concept. *Journal of Language and Education*, *5*(2), 48-66. https://doi.org/10.17323/jle.2019.8481
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Prentice Hall.
Jackson, J. (2020). *Introducing language and intercultural communication* (2nd ed.). Routledge.
Kegan, R. (1994). *In over our heads: The mental demands of modern life*. Harvard University Press.
King, P., Magolda, M.B.B. (2005). A developmental model of intercultural maturity. *Journal of College Student Development*, *46*, 571–592. https://doi.org/10.1353/csd.2005.0060
MacIntyre, P. D. (1994). Variables underlying willingness to communicate: A causal analysis. *Communication Research Reports*, *11*(2), 135-142. https://doi.org/10.1080/08824099409359951
MacIntyre, P. D., & Charos, C. (1996). Personality, attitudes, and affect as predictors of second language communication. *Journal of Language and Social Psychology*, *15*, 3-26. https://doi.org/10.1177/0261927X960151001
MacIntyre, P. D., Dörnyei, Z., Clément, R., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: a situational model of L2 confidence and affiliation. *The Modern Language Journal*, *82*(4), 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x
MacIntyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied Linguistics*, *32*(2), 149–171. https://doi.org/10.1093/applin/amq037
McCroskey, J.C. (1992). Reliability and validity of the willingness to communicate scale. *Communication Quarterly*, *40*, 16-25. https://doi.org/10.1080/01463379209369817
McCroskey, J.C., &Richmond, V.P. (1987). Willingness to communicate. In J.C. McCroskey & J.C. McCroskey & J.A. Daly (Eds.), *Personality and interpersonal communication* (pp. 129-156). Sage.
Miles, M, Huberman, M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd edn). Sage.
Mystkowska-Wiertelak, A., & Pawlak, M. (2017). *Willingness to communicate in instructed second language acquisition: Combing a macro- and micro-perspective*. Multilingual Matters.
Pawlak, M., & Mystkowska-Wiertelak, A. (2015). Investigating the dynamic nature of L2 willingness to communicate. *System*, *50*, 1-9. https://doi.org/10.1016/j.system.2015.02.001
Peng, J. (2012). Towards an ecological understanding of willingness to communicate in EFL classrooms in China. *System*, *40*, 203–213. https://doi.org/10.1016/j.system.2012.02.002
Peng, J. (2014). *Willingness to communicate in the Chinese EFL university classroom*. Bristol: Multilingual Matters.
Peng, J. (2019). Understanding willingness to communicate as embedded in classroom multimodal affordances: Evidence from interdisciplinary perspectives. *Linguistics and Education*, *51*, 59–68. https://doi.org/10.1016/j.linged.2019.04.006
Research Institute for Studies in Education (2017). *Global Perspective Inventory: Theoretical foundations and scale descriptions*. . Iowa State University. https://www.gpi.hs.iastate.edu/documents/GPI%20Theory%20and%20Scales.pdf
Seidlhofer, B. (2011). *Understanding English as a lingua franca*. Cambridge University Press.
Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, *86*(1), 54-56. https://doi.org/10.1111/1540-4781.00136

Yashima, T., MacIntyre, P., & Ikeda, M. (2018). Situated willingness to communicate in an L2: Interplay of individual characteristics and context. *Language Teaching Research*, *22*(1), 115-137. https://doi.org/10.1177/1362168816657851

Zhang, J., Beckmann, N., & Beckmann, J. F. (2019). One situation doesn't fit all: Variability and stability of state willingness to communicate in a Chinese college English classroom. *Language Teaching Research*. https://doi.org/10.1177/1362168819891330

Appendix 1

The Questionnaire

Part I:

The following scales describe how you think, how you view yourself, and how you relate to others who are different. Please circle the number that best describes your opinion: 1 strongly disagree, 2 disagree, 3 neither agree nor disagree, 4 agree, or 5 strongly agree.

1. I consider different cultural perspectives when evaluating global problems.
1 2 3 4 5

2. I rely primarily on authorities to determine what is true in the world. (r)
1 2 3 4 5

3. I am informed about current issues that impact international relations.
1 2 3 4 5

4. I understand the causes of conflict among nations with different cultures.
1 2 3 4 5

5. I know who I am as a person.
1 2 3 4 5

6. I am willing to defend my views when they differ from others.
1 2 3 4 5

7. I am sensitive to those who are discriminated against.
1 2 3 4 5

8. I am open to people who strive to live lives very different from my life.
1 2 3 4 5

9. I think of my life in terms of giving back to society.
1 2 3 4 5

10. I work for the rights of others.
1 2 3 4 5

11. I frequently interact with people from a race/ethnic group different from my own.
1 2 3 4 5

12. I intentionally involve people from many cultural backgrounds in my life.
1 2 3 4 5

Part II:

The following statements describe tasks or situations inside an English class and outside the classroom. Please circle the number that best describes your opinion: 1 definitely not willing, 2 probably not willing, and 3 perhaps not willing 4 perhaps willing, 5 probably willing, or 6 definitely willing.

1. I am willing to present my arguments to the rest of my class in English.
1 2 3 4 5 6

2. I am willing to give a presentation in front of the class in English.
1 2 3 4 5 6

3. I am willing to participate in a small-group discussion in English.
1 2 3 4 5 6

4. I am willing to ask the teacher in English to repeat what s/he said.
1 2 3 4 5 6

5. I am willing to communicate with a foreigner (teacher, friend, etc.) in English on the internet.
1 2 3 4 5 6

6. I am willing to initiate communication with a foreigner I met on the street.
1 2 3 4 5 6

7. I am willing to speak to a foreigner who asks for help (e.g., help him/her find directions).
1 2 3 4 5 6

8. If given the opportunity, I am willing to speak to exchange students in English.
1 2 3 4 5 6

Appendix 2

Interview Prompts

1. How often and in what circumstance are you willing to communicating with others in English?
2. Have you ever tried to communicate with others in English in different situations?
3. How are you willing to communicate with people from different countries in English?
4. What is the role of English in your life now and in the future?

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 1, 2020

Nguyen, L. (2020). A Case Study of Vietnamese EFL Teachers' Conception of Language Output and Interaction. *Journal of Language and Education, 6*(1), 55-71. https://doi.org/10.17323/jle.2020.9777

A Case Study of Vietnamese EFL Teachers' Conception of Language Output and Interaction

Loi Nguyen
Can Tho University

Correspondence concerning this article should be addressed to Loi Nguyen, Can Tho University, Can Tho University, Campus I, 30/4 Street, Can Tho, Vietnam. E-mail: loinguyen@ctu.edu.vn

There exists a gap between what second language acquisition research has to say and what teachers do in the classroom. As part of an attempt to bridge this gap, this study is driven by the motive to understand how pedagogical innovations such as task-based language instruction can be influenced by teacher beliefs. Drawing on the perspective of research on teachers' thinking which aims to inform language teaching pedagogy and teacher education, the study employed multiple data sources (focus group, lesson plan, and stimulated recall interview) to tap into the conception of six Vietnamese EFL instructors regarding language output and interaction. Cross-case analysis showed that most of the teachers geared language output and interaction activities towards achieving a targeted linguistic aim. Further analysis revealed that this view reflects a synthetic, product-oriented conception of teaching and learning by skill-building, and is in line with traditional approaches which emphasise transmission style and form instruction. This finding implies that constructivist perspectives on teaching such as task-based language instruction may run counter to teachers' existing conception of teaching. The implementation of task-based instruction thus needs to consider negotiating between supporting teachers to focus on meaning and the need for form-focused instruction.

Keywords: teacher conception, output, interaction, task-based language teaching, product-oriented

Introduction

The transfer from theory to practice is still generating research interest in language learning and teaching (Mackay, Birello & Xerri, 2018) and contextual influences have become a crucial issue. In Vietnam, English is a school subject, and English teachers are familiar with the transmission style of teaching and form-focused instruction (Nguyen, Le & Barnard, 2015). However, recent curricular innovations have embarked on meaning-oriented, task-based language instruction that facilitates a constructivist approach to language teaching (Barnard & Nguyen, 2010). This innovation has attracted much attention among scholars and researchers. One assumption is that the new approach may run counter to teachers' existing beliefs and practices.

Decades ago, it was cautioned that teachers' beliefs constituted one of the strong forces that restricted the impact of educational reforms (Fullan, 1993). Such awareness has led to a boom of research on teacher cognition in an attempt to unveil the hidden side of teachers, which aims to offer implications for teacher education and development (Borg, 2006). Johnson (2006, p.239) has pointed out "an epistemological gap between how L2 teacher educators have traditionally prepared L2 teachers to do their work and how L2 teachers actually learn to teach and carry out their work".

Theoretically, knowledge about second language acquisition (SLA) principles is argued to offer teachers a tool for trial in their classrooms, and the opportunity to reflect and change (Johnson, 2006; MacDonald, Badger & White, 2001). However, research that focuses on the effect of SLA theory training has revealed little impact on teacher beliefs about language learning and teaching (MacDonald, Badger & White, 2001; Mattheoudakis, 2007; Peacock, 2001). In this respect, according to Berliner (2005), research on teachers' cognition should concentrate on "phenomena that have been found important from the perspective of the process-product research programme" (p.14). Thus, basic SLA principles such as language output and interaction in SLA theory merit an inquiry. Instead of investigating teachers' perceptions of task-based language teaching, this study

explored their responses to these SLA-enabling conditions underlying the task-based approach. This bottom-up approach to understanding teachers' reactions is expected to provide insight into how teachers perceive and uptake task-based language teaching (TBLT). The report in this paper is part of the researcher's doctoral thesis (Nguyen, 2011).

Literature Review

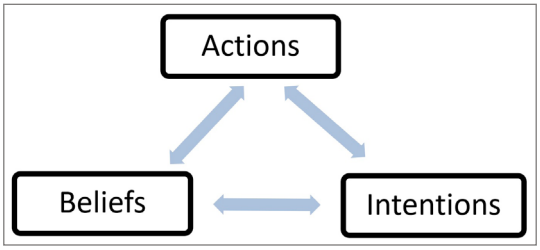
Conceptual Framework of Teacher Conceptions

Teachers' thinking has generated much interest in educational inquiry (Eley, 2006; Day, Calderhead, & Denicolo, 2012; Kember, 1997). However, the development of research on teachers' cognition has led to overlapping terminologies (Borg, 2006). Given this overlap, the current study employs the term 'conception' and Pratt's (1992) conceptual framework of teacher conceptions.

As noted by Entwistle, Skinner, Entwistle, and Orr (2000), in North America, the school education literature prefers such terms as 'beliefs' and 'knowledge', while the higher education literature uses the term 'conceptions' to indicate "researchers' ways of describing different aspects of teaching" (p.8). For example, Freeman and Richards (1993, p.194) referred to conceptions as systematic "bodies of knowledge about an object, idea, or phenomenon", based on which they identified *scientifically-based*, *theory and values-based*, and *art or craft-oriented* conceptions. The term 'conception' in this sense reflects the epistemological belief of how knowledge grows and teaching should be conducted. In contrast, according to Entwistle et al. (2000), in the European literature, the term 'conception' is commonly employed in higher education to denote "teachers' own ways of thinking and their beliefs about teaching" (p.8).

A popular definition of conception in adult education is described by Pratt (1992). Accordingly, "conceptions are specific meanings attached to phenomena, which then mediate our response to situations involving those phenomena" (p.204). Pratt (ibid.) further explained that a conception consists of actions, intentions, and beliefs, together forming a dynamic relationship as presented in Figure 1.

Figure 1
Aspects of Teacher Conception of Teaching



Note. Adapted from "Conceptions of teaching," by D. D. Pratt, 1992, *Adult Education Quarterly, 42*(4), p. 206. Copyright 1992 by SAGE.

Following Figure 1, one's beliefs, intentions and actions are intertwined, and together constitute one's interpretation. Thus, by examining beliefs in conjunction with intentions and actions, a conception can be unveiled. Pajares (1992, p.320) states that "beliefs and concepts are central to a conception". Benson and Lor's (1999) distinction between conceptions, beliefs, and approaches also aligns with Pratt's conceptualization. Accordingly, a belief can be investigated by data interpretation, while conceptions "call for a further level of analysis" (Benson & Lor, 1999, p.464). Conceptions and beliefs are translated into instructional approaches employed by teachers in a specific context. Therefore, based on the interaction between belief statements, intentions, and actions, teachers' conceptions can be identified (Pratt, 1992).

In language teaching, a similar conceptualisation of the term is found. To Tsui (2003), teachers' personal beliefs, assumptions, metaphors, images, and values constitute their conceptions. These conceptions, she contends, have an interactive effect on their teaching and development. This way of conceptualizing teacher conception

is also employed by other researchers (Farrell & Lim, 2005; Freeman, 1991; Mangubhai, Dashwood, Berthold, Flores, & Dale, 1998; Sato & Kleinsasser, 1999; Shi & Cumming, 1995).

Teachers' conception in the current study is, therefore, defined as teachers' personal ways of making sense of new information or knowledge. It can be reflected in teachers' personal established beliefs, pedagogical intentions, and classroom actions. Teachers' conceptions interact with their teaching context and classroom experiences (Borg, 2006).

Language Output and Interaction in Second Language Acquisition

Researchers have recognised the indispensable role of language output and interaction in fostering second language acquisition (Long, 1996; Pica, 1994; Shehadeh, 1999; Swain, 1995). Language production, both non-interactively and interactively, have recognised values. In the light of Output Hypothesis (Swain, 1995), second language learners can test out their assumptions about the target language rules, become fluent, and extend their interlanguage discourse (Ellis, 2003; Skehan, 1998; Swain, 1995). When pushed to produce language, learners become aware of "the relationship of the forms and rules to the meaning they are trying to express" (Swain & Lapkin, 1998, p. 69). Furthermore, the Interaction Hypothesis (Long, 1996) states that through interaction, learners negotiate for meaning and receive feedback which pushes them to modify their language in such a way that makes it comprehensible. In this way, learners actively generated comprehensible input, which is useful for second language learning (Long, 1996; Markee, 1997). Long (1996) stressed the connection of interaction, comprehensible input, and output with selective attention as the facilitative conditions for SLA.

Convergent empirical evidence in SLA research has testified to the relationship between output, interaction and second language learning (Ellis & He, 1999; Ellis, Tanaka, & Yamazaki, 1994; Gass & Torres, 2005; Izumi & Bigelow, 2000; Mackey, 1999; Pica, Young, & Doughty, 1987). After conducting a meta-analysis, Kech, Iberri-shea, Tracy-Ventura, and Wa-Mbaleka (2006), for example, concluded that interaction has some effect on the acquisition of certain lexical and grammatical features, especially when tasks force the use of essential features or elicit useful language for task completion.

Pedagogically, Ellis (2005) proposed ten crucial principles for instructed language acquisition. Two of these principles include creating opportunities for output and interaction in the target language. Proponents of TBLT also place a focus on communication and language use as the ultimate goal and principle of learning and teaching (Foster & Skehan, 1999; Prabhu, 1987; Skehan, 1996; Skehan, 2003; Swain & Lapkin, 2000; Willis & Willis, 2007).

In summary, theoretical, pedagogical and empirical arguments have confirmed that language output and interaction are potential facilitative conditions for second language learning. Therefore, from the SLA perspective, ESL and EFL teachers should engage learners in language use practice (Ellis, 2005), particularly when exposure to the target language is too inadequate to support natural language acquisition (Green, 2005). These theoretical positions, however, need to be examined against teachers' existing beliefs and practices.

Research on Teacher Conceptions about SLA Principles

Studies that explore the effect of SLA and methodology training on teachers' uptake of SLA principles generally point to a limited change in their beliefs about SLA. Mattheoudakis (2007) found that Greece EFL pre-service teachers changed some of their beliefs significantly after a three-year programme. Most of the teachers, for example, changed their belief about focusing on grammar knowledge, which they strongly held. They also thought teachers should correct all errors made by beginner learners, but this belief became obsolete. In Hong Kong, Peacock (2001) investigated 145 ESL pre-service teachers learning in a three-year programme, and discovered some changes in their SLA beliefs; but the majority of teachers still held on to the view that learning a language means mastering vocabulary and grammar rules. Likewise, MacDonald et al. (2001) detected a shift in only some beliefs of 55 TESOL non-native speaker undergraduate and postgraduate students after an SLA course. Overall, they found that the teachers significantly moved away from the behaviourist view, particularly from viewing language input as grammatical knowledge. Nonetheless, these teachers remained doubtful of the benefit of learner-learner interaction to their language learning. MacDonald et al. (2001) attributed this reluctance to the teacher-fronted teaching culture to which these students were accustomed before attending the training.

Such inadequate trust was also expressed by the six Thai EFL instructors in McDonough's (2004) study. They were concerned that interaction cannot push learners to modify their linguistic output or provide useful feedback to peers, and that learners produce less accurate TL forms during interaction. Pair work and group work, for them, were the opportunities for students to practice certain targeted language items, believing that students should focus on intended forms during practice. They also reported some factors which constrained their practice of promoting interaction, namely how to manage and monitor interaction in large classes with fixed desks, and the need for preparing students for high-stakes tests. In a different way, Howard and Millar (2009) responsively introduced Ellis's (2005) ten instructed SLA principles to fifteen South Korean EFL teachers attending a four-week professional development course in New Zealand. The researchers found that the Korean teachers lacked the confidence in using English proficiency to conduct class activities in English. Although all the teachers viewed interaction as an important principle, only two of them reported implementing interaction in their classrooms. Inadequate training, students' limited proficiency and L1 overuse, large and mixed-proficiency classes, and limited instructional time were the major reasons why most of them rarely generated opportunities for output and interaction. The biggest challenge involved the university entrance examination, which focused on grammar and reading rather than communicative skills. These findings, in general, resemble what McDonough (2004) reported.

In Vietnam, scant research has examined teachers' conceptions of SLA principles. Some studies reported findings relevant to the current issue under investigation. Chau (2009), with a survey of 106 EFL university instructors, revealed that the teachers gave more priority to grammatical accuracy than communication skills. Le (2008), in a different way, focused on high school teachers' beliefs regarding the new English task-based curriculum, an approach that concentrates on communication as well as language output and interaction. He concluded that the teachers' beliefs and practices remained the same. The teachers still focused on grammar and reading skills, which were to prepare their students for examinations. Case studies such as Barnard and Nguyen (2010), Nguyen (2014), and Nguyen, Le, and Barnard (2015) further confirmed that Vietnamese EFL teachers had a strong belief in teaching grammar, which restricted their focus on meaning and communication as required by task-based language teaching (TBLT).

The studies reviewed above provide a general understanding of EFL teachers' conceptions of SLA principles. However, most studies (e.g., MacDonald et al., 2001; Mattheoudakis, 2007; Peacock, 2001) concentrate on pre-service teachers, subscribing to an etic perspective, which stresses a straightforward causal effect and treats participants as "objects rather than subjects" (Kalaja & Barcelos, 2003, p.2) without considering the mediational effect of teachers' cognition. Teachers' conceptions and beliefs are strong mediators of their practice and change. A study by Howard and Millar (2009) relied on a context-responsive approach to teacher development to impact change, but the study employed interviews and questionnaires without classroom data. These methods fail to consider the correspondence between teachers' self-report and their real classroom practices (Breen, Hird, Milton, Oliver, & Thwaite, 2001). Although the case studies (Barnard & Nguyen, 2010; Nguyen, 2014; Nguyen, Le, & Barnard, 2015) have contributed some insight into Vietnamese teachers' responses to TBLT, further research using multiple methods is indispensable for gaining further insight. The current study relied on a bottom-up approach to working with EFL teachers at a workshop that focused on output and interaction, following them in their classrooms to record their lessons and interview them. From this, the study sought to answer the following question:

What conception do Vietnamese university EFL instructors hold about language output and interaction as manifested in their self-reporting and teaching of General English?

Method

Design

The present case study employed multiple methods to tap into the Vietnamese teachers' conceptions, following their self-report on group discussions, lesson planning, teaching, and stimulated reflections.

As a case design, the study relied on the working concept of cases as "specific individuals, particular events, processes, organizations, locations, or periods of time" (Stake, 1995, p.2). It aimed to examine a single instance in detail, as this may be useful to predict recurrent characteristics of human systems (Sturman, 1994), and

extend the observation to a wider population (Cohen, Manion, & Morrison, 2000). Contemporariness, real-life context, and boundary are three essential characteristics of a case study (Yin, 1994). Despite the blurred and indistinctive boundary of what a case is (Yin, 1994), as well as the tension in identifying uniqueness, complexity, and patterns, David (2006) contends that the power of case research is an in-depth exploration and description. In a case study, the results are not statistically but theoretically generalizable. According to (Yin, 1994, p.10), evidence from cases can serve to modify or develop theory, thus being generalised as "theoretical propositions" rather than "populations or universes".

In the present study, six individual cases of teachers are treated as subcases within a specific institutional case, a representative university. They serve dual purposes: to compare and reveal a holistic description of teacher conceptions, and to illustrate the meanings the teachers attached to language output and interaction in their specific context of General English (GE) classrooms. The study aimed to gain insight into how SLA precepts can be put to practice, a way of informing the gap between theory and practice. A combination of multiple data sources and cross-case analysis aimed to uncover "the complex "hidden side" of teachers (Freeman, 2002).

Instruments

Three main instruments were used to tap teacher conceptions: focus group interviews, lesson plans and quick interviews, and stimulated recall interviews after teaching.

Focus Group Interviews

Focus groups were selected because this method can elicit experiences, attitudes, and opinions (Wilson, 1997). A focus group discussion was conducted to obtain teachers' initial understanding of output and interaction before the teachers wrote lesson plans and had their lessons video-recorded. A discussion guide was provided for two small groups of three teachers, an ideal number to keep the discussion focused (Cohen, Manion, & Morrison, 2000). Each group discussion lasted around 20 minutes and the teachers' discussions in English were video-recorded. The following questions were posed:

1. What does the term 'output' mean to you?
2. What role does it have in language learning?
3. What is good interaction? What role does it play in language learning?
4. To what extent do you create opportunities for output and interaction in your lessons? Why?

Lesson Plans and Quick Interviews

Lesson plans and interviews were employed to document the teachers' intentions and beliefs about output and interaction. They served as the initial door to look into the teachers' conceptions. A quick individual interview was conducted individually with teachers before their lesson was taught, aiming to clarify the lesson's procedure, activities, and objectives.

Stimulated Recall Interviews Following Video-Recording

Stimulated recall interview (SRI) was employed to engage participants in viewing their lesson recordings, and then recalling and evaluating their actions. This introspective method is suited to elicit data about "thought processes involved in carrying out a task or activity" (Gass & Mackey, 2000, p.1). The issues of reliability and validity in stimulated recall data were addressed. Following Gass and Mackey's (2000) suggestions, to reduce risks of forgetting and ensure 95 percent accurate recall at least, the stimulated recall interview was conducted less than 48 hours after the lesson. The effect of prompt questions on teachers' recall was addressed by designing prompts that focused on their reflection at the time a classroom event happened (Gass & Mackey, 2000) (e.g., *What were you thinking? What was in your mind at that time?*). The interview schedule (see appendix) was based on a sample of Mackey, Gass, and McDonough (Cited in Gass & Mackey, 2000) to assist the interview process. Each interview lasted around one hour and was conducted in Vietnamese by the researcher. The interview was conducted in an arranged empty classroom so that the teachers felt secure. The teachers watched their recorded lessons on the researcher's laptop. While reviewing the lesson, the researcher paused at classroom activities related to output and interaction, and used the interview protocol (see appendix) to prompt the recall.

Participants

Purposeful sampling (Glense, 2005) was applied to select the participants with different teaching experiences, ages, and degree qualifications. Six lecturers from the Department of English of a university in the Mekong Delta of Vietnam agreed to participate. As shown in Table 1, these teachers' workloads for General English (GE) classes varied because they taught students of various majors and of different levels (GE level 1: 45 hours, GE level 2: 45 hours, and GE level 3: 60 hours). They had a range of teaching experiences from less than one year to 12 years. Three of the teachers held a Master's degree: one TESOL, one American Studies, and one Educational Studies. All of them had completed a Bachelor's programme in teaching English as a foreign language.

Table 1
Profiles of Six Vietnamese EFL Teachers

Teacher	Age	Gender	Years of service	Qualifications	Status	*GE hours/ total semester workload
Teacher 1	35	Female	12	BA in TEFL; MA in TESOL	Lecturer	45/175
Teacher 2	34	Female	11	BA in TEFL; MA in American Studies	Lecturer	135/360
Teacher 3	29	Female	6	BA in TEFL; MA in Educational Studies	Lecturer	135/270
Teacher 4	26	Female	4	BA in TEFL	Lecturer	185/330
Teacher 5	24	Female	3	BA in TEFL	Lecturer	180/330
Teacher 6	22	Female	0.5	BA in TEFL	Apprentice Lecturer	225/225

Note. Each GE hour lasts 45 minutes. Reprinted from *Dynamic conceptions of input, output and interaction: Vietnamese EFL lecturers learning second language acquisition theory* (p. 114), by V. L. Ngyen, 2011, University of Waikato. Copyright 2011 by V. L. Ngyen.

Data Collection Procedure

Invitation letters were sent to nine lecturers in the Department of English of the university where the researcher is working. Six of them agreed to participate. A meeting was set up with them to talk about the project in detail and consent forms were signed. Following that, a focus group discussion was conducted in the first workshop about SLA facilitating conditions. They were divided into two groups, and each discussion was recorded with a digital recorder. After the workshop, the teachers planned and taught a lesson that incorporated their views about language output and interaction. Before teaching the lesson, they were asked to clarify their intentions in a quick interview. While the lesson was taught, the researcher recorded it with a handheld camera. To avoid disturbing them, the researcher sat in a corner, using the camera's zoom function. One day after the lesson was taught, a stimulated recall interview was conducted.

Data Analysis

The SRIs were transcribed and translated into English by the researcher. This allowed for a more precise translation because of the researcher's familiarity with the contextual information. Feedback on the translation was obtained from a native speaker and a colleague so that readability could be improved and information about the context could be fully included. Upon obtaining the feedback, the researcher returned to the recordings to check the information reported and fixed the English translation to ensure a match of representations of information.

Data analysis followed the procedure suggested by Braun and Clarke (2006). Data were initially coded with short descriptions and then reduced into themes. A cross-case comparison of themes was made to obtain a holistic

view, and individual cases were examined to gain further insights and develop thick and detailed descriptions. Lesson plans were described in terms of objectives and activities, and documented into tables. A triangulation

Results

An Initial Product-Oriented Conception

The data from focus group interviews showed that all the teachers initially held a perception of output as language production, viewing it as an end-product rather than a process to facilitate language learning. For example, Teacher 4 stated, "The language output is what students can use, can produce after a lesson". Likewise, Teacher 1 said output means "How much progress students make in picking up the language".

The teachers also identified features of good interaction congruent with the literature. All of them viewed effective interaction as two-way communication or information exchange. Teacher 2 said, "If they do not understand, they should ask, and by asking questions to be clarified, that is a good interaction." Besides recognizing negotiation for meaning as the key feature of good interaction, Teacher 1 added effective interaction a communicative goal: "They can achieve the aim that they want." Teacher 6 and Teacher 4 also perceived interaction as an essential condition for peer feedback and learning. Teacher 4 clarified:

> I think good interaction means that students can work effectively with their partners and they can learn from their partners, and their partners can find out their mistakes, adjust themselves, and correct their mistakes themselves.

All the teachers thought that interaction is required for English learning. Teacher 4 emphasised that, "…there must be interaction in the classroom. Interaction helps students to produce or to perform what they have studied; it is to prepare them for language performance outside the classroom".

Most of the teachers agreed they regularly created opportunities for students to practice language output and interact with one another. For instance, Teacher 4 used role cards to encourage students to make a conversation. Teacher 2 employed discussion before her students did main reading or listening tasks. Besides, Teacher 5 shared her way of providing output practice from a controlled to a free manner:

> With controlled practice, I give students handouts, some information gap activities or some charts, so that I can check at least that they can use English under the teacher's control, and then move on to free production. I provide many choices, many questions, or many prompts in order for the students to be able to talk about what they want, based on their acquisition.

Only Teacher 1 said she rarely engaged GE students in group or pair interaction in the classroom because of their limited and mixed-level proficiency, willingness to participate, and especially the time restriction. Teacher 1 explained that with only 30 or 45 hours, with that number of lessons [ten units], she found it hard to create interaction in the classroom. Teacher 2 also reported that for output like writing she usually asked her students to do it at home because the activity is time consuming.

In general, the focus group data showed that all the teachers initially shared a product-oriented conception of language output. They understood the nature of interaction, and believed language output and interaction are important for language learning. However, some contextual factors could affect their practice such as time restrictions and students' characteristics.

Conception of Output and Interaction in Practice

An analysis of the lesson plans and stimulated recall revealed some intentions and actions across four major themes: target linguistic content-oriented, language accuracy-oriented, institutional restrictions, and student characteristics.

Target linguistic content-oriented

Target linguistic content orientation describes the teachers' control of language production activities in such a way that focuses on particular language items they were teaching, although the way output and interaction were manipulated varied slightly among the six teachers. Their lesson plans partially uncovered their intentional focus.

As shown in Table 2, oral output and interaction was implemented in all the lessons, whereas written output was also given in Teacher 1's and Teacher 2's lessons. Most of the lesson objectives, however, showed the production of a particular TL content. Teacher 3, Teacher 2, and Teacher 6 clearly gave a priority to the linguistic objective. Teacher 1's plan put the topical content on top, and the linguistic objective followed as a support, which was confirmed in the lesson plan interview: "The grammar use [final objective listed] supports all the output parts above." Teacher 5's planned focus appeared to resemble Teacher 1's. Teacher 4 did not specify any objectives.

Table 2
Lesson Plans for Output and Interaction

Teachers	Lessons	Time	Objectives	Focus
Teacher 1	How do you get to work? (Level 2)	135 min.	- Read and comprehend a short passage on the topic of travel and transportation. - Talk about how they [students] get to school/work and the trip they have just taken. - Listen to some people's talks on travelling. - Write a paragraph describing how they get to school/work/ travel. - Use grammatical points such as articles, present tense, or past tense in both writing and speaking.	Comprehension Oral & topical Comprehension Written & topical Linguistic, oral and written
Teacher 2	I'm going to save money. (Level 3)	135 min.	- Use *be going to* to talk and/or write about their near future plans and *because* plus a clause to give the reason for the plans. - Use so plus a clause to give a consequence. - Use a number of vocabulary items related to future plans (save money, buy a new bicycle/dictionary/cell phone, move out/in, take a new course, spend less money on clothes/food, invite friends to a party, change sleeping habits, etc.)	Oral/written & linguistic
Teacher 3	There is/ there are (Level 1)	NG	Students will be able to describe things and people using There is/There are	Oral & linguistic
Teacher 4	I'm going to save money. (Level 3)	NG	NG	None
Teacher 5	Can I help you? (Level 3)	135 min.	- Read and listen for specific information. - Say what they want to buy, make decisions as well as the way to express opinions at a shop. - Talk about their shopping habits. - Use collocations for uncountable and countable nouns.	Comprehension Oral & functional Topical Oral & linguistic
Teacher 6	A day in my life (Level 2)	NG	- Use words/phrases about daily activities fluently. - Ask/answer Yes-No and Wh-questions in present simple tense.	Oral & linguistic

Note. NG: not given. Reprinted from *Dynamic conceptions of input, output and interaction: Vietnamese EFL lecturers learning second language acquisition theory* (p. 176), by V. L. Ngyen, 2011, University of Waikato. Copyright 2011 by V. L. Ngyen.

The teachers' intentions were further revealed in the way they organised output and interaction activities in Table 3

Table 3
Procedure of Lesson Plans for Output and Interaction

Teachers	Lessons	Lesson procedure
Teacher 1	How do you get to work? (Level 2)	-T introduces vocabulary on transportation and travel. -Ss practice in pairs, asking and answering questions on how to get to school. -Ss read a passage on transportation. -Ss notice the use of articles in the passage. -T explains and Ss complete practice exercises in the book. -T introduces vocabulary about vacation activities. -Ss listen to people talking about their vacations. -Ss talk about their vacations in pairs. -Ss write a paragraph describing how they go to school/work or their vacation.
Teacher 2	I'm going to save money. (Level 3)	-T presents vocabulary. -Ss read the passage "My New Year's Resolution." -Ss work in pairs and discuss the reasons why some people make their resolutions (given in the textbook unit). -Ss report the reasons and T writes them on the board. -Ss work in pairs to match the reasons with the resolutions -Ss listen to four people talking about their resolutions and take notes. -T translates the grammar points and examples presented in the unit. -Ss recognise the difference between simple present and present continuous tense. -Ss write sentences with the verbs given in the textbook unit. -Ss write about their plans individually. -Ss go around and ask each other about their plans for this school year
Teacher 3	There is/there are (Level 1)	-T presents THERE BE. -Ss practice the structure with a transformation drill. -Ss practice asking and answering the questions in pairs. -Ss do practice exercises in the book. -Ss work in pairs to ask each other about the numbers of things and people in their pictures (given by T).
Teacher 4	I'm going to save money. (Level 3)	-T teaches vocabulary. -Ss match the resolutions with the reasons (given in the textbook unit). -Ss do information gap task, exchanging information about the resolutions of two people. -Ss prepare to talk about their plans for the weekend. -Ss share their plans with a partner. -T correct any common mistakes during pair work.
Teacher 5	Can I help you? (Level 3)	-T teaches vocabulary, presents reflexive pronouns, expressions to say in a shopping situation. -Ss practice grammar activities. -Ss listen to a shopping conversation. -Ss work in pair/group to share decisions about shopping for something. -Ss role-play the conversations.
Teacher 6	A day in my life (Level 2)	-Ss work on the meaning of new phrases given in a list, and ask each other how often and what time they do the activities in the list. -One pair of students demonstrate the practice. -Ss listen to Sam's activities [on an audio] and complete the table/list. -Ss work in pairs, ask and answer questions to complete a chart about three famous persons.

Note. Reprinted from *Dynamic conceptions of input, output and interaction: Vietnamese EFL lecturers learning second language acquisition theory* (p. 176), by V. L. Ngyen, 2011, University of Waikato. Copyright 2011 by V. L. Ngyen.

Table 3 further shows that the lesson procedures of Teacher 3 and Teacher 6 were aligned with their specified aims, while those of Teacher 1, Teacher 2, and Teacher 5 were not. In contrast, in Teacher 1's plan, the language focus was ordered before the language output activity, which was incongruent with her objectives reported in the lesson interview. Teacher 2's lesson apparently proceeded in the way that prioritised meaning expression before the simple past tense was taught, which differed from the objectives listed (Table 2). Teacher 5 clearly indicated a focus on teaching target linguistic forms prior to providing output practice in the subsequent activities. The target forms included 'self' and common expressions used at a shop. In contrast, Teacher 4's procedure spelled out her intention to prioritise meaning; that is, students' speaking about New Year's resolutions and plans for weekends.

Further evidence from the SRIs uncovered a form-oriented linguistic performance. In Teacher 3's video-recorded lesson, the form 'there be' was explicitly instructed first. Then a transformation drill was used to develop students' automatic processing and accurate production of the form. A question-answer drill followed to further prepare the students for an information gap task later that was provided as an opportunity for contextualised practice. The teacher explained that the last task was to evaluate whether the students used the structure accurately. She recalled:

> My purpose at this stage was to give them practice on how to make and answer the questions. Therefore, this was the preparation for them to do the production stage later on. It was a bit inauthentic but useful to them, I think, because they repeated the *structure*, and they would memorise that *structure*. (Emphasis added)

Likewise, Teacher 2 provided opportunities for language production late in her lesson. After presenting the structure 'be going to', she asked her students to write down their own weekend plans, and then move around to share this with their classmates orally. For Teacher 2, these activities allowed students to reproduce the language, or "apply the language they had just learned to talk about their weekend." She explained that this was aimed at helping her students be able to use 'be going to' to talk about their weekend plan, which was in line with her lesson plan.

For Teacher 1, the first language practice, the ask-answer practice on means to go to school, was the opportunity when "they can make a small conversation; for example, they would ask about how to go, how long it takes, why they like to take a certain transport, or so on." This explanation reflects a linguistic focus on certain questions. A more authentic task, she added, was students' talking about their own vacations in pairs, which represents the same linguistic focus. She recalled:

> ...in pairs, they more or less had a chance to speak English, but in terms of using *these questions* in real life, I am afraid that this activity was not enough....So in the next lesson, I planned to give them another situation where they will role play going to a travel agency to ask, for example, if they want to go to Dalat City, they will ask 'How can I go there?' or 'How far?' or 'How much?' so they can know how to ask questions. (Emphasis added)

Teacher 6 provided two language practice activities. The first one was an interview with cue cards where students acted as journalists interviewing the daily routines of a celebrities, followed by a brief oral report. In this activity, she recalled, "students can both ask and answer and finally come to the same outcome," an opportunity for them to practice the target language items taught earlier in the lesson. The second activity was the students' interview about their daily routines where they practised using the questions they had studied, "First, they asked about the activities, asked *what*, and second *how often*. Those who asked would learn how to ask, and those who answered would use *adverbs of frequency*, and then focusing on time, they would use *prepositions of time* to answer questions." (Emphasis added)

However, Teacher 6 believed language should bear a resemblance to real life, instead of what the textbook provided. This urged her to introduce the input close to her students' daily routines. For her, this activity generated more authentic language use because it freed students so that "they could choose to ask and answer about their daily activities". Her explanation evidently revealed the need for teaching certain language items targeted in the lesson, but she adapted to create a meaningful language practice.

Overall, there was obvious evidence that the teachers had a tendency to gear language output and interaction activities towards the linguistic content specified in the textbook in the control-free process. This predisposition is linked to their concern for controlling students' language use to impact accuracy.

Language Accuracy-Oriented

A concern for language accuracy was observed among the six teachers. It was manifested in a variety of ways they ensured students' accurate performance. Table 3 shows that most of the teachers delayed 'free production' until the end of their lessons. In practice, the teachers initially taught students the vocabulary items or grammatical structures necessary for their language output later. Presentation was stressed as being essential for students as it can give them various sources of language input. Teacher 5 said, for the usual utterance 'I need', students can

be taught to say 'I am looking for or I'd like". Teacher 6 revealed this belief when she reviewed some grammatical patterns to guarantee her students' accurate language production: "I wanted them to do it [make questions] by themselves before they did the task because I heard some students asking *inaccurate* questions". (Emphasis added)

The second way to ensure accuracy involved pronunciation practice before pair work practice. This helps students develop both fluency and accuracy in using the target forms in subsequent output activities. Teacher 2 explained why she had students repeat some words aloud, "I wanted to prepare them for the speaking activity". Teacher 6 shared the same worry in her recall session, "I wanted to ensure they had correct pronunciation before they produced output".

Another example lies in teachers' reasons for initiating interactions with their students. Teacher 2 viewed her intention to interact with her students not as a way to maximise learning opportunities for students, but as a classroom management strategy. She commented on an exchange with some students after their pair work talking about daily routines: "It was not to find out about their peers' routines, but to check who had worked and whether they could use *accurate grammatical structures* or pronounce an *accurate final -s*." (Emphasis added)

Teachers' authority is manifested in their control of how students wanted to use the language. The power distance in the cultural context of Vietnam (Hofstede, 2019) has entitled Teacher 2 to dominate her students and forced them to stay on the track she intended. She recounted: "This morning I crossed out a sentence a student said. I wrote a clear model introduction: 'Hello, everyone, I would like to tell something about a US holiday.' The student began with 'Today I introduce to you a holiday.' I wrote it on the board, crossed it out, and told him never to use it again."

In summary, although the six teachers exhibited an initial understanding of the merit of language output and interaction, a common thread among them was their manipulation of lesson activities to target the accurate production of particular planned linguistic forms. Such practice reflects a conception of language accuracy and form-oriented teaching and learning, and is further mediated by various contextual constraints that will be discussed below.

Institutional Constraints

The way the teachers implemented output and interaction was restricted by the institutional syllabus restriction, which was based on the textbook. Their lesson plans represented the textbook units they had to teach. Evidence from lesson interviews further showed that the prescribed content somehow influenced how they taught despite its relevance to the students' lives. Teacher 1, for example, recounted, "From the title 'How do you get to work?' when you introduce the lesson, you can ask this question, so the students would have to notice the question and how to answer it. The question lies in Exercise 3. I turned this exercise into an oral practice."

Time restrictions created difficulties for optimising output and interaction in the GE classroom. According to the GE syllabus, ten units had to be completed within forty-five standard hours, each lasting 45 minutes. Thus, a particular activity like sharing opinions about some statements in pairs was assumed to be potentially time-consuming and was replaced with a teacher-fronted discussion by Teacher 2.

> Before listening, there is a discussion section 'What do you think about these statements?' so I prepared them for this discussion. Here in the book there was 'work in pairs' but I did not do that… If I had had time I would have let them do that.

Also, because of the time limit, accuracy emphasis right from the beginning was thought to be effective. Teacher 2, for instance, recalled a moment in her lesson that "My purpose was to let them listen first, so that they could imitate the conversation, so they could be correct from the start. If I had let them think and do the work by themselves, they would have used Vietnamese or asked me and it would have wasted time."

Student Characteristics

The teachers perceived student characteristics as a barrier to fostering language output and interaction. These characteristics involved students' limited and mixed levels of English, students' passive attitude to learning, and students' failure to initiate communication.

First, students' limited and mixed levels of proficiency made the teachers cautious. As mentioned, all the teachers had a strong belief in language instruction for language performance, mostly manipulating student production by giving a model and delaying activities for free language production until the later phases of a lesson. Teacher 1 said in the recall, "For this level [GE], if you do not give a model and just give a task and ask them to talk… even though I gave them a model, as you see, their output was so bad". The teacher's orientation towards language output accuracy reflects a cost-benefit analysis view in teaching English to mixed-ability classes. Teacher 2 expressed a similar mind-set, "You will become very tired of correcting them. Here they were all prepared, and they just needed to speak." She expressed her strong belief that without a dialogue model, they would have made mistakes. This concern represents a product-oriented perspective on learning where trial-and-error is discouraged. She continued, "They did not have to try and make mistakes. *I was afraid they would make mistakes*… For GE students I think we should introduce the correct form right from the start, so they could follow the track." (Emphasis added)

Students' unfamiliarity with active learning was an additional reason. This implied a common assumption that Vietnamese students in GE classes are often passive learners. Teacher 5 commented, "If they had been more active, I would have asked them to talk in front of the class and make their decisions". Teacher 2 complained, "They just listen and wait to be asked to give an answer, but they seldom ask questions." Teacher 1 described it as "the students' culture of learning, an established study habit that was still not learner-centred". With such a learning culture, Teacher 1 said, "[Teachers] have to give them a model and then give them time to practice," or "teachers need to scaffold lessons and can't allow students complete freedom." This culture is related to students' lack of competence in handling a conversation. Teacher 2 attributed this to the culture where "people do not like to ask questions", and stressed that "students can't communicate or start a conversation, and they will get confused and begin to speak Vietnamese". Such a perception might have induced the teachers to model interaction for their students. For example, Teacher 5 said "I began to ask sample questions, so they would know how to maintain a conversation".

Overall, most of the EFL teachers in the current study had a tendency to control output and interaction activities mainly for the desire of accurate performance. Their responses reflected a language form and accuracy-oriented teaching view. Factors in the educational context, including both institutional factors and students' characteristics, mediated their conception.

Discussion

Unlike the studies of MacDonald et al. (2001) and McDonough (2004), the current study showed that the six Vietnamese teachers believed in the importance of learner language production in second language learning, a similar outcome reported in Howard & Millar's (2009) study. The teachers' tendency to manipulate output and interaction activities for the mastery of certain targeted linguistic content is in line with previous research in a similar context (McDonough, 2004). Although the Vietnamese teachers did not believe that peer interaction could lead learners to make errors as their colleagues did in McDonough's (2004) study, they implicitly signalled disbelief in the contribution of student talk in pairs and groups, given their mind-set of language output accuracy. To foster learner output and interaction, most of the Vietnamese teachers created activities for contextual and communicative practice of certain targeted language items, the kind of activities resembling what Ellis (2003) called linguistically focused tasks designed to support meaningful language practice. Such teaching practice is depicted as a weak approach to using tasks (Skehan, 1996), which is favoured by many Asian teachers in Hong Kong primary schools (Carless, 2003, 2007) and the Thai tertiary classroom context (Todd, 2006).

The conception of the teachers in the current study described above represents Vietnamese EFL teachers' belief in grammatical accuracy over communication skills (Chau, 2009). It also resembles what previous research on curricular innovations regarding the necessity for explicit grammar teaching (Barnard & Nguyen, 2010), and preparing students linguistically before task performance (Nguyen, Le & Barnard, 2015). Such established practice that builds around particular language forms reflects a deep belief about form-focused instruction, and a lack of focus on meaning aligned with TBLT (Nguyen, 2014).

Conceptions of learning have a correlation with and underpin instructional approaches (Goodyear & Hativa, 2002; Lamb & Kember, 2006; Richards & Rodgers, 2001), and conceptions of learning are related to conceptions

of language teaching (Benson & Lor, 1999; Freeman & Richards, 1993). Therefore, we can infer the conception of the Vietnamese EFL teachers about English learning and teaching based on their conception of language output and interaction. In this way, a synthetic approach (Long & Crookes, 1992), a reflection of a product-oriented conception of teaching, could better capture these teachers' views. It represents the perspective of language pedagogy primarily oriented towards the presentation and practice of discrete language items or forms, as opposed to communicative meaning or fluency-oriented pedagogy (Wilkin, 1976). It can also be likened to focus-on-forms (Ellis, 2001), although their conception is not entirely structural at the expense of communicative meaning. The orientation of these teachers is likely to have its roots in their established experiences in teaching GE as well as professional backgrounds. It is further shaped by contextual constraints, namely institutional constrains and student characteristics. Such a conception seems to compromise the use of communicative tasks in that language production opportunities serve to prime the internalization of TL forms.

Conclusion

Despite its limited generalizability, the study offers useful implications. Considering that language pedagogy needs to investigate teaching conceptions to benefit language teacher education and development (Freeman & Richards, 1993), it is imperative that teachers' preconceptions, which could be shaped by prior training or classroom experience, be tackled and brought to explicit discussion and reflections. In that way, theory and practice could be brought closer. The conception of Vietnamese EFL teachers in the current study uncovers a form-focused perspective on language teaching, which is reflected in the present-practice-produce model they are familiar with. This conception seems to fit into their teaching context. Thus, for bottom-up approaches such as TBLT which advocates a constructivist and trial-and-error view of learning to be incorporated in an Asian context like Vietnam, this teacher conception should be addressed if the gap is to be narrowed.

In conclusion, the current study employed multiple sources of data to unpack what six Vietnamese EFL instructors believed about language output and interaction. The case-based thematic analysis revealed that the teachers held a conception of language teaching in the light of skill building theory (McLaughlin, 1987). This empirical evidence is corroborated by similar studies in other Asian contexts. The finding implies that the Western idea about a flexible communicative approach with a strong focus on meaning and a constructivist view of learning needs to work in tandem with the conception of learning and teaching in the light of skill acquisition and building in EFL settings. In these settings, teaching and learning are not supported in terms of an authentic language environment, institutional imperatives, and students' characteristics compatible with a constructivist learning approach. The limitation of the current study is the small sample size biased towards gender. Further research could employ a larger sample which involves a wider range of variables.

Acknowledgement

Thanks are due to blind reviewers and the editors who contributed constructive feedback to shape up this paper.

References

Barnard, R., & Nguyen, G.V. (2010). Task-based language teaching (TBLT): A Vietnamese case study using narrative frames to elicit teachers' beliefs. *Language Education in Asia, 1*(1), 77-86. 10.5746/LEiA/10/V1/A07/Barnard_Nguyen

Benson, P., & Lor, W. (1999). Conceptions of language and language learning. *System, 27*(4), 459-472.

Berliner, D. C. (2005). The place of process-product research in developing the agenda for research on teacher thinking. In P. M. Denicolo & M. Kompf (Eds.), *Teacher thinking and professional action* (pp. 3-15). Routledge.

Borg, S. (2006). *Teacher cognition and language education*. Continuum.

Braun, V., & Clarke, V. (2006). Using thematic analysis in Psychology. *Qualitative Research in Psychology, 3*, 77-101. http://dx.doi.org/10.1191/1478088706qp063oa

Breen, M.P., Hird, B., Milton, M., Oliver, R., & Thwaite, A. (2001). Making sense of language teaching: Teachers' principles and classroom practices. *Applied Linguistics, 22*(4), 470-501. https://doi.org/10.1093/applin/22.4.470

Carless, D.R. (2003). Factors in the implementation of task-based teaching in primary schools. *System, 31*(4), 485-500. https://doi.org/10.1016/j.system.2003.03.002

Carless, D. (2007). The suitability of task-based approaches for secondary schools: Perspectives from Hong Kong. *System, 35*, 595-608. https://doi.org/10.1016/j.system.2007.09.003

Chau, T. H. (2009, April). *Balancing grammar and communication in EFL teaching: A study of teachers' beliefs about grammar instruction and error correction*. Paper presented at the Seventh Asian EFL Journal Conference. Pusan, Korea.

Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education*. Routledge Falmer.

David, M. (2006). Editor's introduction. In M. David (Ed.), *Case study research* (vol. I, pp. xiv- xlii). Sage.

Day, C., Calderhead, J., & Denicolo, P. (Eds.). (2012). *Research on teacher thinking: Understanding professional development*. Routledge

Eley, M.G. (2006). Teachers' conceptions of teaching and the making of specific decisions in planning to teach. *Higher Education, 51*(2), 191-214. https://doi.org/10.1007/s10734-004-6382-9

Ellis, R. (2001). Introduction: Investigating form-focused instruction. *Language Learning, 5*1, 1-43. https://doi.org/10.1111/j.1467-1770.2001.tb00013.x

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Ellis, R. (2005). Principles of instructed language learning. *Asian EFL Journal, 7(3).* https://www.asian-efl-journal.com/main-journals/principles-of-instructed-language-learning/

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition, 21*, 285-301.

Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning, 44*(3), 449-491.

Entwistle, N., Skinner, D., Entwistle, D., & Orr, S. (2000). Conceptions and beliefs about good teaching: An integration of contrasting research areas. *Higher Education Research and Development, 19*(1), 5-26. https://doi.org/10.1080/07294360050020444

Farrell, T.S.C., & Lim, P.C. (2005). Conceptions of grammar teaching: A case study of teachers' beliefs and classroom practices. *Teaching English as a Second or a Foreign Language*, *9*. http://tesl-ej.org/ej34/a9.html.

Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research, 3*(3), 215-247. https://doi.org/10.1177/136216889900300303

Freeman, D. (1991). "To make the tacit explicit": Teacher education, emerging discourse, and conceptions of teaching. *Teaching and Teacher Education, 7*(5-6), 439-454. https://doi.org/10.1016/0742-051X(91)90040-V

Freeman, D. (2002). The hidden side of the work: Teacher knowledge and learning to teach: A perspective from North American educational research on teacher education in English language teaching. *Language Teaching, 35*(1), 1-13. http://dx.doi.org/10.1017/S0261444801001720

Freeman, D., & Richards, J.C. (1993). Conceptions of teaching and the education of second language teachers. *TESOL Quarterly, 27*(2), 193-216. http://dx.doi.org/10.2307/3587143

Fullan, M. (1993). *Changing forces: Probing the depths of educational reform*. Falmer.

Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Lawrence Erlbaum Associates.

Gass, S.M., & Torres, M.J.A. (2005). Attention when? An investigation of the ordering effect of input and interaction. *Studies in Second Language Acquisition, 27*, 1-31.

Glense, C. (2005). *Becoming qualitative researchers: An introduction* (3rd ed.). Pearson Education.

Goodyear, P., & Hativa, N. (2002). Introduction: Research on teacher thinking, beliefs and knowledge in higher education. In N. Hativa & P. Goodyear (Eds.), *Teacher thinking, beliefs and knowledge in higher education* (pp. 1-14). Kluwer.

Green, C. (2005). Integrating extensive reading in the task-based curriculum. *ELT Journal, 59*(4), 306-311. http://dx.doi.org/10.1093/elt/cci059

Howard, J. & Millar, S. (2009). The applicability of principles for instructed second language learning: A South Korean perspective. *Asian EFL Journal, 11*(4). http://www.asian-efl-journal.com/December_2009_jh.php

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly, 34*(2), 239-278. https://doi.org/10.2307/3587952

Johnson, K.E. (2006). The socio-cultural turn and its challenges for second language teacher education. *TESOL Quarterly*, *40*(1), 235–252. https://doi.org10.2307/40264518

Kalaja, P., & Barcelos, A.M.F. (2003). Introduction. In P. Kalaja, & A.M.F. Barcelos (Eds.), *Beliefs about SLA: New research approaches* (Vol. 2, pp. 1-4). Kluwer.

Kech, C.M, Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S. (2006). Investigating the empirical link between task-based instruction and acquisition: A meta-analysis. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 91-132). John Benjamin's Publishing Company.

Kember, D. (1997). A reconceptualisation of the research into university academics' conceptions of teaching. *Learning and Instruction, 7*(3), 255-275. https://doi.org/10.1016/S0959-4752(96)00028-X

Lam, B.H., & Kember, D. (2006). The relationship between conceptions of teaching and approaches to teaching. *Teachers and Teaching, 12*(6), 693-713. https://doi.org/10.1080/13540600601029744

Le, V.C. (2008). Teachers' beliefs about curricular innovation in Vietnam: A preliminary study. In Y. H. Choi, B. Spolsky & R. Barnard (Eds.), *ELT curriculum innovation and implementation* (pp. 191-216). Asia TEFL.

Long, M.H. (1996). The role of linguistic environment in second language acquisition. In W. Ritchie & T. J. Bhatia (Eds.), *Handbook of second language acquisition* (pp.413-468). Academic Press.

Long, M. H., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly, 26*(1), 27-56. https://doi.org/10.2307/3587368

MacDonald, M., Badger, R., & White, G. (2001). Changing values: What use are theories of language learning and teaching? *Teaching and Teacher Education, 17*(8), 949-963.

Mackey, A. (1999). Input, interaction, and second language development. *Studies in Second Language Aacquisition, 21*, 557-587. https://doi.org/10.1017/S0272263199004027

Mackay, J., Birello, M., & Xerri, D. (2018). *ELT research in action: Bridging the gap between research and classroom practice*. IATEFL

Mangubhai, F., Dashwood, A., Berthold, M., Flores, M., & Dale, J. (1998). *Primary LOTE teachers' understandings and beliefs about communicative language teaching: Report on the first phase of the project*. http://eprints.usq.edu.au/890/1/Mangubhai_et_al_CLTREP97.pdf

Markee, N. (1997). Second language acquisition research: A resource for changing teachers' professional cultures? *Modern Language Journal, 81*(1), 80-93. https://doi.org/10.1111/j.1540-4781.1997.tb01628.x

Mattheoudakis, M. (2007). Tracking changes in pre-service EFL teacher beliefs in Greece: A longitudinal study. *Teaching and Teacher Education: An International Journal of Research and Studies, 23*(8), 1272-1288. http://dx.doi.org/10.1016/j.tate.2006.06.001

McDonough, K. (2004). Learner-learner interaction during pair and small group activities in a Thai EFL context. *System, 32*, 207-224. http://dx.doi.org/10.1016/j.system.2004.01.003

McLaughlin, B. (1987). *Theories of second language learning.* Edward Arnold

Nguyen, G. V. (2014). Forms or meaning? Teachers' beliefs and practices regarding task-based language teaching: A Vietnamese case study. *The Journal of Asia TEFL, 11*(1), 1-36. https://doi.org/10.1075/ltyl.00004.new

Nguyen, V.G., Le, V.C., & Barnard, R. (2015). "Old wine in new bottles": Two case studies of task-based language teaching in Vietnam. In M. Thomas & H. Reinders (Eds.), *Contemporary task-based language teaching in Asia* (pp. 68–86). Bloomsbury.

Nguyen, V. L. (2011). Dynamic conceptions of input, output and interaction: Vietnamese EFL lecturers learning second language acquisition theory [Unpublished doctoral dissertation, University of Waikato]. https://hdl.handle.net/10289/5167

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research, 62* (3), 307-332. https://doi.org/10.3102/00346543062003307

Peacock, M. (2001). Pre-service ESL teachers' beliefs about second language learning: A longitudinal study. *System, 29*(2), 177-195. https://10.1016/S0346-251X(01)00010-0.

Pica, T. (1994). Research on negotiation: What does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning, 44*(3), 493-527. https://doi.org/10.1111/j.1467-1770.1994.tb01115.x

Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly, 21*(4), 737-758. https://doi.org/10.2307/3586992

Prabhu, N.S. (1987). Second language pedagogy. Oxford University Press.

Pratt, D. D. (1992). Conceptions of teaching. *Adult Education Quarterly, 42*(4), 203-220. https://doi.org/10.1177/074171369204200401

Richards, J., & Rodgers, T. (2001). *Approaches and methods in language teaching* (2nd ed.). Cambridge University Press.

Sato, K., & Kleinsasser, R. (1999). Multiple data sources: converging and diverging conceptualisations of LOTE teaching. *Australian Journal of Teacher Education, 24*(1), 17-34. http://dx.doi.org/10.14221/ajte.1999v24n1.2

Shehadeh, A. (1999). Non-native speakers' production of comprehensible output and second language learning. *Language Learning, 94*(4), 627-675. https://doi.org/10.1111/0023-8333.00104

Shi, L., & Cumming, A. (1995). Teachers' conceptions of second language writing instruction: Five case studies. *Journal of Second Language Writing, 4*(2), 87-111. https://doi.org/10.1016/1060-3743(95)90002-0

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*(1), 38-62. https://doi.org/10.1093/applin/17.1.38

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Skehan, P. (2003). Task-based instruction. *Language Teaching, 36*, 1-14. https://doi.org/10.1017/S026144480200188X

Stake, R.E. (1995). *The art of case study research*. Sage.

Sturman, A. (1994). Case study methods. In J.P. Keeves (Ed.), *Educational research, methodology, and measurement: An international handbook* (pp.61-66). Pergamon.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125-144). Oxford University Press.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *The Modern Language Journal, 82*(3), 320-337. https://doi.org/10.1111/j.1540-4781.1998.tb01209.x

Swain, M., & Lapkin, S. (2000). Task-based second language learning: the uses of the first language. *Language Teaching Research, 4*(3), 251-274. https://doi.org/10.1177/136216880000400304

Todd, W.R. (2006). Continuing change after the innovation. *System, 34*(1), 1-14. https://doi.org/10.1016/j.system.2005.09.002

Tsui, A. B. M. (2003). *Understanding expertise in teaching: Case studies of second language teachers*. Cambridge University Press.

Wilkins, D. A. (1976). *Notional syllabuses*. Oxford University Press.

Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford University Press.

Wilson, V. (1997). Focus groups: A useful qualitative method for educational research? *British Educational Research Journal, 23*(2), 209-224. https://doi.org/10.1080/0141192970230207

Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Sage.

Appendix

Protocol for stimulated recall interviews

This protocol is adapted from a sample used by Mackey, Gass & McDonough, as attached in Gass and Mackey (2000) for task-based interaction. The protocol is adapted for two purposes: to stimulate participants to recall what they were thinking in their lesson actions and to reflect on aspects of their lesson as related to conditions for effective second language learning.

Instructions

1. Engage in some chitchat for about 1-2 minutes
2. Give the following directions for the task

What we are going to do now is watch the video. I am interested in what you were thinking at the time you were talking or giving an activity. I can see what you were doing by looking at the video, but I don't know what you were thinking. What I'd like you do is tell me what you were thinking, what was in your mind. I am also interested in what you think or perceive about some aspects of your lesson, or your reflection now.

So I am going to pause the video where I want to have some questions. If you are not sure about my questions, please ask me to clarify. If you want to pause at any time and talk about what you were thinking, please feel free to do so.

Demonstrate stopping the video and asking a question for them.

1. If the participant stops the video, listen to what he or she says.
2. Ask the questions on the next page
3. Focus on each condition first; then ask them to tell what they think about/evaluate their practice or lesson.
4. If their response is that they don't remember, do not pursue this because "fishing" for answers that were not immediately provided increases the likelihood that the answer will be based on what the person thinks now or some other memory or perception.
5. Try not to direct participant responses.
6. Try not to react to responses other than providing back-channeling cues or non-responses: *Oh, Mmh, I see, uh-huh, alright.*

Prompt questions for stimulated recall interviews

These indicative questions are based on the question frame employed by Clark and Peterson (1981). Some of them are taken from the researchers (*). Others are adapted (**) and created, but depending on the specific activities in each lesson, they were adapted.

1. What were you thinking when you gave that task/activity? **
2. What were you aiming at when you gave this task/activity?
3. How was the students' response? **
4. Were you thinking about any other alternative actions or strategies at that time?*
5. What were you thinking about students' interaction here?
6. Do you think students had good interaction? Can you justify that?
7. What would you say about their reactions? Why was it so?
8. What would you say about the activity/task you intended?
9. What is your general comment about this lesson?
10. Do you think you have created optimal opportunities for output and interaction? How?

Ojo, O. K., Ayandele, O., & Egbeleye, S. A. (2020). Euphemisms of Corruption among Students of Higher Institutions in South West Nigeria. *Journal of Language and Education, 6*(1), 72-82. https://doi.org/10.17323/jle.2020.10436

Euphemisms of Corruption among Students of Higher Institutions in South West Nigeria

Olumuyiwa K. Ojo, Olusola Ayandele, Sunday A. Egbeleye
The Polytechnic, Ibadan

Correspondence concerning this article should be addressed to Olumuyiwa K. Ojo, The Polytechnic, Ibadan, Sango-Eleyele Road, Ibadan, 200284, Nigeria. E-mail: olumuyiwak.ojo@polyibadan.edu.ng

Corruption is a pervasive practice in Nigeria that is commonly associated with government officials who divert public funds for private use, while minimal attention is paid to acts of corruption in the educational sector. This study, which is part of research on how language is used to drive and conceal corruption in Nigeria, aims at revealing some corrupt practices in Nigerian higher institutions that are concealed because of the euphemistic language used by students to describe and help perpetuate corrupt practices in their relationships with academic and non-academic staff of different institutions. Four institutions of higher education in south-west Nigeria were purposively selected and focus group discussions were conducted with 54 conveniently selected students of these institutions to collect qualitative data on the explanation of linguistic codes derived from the first phase of this study. The findings revealed extensive usage of 'runs' as a superordinate code for diverse acts of corruption including: sex for marks, cash for marks, sex/cash for grade alterations, examination malpractice, and the use of fake documents. Parents and guardians need to listen closely to the language of students in higher education for early detection of assimilation and acceptance of corrupt practices as a way of life.

Keywords: runs, euphemism, corruption, academic fraud, coded language

Introduction

Corruption is a global social malady that distorts the established order of society by conferring undeserved benefits on individuals or groups. Transparency International (2007, p.15) defines corruption as "the abuse of entrusted power – by a public official or a businessperson for example – for private gain". It classifies corruption into different categories as grand and petty, based on the amount of money lost and the sector involved. Grand corruption, according to this global watchdog on corruption, refers to corrupt practices perpetrated by high-ranking government officials to alter policies or functioning of the state. Petty corruption refers to the daily abuse of delegated power by low- and mid-level public officials in charge of providing access to basic amenities such as schools, healthcare services, security, and other services to the citizens.

Corruption in higher education refers to a "lack of academic integrity" (Denisova-Schmidt, 2017, p.1). Osipian (2009) defines corruption in education as a system of all informal relations aimed at regulating "unsanctioned access to material and nonmaterial assets." Academic corruption is described as all wrong doings and dishonest acts perpetrated within colleges and higher institutions of learning by institutional administrators, lecturers or teachers, students, and other stakeholders in the academic sector that have negative consequence on the standard of education and the academic goals of the institutions (Adebisi, Adebisi, & Arogundade, 2012). Chapman and Lindner (2016) opined that academic corruption is a threat to the integrity of certificates obtained from the educational sector; this unethical practice is harmful and capable of destroying the very purpose of education. Corruption has escalated in Nigeria and its practices extend to the university system (Okojie, 2012 as cited in Uche, 2014).

The concept of corruption is usually learnt from others through the process of socialization. Learning the culture of corruption cannot be discussed only in terms of language since language is only one of the major instruments

of communicating and transmitting culture; thus the risk of transmitting the culture of corruption via linguistic resources to the younger generation becomes inevitable (Dumbili & Sofadekan, 2016). Language is a human system of communication that uses arbitrary signals, such as speech sounds, gestures, or written symbols. Since it is primarily used in the transfer of meaning (messages) in the process of communication, language cannot be separated from the culture of the people using it, the period of its usage, and the environment of its usage. Bloch and Trager (1942, p. 2) describe language as "a system of arbitrary vocal symbols by means of which a social group cooperates." Thus, the lexicon of a language can give significant insight into the philosophy and values of the people who use the language. It is in this context that the euphemisms employed by students in descriptions of corruption become useful for understanding Nigerian students' attitude towards corruption. Devitt and Sterelny (1987) as cited in Adetugbo (1997, p. 6) define language as "a uniquely powerful communication system that is stimulus-and-medium-independent, abstract, arbitrary and productive." The productive and arbitrary characteristics of language allow users to creatively invent new words to refer to ideas, thoughts, and behaviour. It allows users to invent euphemisms for ideas that others may find offensive.

Euphemism is a figure of speech that cleverly hides the truth of its reference and is designed to avoid confrontation, hurting people's feelings, or as a substitute for profanity (Rittenburg, Gladney, & Stephenson, 2016). They are usually used to make a sensitive, unpleasant, or offensive event, behaviour, or language more acceptable to the listener (Chi & Hao, 2013). Euphemisms are coined to also enable corrupt people to pass across their messages tactfully by obfuscating the real meaning of what is being said and rendering it incomprehensible to non-users of the code so much so that a person who is not familiar with the code will face difficulties grasping what is being said (Meutem Kamtchueng, 2017). In the context of describing or referring to corrupt practices, euphemisms are used to make the behavior more acceptable to the casual listener and the society at large in many countries across the world. Corruption and corrupt acts by public servants are labeled as *Chorizo,* 'sausage' in Costa Rica and across Central America; *Coima*, 'gifts', in Argentina; *Fakelaki*, 'envelope', in Greece; and *Kula*, 'eating', in Tanzania; while euphemisms for a bribe in Italy is *Bustarella*, 'little envelope'; *Hongbao, 'envelope',* in China; *Mordida*, 'to bite', in Mexico; and *Magharich', 'gift-giving',* in Armenia as described in the Global Encyclopedia of Informality (Ledeneva, 2018).

Corruption has been euphemized in the indigenous languages spoken by Nigerians in their daily transactions, the Igbos call it *Igbuozu,* the Yorubas call it *Egunje,* and the Hausas call it *Chuachua*[1]. The language used in the business of corruption has been a source of interest to different scholars. In Nigeria, Agbedo (2012); Bassey and Bassey (2014), Adegoju and Raheem (2015), Dumbili and Sofadekan (2016) and other researchers have examined how corruption thrives on the wings of some ingeniously crafted words to suit the goals of perpetrators in given settings. Baez-Camargo and Ledeneva (2016) revealed that members of the elite used *stomach infrastructure* to allow ordinary citizens access to scarce basic amenities in order for the ruling class to maintain their grip on power so as to use government funds and facilities to achieve personal goals. Adegoju (2007) identifies the debasement of Nigeria's value system as a reason for the creative use of words when engaging in corruption. Aribisala (2014) pointed to how ex-president Goodluck Jonathan employed euphemisms to argue that "what many Nigerians refer to as corruption is actually stealing, stealing is not the same thing as corruption"[2]. Bangura (2007) uses metaphorical language to discuss corruption in Sierra Leone, while Desta (2016) studies euphemisms for corruption in Ethiopia. Safotso (2015) and Meutem Kamtchueng (2015, 2017) researched the language of corruption in Cameroon.

According to Denisova-Schmidt (2017) corruption in the higher education sector takes various forms, whether monetary or nonmonetary, and might range from bribery and funds misappropriation to fake degrees, plagiarism, ghostwriting, and cheating. Students in Nigeria have conceptualized absenteeism, armed robbery, boycotts of lectures, bribery, cheating, cultism, lecturers dating students, dereliction, destruction of property, educational malpractice, electoral malpractice, embezzlement, examination malpractice, extortion, favouritism, fighting, forgery, fraud, gangsterism, immorality, impersonation, intimidation, lateness, littering, lying, materialism, misappropriation, murder, plagiarism, prostitution, rape, rioting, robbery, the sale of (school) property, seduction, sexual harassment, sorting (inducement), smuggling, stealing, theft, unruly behaviour, vandalism, and victimization as corrupt behaviours (Idoniboye-Obu, 2014). These acts of corruption were covered up by students with the use of euphemisms that may exclude the uninitiated (Ojo, Egbeleye, & Ayandele, 2017).

[1] Ndokwu, J.N (2004, May 26). Corruption, corrupt practices and Honest leadership in Nigeria. http://www.corruption.newscorrupt.org/printer

[2] Aribisala, F. (2014). Nigerian politicians are thieves, but they are not corrupt. *Vanguard Newspaper*, 2014. Retrieved from http://www.vanguardngr.com/2014/ 07/nigerian-politicians-thieves-corrupt/

Runs (used in the plural form), for example, is an euphemism used by students to refer to different subversions of established rules and regulations during a student's stay on a campus of a higher institution in Nigeria, from admission to graduation (Ojo et al., 2017). It is a general term for all corrupt practices done in collusion with other corrupt members of the academic community in which such students operate: non-academic staff, lecturers, business centre operators, students, etc. Although *run* exists as a verb in the English language, its euphemistic usage has a completely different meaning. *Runs* in this sentence for example: "Let me do some *runs* about the coming examination." means to make some illegitimate plans/arrangements that will enable the speaker to pass the coming examination without studying. *Runs* is therefore a euphemism that cleverly hides a plethora of borderline illegal and immoral activities among students at higher institutions in Nigeria. Our search for literature that link *runs* with corrupt practices on Nigerian campuses did not yield much fruit. Scholars like Ojebode, Togunde, and Adelakun (2010) see it as a one-off sexual encounter with an older man; they defined *runs* as an ultra-casual encounter with sex often taking place in a house, hotel room, nightclub, or even in a car, with the partners knowing nothing or little about each other. Ekundayo (2013) discovers the entrenchment of *runs* among educated Nigerians to denote shady deals, socio-economic activities, sexual engagement, etc. done within certain time limit and some level of privacy. A *runs girl* means a girl into sex for quick money while a *runs guy* is a guy involved in illicit activities. Similarly, Okorie and Bamidele (2016) identify a *runs girl* as a euphemism for the term prostitute. Some young people engage in *runs* because they either do not know that it is wrong, refuse to accept it is wrong, or could not understand why it is wrong (Okoli, 1997).

An appropriate theoretical framework for this study is the Sapir-Whorf hypothesis of linguistic relativity. Linguistic relativity as a strand of the Sapir-Whorf Hypothesis proposes that language influences how we think, the way we perceive and remember, and generally, it predisposes us to conceptualize the world in a certain way (Penn, 1972). There is a two-way influence between 'the kind of language we use' and 'the way we see the world'. The hypothesis places emphasis on the social context of language use rather than strictly linguistic considerations, such as the social pressure in particular contexts to use language in one way rather than another (Chandler, 1994, as cited in Adegoju, 2007). Hill and Mannheim (1992) and Thierry (2016) agree that one's language can connect and influence one's thought and one's world view.

The current study collected qualitative data on the variant usage of *runs,* which is the dominant linguistic code for corruption derived from students' responses in a previous study by Ojo et al (2017). This current study is a follow-up to their study, which used a sample of 200 undergraduate students in Nigeria and found most of the respondents to be highly knowledgeable about the various euphemisms used for corruption in Nigeria (Ojo et al., 2017). The respondents identified *runs* as an umbrella term for the daily abuse of entrusted power in the academic environment by low- and mid-level officials (academic and non-academic) in their interactions with students. In this paper, we seek to understand the forms of *runs*, provide additional explanations for the other euphemisms of corrupt practices that *runs* denotes in the Nigerian academic context, and recommend possible ways of curbing this ugly trend.

Method

General Background

A cross-sectional descriptive study was conducted among undergraduates of four purposively selected public higher institutions in two cities in Oyo State, Nigeria: (a) one federal university; (b) one state polytechnic; (c) one federal college of education; and (d) one state college of education. The study respondents were representatives of Nigerian students, particularly those in south-west Nigeria. Locating our study in this part of Nigeria was purely a matter of convenience, rather than an attempt to suggest that *runs* was more prevalent there than elsewhere within or outside Nigeria.

Participants

The respondents were sampled according to gender, level, and type of institution (Appendix A). There were twenty-eight (52%) male students and twenty-six (48%) female students among the respondents. Eight (15%) students were in 100 level/National Diploma 1/Year 1, thirteen (24%) were in 200 level/ National Diploma 2/

Year 2, nineteen (35%) were in 300 level/ Higher National Diploma 1/Year 3, while fourteen (26%) were in 400 level/ Higher National Diploma 2 classes. Fifteen (28%) of the respondents were students in state polytechnic while the federal university had thirteen (24%), the federal college of education thirteen (24%), and the state college of education thirteen (24%) respondents.

Materials

Data collection using the qualitative technique of Focus Group Discussions (FGDs) was employed to elicit reliable responses from the respondents and this took two months (March-April 2018). The focus group discussions was aided by the use of an FGD guide (Appendix B), which contains the topic of discussion and the respondent's sex, level, and name of institution. Some of the questions discussed were: "What is corruption? What is *runs*?" and probing respondents for corrupt practices like bribery, examination malpractices, and other types of *runs* they had experienced in school.

Procedure

Two focus group discussions were conducted in each of the four selected schools, in the evenings, at locations where students usually congregate to relax.

The researchers introduced themselves to the students and explained the purpose of the study. The consent of students was sought prior to their participation and their privacy was guaranteed to the extent that information would never be traced to these respondents. Each FGD session started with a brief introduction of the members and the topic of discussion, followed by ground rules and a discussion of the questions (Appendix B). They lasted a minimum of 50 minutes, with respondents ranging from five to eight (in all, 52 respondents engaged in the FGDs). The FGD guide contained questions relating to the corrupt practices that *runs* is used to represent. Respondents were asked to conceptualize *runs*, mention its euphemistic meanings, and explain how students use those euphemisms when describing corrupt practices on their campuses. Later, the researchers summarized the main points for respondents to make comments on or ask questions about. The researchers applied ethical guidelines as specified by the Department of Psychology, University of Ibadan, Nigeria, obtained consent from the respondents, and ensured that their responses were treated confidentially and used purely for academic research.

Data Analysis

The audio recordings of the FGDs were transcribed, and we employed constant comparison analysis (Leech & Onwuegbuzie, 2008) for coding and grouping the responses by paying attention to terms, contexts, meanings, and examples used by respondents. Responses in Yoruba (a Nigerian indigenous language) and Nigerian Pidgin English are retained in their original forms for authenticity although they were translated to English during data analysis for communicative purposes.

Results

Responses during the focus group discussions showed that corrupt practices described as *runs* on campuses can be categorized into four categories, namely:

1. Trading sex for marks (having sex with or providing other women to have sex with lecturers/school officials)
2. Examination malpractice
3. Bribing school officials with money for marks or changes of grades
4. Using fake documents (results/certificate) for admission/clearance.

The literal and contextual meaning of the euphemisms for each category of *runs* is discussed below.

Trading Sex for Marks

Table 1a reveals the prevalence of sexual immorality or corruption of established standards in the institutions of higher learning. Nine euphemistic terms are used to describe various shades of exchanges of sexual satisfaction for marks or alteration of grades.

Table 1a
Trading Sex for Marks

Euphemism	Literal Meaning	Contextual Meaning
Fifty/fifty	A sharing arrangement where both parties shared the items equally.	An arrangement where a female student engages in sex with a male lecturer for marks.
Praise and Worship	This is a part of Christian liturgy where worshippers praise God with music of different kinds.	A situation where a female student sets out to satisfy a male lecturer in a sexual act.
Trade by barter	A system of trade where an item is exchanged for another item having equal value.	This is used to depict the act of offering sex to a lecturer in exchange for marks.
Settle in kind	This is ordinarily used to mean giving someone a moral support instead of cash /money	A situation where a female student offers her body to a lecturer as settlement to obtain marks.
Farase (Verb)	A Yoruba expression for manually engaging in a task one could have paid others to do.	This is when a female student offers her body to a lecturer to obtain marks.
Control + Shift + Delete	A computer keyboard command to restart a system.	This refers to an operation originating from a student who arranges a female student for a willing lecturer who in turn changes the poor examination mark of the organiser.
Bait (Noun)	Any food item used in catching fish or other animals by attracting them to a hook or net	An arrangement where a student organizes a female student who fits the specifications of a lecturer in exchange for marks.
Swap (Verb)	An exchange of two comparable things	This is when a male student releases his girlfriend to a male lecturer for sexual intercourse in exchange for marks.
Gbomofun (Verb)	This is a Yoruba word which means give the baby to someone	Omo is used here to mean babe/girl/lady. This is a situation where a student arranges a girl for a lecturer in order to get marks.

Examination Malpractice

Table 1b reveals the euphemistic terms used for different acts of examination malpractice. Six such terms were identified and explained in the table. They reveal the extent to which corrupt students will go in order to cheat during examinations. The innovative terms are designed to shift attention away from the negative connotation of fraud to a more linguistically neutral or positive meaning.

Table 1b
Examinational Practices

Euphemism	Literal Meaning	Contextual Meaning
Panpa (Noun)	Yoruba word for a kind of charmed amulet used by traditional African hunters and kept hidden in parts of the body where it could be easily retrieved when needed	An already prepared material consisting of answers to likely examination questions hidden in part of the student's body when going into examination hall
Follow-come (Noun)	A Nigerian Pidgin expression for original components of an item from the manufacturer.	Pages of textbooks or other publications smuggled into examination halls by student
Expo (Noun)	From the root of the word 'exposition'; to unveil an item.	A leaked examination question.
Siwasile (Noun)	A Yoruba compound word expressing a request to be set free.	A coded request to invigilators to accept gratification in exchange for creating a free atmosphere for cheating.
Tattoo (Noun)	Ornamental decoration on human body.	Writing likely examination answers on any part of the body, especially on the lap to be used during the examination.
Formation (Noun)	Arrangement of human beings in a specific order to archive a predetermined objective.	A preplanned sitting arrangement that would allow cooperative exchange of materials or ideas by students during examination.

Bribing School Officials with Money for Marks or Changes of Grades

Table 1c reveals the euphemisms for bribery among students. Six such terms were identified and explained. They refer to offering cash to school officials for marks and the alteration of grades. The euphemisms are suggestive of their meanings.

Table 1c
Bribing School Officials with Money for Marks or Change of Grades

Euphemism	Literal Meaning	Contextual Meaning
Sort (Verb)	To fix a problem or to handle a task.	Giving money to a lecturer to pass an examination or a failed course.
Subscribe (Verb)	To pay for provision of a service, such as internet access.	Money given to a lecturer so as to be among those that will be favoured or get a good grade.
Followup (Noun)	To follow through an action to ensure its success.	Lobbying a lecturer to change marks scored by a candidate
Short-cut (Noun)	A shorter route to get to a destination.	This is used to mean paying a lecturer to escape taking part in assignments and class tests while still passing the course with good grades.
Upgrade (Verb)	To improve the quality or performance of an equipment or machine.	This refers to the payment made by students to data clerks to alter the original CGPA points of students to a better one.
Aja 4 gbera	This is slang used in virtual games to encourage race dogs by patrons. It was popularized in the song of a Nigerian hip-hop artist.	This is when a student pays money to change the level/ status of his or her CGPA.

Using Fake Documents

Table 1d reveals the euphemisms used by students for the use of fake documents. Four such terms were identified. These show the variety of forged documents that students use for gaining admission to institutions and sustaining their studentship.

Comparison of Responses Based on Institution, Gender, Age, and Level

The responses of the students in the federal university show that they were more conversant with euphemisms "*runs*" dealing with examination malpractice and trading sex for marks compared to those from the polytechnic and colleges of education whose students extensively discussed the four categories of "*runs*". Similarly, female respondents spoke more about trading sex for marks and examination malpractice as the types of "*runs*" they were aware of, while the majority of the male respondents were familiar with the four categories of "*runs*". Older students and those in higher classes were more exposed to these euphemisms of corrupt practices than students who were younger and in lower classes.

Table 1d
Using Fake Documents

Euphemism	Literal Meaning	Contextual Meaning
Scam	A fraudulent act or deal	The act of using fraudulent means to get admission or do registration.
Ike	This is a Yoruba world for plastic but often used as slang for credit/debit cards.	This refers to using fake result to process and gain admission into higher institutions.
Fk	This is an abbreviation of the word fake used by certain people engaging in fraud or corruption.	This means using fake documents to gain admission.
Ro'dan	A Yoruba expression which depicts getting something done in a somewhat magical manner.	This refers to using fraudulent or corrupt means to achieve a purpose

Discussion

Trading Sex for Marks

The euphemisms under this category are drawn from various registers like business (fifty/fifty, trade by barter), religion (praise and worship, bait), computer technology (swap, control+shift+delete), Yoruba popular culture (*gbomofun, farase*). They describe various acts of immoral sexual relationship between students and school officials. Nine such terms were identified and the context of usage is explained in Table 1a. Studies have revealed that sex has become a frequent feature of relationships between male lecturers and female students in Nigeria (Fayankinnu, 2004; Ezumoh, 2004). Fifty/fifty, praise and worship, trade by barter, settle in kind and *farase* are *runs* by female students who exchange sex for marks, which support Erinosho's (2004) notion that sex is being traded for marks in some Nigerian higher institutions. Okeke (2011) reveals that some lecturers request sex from students in exchange for marks or make sexual gratification a condition to get good grades. This is illustrated in a recent case involving a professor at Obafemi Awolowo University and a student. The professor allegedly gave the female students two options: have sex with him five times in exchange for upgrading her grades to a pass mark or fail the class[3] . A BBC Africa Eye undercover documentary on "sex for grades" captured four lectures blackmailing female students for sex in exchange for marks and admission at two universities in Nigeria and Ghana[4]. Similarly, Nwaopara, Ifebhor, and Ohiwerei (2008) revealed that sexual harassment of female students by their male lecturers is rampant in different institutions and many lecturers ignore the global condemnation of such practices. Ojebode, Togunde, and Adelakun (2010) revealed that some ladies engage in strategic relationships to further their academic pursuits. Likewise, Okwu (2006) explained that some female students use their sexual prowess to pass their examinations.

Examination Malpractice

Six euphemistic terms were identified in this category. They all relate to corrupt practices by students in the context of examination halls. The terms are drawn from different registers like technology (follow-come, expo), Yoruba popular culture (*siwasile, pampa*), make-up (tattoos), and sport (formation). The purpose of this usage is to deceive invigilators, escape scrutiny, and avoid detection. Copied notes, textbooks, and lecture materials smuggled into examination halls known as "*panpa"*, "follow-come", "expo", and "tattoos" are used as aids during examinations. Unauthorized materials smuggled into examination halls by students at educational institutions in European countries are similarly euphemized as "crib sheets" or "ponies" by English speakers, *spickzettel* by German speakers, and *shpargalka* or *shpora* by Russian speakers (Denisova-Schmidt, 2017). Their use by students demonstrates a lack of preparation, a fear of failure, and the admission of unqualified candidates (Ruwa, 1997, as cited in Kawugana & Woyopwa, 2017). "*Siwasile"* and "formation" describe a strategy for students to sit in positions that will allow weak students to copy from brilliant ones with the connivance of crooked invigilators while avoiding detection by incorruptible proctors. Students in corrupt education systems don't acquire the skills and knowledge that would enable them to impact their society meaningfully. They learn from a young age that a lack of integrity is an acceptable way of life[5].

Bribing Officials with Money for Marks or Change of Grades

Six terms were identified in this category. They describe various acts of corruption perpetuated by students in connivance with academic and non-academic staff. The terms are traceable to the following registers: business (sort, follow-up, subscribe), transportation (shortcut), technology (upgrade), and gambling (*aja 4 gbera*). Sorts, settle, follow up, peace offering, and short-cut represent bribing with money for marks. Okoye (2006) argued some lecturers are engrossed with the businesses of selling marks in their offices. "Upgrade" and "*aja 4 gbera"* point to the involvement of both academic and non-academic teaching staff in charge of results processing. Examination officials engage in the deliberate alteration of a candidate's original marks for a fee (Eze 2012; and Olanipekun, 2013, as cited in Oko & Adie, 2016). Female students without the financial capacity to pay for grades may engage in prostitution to raise funds while larceny and similar vices are perpetuated by male students desperate to raise money to pay for marks (Adewale, 2014; Uzochukwu, 2015, as cited in Oko & Adie, 2016).

[3] Adebayo, B., & Busari, S. (2018, May 23). Lecturer demanded sex in return for better grades, Nigerian student says. *CNN International*+. http://edition.cnn.com/2018/05/23/africa/sex-for-grades-university-nigeria-intl

[4] Mordi, K. (2019, October 7). Sex for Grade: undercover inside Nigerian and Ghanaian universities – BBC African Eye Documentary. Retrieved from https://www.bbc.com/news/av/world-africa-49907376/sex-for-grades-undercovers-in-west-african-universities

[5] Transparency International. (2007). *Report on the Transparency International Global Corruption Barometer 2007*. Transparency International.

Using Fake Documents

Four terms were identified in this category. They represent euphemisms for different types of fake documents. *Ike*, is the euphemism for all fake documents ranging from certificates, receipts for payments, or tuition fees. It is derived from the word for plastic, which could have the extended meaning of an inferior material. *FK* is a phonemic abbreviation of the word "fake". *Ro'dan* is a Yoruba word whose roots can be traced to the acts of deceptive magic. "Scam" is a direct word for fraudulent acts. Some business centre operators, lecturers, and non-teaching staff are involved in the forgery of documents, admissions fraud, and certificate racketeering (Kanno, 2004). Trushin (2015), cited in Denisova-Schmidt (2017), reported the existence of a black market for fake diplomas in Russia. A former vice-chancellor of the University of Port Harcourt disclosed that "many students had been admitted into universities with falsified secondary-school certificates"[6]. Students who used "*Ike*" or "*Fk*", or who are involved in "*Ro'dan*" and "Scam", are building their life on a falsehood that could result in failure in their chosen career if/when they are caught. Tallylo (2006) reported that many students were dismissed from the University of Lagos because of certificate scandals.

Conclusion

Given that a number of illegal behaviors are collectively referred to as corruption, the findings regarding the existence of different euphemisms for *runs* is a new addition to the literature on corruption and euphemisms describing corruption. Understanding other euphemisms for *runs* provides insight into the operations, acts, and motivation for corruption among Nigerian students. This should have implications for the war against corruption, theorizing academic corruption, and intervention strategies. This study is limited by the use of a small number of respondents and the obtained results are only representative of the conveniently sampled individuals in the focus group. They authors recommend that research on academic corruption continue to uncover new forms of corruption and suggest replications with larger samples and with participants from different areas of Nigeria.

Students and parents should clearly define and understand every act of corruption in the academic environment. Higher education institutions should introduce codes of conduct for students and staff, and adopt whistleblowing mechanisms to encourage the reporting of academic corruption without fear of victimization. The Nigerian society should consciously shift its attention away from mere certification to competence in order to encourage students to seek knowledge rather than employ fraudulent means to obtain unmerited certification. Reported cases of corruption should be investigated and the perpetrators prosecuted. The use of *runs* in the educational system is a monster that threatens the state of education and the credibility of certificates awarded in Nigeria.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Adebisi, P. A., Adebisi, O. S., & Arogundade, K. K. (2012). Academic corruption and the challenge of unemployable graduates in Nigeria: Implications for entrepreneursh ip development and economic growth. *Journal of Commerce*, *1*(4), 1 – 12.

Adegoju, A. (2007). Corruption of language and Nigeria's debased value system. *Nebula, 4*(3), 339-356.

Adegoju, A., & Raheem, S. O. (2015). Gone are the days of 'kola(nut)': New trends in language habits and coding of corrupt practices in Nigeria. *Journal of Language and Literature, 26*, 155-171.

Adetugo, A. (1997). *English phonetics: A course text*. University of Lagos Press.

Agbedo, C.U. (2012). 'Verbal carpentry' as language of political corruption in Nigeria. *OSR Journal of Humanities and Social Science, 6*(2), 27-36. http://doi.org/10.9790/0837-0622736

Alutu, O. E., & Alutu, A. N. G. (2003). Examination malpractice among undergraduates in a Nigerian University: Implications for academic advising. *Guidance and Counselling Journal, 18*(4), 149-152.

Baez-Camargo, C., Ledeneva, A. (2016). Where does informality stop and corruption begin? Informal governance

[6] Kigotho, W. (2004, November 1). Nigerian University revokes thousands of diplomas in crackdown on academic fraud. *The Chronicle of Higher Education*. Retrieved from http://www.utexas.edu/conferences/africa/ads/19.html

and public/private crossover in Mexico, Russia and Tanzania. *The Slavonic and East European Review , 95*(1), 49-75. http://doi.org/10.5699/slaveasteurorev2.95.1.0049..

Bangura, A. K. (2007). Understanding and fighting corruption in Sierra Leone: A metaphorical linguistic approach. *Journal of Third World Studies, 24*(1), 59-72.

Bassey, E., & Bassey, V. (2014). Language of corruption and anticorruption in Nigeria. *Journal of Contemporary Research, 11*(2), 45-59.

Bloch, B., & Trager, G.L. (1942). *Outline of linguistic analysis*. Linguistic Society of America.

Chapman, D. W., & Lindner, S. (2016). Degrees of integrity: The threat of corruption in higher education. *Studies in Higher Education, 41*(2), 247–268. https://doi.org/10.1080/03075079.2014.927854

Chi, R., & Hao, U. (2013). Euphemism from sociolinguist tics perspective. *Studies in Sociology of Science*, *4*(4), 45-48. http://dx.doi.org/10.3968/3898

Denisova-Schmidt, E. (2017). The Challenges of Academic Integrity in Higher Education: Current Trends and Prospects. *CIHE Perspectives, 5.* The Boston College Center for International Higher Education.

Desta, H. H. (2016). Euphemism for corruption in Ethiopia. http://aigaforum.com/article2016/Euphemism_for_Corruption_in_Ethiopia_Desta_Haileselassie_Hagos.pdf

Dumbili, E.W., & Sofadekan, A. (2016). "I Collected Money, not a Bribe": Strategic ambiguity and the dynamics of corruption in contemporary Nigeria. *Social Science*, *5*(3), 36-42. http://dx.doi.org/10.3390/socsci5030036

Ekundayo, S.B. (2013). Lexico-semantic 'intraference' in educated Nigerian English (ENE). *International Journal of English Linguistics*, *3*(6), 17-30. http://dx.doi.org/10.5539/ijel.v3n6p17

Erinosho, L. (2004). Challenges in higher education in Nigeria. *The Nigerian Social Scientist, 7*(2), 31-38.

Ezumah, N. (2004). Understanding sexual harassment in institutions of higher learning in Nigeria. *The Nigerian Social Scientist,* 7(2),19-23.

Fayankinnu, E. A. (2004). Ethics and sexual harassment in staff / students relations: Who harasses who? *The Nigerian Social Scientist,* 7(2), 13-18.

Hill, J. H., & Mannheim B. (1992). Language and world view. *Annual Review of Anthropology, 21*, 381-406. http://psycnet.apa.org/doi/10.1146/annurev.an.21.100192.002121

Idoniboye-Obu, S.A. (2014). Corruption in higher education in Nigeria: Prevalence, structures and patterns among students of higher education institutions in Nigeria [Unpublished doctoral dissertation]. University of KwaZulu-Natal.

Kanno, T. N. (2004). Changing values in Nigerian society and curriculum implications. In B. B. Oderinde & O. Ekpo (Eds.), *Values education* (pp. 12-17). Nigerian Academy of Education

Kawugana, A., & Woyopwa, A. K. (2017). Impact of exam malpractice on the quality of graduates in Nigeria. *International Journal of Education and Evaluation*, *3*(6), 45-51.

Ledeneva, A. (Ed.). (2018). *The global encyclopaedia of informality: Understanding social and cultural complexity* (Vol. 2). UCL Press. https://doi.org/10.14324/111.9781787351899

Leech, N. L., & Onwuegbuzie, A. J. (2007). An array of qualitative of qualitative data analysis tools: A call for qualitative data analysis triangulation. *School Psychology Quarterly*, *22*, 557-584.

MeutemKamtchueng, L. M. (2015). C'estçaqueje mange? / Is that what I eat? - Examining the language of corruption in Cameroon. *International Journal of Language Studies, 10*(1), 125-148.

Meutem Kamtchueng, L. M. (2017). Give an envelope for the boss / Give a brown envelope for the boss - On the semantics and characteristics of the metaphors for bribe across cultures: Focus on Cameroon and Nigeria. *Journal of Language & Education*, *3*(1), 45-57. https://doi.org/10.17323/2411-7390-2017-3-1-45-57

Nwaopara, A., Ifebhor, A., & Ohiwerei, F. (2008). Proliferating illiteracy in universities: A Nigerian perspective. *International Journal for Educational Integrity, 4*(1), 31-42.

Ojebode, A., Togunde, D., & Adelakun, A. (2010). Beyond money and gifts: Social capital as motivation for cross-generational dating among tertiary school female students in south west Nigeria. *The International Journal of Interdisciplinary Social Sciences, 4,*169-182. https://doi.org/10.18848/1833-1882/CGP/v05i04/51673

Ojo, O.K., Egbeleye, S.A., & Ayandele, O. (2017). A psycholinguistic analysis of the euphemism of corruption in Nigeria. Paper presented at the 2017 National Conference of the Faculty of Business and Communication Studies, The Polytechnic, Ibadan, Nigeria.

Okeke, C.M.A. (2011). Impact of sexual harassment on women undergraduates' educational experience in Anambra state of Nigeria [Unpublished doctoral dissertation]. Seton Hall University.

Oko, S. U., & Adie, R .I. (2016). Examination malpractice: Causes, effects and possible ways of curbing the menace. A study of Cross River University of Technology. *International Journal of Managerial Studies and Research, 4*(1), 59-65. http://dx.doi.org/10.20431/2349-0349.0401006

Okoli, C.E. (1997). Examination malpractice, the bone of our society. In A. Ejiogu & K. Alani (Eds.), *Emergent*

Issues in Nigeria Education (pp. 343-348). Unilag Consult Publication.

Okorie, M. M., & Bamidele, O. (2016). Language and class resistance in Nigeria: A foucauldian perspective. *Africology: The Journal of Pan African Studies, 9*(6), 4 –15.

Okoye, R. O. (2006). *Special examination centers: Their activities and measures for controlling them*. NISEP publications.

Okwu, O. J. (2006). A critique of students' vices and the effect on quality of graduates of Nigeria tertiary institutions. *Journal of Social Science, 12(3),* 193-198.

Penn, J. M. (1972). *Linguistic relativity versus innate ideas: The origins of the Sapir Whorf hypothesis in German thought*. Mouton.

Rittenburg, T. L., Gladney, G. A., & Stephenson, T. (2016).The effects of euphemism usage in business contexts. *Journal of Business Ethics*, *137*(2), 315–320. https://doi.org/10.1007/s10551-014-2501-4

Safotso, T. G. (2015). The metalanguage of corruption in Cameroon-part I: The registers of general administration, transport and education. *International Journal of English Linguistics, 5*(2), 47-54.

Tallylo, N. L. (2006). *Ghost voice malpractice.* Third Eye Publication.

Uche, R. D. (2014). Students' perception of corrupt practices among Nigerian universities lecturers. *American Journal of Humanities and Social Sciences*, *2*(1), 66-69. https://doi.org/10.11634/232907811402511

Appendix A

100 level refers the first year of undergraduate studies in the university. 200 level refers to the second year of undergraduate studies in the university. 300 level refers to the third year of undergraduate studies in the university. 400 level refers to the fourth year of undergraduate studies in the university. Universities award Bachelor degrees after four to six years of post-secondary studies.

National Diploma 1 (ND I) refers to the first year of national diploma studies in the polytechnic. National Diploma 2 "(ND II) refers to the second year of national diploma studies in the polytechnic. Higher National Diploma 1 (HND I) refers to the first year of higher national diploma studies in the polytechnic. Higher National Diploma 2 (HND II) refers to the second year of higher national diploma studies in the polytechnic. Polytechnics award National Diploma after two years of post-secondary studies and Higher National Diploma is awarded after two years of post-national diploma studies.

National Certificate of Education (NCE) Year 1 refers to the first year of national certificate of education studies in the college of education. National Certificate of Education (NCE) Year 2 refers to the second year of national certificate of education studies in the college of education. National Certificate of Education (NCE) Year 3 refers to the third year of national certificate of education studies in the college of education. Colleges of education train teachers for primary and secondary schools and award National Certificate of Education after three years of post-secondary studies.

Federal institutions of learning are owned by the Federal Government of Nigeria while state institutions are owned by state governments.

Appendix B

Focus Group Discussion (FGD) Guide

Introduction: (Name of Moderator, note taker, topic and modality of conducting the FDGs)

Sex: Male [ ] Female [ ]

Level:_____

Name of Institution:_____

Topic: Euphemisms of Corruption among Students of Higher Institutions in Southwest Nigeria

Questions:

1. What is corruption?
2. What is *runs*?
3. Probe for types of *runs*
4. Probe for names (euphemisms) given to *runs* by students
5. Probe for *runs* relating to sexual activities for marks
6. Probe what *runs* is used for bribing officials for marks
7. Probe for *runs* that takes place before, during and after examinations
8. Probe for *runs* during the admission, clearance and graduation process
9. Ask respondents any other question as you round off discussion.

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 1, 2020

Saghafi, K. , & Elahi Shirvan, M.E. (2020). Rapid Changes in Foreign Language Learning Anxiety Caused by a Multiplicity of Topics: An Idiodynamic Approach. *Journal of Language and Education, 6*(1), 83-102. https://doi.org/10.17323/jle.2020.9684

Rapid Changes in Foreign Language Learning Anxiety Caused by a Multiplicity of Topics: An Idiodynamic Approach

Khatereh Saghafi[1], Majid Elahi Shirvan[2]

[1] Hakim Sabzevari University
[2] University of Bojnord

Correspondence concerning this article should be addressed to Majid Elahi Shirvan, Foreign Languages Department, University of Bojnord, Bojnord, Iran. E-mail: elahishmajid@gmail.com

Considering the dynamicnature of foreign language anxiety (FLA), we applied an idiodynamic method to explore topic-based variations of FLA. Before the study was conducted, a class of 20 female intermediate English as foreign language learners were assessed using the foreign language classroom anxiety scale (FLCAS). Two low-anxiety learners and two high-anxiety learners were selected to participate in this study. The idiodynamic method involved videotaping the participants' responses to four topic-based questions, their self-ratings of fluctuations in FLA while answering the questions, and drawing attributions for topic-based changes in FLA. The results demonstrated both within-individual and between-individual stability and variation in FLA. Linguistic block, topic familiarity, topic interest, and topic-related emotional loading were revealed as the major factors affecting the dynamics of FLA. The pedagogical implications of the findings are discussed.

Keywords: foreign language learning anxiety, idiodynamic method, change, topic, complex system

Introduction

Foreign language anxiety (FLA) has been a highly investigated construct in the field of second language acquisition (SLA). The bulk of the studies addressing this construct have mainly provided us with the underlying factors as well as its association with other language-related constructs from a linear perspective. However, the emergence of complex dynamic systems theory has urged researchers to explore FLA from a complex dynamic perspective. This shift in perspective towards research on FLA has pivoted on the issue of the ergodicity problem in individual differences in SLA (see Lowie & Verspoor, 2019). That is, the average FLA level of a sample of language learners cannot represent the FLA trajectory of each individual in that sample. Considering this shift, in this study we aimed to explore the dynamics of FLA under the influence of conversational topics via an idiodynamic approach.

Literature Review

The literature on FLA indicates three main approaches to research on FLA in terms of its underlying assumptions and historical trends. These approaches are confounded, specialized, and dynamic approaches (MacIntyre, 2017). The confounded approach encompasses ideas regarding language anxiety as well as its influence on learning a foreign or second language from an amalgamation of numerous resources, which are not necessarily associated with language learning. The specialized approach explores and identifies experiences of anxiety that are specifically related to the process of language learning. The dynamic approach regards anxiety experiences in second language learning as situated complex processes in association with other contextual experiences under the influence of a network of ecological factors (Gkonou, Daubney, & Dewaele, 2017).

The main challenge of the research on the confounded approach (e.g. Chastain, 1975; Kleinmann, 1977; Spielberger, 1966) was related to the scales for measuring anxiety in the language learning process, derived from general psychology, which were rarely in connection with the language itself (MacIntyre, 2017).

The literature of the studies applying the specialized approach can be categorized into the following aspects. First, some studies have dealt with the nature of FLA (e.g. Aida, 1994; Cohen & Norst, 1989; Dewaele, Petrides, & Furnham, 2008; Gardner & MacIntyre, 1993;Horwitz, 2001;Horwitz, Horwitz, & Cope, 1986; Horwitz & Young (eds.), 1991; ; MacIntyre, 1995; MacIntyre, 2007; Sparks, Ganschow & Javorsky,2000; Tran, Baldauf Jr, & Moni, 2013).

Second, a series of studies have addressed the influence of FLA on learners' achievement (e.g. Aida, 1994; Chastain, 1975; Dulay & Burt, 1977; Elkhafaifi, 2005; Gardner & MacIntyre, 1993;Horwitz, 2001; Horwitz, 1986; Huang, & Hung, 2013; Kleinmann, 1977;MacIntyre, 1995; MacIntyre, 2007; MacIntyre & Gardner, 1994; Saito & Samimy, 1996; Sparks & Ganschow, 1991). Third, some studies have encompassed the sources of FLA and the influence of socio-cultural or instructional conditions on FLA (Bailey, 1983; Dewaele, Petrides & Furnham, 2008;Gregersen & Horwitz, 2002; Horwitz, 2001;Horwitz & Young (eds.), 1991; Koch & Terrell, 1991; Young, 1990).

Furthermore, numerous studies have considered the relationship of FLA with other learner-related factors (e.g. Bailey, 1983; Dewaele, 2019; Dewaele, Petrides, & Furnham, 2008; Elahi Shirvan, Khajavy, Nazifi, & Taherian, 2018; Elahi Shirvan, Khajavy, MacIntyre, & Taherian, 2019;Gregersen & Horwitz, 2002; Horwitz, 1996; MacIntyre, 1995; Matsuda & Gobel, 2004;Sparks & Ganschow, 1991; Saito & Samimy, 1996).MacIntyre, Dörnyei, Cl´ement, and Noels (1998) reported a negative association between L2 willingness to communicate (WTC) and FLA. Examining this correlation among Chinese second language learners of English, Liu and Jackson (2008) also found a significantly negative association between their WTC and FLA. Considering this finding, they suggested that both WTC and FLA should be regarded as complementary constructs to understand learners' emotional feedback to their second language acquisition. This strong association between the two variables has been also confirmed by a recent meta-analysis of L2 WTC and its high evidence correlates (Elahi Shirvan et al., 2019) and a regression model of independent variables influencing L2 WTC (Dewaele, 2019).

Recently, given the tripartite of cognition-affect-motivation, using latent growth curve modeling, Elahi Shirvan et al. (2018) found an increase in the strength of association between adult second language learners' FLA and their self-efficacy during a university general English course. They also found that the long-term link between the two variables is directly associated with the learners' ideal self-motivation in the first session of the course.

Regarding the contribution of second language learners' age to their feeling of anxiety, Dewaele, Petrides, and Furnham (2008) concluded that by starting to learn a second language at younger age, learners experience lower levels of anxiety. They also reported that second language learners with high levels of trait emotional intelligence experienced lower levels of anxiety. Furthermore, with respect to the social factors involving second language learners' feeling of FLA, Spitalli (2002) reported a significantly negative correlation between second language learners' attitudes toward the second language culture and their levels of FLA. Moreover, Matsuda and Gobel (2004) reported that learners with overseas experience had lower levels of anxiety than those without it.

Finally, a few studies have focused on skill-specific (listening, writing, and reading) FLA (Hilleson, 1996; Horwitz, 2001; Elkhafaifi, 2005; Oh, 1992). For instance, examining learners' anxiety in reading tasks, Oh (1992)reported that second language learners' anxiety in these tasks can be influenced by the familiarity of the learners with the tasks, the difficulty of the texts, and the learners' perceptions of the validity of the tasks. Exploring the experiences of second language learners' anxiety in language specific skills, Hilleson (1996) also found anxieties in relation to writing and reading skills. Likewise, introducing the concept of foreign language reading anxiety (FLRA), a scale for its measurement called the foreign language reading anxiety scale (FLRAS) was developed by Saito et al. (1999).They regarded FLRA as quite distinct from FLA. Similarly, learners' anxiety in reaction to reading skills was discussed by Horwitz (2001). Furthermore, foreign language listening anxiety (FLLA) and FLA were described as quite distinct constructs by Elkhafaifi (2005), exploring second language learners' anxiety in listening comprehension as well as their FLA in a critical language.

The dynamic approach to FLA suggests that anxiety should be regarded as an emotion that is in a fluctuating state over time, even on small time scales like minutes and seconds, and that the (intra- or interpersonal)

stability or fluctuation of this emotion is susceptible to contextual factors (Gkonou, 2017; Saghafi, Adel, & Zareian, 2017). This is due to the fact that foreign language learners are not ergodic ensembles (see Lowie & Verspoor, 2019); that is, the patterns of anxiety experienced by each individual learner is different from those experienced by the average learners in the class. Considering this approach, we can conjecture that although the previous findings of FLA research, adopting the specialized approach, have verified traces of anxiety in students' language learning, they have not kept the multi-faceted nature of FLA in perspective. Thus, a dynamic state-oriented approach should be employed to explore various processes contributing to changes in FLA, compensating for the previous trait-oriented ones. One of the recent approaches to explore the dynamics of individual learners' FLA experiences has been the idiodynamic approach.

The Idiodynamic Method

The idiodynamic method can be applied to study affective and cognitive states of human communication (MacIntyre, 2012). Individual characteristics rapidly deal with situational features to regulate communication behavior (MacIntyre, 2012). The affective and cognitive contexts accompanying communication behavior are being continuously shaped, modified, and reinforced simultaneously in a constant ebb-and-flow mood (Fogel, 2006). The lack of attention on the contextual factors has rendered the modus operandi of common nomothetic methodologies unsuitable for exploring the dynamics of change (MacIntyre, 2012). Regarding the dynamic nature of FLA, the network of subtle changes inducing the domino effect on FLA have been poorly addressed and, consequently, traces of influential features influencing the complexity of FLA have either been taken for granted or escaped the eyes of research on anxiety (MacIntyre, 2012). With the affective dimensions of communication in the spotlight, the idiodynamic method and its microscopic scope, focusing on the interaction between the situation and the person, set the stage for the dynamics of change in the variables under study (MacIntyre, 2012).

The term idiodynamic refers to the passage of changes enacted by the environment within an individual in a given situation "as an event unfolds" (MacIntyre, 2012, p. 362) contrasting ''idiographic'' approach to personality by Allport (Rosenzweig & Fisher, 1997). Opposing an approach scanning individual-level traits (idiographic) or group-level traits (nomothetic) or, the idiodynamic method maps out individuals' behavior during an episode of one's life (MacIntyre, 2012). In this regard, the idiodynamic method encompasses an oral interview or a conversation which takes a short period of time in terms of minutes or seconds, called a communication episode (MacIntyre, 2012).

This method employs a process-oriented individual-formative approach of exploration. In this method a participant's communication, which is video recorded, is displayed right after it is completed by asking the participant to rate him/herself or another participant on a cognitive or affective variable under study using a software program by which a rating graph is provided right away and applied to the interview to justify the participant's thoughts and feelings underlying the ratings regarding the variable being studied (MacIntyre, 2012). The video is documented for further analysis. The researcher and the respondent can pause or play the video back in case of any valley or peak observed in the graph or the need for further explanation during the interview. Afterwards, the respondents' justifications are recorded (MacIntyre, 2012). In addition, the idiodynamic method can be employed in experimental studies manipulating contextual features. The dynamic systems theory (DST), addressing communication as being an apex of a net of interrelated fluid systems, theoretically underpins the idiodynamic method (MacIntyre, 2012).

Research on the DST of human development supports the application of the idiodynamic method to SLA via the conceptualization of some key terms (Howe & Lewis, 2005; van Geert, 2011). Four main properties of dynamic systems were proposed by deBot, Lowie, and Verspoor (2007):

- *Each state feeds into its previous state and grows up to another state under the light of some influences (Howe & Lewis, 2005, especially pp. 248–249).*
- *Dynamic systems are chained up each-to-all and all-to-each, falling under each other's influence.*
- *Dynamic systems co-ordinate to converge into preferred/attractor states diverging from unsetteled/repeller or shifting states.*
- *Even the most minor changes in one part of the system might have a far-reaching revolutionary effect in the whole system calling attention to the theory of the butterfly effect. On the other hand, it is also likely that a significant change causes no big change but a cosmetic or small one in the system.*

Furthermore, communication behavior, cognition, and affect can be viewed from the standpoint of the DST. Different systems inter-connectedly affect one another either to serve or strangle communicative goals and intentions. The following list refers to some examples of these systems:

- *Core physiology*
- *Affective states*
- *Cognitive events*
- *The local social system*
- *The global pattern of intergroup relations*

The term "attractor state" demonstrates a state in which the system tends to be settled (de Bot e t al., 2007). For instance, a system is in the attractor state if two positive contextual characteristics correlate positively. Taking the DST into perspective, we can trace minor and major fluctuations in one section of the system that directs the rest of the system for better or worse. The more the rising and falling of the two states correlate, the more distinctly they weave into particular communicational patterns. Traditional approaches holding a trait-oriented view towards psychological constructs like FLA have not been able to explain the dynamic nature of these constructs (Parlade´ & Iverson, 2011; Thelen & Smith, 1994). Contrasting traditional research approaches, the idiodynamic method, holding a sensitive perspective to the initial conditions, unfolds communication processes via a state-oriented view towards psychological constructs, which have been neglected by traditional approaches to the trait research (de Bot et al., 2007). Primary conditions, susceptible to any sort of change, trigger a wave of conditional violations throughout the whole system, which make initial predictions crumple even in cases of in-advance familiarity with the component parts of a given system.

Current Research Examples Using the Idiodynamic Method

The first study that introduced and utilized idiodynamic method was released in 2011 by MacIntyre and Legatto investigating L2 WTC. Outcomes of the study took fairly divergent paths, both within the individual on a time continuum and over the different questions of the interview. Thus, both consistency and variation among a small group of homogeneous participants were reported. This study also confirmed that linguistic block is the most critical process affecting WTC. As MacIntyre and Legatto (2011) asserted, WTC is a representation of the dynamically fluctuating nature of communication system.

Inspired by MacIntyre and Legatto (2011), in research on the dynamics of FLA, given the influence of the different contextual factors, researchers have recently applied the idiodynamic approach. For example, Gregersen, Meza, and MacIntyre (2014) studied learners' FLA at an individual level, its dynamicity, and elucidation of frequent changes in the participants' affective reactions. The study provided evidence of the fact that experiences of anxiety in foreign language learners should be studied at the individual level. Moreover, considering individual experiences of anxiety, Elahi Shirvan and Talebzadeh (2017), via an idiodynamic approach, concluded that the status of the interlocutors in conversations and the familiarity level of the learners in their study with these interlocutors contributed to the fluctuations of FLA in classroom conversations. Furthermore, Gregersen, MacIntyre, and Olson (2017), via the application of the idiodynamic method, reported that learners' anxiety of FLA might be partially transparent to their peers and teachers in the classroom. Inspired by these recent studies using the idiodynamic approach to the investigation of the individual learners' patterns of anxiety, in this study we intended to explore the influence of conversational topics on the emergence of dynamic patterns of anxiety in individual learners of English as foreign language.

The Present Study

Given the lack of studies exploring the dynamics of EFL learners' anxiety under the influence of conversational topics through the lens of dynamic systems theory, the present study takes a qualitative approach, drawing on the idiodynamic method to explore topic-related fluctuations in FLA over four topic-based questions. In the present study we examine the following research questions:

1. To what extent do we see within-individual variations in FLA over time under the influence of different conversational topics?
2. To what extent do we see between-individual variations in FLA over time under the influence of different conversational topics?
3. What attributions do the individual participants make for rises and falls in FLA?

Method

Participants

Using maximum variation sampling, we selected two high-anxiety and two low-anxiety Iranian female students aged 15-17 from an intermediate English course with 20 students at a private institute in Mashhad, Iran. To measure the FLCA of the participants, we applied the foreign language classroom anxiety scale (FLCAS) (see Table 1).

Table 1
The Participants' Profiles

Pseudonym	Gender	Age	Proficiency level	Anxiety mean*
Ziba	Female	16	intermediate	50
Melika	Female	15	intermediate	90
Zahra	Female	16	intermediate	80
Azar	Female	17	intermediate	30

Instruments and Materials

Persian Version of the Foreign Language Classroom Anxiety Scale

The Persian version of FLCAS (Elahi Shirvan, Khajavi, & Taherian, 2016) was used to measure the participants' level of FLA prior to data collection. The psychometric properties of this scale were measured via Rasch analysis. The item reliability of the scale was .93 and its person reliability was.95. The scale included 26 items of the 33 items of the original FLCAS (Horwitz et al., 1986) because seven items misfit the unidimensionality hypothesis, underlying the Rasch model, within the context of Iran.

EFL Topic-Based Questions

Four oral tasks were undertaken in the following order using the "anecdote" tasks from the book of American Inside Out, intermediate level:(T1) Describe something that has special significance for you; (T2) Describe a movie that has disappointed you; (T3) Describe a party you did not enjoy; (T4) Describe your favorite celebrity. The respondents were also provided with leading hints on how to give responses to the questions with accurate information, with no sense of confusion. Before the interview was carried out, the participants were given a short period of time, two minutes, in order to comprehend the topic, review the hints, and dispel confusion over the task instruction. All four questions were at the same difficulty level. As soon as they felt ready, the interviews were held and video recorded. Moreover, the participants were provided with regular constructive verbal and non-verbal feedback during the conversations.

Video Recording Camera

Using a mobile camera, we video recorded each interview and fed them through an anxiety rating software program (Anion Variable Tester, V2) right after each interview was conducted.

Anxiety Rating Software

For the idiodynamic rating of the participants' anxiety, we used Anion Variable Tester, version 2, (MacIntyre, 2012). As soon as the video-taped interview of each topic discussed was entered into the software program, each participant watching the video recording of her interview, rated her anxiety level on a clear graph with a vertical axis indicating an anxiety level ranging from -5 to +5 and a horizontal axis indicating time in terms of seconds. Windows-based software was developed to plot a graph of the participant's ratings on her anxiety level while displaying her recorded interview and keeping a record of her anxiety level ratings in two different windows simultaneously on the same screen. Respondents watching the video clicked on the computer mouse and rated their anxiety level on a scale of -5 to +5. It should be noted that in case no anxiety level rating was assigned, the software assigned a level of zero for the explored construct. Software outcomes were released in

the form of bitmapped graphs of dynamic anxiety ratings and an Excel spreadsheet, which were compatible with time and anxiety data.

Graph of Changes in FLA

A printed graph of dynamic anxiety ratings was provided for each participant immediately after each interview was conducted and rated via the anxiety rating software program. The second researcher in this study played the video-recorded interview once more and paused it each second the graph was shown to the respondents to explore the explanations of the participants regarding the dynamics of their anxiety as presented in their graph of FLA.

Observer's Ratings

Plotting the respondents' affective feedback, the first researcher in this study, viewing the video recordings, drew up a narrative report.

Procedure

This study was conducted in five phases. Before the study was conducted, the Persian version of the FLCAS was administered to an intermediate English class of 20 students out of which the participants of the study were selected (phase1).This questionnaire including 26 items assessed EFL learners' anxiety level taking the course on a Likert-type scale of five ranging from strongly agree to strongly disagree. After this phase, concerning the participants' anxiety level, four students, two of high and two of low anxiety level, were selected, suiting the purpose of the study. The participants took four topic-based questions and were videotaped at the same time (phase 2). Afterwards, they answered each question, watched their recorded video, and self-rated their anxiety level (phase3). The second researcher in this study documented his understanding of the respondents' verbal and non-verbal reactions on a piece of paper (phase 4, which will be presented in the results section). Later, as soon as the participants answered all the questions and self-rated their anxiety levels, they justified their anxiety self-ratings (phase 5, which will be presented in the results section and the appendix) that were analyzed after being recorded, translated, and transcribed.

Results

The results of the fluctuations in anxiety in response to the four topics are presented for each participant separately. First, the observation of each participant's dynamics of anxiety while answering the questions is reported, and second, the findings of the interviews with each participant in terms of her explanations for her anxiety levels in response to the questions are presented.

Observation of Ziba's Anxiety Dynamics

Topic 1-4

Ziba was one of the low-anxiety participants. Going through verbal and non-verbal reactions to different topic-based questions, she smoothly vaulted into her first speech. Looking at the tranquility in her eyes, we noticed that she truly channeled the way she felt about her necklace to us attentively. After her first sentence, she uttered "a very special kind of stone" she sat up, narrowed her eyes and peacefully reflected in order to recall the name of stone and suddenly, her glittering eyes flying open, came into her own just in time to recalling name. As she started revealing the story behind the necklace, dredging up the memory of her brother leaving her family to head to university, she blushed She went red in the face, rubbed her hand together, and she was close to tears every time she discussed this memory. Although she was emotional, she never lost her string of thought or failed to deliver her intended meaning in English. It seemed that words and sentences with degrees of emotion such as "memories", "necklace", "my brother", "him", and "sadness" were challenging to discuss.

[1]The transcript of Ziba's first conversation is as follows. It should be noted that anxiety ratings appear above the text ranging from -5 to 5 and the number on each word indicates her clicks representing her level of anxiety):

I[0] have[0] a necklace[0]. It's[0] green[0]. It's[0] made[0] of very[1] special[1] kind[1] of stone[1]. It's[0] called[0] Antovenie[0]. When[0] my brother[1] wanted[0] to leave[1] home[1] and go[0] to university[-1] in another[1] city[0], he[0] gave[0] me[0] that[0] necklace[0] to[0] remind[1] me of him[1] and he[1] told[0] me[0] it[2] makes[2] you {me} [2] happy[2] because[0] it's[0] a[0] special[0] stone[0] and every[0] time[0] you're {I'm} [0] sad[0] I just[0] look[0] at[0] this[0] stone[0] and remember[1] our[1] happy[1] memories[1] together[1]. So, anytime[0] I'm[0] sad[1] I[0] look[0] at that[0] necklace[0] and after[0] 2 minutes[0] I[0] forget[-1] my[0] sadness[2]. I[0] always[-1] carry[-2] it[-3] with myself[-4] everywhere[-5]. If[0] I[0] die[2], I[1] would[1] return[2] it[2] my[2] brother[2].

Answering the second question, she made up her mind quickly and announced that she was ready. She started answering the second question with no signs of anxiety. As before, the words' emotional load caused a high level of anxiety. While narrating her story, she suddenly became angry. As she continued her story, she started rubbing her face and nodding her head and kept digging her nails into her lap, which indicated a high level of anxiety. Next, as she was responding to question 3, she became calmer, signaling lower levels of anxiety. But from the moment she began recounting a horrible incident that took place after a party, a sense of anxiety came over her again. Furthermore, Ziba's speech on her favorite actor stood out among her answers as she fluently and accurately started off her speech and remained at ease, which continued until her last word.

Ziba's Explanations of Her Anxiety Dynamics

During the discussion of dynamic anxiety ratings, Ziba explained that the first question haunted her and jogged her memory of her brother leaving her. As we put our question to her, she spoke honestly about her attachment to her brother and her intense grief caused by his not being around anymore. She concluded that her anxiety was allayed during her speech under the influence of the words' and sentences' emotional load on her. She also explained that she could easily overcome her anxiety. Moreover, she added that a high proficiency level and high self-confidence could make up for her high anxiety level to allow her to deliver a flawless speech. Moreover, Ziba explained that her anxiety during her second speech was increased by the heartrending story of the movie that she reflected on and she did not experience a linguistic block on the words 'rape' and 'abuse'. She decided that if she could not make herself understood one way, she would either find another way to explain or describe the words she did not know.

Discussing the second topic, Ziba stated that thinking back on the moment of the incident exacerbated her anxiety as it emotionally overloaded her, yet it did not distract her from focusing on her English wording and structure. She claimed stressing over using the right collocations had never occupied her mind as she believed everyone had to take the risk of committing mistakes when applying English in communication. On topic 4, she explained that her high-anxiety moments fell back on her personality trait of jealousy when she said, "He has gray eyes and blonde hair" as well as for the word "fan". She also discussed that while answering the fourth question, she had experienced less anxiety as she had built more confidence over the first three questions and she had gotten used to the commonalities of the practice (see Figure 1).

Observation of Melika's anxiety dynamics

Topic 1-4.Phase 4

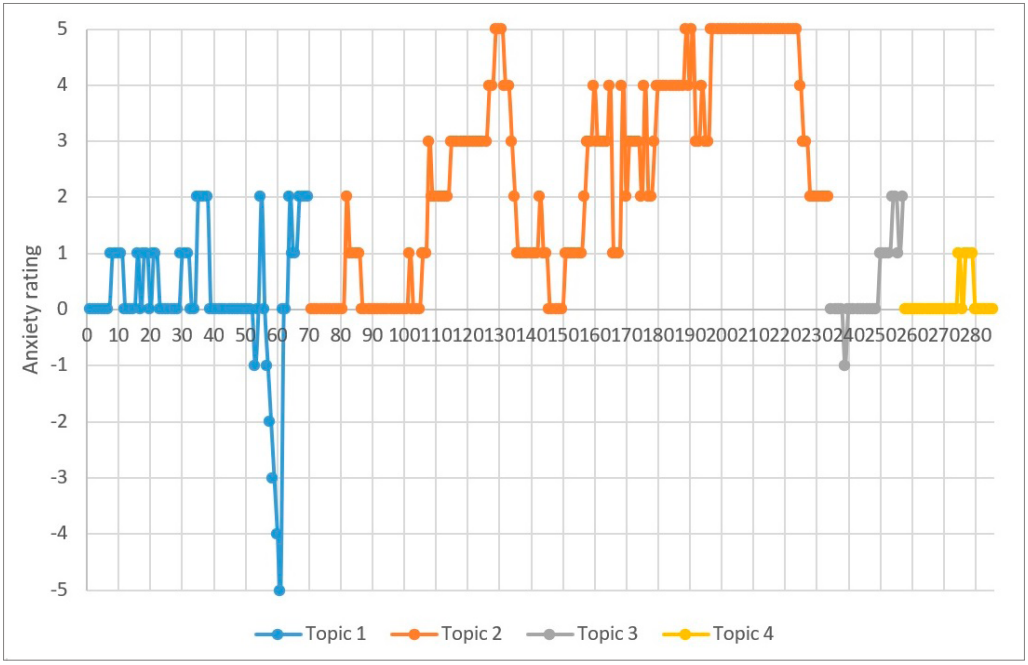
Melika is one of the high-anxiety students. We observed that she went through moments of high anxiety as she was answering question 1 and 2since she was gasping for breath and her face turned pale. She was also observed taking many awkward pauses, some momentary and some pretty lengthy, and he was observed lifting, squeezing, shutting, wiping, and averting her eyes from us during the first two questions. Even after her code-switching, asking for the English synonym of a Persian word from us, her anxiety did not fall because, as we guessed, she could not use the word "remind" in the right form.

Despite her moments of high anxiety while answering the first two questions, she started answering the third one

[1] The self-rating transcripts of the first participant for the first topic are presented in the article but to save space, Ziba's other self-rated transcripts for topic 2-4 and other participants' self-rated transcripts are presented in the appendix. A table of participants' anxiety provoker, reducer, and flat stimuli is also provided in the appendix.

with no sign of anxiety. But she looked anxious after that. Uttering the word "player", she gave us a questioning look at us and produced the made-up word "sporter". As soon as we reminded her of the word "player", she came to feel at ease for a short period in time. With she struggled talking about her favorite volleyball player, she seemed to suffer from a high level of anxiety. She delivered her fifth speech on and off, while she was highly anxious. She seemed puzzled about the topic as she did not follow a structured speech plan.

Figure 1
Ziba's Anxiety Rating



OMelika's explanations of her anxiety dynamics

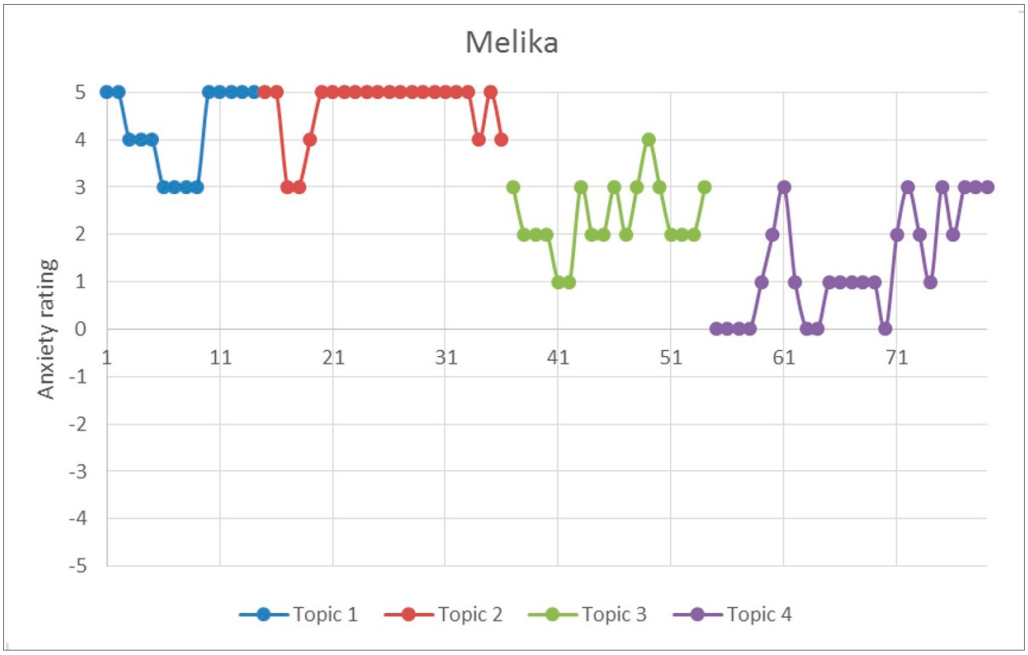
Melika explained that linguistic block was mostly the main hindrance to her affective self-control. Besides, her fear of being recorded made her go blank, stumbling over her English. She also reported difficulty making up her mind around the topic, which reduced her peace of mind as she was given two minutes to think about her answer. This time seemed too short for her to come up with a flash of inspiration; thus, she had no time left to plan her answer in terms of using the right words and structures. Following her unsettled state of mind over the topic, she claimed that time dragged while she was answering the question. On topic 2, Melika explained that she could not provide much information on the topic since she usually did not watch much TV and she was not also a cinema fan; thus, she could not think of a movie she knew everything about. Moreover, she explained that her confusion regarding the topic was accompanied with confusion over the English words and structure.

Furthermore, she noted that her education background was not sufficient to help her with a suitable response to the question. She also reported that she was afraid she would lose face if she revealed her difficulties speaking English. As a result, she was deep in her own thoughts, which resulted in either pointless or poor speech. She also declared that she was never given the chance to solve her problems when speaking. Another point she addressed was that her dazzling performance in class achievement tests was a result of a number of multiple-choice and true-false questions regarding vocabulary, grammar a by which they were graded, excluding their all-inclusive performance with respect to the absence of speaking and listening assessment as well as reading and writing in a productive way. She discussed that her good grades in such unreliable achievement tests always seemed satisfying to her and her family since it was enough to pass the entrance exam. Furthermore, Melika explained that after she was recorded twice, she did not feel uncomfortable being recorded the third time. But she claimed that she had difficulty describing her favorite volleyball player because she had never been interested in learning about famous people's lives. Therefore, it was difficult to organize her speech.

She also explained that under the influence of this lack of interest in the topic she went through a high anxiety level. Deep in contemplation, she explained that she could have been able to have a better performance on

the questions if she had been given more time to utilize the hints and plot her speech in her head. Thus, she asked us to give her a longer period of time, about an hour, to be interviewed again. Her request was rejected as it would have put the research authenticity of our study into question. Furthermore, she noted that her poor performance and her high level of anxiety worked cyclically as she used "is" instead of "was" and committed the same mistake afterward because her anxiety was aroused by her first mistake and reinforced by its repetition (see Figure 2).

Figure 2
Melika's Anxiety Rating.



Observation of Zahra's anxiety dynamics

Topic 1-4. Phase 4

Zahra is one of the two high-anxiety students. She began answering the first question rubbing her left hand against the other, avoiding eye-contact with us, breathing hard, biting her nails, staring at the wall, and rolling her eyes. She also used a code-switching strategy to compensate for her linguistic block. Moreover, she looked to be under a great level of anxiety while pronouncing the word "turquoise", but her anxiety dropped after we repeated the right pronunciation of the word. To our surprise, she started answering the second question with a peaceful look on her face. Although she took lots of pauses, she seemed to enjoy being engaged in the topic. Despite her being in a stress-free state answering the third question, she looked puzzled while answering. As a result, she took lots of short pauses. However, she seemed at ease only after we helped her with the word "spot". She started answering the fourth question with her face swelling with interest in the topic. She spoke simple English and used short sentences. She also took short pauses to reflect on the information she had heard or read about her favorite celebrity.

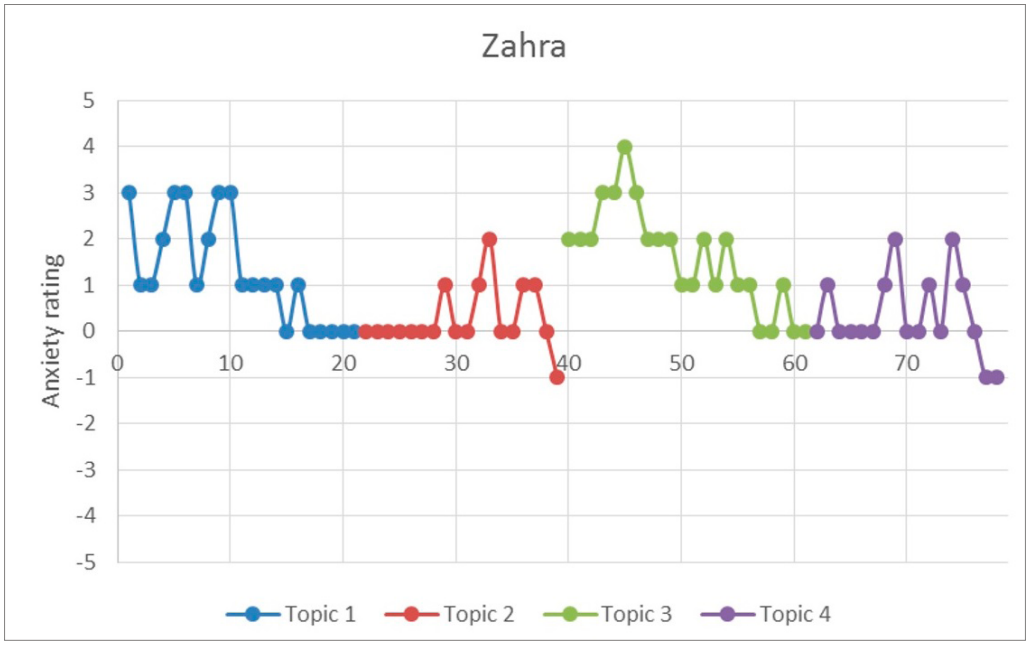
Zahra's explanations of her anxiety dynamics

Zahra reported that she began her first answer with a high level of anxiety because she did not know what "Firoozeh" was in English and her whole speech was going to be built on and around this word. She explained that she was fortunate that we could note her anxiety and translate her code-switched word, Firoozeh, to English. Moreover, she reported that her anxiety was about to decrease but after she failed to pronounce the new word, she felt embarrassed , slow to speak about the topic, and highly anxious. Zahra explained that, as we had an understating of her affective state, she repeated the word with less anxiety as she felt understood. She also explained that before the interview she could easily imagine herself going through moments of anxiety

and shaking in her boots because she knew that she did not own a rich vocabulary and could not fully employ all of the grammatical rules she had learned to talk about different topics. But the moment we raised the first question, as she claimed, she felt a load was taken off her shoulders and this helped reduce her anxiety.

She noted that she had never been given the chance to express her personal ideas. Thus, she felt pleased while she was talking about her thoughts on movies and being videotaped. She explained also that she felt anxious when she was recalling the words and names in response to the first topic. She also reported her experiences of anxiety on the word "if" when producing her last sentence because it required a long pause for her to structure her sentence. She explained that she could not make up her mind answering the third question because an unrelated family issue ran through her mind during the preparation time; thus, her search for right words and structures to fabricate her story spoiled her speech and she was unable to recall the simplest words such as "uncle" or "knife". She also found code-switching a soothing way to make herself understood. She also said that she was highly anxious due to the chunks of speech that required her to recall difficult events in her past. She noted that her interest in the fourth topic and her primary preparation could help her achieve peace of mind during her speech. She also explained that her personal characteristics of high self-confidence and sociability made up for her occasional linguistic blocks. She decided not to focus on the reliable delivery of information but merely share her opinion (see Figure 3).

Figure 3
Zahra's Anxiety Rating



Observation of Azar's anxiety dynamics

Topic 1-4

Azar was one of the low-anxiety students. Despite her performance on the FLCAS, she anxiously asked us to extend the preparation time. She started answering her question while slurring her words. After some time passed, she began to feel more confident and maintain the topic with less anxiety than before. However, her anxiety was higher saying the word "cage". Announcing that she was ready before her preparation time was up, she appeared interested in the second topic, although she experienced moments of anxiety. She started answering the third question and seemed that she was rather unsure about the topic. Her anxiety fluctuated, rising on the sentence "you know" and on the word "cold". She also made a number of grammatical errors. In response to the fourth question, she felt overcome with anxiety and her chest tightened in fear, and she inhaled deeply. She made a number of errors in terms of grammar and vocabulary. Finally, she slumped down in her chair in relief as the interview finished.

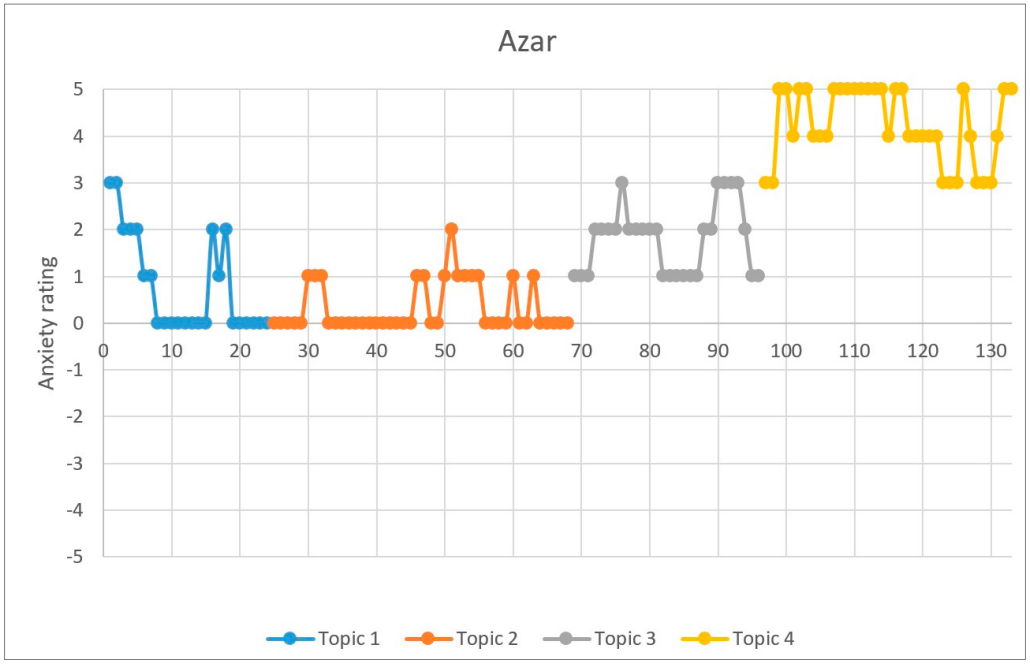
Azar's explanations of her anxiety dynamics

Azar explained her additional-time request as she could not overcome her nervousness or her lack of interest in the topic. Moreover, she reported that her anxiety dropped to zero in a short amount of time. Furthermore, she added that she felt anxious about the word "cage" because she was not sure whether she was using the right word or not. She explained that our supportive reaction to her questioning gesture reduced her anxiety. She concluded that her anxiety was mostly induced by a linguistic block. She reported that she could project her answer immediately for the second topic, which raised her interest in the topic. In addition, she reported that she felt quite anxious as a result of collocation-related linguistic block. Moreover, her difficulty in recalling words and names contributed to her nervous pauses while talking about the topic.

She also added that she felt a little anxious when producing her last sentence, which started with "so", since she found herself so deep in reflection that she temporarily lost her train of thought. She also explained that having no idea how to answer to the third question, she experienced high anxiety, confusion, and obliviousness to her errors from the very first moment. In addition, she addressed her perceived low English proficiency as an anxiety provoker as she could not think about the topic and grammatically formulate the sentences simultaneously. Moreover, she noted that accounting for her unstable state of mind regarding the topic, she decided to cheat from one of the hints (were the guests cold?) and make up a fake story. She noted that she suffered from a high level of anxiety since the third question did not elicit any of her real-life experiences.

Furthermore, she discussed that she could barely recall the word "cold" from the hint, which raised her anxiety, but our supportive feedback to her common questioning gesture at the word "cold" assuaged her anxiety. She also reported that she used the conjunction "and" and the phrase "you know" to take some time to plot her fake story and avoid lengthy pauses; however, they could not alleviate her anxiety level because she knew that her speech was not very informative for us. She revealed that although she was interested in the topic, her excitement did not allow her to concentrate on the topic. She explained that when thinking of her favorite celebrity, her blood pressure dropped and she felt cold. She could neither think clearly how to produce flawless speech in terms of grammar and vocabulary nor recall the wealth of information she had read about her favorite singer. She concluded that she lost her train of thought due to her strong affection for her favorite singer (see Figure 4).

Figure 4
Azar's anxiety rating



Discussion

The idiodynamic method casts light on emerging information about the dynamics of FLA that nomothetic methods could not disclose in previous studies. The first research question sparked off the debate over the issue of within-individual variations in FLA during the conversational topics. Regarding the within-individual changes in the reactions set off by the same stimuli during conversational topics, it was demonstrated that the anxiety provoking and reducing stimuli were both consistent and inconsistent within the individuals. This means that due to the content of each question, each individual went through different moments of anxiety. That is, as a complex system, the anxiety trajectory of each participant was susceptible to each topic as an attractor, which contributed to the stability of anxiety states for each participant. Ziba's anxiety provokers and reducer sas well as the reaction they sparked did vary from one question to another. Likewise, Melika, Zahra, and Azar's anxiety provokers and reducers as well as the reactions they induced varied from one question to another. For example, Zahra pointed to the issue of linguistic block as an anxiety provoker while answering the fourth question, yet she claimed the same issue as an anxiety reducer when answering the first and second questions. Besides, Melika was gripped by the deep-seated fear of being videotaped while replying to questions 1 and 2, yet she felt totally indifferent to the same stimulus when replying to questions 3 and 4. Moreover, Azar's within-individual variation in terms of topic interest was one of the shocking outcomes of the study underlying the free-floating nature of human beings as her anxiety was fueled by a burst of affection for her favorite singer.

This provides evidence for the topic-sensitive nature of the learners' states of anxiety, and thus the *complex nature* of each individual learner's anxiety. Topics shape an inseparable part of a classroom interaction. Given the interplay of each individual learner's emotional repertoire in terms of her background experiences with the emotion-laden words in each question, her path of feeling anxiety might go through different states. For instance, compared to her state of anxiety under the influence of the first question, the words "disappoint" and "not enjoy" triggered Ziba's past experiences, which led to her higher states of anxiety. For Azar, the words "favorite celebrity" reminded her of her idol singer and the strong affective bond she felt with her. Under the influence of this tight emotional connection with her favorite singer posed by the fourth question, she experienced moments of higher anxiety than her those in her exposure to the second and third questions.

The second research question debated the issue of between-individual variations in FLA over time. The present outcomes indicated that anxiety provoking and reducing stimuli were both consistent and inconsistent between individuals. The findings demonstrated that the reactions set off by these stimuli varied from one individual to another. That is, contributing to the emergent nature of FLA, each question paves the way for the emergence of a different state of anxiety for each individual. This means that each topic played a different attractive role and; as a result, a different pattern of anxiety for each participant. For instance, in response to the fourth question, Zahra and Azar experienced different patterns of anxiety. Zahra's interest in the topic and her primary preparation brought her peace of mind. This peace of mind led to her feeling a low state of anxiety.

However, the strong affective connection with her favorite singer raised by the topic in the fourth question contributed to a climax of experienced anxiety for Azar. It should be noted that the emergence of different states of experienced anxiety for each participant was not limited to the existence of a specific topic, but a glance at the four participants' trajectories of anxiety indicates that under the influence of the four topics, the states of anxiety emerged differently for each participant. This supports the fact that "L2 learners do not form ergodic ensembles" (Lowie & Verspoor, 2019, p. 184) and is in alignment with Gregersen, MacIntyre, and Meza (2014) who said it should be explored from an individual level. The patterns of anxiety under the influence of the four topics for each individual participant are different from their average experience of anxiety. In addition, consistent with Lowie and Verspoor (2019), the participants' anxiety ratings were surprisingly not in line with their scores on the FLCAS, representing potential spontaneity in light of individual differences under the influence of the four topics.

With regard to the third research question, the findings indicated that the participants attributed their feelings of anxiety in the conversations to the different topics in the questions raised to them. This means that given the situated nature of FLA, in line with Gkonou (2017), the stability and fluctuations in anxiety depend on interactional factors and one of these interactional factors is the range of topics in the interactions. Particular to the content of each topic, each participant's anxiety was triggered differently. For instance, the existence

of the word “not” and the emotionally negative laden word “disappoint” for Zahra induced high anxiety in her while responding to the second and the third questions. On the other hand, for Melika and Zahra, the emotionally positive laden word “favorite” in the fourth question provided them with a sense of interest in the topics, helping them experience the lowest levels of anxiety in their trajectories of anxiety. However, the fourth question triggered a strong affect in Azar with her favorite singer, leading to a high anxiety state in her anxiety trajectory.

With respect to the pedagogical implications of this study, the results showed that the use of the idiodynamic method could provide us with documents regarding the dynamics of FLA under the influence of conversational topics. A body of ample evidence supporting four main features of dynamic systems (de Bot et al. 2007) came to light through the insight gleaned from the current study. Our methodology tested the first property of dynamic systems providing a continuous graph of individual variations in FLA on which the participants’ ratings per second revealed FLA as a fluctuating state within a short period of time, which is in line with the respondents’ justifications for the rise and fall of their anxiety. For instance, Zahra explained how a lack of initial preparation hit her with a stabilized ripple effect that spoiled her speech. This provided evidence for the within-topic consistency of the first property of dynamic systems although we cannot assert that our findings demonstrate total between-topic consistency of this property as Zahra’s low anxiety ratings while answering the first question can neither explain nor predict her high anxiety ratings when replying to the second question. On the other hand, Melika reported that her anxiety over her fear of being videotaped dropped as time went by. Hence, the between-topic consistency of the first property was reflected to a small extent.

Moreover, we observed interconnectedness between the affective state of anxiety and the participants’ educational background, consistent with Melika’s explanations, the cognitive state of linguistic block that was inconsistent with Ziba’s within-topic and between-topics anxiety, and Zahra’s between-topics anxiety. Furthermore, we saw interconnectedness between individuals’ personality traits and their affective state of anxiety in this study. For example, Ziba reported that her high self-confidence made up for her linguistic block, decreasing her anxiety, and Azar noted that she got stuck in a state of high anxiety at the beginning of her first speech due to her shyness and low ambiguity tolerance. The findings also showed that difficulty in vocabulary retrieval and a low fluency level were two of the main anxiety provokers for the participants in this study, contributing to their moment-to-moment fluctuations of anxiety.

Documenting the third property of dynamic systems, we see both variations and stability in the anxiety of the participants reflecting on both within-topic and between-topic explorations carried out during the within-individual investigation of their FLA. For instance, Melika noted that linguistic block was always deleterious to her, yet the same stimulus induced both flat and anxious reactions when answering different questions as Azar’s interest in the topics was once reported as facilitating and once debilitating in her conversations.

In addition, our study provides evidence for the fourth major property of the dynamic systems, the butterfly effect, as Zahra experienced a short period of distraction during the preparation time, which caused her not to be able to map her speech properly, and she consequently suffered from high levels of anxiety when answering the second question. Ziba also reported a major problem of linguistic block. It was likely that her curve would continue to fluctuates, but instead her anxiety curve remained stable since she immediately decided to give the definition of the words she had difficulty recalling.

Conclusion

The recent dynamic perspective towards anxiety in learning a foreign language as well as the issue of ergodicity in research on individual differences inspired us in this study to explore the fluctuations of four participants’ anxiety in their conversations under the influence of different conversational topics. The findings indicated that the patterns of anxiety emerged differently both intra-individually and inter-individually; that is, from a complex dynamic perspective, each topic attracted each individual’s anxiety to different attractor states, contributing to the dynamics of their anxiety. Furthermore, being sensitive to the topics in the conversations, each individual attributed the dynamics in her anxiety trajectory to different emotion-laden words in the topics.

Finally, the use of the idiodynamic method in this study provided us with insights into the complex nature of anxiety under the influence of different topics for conversation. With regard to the pedagogical implications of the findings of this study, we can note that foreign language teachers should be sensitive to the types of topics they use in their classroom interactions with their students. That is, the same question might have a different effect in terms of its contribution to the students' anxiety, creating inter-individual differences of FLA. Furthermore, the series of questions for one student might lead to fluctuations in anxiety for that student, causing intra-individual variations in FLA, even in short-term conversations. It should be noted that due to the qualitative nature of this study, the findings cannot be generalized to other settings of experiencing FLA by learners of a foreign language. This means that since language learners are not ergodic ensembles, further research on the dynamics of FLA under the influence of conversational topics can shed more light on the dynamic nature of FLA.

Conflict of Interest

The authors declare that they have no conflict of interest.

References

Aida, Y. (1994). Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. *The Modern Language Journal*, *78*(2), 155–168. http://dx.doi.org/10.1111/j.1540-4781.1994.tb02026.x

Arnold, J. (Ed.) (1999). *Affect in language learning*. Cambridge University Press.

Bailey, K. M. (1983). Competitiveness and anxiety in adult second language learning: Looking at and through the diary studies. In H. W. Seliger & M. H. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 67-102). Newbury House.

Chastain, K. (1975). Affective and ability factors in second language acquisition. *Language Learning*, *25*(1), 153–161. http://dx.doi.org/10.1111/j.1467-1770.1975.tb00115.x

Cohen, Y., & Norst, M. J. (1989). Fear, dependence and loss of self-esteem: Affective barriers in second language learning among adults. *RELC Journal*, *20*(2), 61–77. http://dx.doi.org/10.1177/003368828902000206

de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, *10*, 7–21. http://dx.doi.org/10.1017/S1366728906002732

Dewaele, J. M. (2019). The effect of classroom emotions, attitudes toward English, and teacher behavior on willingness to communicate among English foreign language learners. *Journal of Language and Social Psychology*, 0261927X19864996. https://doi.org/10.1177/1362168817692161

Dewaele, J-M., Petrides, K.V., & Furnham, A. (2008). Effects of trait emotional intelligence and sociobiographical variables on communicative anxiety and foreign language anxiety among adult multilinguals: A review and empirical investigation. *Language Learning, 58*(4), 911–960. http://dx.doi.org/10.1111/j.1467-9922.2008.00482.x

Dulay, H., & M. Burt (1977). Remarks on creativity in language acquisition. In M. Burt, H. Dulay & M. Finnochiaro (Eds.), *Viewpoints on English as a second language* (pp. 95-126). Regent.

Elahi Shirvan, M., & Talebzadeh, N. (2017). English as a foreign language learners' anxiety and interlocutors' status and familiarity: An idiodynamic perspective. *Polish Psychological Bulletin*, *48*(4), 489-503. https://doi.org/10.1515/ppb-2017-0056

Elahi Shirvan, M., Khajavy, G.H., & Taherian, T. (2016). Psychometric analysis of Persian adaptation of foreign language classroom anxiety scale. In I. Hakkı Mirici, I. Hakkı Erten, H. Öz, & I. Vodopija-Krstanović (Eds.). *Research papers on teaching English as an additional language* (pp. 195-205). University of Rijeka.

Elahi Shirvan, M., Khajavy, G. H., Nazifi, M., & Taherian, T. (2018). Longitudinal examination of adult students' self-efficacy and anxiety in the course of general English and their prediction by ideal self-motivation: Latent growth curve modeling. *New Horizons in Adult Education and Human Resource Development*, *30*(4), 23-41. https://doi.org/10.1002/nha3.20230

Elahi Shirvan, M., Khajavy, G. H., MacIntyre, P.D., & Taherian, T. (2019).A Meta-analysis of L2 willingness to communicate and its three high-evidence correlates. *Journal of Psycholinguistic research, 49*(2), 1-27.

https://doi.org/10.1007/s10936-019-09656-9
Elkhafaifi, H. (2005). Listening comprehension and anxiety in the Arabic language classroom. *The Modern Language Journal, 89*(2), 206-220.http://dx.doi.org/10.1111/j.1540-4781.2005.00275.x
Fogel, A. (2006). Dynamic systems research on inter-individual communication: The transformation of meaning making. *Journal of Developmental Processes*, *1*, 7–30.
Gardner, R. C.,& MacIntyre, P.D. (1993). On the measurement of affective variables in second language learning. *Language Learning*, *43*(2), 157–194. http://dx.doi.org/10.1111/j.1467-1770.1992.tb00714.x
Gkonou, C. (2017). Towards an ecological understanding of language anxiety. In C. Gkonou, M. Daubney, & J. M. Dewaele (Eds). *New insights into language anxiety: Theory, research and educational implications* (pp.136–156). Multilingual Matters.
Gkonou, C., Daubney, M.,& Dewaele, J. M. (Eds.) (2017). *New insights into language anxiety: Theory, research and educational implications*. Multilingual Matters.
Gregersen, T. S.,& Horwitz, E. K.(2002). Language learning and perfectionism: Anxious and non-anxious language learners' reactions to their own oral performance. *The Modern Language Journal*, *86*(4), 562–570. http://dx.doi.org/10.1111/1540-4781.00161
Gregersen, T, Meza, M. D. & MacIntyre, P. D. (2014). The motion of emotion: Idiodynamic case studies of learners' foreign language anxiety. *The Modern Language Journal*, *98*(2), 574-588. http://dx.doi.org/10.1111/modl.12084
Gregersen, T., MacIntyre, P. D., &Olson, T. (2017). Do you see what I feel? An idiodynamic assessment of expert and peer's reading of nonverbal language anxiety cues. *New insights into language anxiety: Theory, research and educational implications* (pp.110–134). Multilingual Matters.
Horwitz, E. K. (1986). Preliminary evidence for the reliability and validity of a foreign language anxiety scale. *TESOL Quarterly*. *20*(3), 559–562. http://dx.doi.org/10.2307/3586302
Horwitz,E.K.(1996).Even teachers get the blues: Recognizing and alleviating language teachers' feelings of foreign language anxiety. *Foreign Language Annals*, *29*(3), 365–372. http://dx.doi.org/10.1111/j.1944-9720.1996.tb01248.x
Horwitz, E. K. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics*, *21*, 112–126. http://dx.doi.org/10.1017/S0267190501000071
Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching, 43*(2), 154-167.http://dx.doi.org/*10.1017/S026144480999036X*
Horwitz, E. K., Horwitz, M.B., & J. Cope (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125-132. http://dx.doi.org/10.1111/j.1540-4781.1986.tb05256.x
Howe, M. L., & Lewis, M. D. (2005). The importance of dynamic systems approaches for under-standing development. *Developmental Review*, *25,* 247–251. http://dx.doi.org/10.1016/j.dr.2005.09.002
Huang, H. T. D., & Hung, S. T. A. (2013). Comparing the effects of test anxiety on independent and integrated speaking test performance. *Tesol Quarterly*, *47*(2), 244-269.
Kleinmann, H. H. (1977). Avoidance behavior in adult second language acquisition. *Language Learning*, *27*(1), 93–107. http://dx.doi.org/10.1111/j.1467-1770.1977.tb00294.x
Koch, A. S., & T. D. Terrell (1991). Affective reactions of foreign language students to natural approach activities and teaching techniques. In E.K. Horwitz & D.J. Young (Eds.). *Language anxiety: From theory and research to classroom implications* (pp. 108–126). Prentice Hall.
Krashen, S. D. (1981). *Second language acquisition and second language learning*. Pergamon.
Larsen-Freeman, D. (2007). Reflecting on the cognitive–social debate in second language acquisition. *Modern Language Journal*, *91*(1), 773–87. http://dx.doi.org/10.1111/j.1540-4781.2007.00668.x
Liu, M.,& Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The Modern Language Journal*, *92*(1), 71–86. http://dx.doi.org/10.1111/j.1540-4781.2008.00687.x
Lowie, W. M., & Verspoor, M. H. (2019). Individual differences and the ergodicity problem. *Language Learning*, *69*(1), 184-206. https://doi.org/10.1111/lang.12324
MacIntyre, P. D. (1995). On seeing the forest and the trees: A rejoinder to Sparks and Ganschow. *The Modern Language Journal*, *79*(2), 245-248. http://dx.doi.org/10.1111/j.1540-4781.1995.tb05438.x
MacIntyre, P. D. (1995). How does anxiety affect second language learning? A reply to Sparks and Ganschow. *The Modern Language Journal*, *79*(1), 90–99. http://dx.doi.org/10.1111/j.1540-4781.1995.tb05418.x
MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *The Modern Language Journal,91*(4), 564–576. http://dx.doi.org/10.1111/j.1540-4781.2007.00623.x

MacIntyre, P. D. (2012). The idiodynamic method: A closer look at the dynamics of communication traits. *Communication Research Reports*, *29*(4), 361-367. http://dx.doi.org/10.1080/08824096.2012.723274

MacIntyre, P.D. (2017). An overview of language anxiety research and trends in its development. *New insights into language anxiety: Theory, research and educational implications* (pp.11–30). Multilingual Matters.

MacIntyre, P. D., & R. C. Gardner (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning*, *44*(2), 283–305. http://dx.doi.org/10.1111/j.1467-1770.1994.tb01103.x

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

MacIntyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied Linguistics*, *32*, 149–171. http://dx.doi.org/10.1093/applin/amq037

MacIntyre, Dörnyei, P., & Z.D. Cl´ement, R. & Noels, K. A. (1998), Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, *82*(4), 545–562. http://dx.doi.org/10.1111/j.1540-4781.1998.tb05543.x

MacIntyre, P. D., Babin, P. A., & Cle´ment, R. (1999). Willingness to communicate: Antecedents and consequences. *Communication Quarterly*, *47*, 215–229. http://dx.doi.org/10.1080/01463379909370135

MacIntyre, P. D., Noels, K. A., & Moore, B. (2010). Perspectives on motivation in second language acquisition: Lessons from the Ryoanji Garden. In M. T. Prior, Y. Watanabe, &S.-K. Lee (Eds.), *Selected Proceedings of the 2008 Second Language Research Forum* (pp.1-9). Cascadilla Press. http://www.lingref.com/cpp/slrf/2008/index.html

Mates, A. W., & Joaquin, A. D. L. (2013). Affect and the brain. In J. Herschensohn & M. Young-Scholten (Eds.), *The Cambridge handbook of second language acquisition* (pp. 417–435). Cambridge University Press.

Matsuda, S., & Gobel, P. (2004). Anxiety and predictors of performance in the foreign language classroom. *System, 32*(1), 21-36. http://dx.doi.org/10.1016/j.system.2003.08.002

Maxwell, J. A. (2005). *Qualitative research design: An interactive approach* (2nd ed.). Sage.

Oh, J. (1992). The effects of L2 reading assessment methods on anxiety level. *TESOL Quarterly*, *26*(1), 172–176. http://dx.doi.org/10.2307/3587386

Parlade, M. V., & Iverson, J. M. (2011). The interplay between language, gesture, and affect during communicative transition: A dynamic systems approach. *Developmental Psychology*, 47, 820–833. http://dx.doi.org/10.1037/a0021811

Rosenzweig, S., & Fisher, S. L. (1997). ‘‘Idiographic’’ vis-a`-vis ‘‘idiodynamic’’ in the historical perspective of personality theory: Remembering Gordon Allport, 1897–1997. *Journal of the History of the Behavioral Sciences*, *33*, 405–419.

Saghafi, K., Adel, S. M. R., & Zareian, G. (2017). An ecological study of foreign language writing anxiety in English as a foreign language classroom. *Journal of Intercultural Communication Research*, *46*(5), 424–440.

Saito, Y., Horwitz, E. K., & Garza, T.J. (1999). Foreign language reading anxiety. *The Modern Language Journal*, *83*(2), 202–218. http://dx.doi.org/10.1111/0026-7902.00016

Saito, Y. & Samimy, K. (1996). Foreign language anxiety and language performance: A study of learner anxiety in beginning, intermediate, and advanced-level college students of Japanese. *Foreign Language Annals*, *29*(2), 239–251.

Sarason, I. G. (1978). The test anxiety scale: Concept and research. In C. D. Spielberger & I. G. Sarason (eds.*), Stress and anxiety* (vol. 5, pp. 193–216). Hemisphere.

Sparks, R. L.,& L. Ganschow (1995). A strong inference approach to causal factors in foreign language learning: A response to MacIntyre. *The Modern Language Journal*, *79*(2), 235-244. http://dx.doi.org/10.1111/j.1540-4781.1995.tb05436.x

Sparks, R. L., Ganschow, L., & Javorsky, J. (2000). D ´ej `avuallover again: A response to Saito, Horwitz, and Garza. *The Modern Language Journal, 84*(2), 251–255.http://dx.doi.org/10.1111/0026-7902.00066

Spielberger, C. (1966) *Anxiety and behaviour*. Academic Press.

Spitalli, E. J. (2000). *The relationship between foreign language anxiety and attitudes toward multiculturalism in high-school students* [Unpublished Master’s thesis]. Benedictine University.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. MIT Press.

Tran, T. T. T., Baldauf Jr, R. B., & Moni, K. (2013). Foreign language anxiety: Understanding its status and insiders’ awareness and attitudes. *Tesol Quarterly*, *47*(2), 216-243. https://doi.org/10.1002/tesq.85

van Geert, P. (2011). The contribution of complex dynamic systems to development. *Child Development Perspectives*, *5*, 273–278. http://dx.doi.org/10.1111/j.1750-8606.2011.00197.x

Young, D. J. (1990). An investigation of students' perspectives on anxiety and speaking. Foreign *Language Annals*, *23*(6), 539–553. http://dx.doi.org/10.1111/j.1944-9720.1990.tb00424.x

Young, D. J. (Ed.) (1999). *Affect in foreign language and second language learning*: *A practical guide to creating a low-anxiety classroom atmosphere*. McGraw-Hill College.

Appendix

Ziba

Topic 1

I[0] have[0] a necklace[0]. It's[0] green[0]. It's[0] made[0] of very[1] special[1] kind[1] of stone[1]. It's[0] called[0] Antovenie[0]. When[0] my brother[1] wanted[0] to leave[1] home[1] and go[0] to university[-1] in another[1] city[0], he[0] gave[0] me[0] that[0] necklace[0] to[0] remind[1] me of him[1] and he[1] told[0] me[0] it[2] makes[2] you {me} [2] happy[2] because[0] it's[0] a[0] special[0] stone[0] and every[0] time[0] you're {I'm} [0] sad[0] I just[0] look[0] at[0] this[0] stone[0] and remember[1] our[1] happy[1] memories[1] together[1]. So, anytime[0] I'm[0] sad[1] I[0] look[0] at that[0] necklace[0] and after[0] 2 minutes[0] I[0] forget[-1] my[0] sadness[2]. I[0] always[-1] carry[-2] it[-3] with myself[-4] everywhere[-5]. If[0] I[0] die[2], I[1] would[1] return[2] it[2] my[2] brother[2].

Topic 2

Two[0] years[0] ago[0], I[0] saw[0] a movie[0]. The[0] name[0] of that[0] movie[0] was[0] "Hiss DokhtarhaFaryad Nemizanand."[2] I[1] saw[1] this[1] movie[1] because[0] I[0] heard[0] it's[0] an[0] interesting[0] movie[0], but[0] one[0] should[0] watch[0] it[0] many[0] times[0] to get[0] the story[1] because[0] it's[0] a little[0] confusing[1]. I[1] hate[3] this[2] movie[2] because[2] it[2] shows[2] that[2] what[3] happened[3] in the[3] movie[3] has[3] happened[3] to many[3] girls[3]. It[3] also[3] shows[3] that[3] men[4] can[4] force[5] little[5] girls[5] to have[4] sex[4] with them[3] (child abuse). If[2] there[1] would[1] be[1] another[1] movie[1] with[1] the same[1] story[2] I[1] would[1] never[0] watch[0] it[0] because[0] it[0] made[1] me[1] sad[1] and[1] made[1] me[1] think[2] why[3] some[3] men[4] do[3]such[3]things[3] to good[3] girls[4]. In[1] this[1] movie[1], a man[4] babysat[2] a little[3] girl[3]. Her[3] parents[3] trusted[2] him[4] with[2] their[2] child[3]. One[4] night[4] when[4] her[4] parents[4] went[4] out[4] and left[4] their[4] baby[5] with[4] the man[5] to take[3] care[3] of her[4]. But[3] that[3] man[5] forced[5] the[5] girl[5] to[5] have[5] sex[5] with[5]him[5] (raped her) and took[5] pictures[5] of her[5] and told[5] her[5] if[5] you[5] tell[5] your[5] parents[5] what[5] I[5] did[5], I[5] will[5] show[5] this[5] pictures[5] to your[5] parents[5]. So[4] the girl[3] didn't[3] say[2] anything[2] to anyone[2] to[2] save[2] her[2] face[2].

Topic 3

When I[0] was 12[0], we[0] went[0] to Ghoochan[-1] with[0] my family[0] for my cousins[0] wedding party[0]. It[0] was a great[0] party. The food[0] was fine. We ate rice[0] and kebab. We danced[0] a lot and took a lot[0] of pictures. After[1] the party, we[1] went to the parking lot[1] to take our car[1] and go home but[2] the car didn't turn[2] on (start). Although[1] the party was fine but it didn't[2] end very well.

Topic 4

My favorite actor[0] is. When I[0] was 12, my[0] brother suggested[0] me to his movie[0] and he said it[0] was very suitable[0] for your age[0]. So I watched[0] it and I really[0] enjoyed it. Stared[0] in "Harry Potter"[0]. This movie[0] won many prizes[0]. He continued [0]acting in fiction[0] movies. He[0] has a lot of[1] fan. He is[0] handsome. He [1]has gray eyes [1]and blonde hair[1]. I searched a lot[0] about him on[0] the internet so[0] I have a lot of[0] information about [0]him and I'd like[0] to see him[0].

Melika

Topic 1

I have[5] a… notebook. Two years[5] ago… I bought[4] it… alone[4] and … it is very personal[4]… what is "be yadavordan in English?"[5] Aha…It reminds[3] (me) of my friend… it is on my[3] drawer and if I die[3]… I give it to my friend[3] because… [5] because I write… no[5]… because we wrote…[5] write our memories in[5] this … memory book?! [5] (diary) together.

Topic 2

I [5]saw "AsreYakhbandan" movie…[5] in cinema… I went to [3]cinema with my friends[3] to see it because[4]… because they[5]… they wanted to see[5] … watch this[5] movie. I think it was …[5] awful…because I didn't like[5] … the[5]… characters[5] … director of the movie were[5]… was[5]… I don't know [5]the director and the[5]… the actors… of the [5]movie… but…. Because[5] … because[5]… I … don't watch[4] movies many[5]…I mean…[4] a lot….

Topic 3

I love[0] my favorite volleyball[0] player, Shahram[0]Mahmoudi. Because he is[0]one {of} the [1]… the best[2] … sporter?![3] Aha[1] …aha[0] … player in the[0] world… and[1] … I like [1]him…since[1]… three years[1] ago… I saw[1] him in… Milad[0] tower…and…she…[2] he is[3] 20…6 years[2] old…and she is[1] married… no[3]… he is married[2]… uhmm[3]…hamin[3]… finished[3]…

Topic 4

Last year[3] I went to… my friend's[2] party… because[2] (of)… her [2]birthday. I bought one[1] (a) doll for her… and I went to[1] (the) party alone…he[3] … her birthday (was) [2] in her house… and the [2]decoration is…[3] was … very [2]awful…the guests[3]…are…[4] were…. Very cold[3]… and… I didn't [2]know them… the food [2]was terrible…. And…the music[2] was … [3]very… antique.

Last year[3] I went to… my friend's[2] party… because[2] (of)… her [2]birthday. I bought one[1] (a) doll for her… and I went to[1] (the) party alone…he[3] … her birthday (was) [2] in her house… and the [2]decoration is…[3] was … very [2]awful…the guests[3]…are…[4] were…. Very cold[3]… and… I didn't [2]know them… the food [2]was terrible…. And…the music[2] was … [3]very… antique.

Zahra

Topic 1

Two years[0] ago… I went to the cinema[0] with my friends[0]… and saw [0] "Shiyare 143". It was very[0] beautiful and… the first character of the [0]movie was lovely because[0] she was very[1]…patient… [0]but I didn't like[0] the movie… because[1]… because it made[2] me sad… I don't like to see[0] movies with such[0]…stories…I didn't like Gelareh[1] … Abbasi's playing (acting) and…if[1]… if she plays in another[0] movie, I wouldn't see it[-1]…

Topic 2

Five years[2] ago, I went to the[2]… to my aunt's[2] house…on (in) my aunt's[3] house… my husband ands' [3]of my aunt (my uncle, or my aunt's husband) dead[4]… died…so [3]went to my aunt's[2] house… At the party[2] (ceremony)… when my little sister[2] was running… and a "chagoo" (knife) [1]cut her feet… in {at} the [1]moment… she didn't know[2] (notice) it…and… continued running [1]but after some minutes[2] we saw bloody "lake"[1]… Yes[1]…yes[0]…blood spots on the [0]carpet and she's took…[1] we…took her to a [0]hospital… now she's[0] fine…

Topic 3

Rambod Javan[0] is my favorite actor… because[1]… because (of) his[0] acting… I know her[0] (him) (since) four years[0] ago… he's very[0] active… I [1]guess he's 40 years[2] old… and…she[0] (he) has married 2 times… but got divorce… after[0] that… she[1]… he got [0]married again for the three[2]… third time[1]… I haven't see[0] (seen) him … but I'd [-1]like to see him[-1] once…

Topic 4

Two[3] years ago… we traveled to[1]Neishaboor… and[1] … my mother[2] bought me a[3]…"firoozeh" [3]ring… Aha…[1] turquoise ring[1]… it's very very[1] beautiful… I like it[1] a lot… I always wear[0] it because when I wear it[1]…I …[0] remember all of the good[0] memories of Neishaboor … and if I die[0], I would give it to[0] my sister to remember[-0] me {by}…Aha[1]… yes[2]…a…a[3]….a…[3]

Azar

Topic 1

My[3] pet is my favorite[3]. It's a [2]parrot. My [2]father bought it[2] for me a[1] year ago[1]. It's a "he"[0]. I like it very[0]

much[0] because I[0] have a lot[0] of fun with[0] her. I[0] keep him[0] in a cage[2]?! He[1] eats “tokhmeh”[2], and fruit[0]. He can[0] sing, swing[0], and repeat[0] some simple[0] Persian words[0].

Topic 2

One[0] of the movies[0] I don’t like[0] is “Zapas”[0], because I[0] didn’t get[1] the movie… [1]subject?![1] Aha[0]… yeah[0]… the movie[0] plot. I[0] saw the preview[0] of the movie[0] on TV and[0] watched it[0] in cinema[0] with my cousins[0]… the actors[0] of the movie[0] were Rima Raminfar[0], mehran[1]… ops… [1]Ahmad Mehranfar[0], and Amir Jaafari[0]… and[1]… [2]but I don’t [1]remember the [1]director’s and [1]the other actors’[1] names…[0]I like the actors[0], but the story[0] of the movie[0] was …[1] poor[0] … I think[0]…[1]so if these actors[0] act in another[0] movie, I[0] will[0] definitely watch[0] it.

Topic 3

I[1] went to my[1] friend’s party[1]… they[2] don’t bought[2] anything[2] for our friend[2]…and … [3]you know…[2] they went to[2] the party with[2] my mother to[2] have fun[2]… we [1]eat sandwiches[1]…and then[1]… we don’t[1] dance or [1]listen to music[1]…and[2]…I don’t like[2] the party [3]… because[3]… you [3]know…the quests are[3] cold?! And… [2]I didn’t know[1] all of them[1]

Topic 4

My[3] favorite singer[3] is Sirvan Khosravi[5]… I love [5]his sound[4] {voice} … [5]you know…[5] His songs[4] are beautiful[4]…in the past[4]… you [5]know… I[5] didn’t like[5] him… [5]you know… [5]because…[5] you know[5]… I don’t[5] hear her[4] music… [5]you know…[5]one of my[4] friends said[4] that hear her[4] music maybe[4] you like that[4]… so I hear [3]his music and[3] I liked it…[3] you know[5]… I don’t[4] see him but[3] I like to see[3] him… he’s[3]single… [4]and … he’s 30[5]…34 years [5]old…

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 1, 2020

Solnyshkina, M. I., Harkova, E. V., & Kazachkova, M. B. (2020). The Structure of Cross-Linguistic Differences: Meaning and Context of 'Readability' and its Russian Equivalent 'Chitabelnost'. *Journal of Language and Education, 6*(1), 103-119. https://doi.org/10.17323/jle.2020.7176

The Structure of Cross-Linguistic Differences: Meaning and Context of 'Readability' and its Russian Equivalent 'Chitabelnost'

Marina I. Solnyshkina[1], Elena V. Harkova[1], Mariia B. Kazachkova[2]

[1] Kazan Federal University
[2] MGIMO University

Correspondence concerning this article should be addressed to Elena V. Harkova, Department of Theory and Practice of Teaching Foreign Languages, Institute of Philology and Intercultural Communication, Kazan Federal University, 2 Tatarstan str., Kazan, 420021, Russian Federation. E-mail: halenka@rambler.ru

The article presents the results of an original study aimed at finding (1) frequency fluctuations of the term 'readability' in American discourse and its Russian equivalent 'chitabelnost' in Russian discourse over the period from 1920s to the present; and (2) semantic similarities and differences between the English term 'readability' and its Russian equivalent 'chitabelnost' over the same period of time. A contrastive analysis of the words testified to inconsiderable differences in the semantic structures of the terms in the period under study: the term 'readability' has been used with the following meanings: (1) 'the quality of being legible or decipherable' and (2) 'the quality of being easy or enjoyable to read'. The Russian equivalent 'chitabelnost' has two contemporary meanings similar to the aforementioned English meanings as well as the obsolete 'library book checkouts'. With the help of the Google NgramViewer, we identified the 1980s frequency peak of both terms when the modern notion of the concepts was formed. The research into the topical context of readability as 'the quality of being easy or enjoyable to read' demonstrated empiricist tendencies in American studies focused on two types of parameters, i.e. the 'objective' parameters of texts, i.e. sentence length, word counts, number of high/low frequency words, ratio of high/low frequency words to total words, sentence complexity, etc. and 'individual' variables affecting a potential reader, such as 'word familiarity', cognitive and linguistic abilities, cultural and topic knowledge, etc. The Russian school's view, until the 1970s, had traditionally been more holistic and 'biased' towards an individuals' factors. The results of the study have the potential to contribute to cross-linguistic research in the area of text readability assessment, semantics, and scientific literature searches.

Keywords: readability, ngrams, topical context, semantic analysis, text parameters, sentence length, word counts, readability formulas

Introduction

Communication as 'the imparting or exchanging of information by speaking, writing, or using some other medium' (Online Oxford English Dictionary, 1996)[1] implies either generating or comprehending a text, which may be handwritten, printed, electronic, or oral. Successful communication in its turn largely depends on whether the amount, content, and structure of the quanta of the information sent by its generator in the text and received by the addressee are similar or, in an ideal situation, the same. Thus, for the information of any text to be elicited, processed, and stored in a recipient's mind, it is important that the recipient has sufficient cognitive and linguistic abilities.

In the modern world, matching a text (both oral or written/electronic) to the target audience is a problem relevant in a number of spheres: education, PR, the military, government, law, advertising, business, publishing, medicine, and social relations, as these are areas where communication is the foundation of success. The

[1] Online Oxford English Dictionary (OED) (3rd ed.) (1996). Retrieved from https://en.oxforddictionaries.com/

research shows that companies suffer damages and take financial hits if the texts they expose their customers to are hard for an average reader to comprehend (Klare, 2000). On the other hand, editors and educators claim that if a text is too easy, i.e. below the audience' reading expertise, readers lose interest and stop reading it (Randall, 2013).The success of reading depends on the degree of reading comprehension at a certain reading speed and maintaining interest in the content of the text. In an ideal situation, if a text is beyond its target audience's reading level, it needs to be altered or leveled to match the reader. Modern text leveling procedures imply measuring two parameters: (1) the level of cognitive and linguistic abilities of the target audience and (2) text readability (Reading A-Z, 2018)[2]. According to Fernbach (1990), the latter is nowadays interpreted as "the ease with which a text can be read quickly, understood, and memorized". As for the concept of 'text readability' itself, it dates back to medieval times, when word counts were used to estimate the difficulty of the Talmud (see Taylor & Wahlstrom, 1986). In the 1880s, Professor L.A. Sherman conducted a research on the length of the average English sentence in different centuries and concluded that shorter sentences and more concrete terms in a text make the text easier for a reader. Sherman was also the first to argue that readability can be evaluated based on statistical analysis (Sherman, 1893). At about the same time, in 1889, Russian writer Nikolai A. Rubakin published a list of 1500 words 'known by all Russians' that he derived from over 10000 texts written by common people. Rubakin argued that reading comprehension is hampered by unfamiliar words and long sentences (see Choldin, 1979). Unfortunately Rubakin's ideas were soon forgotten in Russia, but text 'readability' studies have since been actively conducted in the USA, UK, and Germany.

Nowadays researchers worldwide are working to address the problem of 'text – reader correspondence' in two ways: (a) from the point of the reader and their subjective characteristics: age, education, background knowledge, memory span, etc.; and (b) from the point of view of the text and its objective parameters. Text objective parameters are generally classified into two types of categories: 'extra-textual' parameters, which include illustration support and graphic features, such as font, spacing, and indentation, and 'inter-textual' parameters, which comprise the following: total word count, number of different words, ratio of different words to total words, number of high frequency words, ratio of high frequency words to total words, number of low frequency words, ratio of low frequency words to total words, sentence length, sentence complexity, etc. (Reading A-Z, 2018[3]; Ivanov, 2013; Hiebert, 2012).

In this paper we aim at two research questions: (1) How different or similar were the frequencies of the word 'readability' in American discourse and its Russian equivalent 'chitabelnost' in Russian discourse over the period from the 1920s to the present? (2) How different or similar were readability constructs and meanings of the English-Russian equivalents 'readability' and 'chitabelnost' over the period from the 1920s to the present? As the Russian word 'chiltabelnost' contains a borrowed suffix '-beln-' (Sologub, 2002), we may anticipate, and that is the first *hypothesis* of the study, that the word began functioning in Russian discourse later than its equivalent in English and its frequency has been considerably lower over the period under study. The implied *hypothesis* behind the second research question is that changes in the meaning of the term become evident once we contrast topical contexts in the specialized discourse. The focus is on the qualitative analysis of the topical context in the discourse, thus revealing diachronic changes in the semantic range of the words and structures of the corresponding concepts. The ultimate goal of the study is to define possible conceptual differences in the terms 'readability' and its Russian equivalent 'chitabelnost' over a period of time between the 1920s and 2019. To this end we synthesize ten historic meanings of the equivalents studied in each decade from the 1920s to 2019.

Method

Background Information

The 'linguistic diversity' of the word 'readability', naming a scientific concept, entails making a distinction between the language-specific semantic values or meanings of the word and the encyclopaedic concept(s) attributed to the word (see Willems, 2012). Acknowledging the two entities, Wierzbicka argues that 'meanings' are supposed to be conveyed in dictionaries while knowledge inherent to concepts is to be included in encyclopaedias (Wierzbicka, 1996). Thus, we resorted to unabridged monolingual dictionaries for the meaning

[2] Reading A-Z. (2018). Printable teacher materials. Retrieved from http://www.readinga-z.com
[3] Ibid.

of 'readability' in the English language[4],[5]and its equivalent 'chitabelnost' in the Russian language[6] to derive all the registered meanings of the words corresponding to separate entries in dictionaries' definitions. Based on the assumptions that "a semantic value is intuitively available to the speaker" and "the context can provide enough evidence on its own to reveal separate senses of a word" (Willems, 2012, p.668) we also tested the registered 'senses' of the words 'readability' and 'chitabelnost' "in naturally occurring utterances" (Willems, 2012, p.665), i.e. contexts. The idea of using the context to derive senses goes back to J.R. Firth's dictum "you shall know a word by the company it keeps" (Firth, 1957, p.11) and in our case the goal was to verify the semantic range of the word used in the discourse of a particular period.

For the purposes of the research, contexts such as "linguistic environments <…> in which a particular word occurs" (Dash, 2008, p.22) were classified into local and topical (Ravin & Leacock, 2000). We predominantly used the local context, i.e. 1–5 words immediately before and after the words under study to verify, extend or narrow the meanings registered in dictionaries. The topical contexts as "the wider circle beyond the sentence level" (Dash, 2008, p.22) were essential in synthesizing the structure of the concepts.

For our future discussions it is also important that, as with any scientific concept, readability is designed and developed primarily for research purposes and, thus, related to other constructs in theoretical schemes. However, the area of implementing the findings on similarities and differences in constructing the concept of 'readability' is quite wide as, obviously, the way a piece of information is presented influences how and how well a reader comprehends it. Reading is a complex task in which text writers are supposed to determine the best means to promote comprehension. We also emphasize that, as with any scientific construct composed of a number of constituents (Kerlinger, 1973), we expect readability to become observable through indicators or manifestations of what researchers have agreed. Based on this, the strategy employed in the exploration of the concept under study was an in-depth analysis of the existing knowledge in the area that entails reviewing the published research on readability in American and Russian discourses in the period between the 1920s and 2019. We were mostly exploring the compiled topical contexts from the Google books corpus, which proved to be sufficient for extracting topical contexts and help discriminate between the historic paradigms of the concepts (1920s – 2019). It was this type of context that opened the major avenue of the research.

Research Design

Synthesizing a concept based on the context in a text is a common approach practiced in modern corpus linguistics (see Zakharov et al., 2014). To define the trends in changes in the frequency of the words 'readability and 'chitabelnost' between 1920 and 2019 we applied distributional semantic models (Firth, 1957), which allowed us to induce the meanings of words from texts. The algorithm of diachronic studies of words and concepts with Google Books Ngram, i.e. a database of 67 billion words in Russian Viewer and 361 billion words in English, has been successfully implemented by numerous researchers in Russia and abroad (see Solovyev, 2013, Zakharovet al., 2014, Hai-Jew, 2014).

The research was conducted in three stages:

- Stage 1, a lexicographic analysis, was aimed at defining the meanings of the two contrasting words ('readability' and 'chitabelnost') based on the data registered in dictionaries.
- Stage 2 is a frequency-based contrastive discourse analysis of the two words 'readability' and 'chitabelnost' using the words' recurrence in the following decades: 1925 – 1929, 1930 – 1939, 1940 – 1949, 1950 – 1959, 1960 – 1969, 1970 – 1979, 1980 – 1989, 1990 – 1999, 2000 – 2009, and 2010 – 2019.
- Stage 3, a conceptual or Topical Context Analysis, was conducted to synthesize and contrast 10 historical concepts of 'readability/chitabelnost' over the ten decades, i.e. from 1925 to 2019.

Procedure

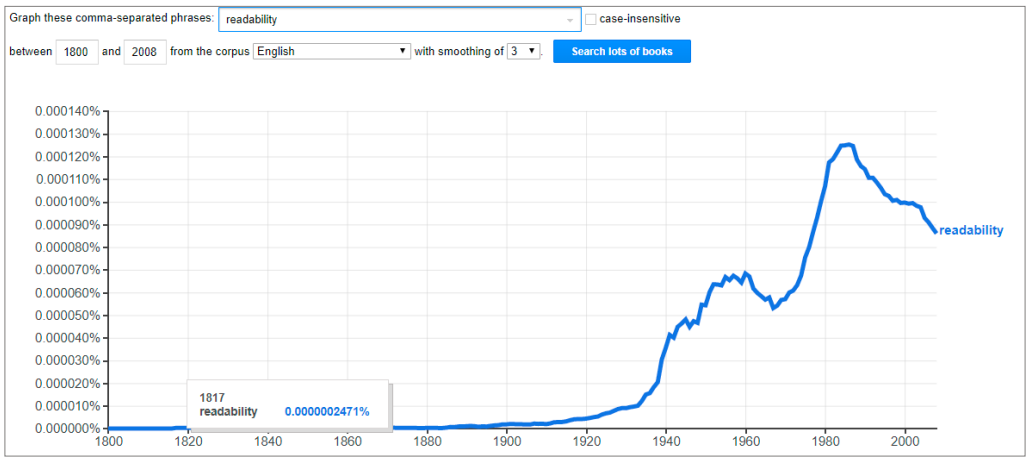
To contrast (1) the semantic volumes of the English 'readability' and its Russian equivalent 'chitabelnost' in different periods of time (the 1920s to 2019) and (2) the frequency of fluctuations in the terms in corresponding

[4] Fellbaum, C. (Ed.) (1998). WordNet: An electronic lexical database. MIT Press.
[5] Merriam Webster's collegiate dictionary. (1993). https://www.merriam-webster.com/dictionary/readability.
[6] Evgen'eva, A. P. (Ed.) (1999). MAS - Slovar' russkogoyazyka [The dictionary of Russian language]. http://feb-web.ru/feb/mas/mas-abc/default.asp.

discourses, we implemented the Ngram Viewer[7], an online tool based on a 'bag of words approach'.[8] We computed graphs of relative frequencies of the words 'readability' and 'chitabelnost', which demonstrate how often they were used in the Google Books corpora (Figures 1, 2 below).

The words 'readability' and 'chitabelnost' are viewed in the study as '1-grams', i.e. strings of characters uninterrupted by a space. The x-axes in Figures 1 and 2 show the periods of the study, i.e. from 1800 to 2019.

Figure 1
Ngram of Publications with the Word 'Readability', 1800 – 2019

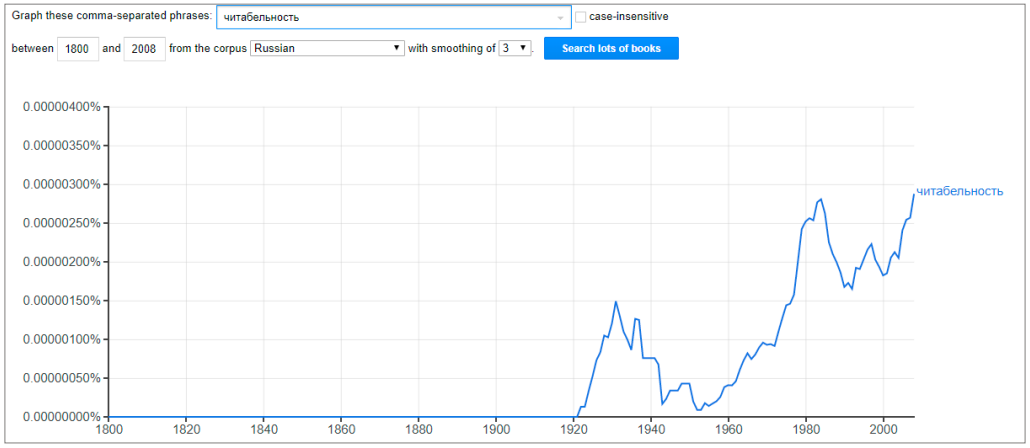


The y-axes show the relative frequency of the word (RF), i.e. the percentage of the specified ngram of all ngrams in the corpus of books of that particular year. For the word 'readability' in year 1817 it is 0.0000002471 % (see Figure 1 above).

The graph in Figure 1 shows that the frequency of the word increased dramatically in the 1930s and in the 1970s as those were the periods when the term was actively used in many more printed sources than in previous decades.

Figure 2 below provides insights into the fluctuating periods of popularity and decline of the corresponding Russian term, 'chitabelnost', over time in Russian discourse: the term became popular in the 1920s, the 1960s – 80s and then, in the 2000s it experienced an unprecedented rise of recurrence.

Figure 2
Ngram of Publications with the Word 'Chitabelnost' (Readability) in Russian Discourse, 1800 – 2019



[7] Google Ngram Viewer (2013). Professional web tool for calculating and tracing relative and absolute frequencies of a word or a phrase. https://books.google.com/ngrams

[8] Ong, M. (2014). Bag of Words (BoW) -Natural language processing [Blog]. https://ongspxm.github.io/blog/2014/12/bag-of-words-natural-language-processing/

We share the opinion that the term 'frequency counts' provides an indicator of trends in the corresponding research area (Hai-Jew, 2014) and the conceptual part of the study was based on the premise that the rate of knowledge growth in the area, as well as paradigm changes, are reflected in publications (Macias-Chapula, 1999).

Limitations

Although the limitations of the Google Books corpus and the Ngram viewer as indicators "of the 'true' popularity of various words and phrases" have been discussed more than once in literature (Pechenick, Danforth, & Dodds, 2015), we advocate for the reliability of the tool based on the assertions presented below. The main criticism against the Google Books corpus revolves around it being a more lexicon-like than text-like dataset and designed as "a reflection of a library in which only one of each book is available" (Pechenick, Danforth, & Dodds, 2015, p.2/24). Another argument raised by Google Books corpus opponents is that it is imbalanced because "scientific texts have become an increasingly substantive portion of the corpus throughout the 1900s" (Pechenick, Danforth, & Dodds, 2015, p. 1/24. However, as we aim at the meanings and scientific constructs as well as the frequency of publications with the words under study in certain years and at certain periods (not the popularity of the words, which depends on the number of printed copies), this skewness to scientific discourse makes the Google Books corpus more scientifically representative and thus preferable for our research.

There were two problems we faced working with the Ngram Viewer. The first one was with optical character segmentation when the system recognized character sets of Russian 'рен' as 'чи' which resulted in referring to sources with the word 'rentabelnost' (Rus. profitability) not 'chitabelnost' (for optical character recognition errors in English texts, see Zhang, 2015). Thus, we had to manually check every snippet of the word 'chitabelnost' in the Russian version of the Google Books corpus.

The second problem was a lack of available resources published in a certain year. E.g. Ngram Viewer graph of readability (see Figure 1) starts in 1817, but the resource does not provide a reference to texts with the word 'readability' published until 1916. In the Russian corpus, the word 'chitabelnost' was used for the first time in 1922 with a relative frequency of 0.0000001530 % (see Figure 4. below). This was the reason we chose the 1920s as the starting point of the 'anthology'. As for the latest sources registered in both versions of the Google Books corpora in Russian and English, and ultimately in Ngram Viewer, they were released in 2019.[9] Thus, although the Google Books corpora designers argue that they provide graphs for the period of 1800-2009, in fact the data elicited covers the period of 1916-2019 for the word 'readability' and 1922-2019 for the word 'chitablnost'.

Materials

As the Ngram Viewer typically provides texts of publications with an ngram under study, we expected to get access to the first registered resource in the English Google Books corpus, i.e. a book published in 1817 as that is the year when the graph starts (see Figure 1 above). Unfortunately, in our case, the earliest publication available in the Google Books corpus is the article "Development and Validation of a New Instrument to Assess the Readability of Spanish Prose" in Volume 65 of 'The Modern Language Journal' published in 1916 by Patricia Vari-Cartier. The author refers to the two variables used to calculate readability, i.e. the average sentence length and the number of syllables: "Each new graph would require a different set of parameters (minimum and maximum sentence and syllable count) and *readability* designations" (Vari-Cartier, 1916, p. 145). Those were studies that were greatly influenced by the works of Sherman mentioned earlier in this paper (Sherman, 1893).

Thus, based on the resources available, we focused on the time frame from 1925 to 2019. For an in-depth analysis of the concept, we sampled 10 topical contexts of the words 'readability' and 'chitabelnost' from 10 different texts of each decade from 1925 to 2019, thus compiling 10 sub-corpora of topical contexts for each decade of the period under study. The topical contexts were selected with the purpose of making the constituents of the concepts under study visible, thus the length of the topical contexts varied from nine sentences to a paragraph of 21 sentences. A compiled sub-corpus of 10 topical contexts contains on average 1912 tokens with the smallest being 1281 tokens (Russian Sub-Corpus, 1920s) and the biggest having 2357 tokens (Russian Sub-Corpus, 1990s). The data were extracted with the help of the Ngram Viewer application,[10] manually filed and marked as English

[9] Readability. (2019). https://clck.ru/JAtwq
[10] Google Ngram Viewer (2013).

Sub-Corpus, 1920s; Russian Sub-Corpus, 1920s; English Sub-Corpus 1930s; Russian Sub-Corpus, 1930s, etc. The topical contexts were elicited from query pages or Google Books pages. We resorted to Google Books pages only in cases when the topical context on the query page was not sufficient to reconstruct the structure of the concept. We compiled the corpus from the first editions of the books exclusively and excluded any re-editions on the assertion that they may present a scientific paradigm of another decade.

Table 1
Size of Topical Context Corpus

Period	American Sub-Corpus, tokens	Russian Sub-Corpus, tokens
1925 -1929	2264	1281
1930-1939	2216	unavailable
1940 - 1949	1991	unavailable
1950 -1959	1936	unavailable
1960 - 1969	1755	unavailable
1970 - 1979	1276	2096
1980 - 1989	1886	2182
1990 -1999	1709	2357
2000-2009	1913	2411
2010-2019	2189	2163
Total	**19135**	**12490**

Results and Discussion

Stage 1. Dictionary Meanings of 'Readability' and 'Chitabelnost'

Modern English dictionaries register 'readability' either as a polysemous word defined as "(1) the quality of written language that makes it easy to read and understand; (2) a quality of writing (print or handwriting) that can be easily read"[11] or refer potential readers to the adjective 'readable'[12]. The latter is defined as "able to be read easily: such as a: legible; b: interesting to read[13]".

Dictionaries from earlier periods do not register the noun 'readability' but the adjective 'readable': "RE'ADABLE, *adjective.* That may be read; fit to be read",[14] "RE'ADABLE The state of being readable; readableness"[15]. No substantive changes were made to the definition of 'readable' since that time, although the 1989 edition of Webster's Dictionary separates the meanings of 'legible' and 'readable', which in modern dictionaries are very often defined as synonyms.

illegible, unreadable

This distinction has been noted by many commentators, dating back as far as Utter 1916. Several commentators, including Evans 1957, and Shaw 1975, have also noted that unreadable can sometimes mean "indecipherable," Fowler 1926, Krapp 1927, Partridge 1942, and Phythian 1979 will not allow this sense of unreadable, but it is treated as standard in dictionaries. It was first recorded in 1830. According to our written evidence, its use in current English is extremely rare. In fact, the closest thing we have to recent evidence of its use is a single citation from the magazine Infoworld, in which its meaning is not so much "impossible to decipher" as it is "impossible to see clearly enough for reading": . permitting them to function in reduced light conditions where LCDs [liquid crystal displays] are unreadable —Nancy Groth, Infoworld, 27 Jan. 1986.

[11] Fellbaum, C. (Ed.) (1998). WordNet: An electronic lexical database. MIT Press.
[12] Merriam Webster's collegiate dictionary (1993). https://www.merriam-webster.com/dictionary/readability.
[13] Ibid.
[14] Webster's Dictionary (1828). http:// webstersdictionary1828.com/Dictionary/readable
[15] Online Webster's Dictionary (1913). https:// www.websters1913.com/words/readable

readable

1. interesting or easy to read
2. legible[16]

> To verify the dictionary meanings we used either local or topical contexts of all the periods of study. Over the period 1925–2019, in full correspondence with the meanings registered in American dictionaries, the word readability was used as (1) "the property of being easy or engaging to read": "Our editorial policy is based on a desire to promote readability without sacrificing either scholarly accuracy or a sense of the ease and irregularity of informal correspondence" (Blom & Blom, 1983, p. xix); (2) the quality of being legible or decipherable, the property of print that affects the ease with which a text can be read': "Readability of technical training materials presented on microfiche versus offset copy" (Baldwin, & Bailey, 1971, p. 37).

In the Russian lexicography of 1925–2019, the word 'chitabelnost' was registered for the first time in Volume IV of Explanatory Dictionary of the Russian Language, also called Ushakov's Dictionary, in 1940, although the lexicographer did not provide the definition but referred readers to the derivative adjective, 'chitabelnyi': "CHITABELNYI (coll). Easy, pleasant to read"[17].

'Chitabelnyi', although mistakenly viewed by some linguists as neologism in the Russian language, was for the first time registered in Russian discourse as early as 1910, and the case is very well documented in the Google Books corpus:

> What is this? Are they bad poems? No, not that bad, but only, as they say, in the editorial jargon "readable": you can read them without resentment for the time spent, but they are not kept in one's memory, and could be written by someone else, not Sasha Chernyi. A. V. Amfiteatrov, "On Sasha the Black", 1910[18].
> Modern Russian dictionaries also define 'chitabelnost' as the noun formed from the adjective 'chitabelnyi' (readable)[19].

> > Chitabelnost, coll. Feature attributed to the adjective 'chitabelnyi', suitableness for reading I myself am not young, I am already fifty-six years old, and I still read it (although the material there is much less interesting than before – it used to be 100% "readable" for me, and now it is 70 in the best case) "Knowledge is Power" 1987[20].

The modern entry of the word 'chitabelnyi' in Wiktionary (2018), the free dictionary, registers two meanings which are also marked as colloquial: "1. coll. suitable for reading, worth reading; 2. coll. easy, pleasant to read; readable" [21].

The word is also registered in the electronic Dictionary of Russian ARGO with the meaning 'something you can read (about books, etc.)'[22]. Thus, Russian dictionaries do not register the meaning 'legible', even though we find contexts where it is realized from the late 1970s: "Factors hampering the readability of the "narrow" ("viczo") font, are the following: density, insufficiency of the intra-letter clearance (Trudy Tiflisskogo gosudarstvennogo universiteta, Proceedings of the Tbilisi State University, 1977) "readability test: can I read the text on your business card in poor lighting conditions" (Rezak, 2008).

The context analysis proved that the first registered use of the word 'chitabelnost', which took place around 1920 (see Figure 2), denotes "library book checkouts": "A common method applied by the majority of those studying a reader is based on counting digits demonstrating readability of this or that author in the library".[23]"First

[16] Webster's Dictionary of English Usage (1989). Merriam-Webster
[17] Ushakov, D. N. (1940). Tolkovyj slovar' russkogo yazyka (IV th v.) [Explanatory dictionary of the Russian language]. Gosudarstvennoe izdatel'stvo inostrannyh i nacional'nyh slovarej.
[18] Wiktionary (2018). Chitabel'nost [readability]. https://clck.ru/HSPSi .Chitabel'nyj [readable]. Retrieved from: https://clck.ru/HSPTt.
[19] Slovar' Akademik [Academician dictionary] (2017). https://dic.academic.ru/dic.nsf/ushakov/1088802.
[20] Wiktionary (2018).
[21] Ibid.
[22] Elistratov, V.S. (2000). Slovar' russkogo ARGO [The Dictionary of Russian ARGO]. Russkie slovari. https://clck.ru/HSPkc.
[23] Golos Rabochego Chitatelya [Voice of a Working Reader]. (1929). Krasnaya Nov [Red New], 5, 212. https://clck.ru/HSPG3.

and foremost, the writer's readability is worth mentioning. For instance, one of the mobile libraries reports 21 copies of Lavrenev books to have been read almost 87 times within a month period – that is from January the 10th until February the 14th"[24].

Stage 2. Absolute and Relative Frequencies of the Words 'Readability' and 'Chitabelnost'

Ngram Viewer also provides an opportunity to calculate the absolute frequency of an n-gram (AF) in some particular year in Ngram Viewer with the following formula:

$AF = RF \times 0.01 \times T$ (Google Ngram Viewer, 2013),

where AF is Absolute Frequency,
RF is Relative Frequency,
T is the total number of tokens in the corpus of that particular year.

The total numbers or the raw data for the English corpus are available online in the file 'total_counts' (Risi, 2016). As the total counts for year 1922 is 1413237707 and the relative frequency is 0.0000050374 (see Figure 3) we compute Absolute Frequency of 'readability' (AF^R) as follows:

AF^R = 0.0000050374 x 0. 01 x 1413237707 = 711.90≈712.

In other words, the data point of year 1922 on the graph is caused by 712 appearances of 'readability' in texts published in 1922.

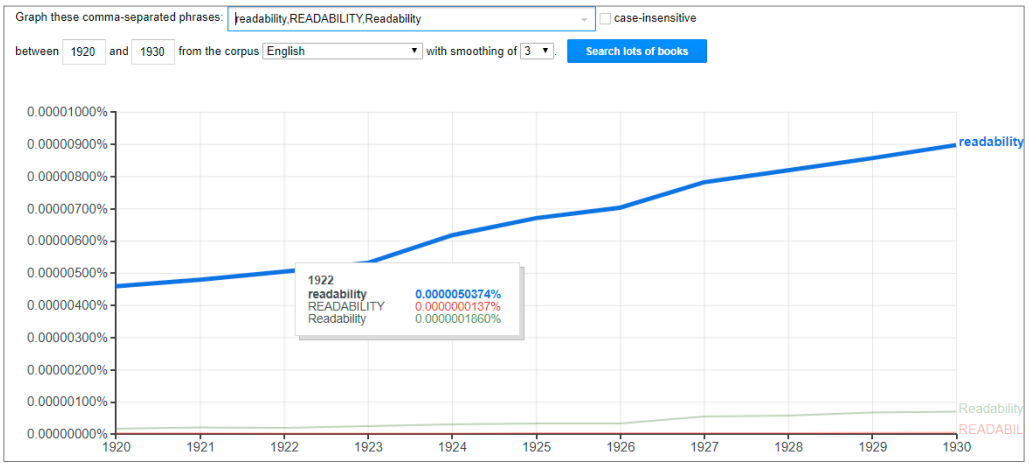
But smoothing makes frequencies look more stable and with the standard setting of 3 for smoothing[25], 712 is in fact the average number of appearances for seven years: three years before 1922 (1919, 1920, 1921), year 1922 and three year after (1923, 1924, 1925).

The same is true for the Russian version of the Ngram Viewer.

The relative frequency of the word 'readability' in American discourse in 1922 was 0.0000050374% (Fig.3 above) i.e. 38.37 times higher than that of the Russian term – 0.0000001530% (see Figure 4. below).

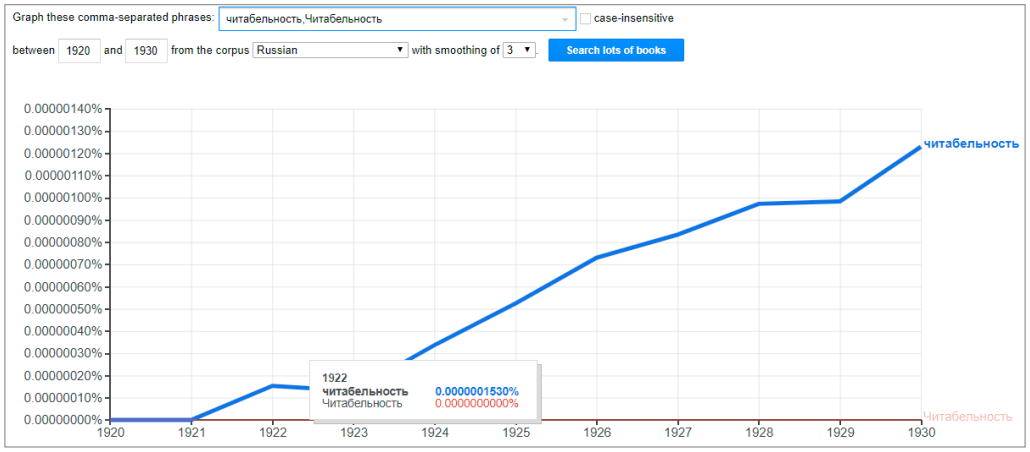
For the purposes of our research we set smoothing to 0, which gave the yearly, not average, values of the relative frequencies of 'chitabelnost' (RF^C) and 'readability' (RF^R) (see Figure 5, Figure 6 below). The procedure also made the peaks on the graph look higher and the pits lower.

Figure 3
Relative Frequency of Readability in the English Discourse in 1922, Standard Smoothing 3



[24] Shteiman, Z. (1928). Literatyrnie episody [Literary episodes]. https://books.google.ru/books?id=3uoGAQAAIAAJ
[25] Google Ngram Viewer.

Figure 4
Relative Frequency of ‘Chitabelnost’ (Readability) in Russian Discourse in 1922, standard Smoothing 3



Another finding of the procedure is the absence of any recorded Russian texts before 1925. In 1925, RF^C was as low as 0.0000009182 (Figure 6 below). The corresponding value on the ‘readability’ graph – RF^R – reached the level of 0.0000065810and that is 6.8 times higher than RF^C. With 1113107246 tokens registered in 1925 (Risi, 2016), we calculate AF^R in the year 1925 as follows:

AF^R in 1925 = 0.0000065819 x 0.01 x 1113107246 = 699.99 ≈700.

As Google Viewer does not provide the raw data for the Russian corpus, we put the actual figures of RF^R and RF^C in ratio terms and simplify the ratio as follows:

$$0.0000065819 : 0.0000009182 = 7.1 : 1.$$

As absolute frequencies are to be in the same ratio, we calculate AF^C:

$RF^R : RF^C = AF^R : AF^C$; $\quad 7.1 : 1 \approx 700 : 102.9 \approx 103.$

Thus, in 1925 we may expect to find 700 records of the word ‘readability’ in American discourse and about 103 records (6.8 times fewer) of the word ‘chitabelnost’ in Russian discourse.

Figure 5
Relative Frequency of the Word ‘Readability’ in English Discourse 1925 – 2000s, Smoothing 0

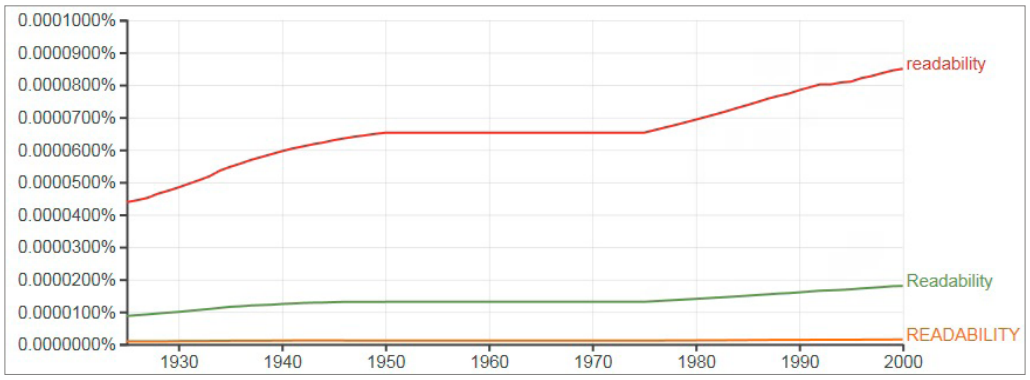
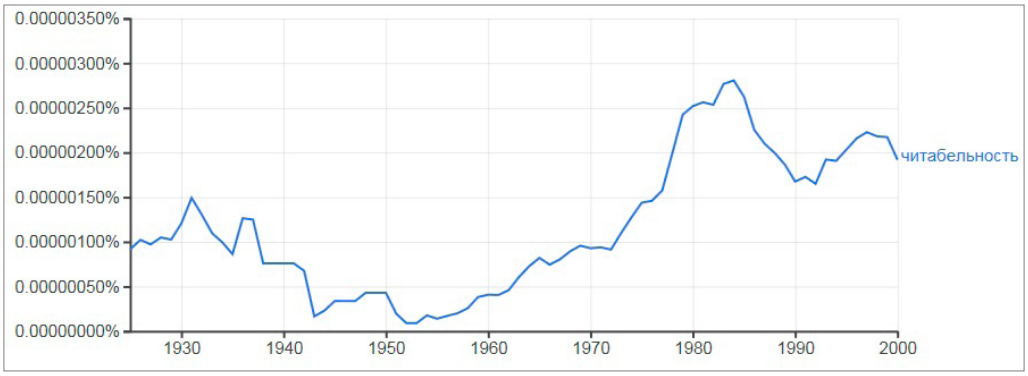


Figure 6
Relative Frequency of the Word Chitabelnost' (Readability) in Russian Discourse in 1925, Smoothing 0



Stage 3. Topical Context Analysis, 1920s–2019

From the late 19th century to the present, Russian and American research schools experienced a number of explosions of interest in the area: in the 1950s and 1980s in the USA and in the late 1920s, 1980s, and 2000s in Russia (see Figures 1, 2 above), representatives of these two schools developed original theories and provided a number of tools to measure text readability. In different years, readability studies focused on various aspects of the effects of textual variables on readers' comprehension.

The rise of the frequency of the word 'readability' in English discourse in the early 1920s (see Figure 1) reflects 'the revolution' in the research area, when the first readability formulas based on linear regression model were published. Those were the years when "librarians and educators realized the need of providing appropriate reading material to people of various reading levels" (Vieth, 1988, p. vii). In those years, close attention was paid to the legibility and illustrations of reading materials as well as vocabulary and sentence length of reading texts (Kitson, 1921). Readability formulas were also developed in the 1920s: the first formula – by Lively and Pressey (1923), then – by Vogel and Washburne (1928). In those years, the concept of readability was viewed as a function of two metrics: word count and sentence length.

A noticeable rise of interest in the problem of text readability in Russian research writing was stimulated mostly due to numerous illiteracy eradication projects, when adults learning to read found children's textbooks offered to them too difficult. In those days, the problem of matching a text and a reader was mainly solved with the help of experts' judgment: "metrics of readability" of a text were aligned in the collection annotated by experts, with the level of text readability being assessed by an expert (see Karpov, 2014). We did not find recordings of any empirical research on the concept of readability in the Russian corpus of Google Books. Paradoxically the adjective 'chitabelnyy', with the meaning of 'readable, easy or enjoyable to read'[26], was used by Lenin in his letter to Karpinsky, probably in 1917: "We decided to publish the attached manifesto instead of not readable theses" (Lenin, 1917, p.10).

In the1930s, Thorndike published a number of profound studies on frequency of words in English discourse: his "Teacher's Word Book of 20,000 Words" (1932) and "Teacher's Word Book of 30,000 Words" (1944) were used as instruments of assessing text readability by generations of educators (Thorndike 1916, 1921, 1932, 1944, 1974). In the English tradition of the 1930s, reading was viewed as one of the aspects of mass communication and Miller developed pros and cons of illustrations in reading texts (see Miller, 1937).

In the 1940s, Johnson made a list of readability factors: "sentence length, difficulty of words, personal pronouns, prepositional phrases, monosyllables, and affixes have an advantage over some of the other factors influencing readability (Johnson, 1947). When evaluating the readability of educational material, scholars in the 1940s also addressed the impression produced by illustrations in children's reading materials, thus interpreting readability as a category related to the quality and number of pictures used in the text (Halbert, 1944; Strang, 1941). In the 1940s, Paterson and Tinker also introduced the term 'relative readability', which was in those times used in reference to the type of font and print that make a text legible or decipherable: "The relative readability of newsprint and book print" (Paterson & Tinker, 1946).

[26] Online Oxford English Dictionary (OED) (3rd ed.) (1996). https://en.oxforddictionaries.com/

In the 1950s, researchers continued to explore the effectiveness of the readability formulas developed in the 1920s: "This evaluative study of readability formulas is based on ratings of 52 books and 3 reading tests <...>. Suggestions are given for making adequate ratings of readability without excessive expenditure of time" (Klare, 1952, p.385). "Recall and prediction scores were correlated with Flesch and Dale-Chall readability scores. All the correlations were positive and high. Both readability formulas showed a higher correlation with learning than with prediction, but the difference was not significant" (Rubenstein & Aborn, 1958, p.28). The range of the metrics offered in those times was rather limited: "Authors of recent formulas have emphasized certain structural aspects of readability: (1) vocabulary level, (2) sentence length and structure, and (3) human-interest words" (Peterson, 1955, p.455).

In the 1960s, readability studies were accelerated by the two outstanding investigations: (1) the Readability Graph named after its designer, Fry, who claimed the Readability Graph to be suitable for estimation of text readability "for all ages, from infant to upper secondary" (Fry, 1968, p. 514); and (2) the SMOG Readability Formula developed in 1969 by a professor at Syracuse University, G. Harry McLaughlin. The formula estimated the age of a reader of a prosaic work based on the calculation of the square root of the proportion of polysyllabic words in the text. The research, known as the SMOG Readability Formula, triggered a number of studies published in English (see Figure 1) and strengthened the notion of readability as "the degree to which a given class of people find certain reading matter compelling and comprehensible" (McLaughlin, 1969).

Thus, the interest in text readability in American discourse was stable throughout the period from the 1930s to the 1960s, as those were the years when researchers widened the range of the parameters influencing text readability and validated readability formulas (see Dale, 1967; Flesch 1949, 1964; Gunning, 1952; Klare & Buck, 1954). Among the 'subjective' factors added to the notion in those years were readers' reading ability, cultural background and knowledge, and readers' motivation.

The Google Books corpus does not provide Russian texts exemplifying the development of the concept 'chitabelnost' from the 1930s to 1960s. We may also argue that American and Russian research paradigms of those times had limited connections: neither reflection on the Flesh readability index nor the SMOG Readability Formula or publications of McLaughlin (1969) are found in Russian discourse before the 1970s.

In the 1970s, Estonian researcher Y. Mikk conducted a series of well-acknowledged studies on the readability of textbooks translated from Russian into Estonian. He grounded three levels of noun abstractness and argued that readability is to be viewed as a function of two variables: the average length of a sentence in the text and noun abstractness. His readability formula, adapted for the Estonian language, had two variables: Readability = (0.131 x average sentence length in characters) + (9.84 x average abstractedness of repeated nouns) – 4.59 (Mikk, 1974).

Mikk's works (1974, 1981) contributed greatly to the development of the concept 'chitabelnost' in Russian discourse. Although in the early 1970s, the term 'chitabelnost' was still functioning with the meaning "the quality of being legible or decipherable" (see Aron, 1972), those were the years when the first empirical studies of readability were successfully conducted in Russia. In 1976, Mackovskij developed and introduced the first readability formula for the Russian language: Readability = (0.62x average sentence length) + (0.123 x % of ⩾3-syllable words) + 0.051 (Mackovskij, 1976).

Unfortunately after the publications of Matskovskiy (1976), Mikk (1981), and Tuldava (1975), interest in the subject gradually faded, the developed readability formulas were not validated on texts of different genre, and for another decade reports on readability experiments were not published. The list of quantitative parameters of Russian texts readability suggested by the Russian scholars of that period was much shorter than the corresponding English one and included the following: the average sentence length (word counts), percentage of words which have more than three syllables, percentage of words of 11 or more letters, noun abstractness (Mackovskij, 1976; Mikk, 1981; Tuldava, 1975).

American readability discourse during this decade developed rapidly, widening the range of text readability variables. The Coleman-Liau index of readability, which appeared in 1975, caused another rise of publications as seen on the graph in Figure 1. The index estimates the number of years of formal education a reader requires to comprehend a text on their first reading and is computed based on two variables, (1) L. i.e. the average number

of letters per 100 words and (2) S, i.e. the average number of sentences per 100 words. Thus, the readability of a text is measured based on the following formula: R= 0.0588 L – 0.296S – 15.8 (Coleman & Liau, 1975).

In the middle of the 1970s, extensive studies were also conducted on the impact of syntax on readability of English texts. Empirical data from Dawkins argued that "because oversimplistic syntax violates basic principles of clear writing, it would be fair to conclude that it is actually difficult to process" (Dawkins, 1975, p.36), thus predicting the introduction of referential and deep cohesion as text readability variables in the 2000s. Although characterizing the state of affairs in the area the scholar also admitted that "Even in our area of syntax (an aspect of readability about which we do know something) we are uncertain about many analyses, lacking in empirical data, amazed by the complexity and variety of elements, clumsy in our methods, and doubtful of our oversimple results" (Dawkins, 1975, p.44).

In the discourse of the mid 1970s, the concept 'chitabelnost' was also actively discussed in the contexts of famous publications of Russian physiologists Zimniya and Leontyev (Zimnyaya, Dridzhe, & Leontyev, 1976). In 1976, Leontyev wrote: "All the variety of factors affecting the readability of printed publications can be reduced to four main groups. These factors are related to: 1) the content of this material; 2) the form or style of its presentation; 3) the organization of the material (the sequence of basic and secondary positions, division into paragraphs, concluding phrases, etc.) and 4) its external design (font, illustrations, cover etc.) (Leontyev, 1976). It was this research that was followed by numerous studies exploring the idea (see Figure 2 above).

The topical context of Russian discourse of *the 1980s* demonstrates the focus of Russian researchers on both qualitative and quantitative text parameters: "Linearity and connectivity are therefore the most essential characteristics of any text, providing its 'readability'. As a guarantee of the existence of the text, 'readability' thus acts as a function of the context" (Filologicheskie nauki [Philological Sciences], 1988, p.160). There were numerous studies on adapting the Flesch-Kincaid formula for the Russian language: "The average readability of text A measured with the Flesch formula is 27.7, text B - 36.9. Thus, the average readability of the textbook is effective, since according to Flesch, texts are readable if their readability is 15-20 or higher" (Mutt, 1984). On the other hand, it was admitted that "In the Russian literature the term 'chitabelnost' is not generally accepted" (Leonov & Elepov, 1986) and researchers used the terms 'trudnost' (Rus. text difficulty), 'prostota' (Rus. text ease), or 'sloznost' (Rus. text complexity) interchangeably.

The peak in the occurrence for the word 'readability' (Figure 1) was registered in the 1980s, with the highest number (~0.000125) in 1986. In 1984, Barr, Pearson, and Kamil published Volume 1 of "The Handbook of Reading Research" where they defined readability as a notion with three separate meanings: "1. Legibility, of either the handwriting or the typography. 2. Ease of reading, owing to the interest-value of the writing. 3. Ease of understanding, owing to the style of writing" (Pearson, 1984, p.681). The scholars argued that "Though the first and second meanings still occur, usage now clearly favors the third meaning, especially in the field of reading" (Pearson, 1984, p.681). Illustrations of that particular sense in the Google Books corpus are numerous: "his novels have few equals in readability, a sometimes deceptive readability" (Killam, 1984, p.192). "Queneau's novels are as remarkable for the sheer readability of their stories as for their other qualities" (p.75) (Shorley, 1985). It was also in the 1980s when P. Fries (1986) equated readability and coherence stating that "mere counting of the language forms contained in a text will not lead to useful judgments of the readability or coherence of that text" (Fries, 1986, p.13).

In the 1990s and early 2000s, the peak from the 1980s was followed by a decrease similar to that observed in the late 1970s. In the 1990s, readability researchers addressed paragraph length as a function of text readability and proved that an average reader has a more positive attitude to texts with short paragraphs of fewer than 100 words than to longer paragraphs (Markel, Vaccaro, & Hewett, 1992). Another study quite extensively quoted in the 1990s was "The effect of syntactic simplification on reading EST texts as L1 and L2" (Ulijn & Strother, 1990). Their empirical study validated the hypothesis that syntactic complexity "does not significantly affect the level of reading comprehension for both expert and novice readers in a particular professional field" (Ulijn & Strother, 1990, p.54). In the 1990s, American discourse registered two main meanings of the word: 1) 'the quality of being legible or decipherable' as in "As illuminance contrast decreased, readability also decreased. However, the relationship between illuminance contrast and readability was direct" (Lomperski, 1995, p. iii); and 2) 'the quality of being easy or enjoyable to read' as in "He examines what each formula is good for and based on vocabulary grading tentatively recommends the Gunning and Fry Readability formula which he successfully applies to

<…>" (Chia, 1998, p.37). Researchers continued discussing readability formulas: "Formally determining grade level according to a standard, computer-based readability formula would heighten the awareness of the those responsible for writing the impartial analysis to the problem of public understanding" (Dubois & Floyd, 1998, p. 178) and Microsoft Word 97 was the first to display readability statistics: Flesch Reading Ease score, Flesch-Kincaid Grade Level score, and passive sentence occurrences. In 1998, Sides uses 'usability' as a synonym of 'readabilty' and made the revolutionary conclusion that "Correctly structured with a feel for rhythm the ebb-and-flow of a phrase, even the long sentence is readable" (Sides, 1999, p.11), thus extending the list of potential metrics affecting readability.

Russian discourse from the 1990s provides examples of 'chitabelnost' in institutional IT discourse where the term is used to modify the word program: "It is also worth noting such advantages of ADA language as modularity, structurality, readability, and documentability of the programs" (Gricenko, 1991, p.65). Topical contexts of the word 'chitabelnost' identify the meaning "the quality of being easy or enjoyable to read[27]", but no modifications to the concept in the 1990s were registered. A possible explanation to the situation could be the reluctance of Russian researchers to use a borrowed word and prefer 'trudnost' (Rus. text difficulty) and 'sloznost' (Rus. text complexity), although verification of this hypothesis is beyond the scope of this article.

In the 2000s, the highest peak of 'readability' frequency took place in 2005 when RFR reached 0.0000998954%, which means the Google Books corpus compiled about 125 texts with the word 'readability' (AFR = 0.0000998954 x 0.01 x 12519922882 = 125.06 ≈ 125). The 2005 vertex was mostly obtained due to "The Principles of Readability" published by DuBay (2004) in August 2004. It soon became one of the most cited articles in the area, which gave a profound boost to studies of readability and readability formulas. Researchers from that time also addressed the disadvantages of using readability formulas[28] and offered online tools to measure readability[29] based on multiple correlation analysis[30].

The Russian research on 'chitabelnost' in those years was mostly focused on websites and texts being legible[31] (http://www.bhv.ru/books/full_contents.php?id=419). Another popular research object was text and technical drawings layout[32,33].

It was also in the *early 2010s* when school textbooks began being assigned 'reading levels'[34] and automatic readability tools started computing readability levels, not only of public speeches (Kayam, 2018; Schumacher & Eskenazi, 2016) but students' writings as well (Peng, 2015). The theme of the decade was the readability of books[35] and webpages (Taylor, 2017; Boztas et al., 2017) with the goal for writers defined by Kirk as "to practice composing a sentence that requires only one reading to decipher the intended message" [36] and their the intended audience.

In the 2010s, a number of researchers in the area validated correlations between companies' documents readability and their business performance "indicating that companies with stronger CSR[37] performance are more likely to have CSR reports with higher readability" (Wang, Hsieh, & Sarkis, 2018, p.66, see also Kim, Wang, & Zhang, 2019; Bonsall & Miller, 2017). Another area of interest at that time was the readability of health information for patients, the representatives of which indicated that writers of medical information on diseases and possible treatment "overestimate the reading ability of the overall population which may "have its greatest impact among those with low literacy and limited access to health care" (Storino et al., 2016, p.831). However, the primary focus of the research *in the late 2010s* was on "the identification of the linguistic features that predict text readability judgments, and how these features perform when compared to traditional text readability formulas such as the Flesch-Kincaid grade level formula" (Crossley et al.,2017, p.340). The new

[27] Lexico. Oxford English Dictionary. https://www.lexico.com/definition/readability
[28] Readability formulas. (2004). http://www.readabilityformulas.com/articles/advantages-and-disadvantages-of-readability-formulas.php
[29] Free readability formulas. (2005). http://www.readabilityformulas.com/articles/advantages-and-disadvantages-of-readability-formulas.php
[30] Methods for measuring text readability. (2005). http://www.standards-schmandards.com/2005/measuring-text-readability/
[31] BHV. Web site svoimi rekami [BHV. Website. DIY]. http://www.bhv.ru/books/full_contents.php?id=419.
[32] Google Books Corporate Social Responsibility. (2004). Readability. https://clck.ru/JBUwT
[33] Google Books. (2004). Readability. https://clck.ru/JBUwp
[34] Google Books. (2005). Readability. https://clck.ru/JBUxD
[35] Google Books. (2015). Readability. https://clck.ru/JBUyF
[36] Kirk, K. (2010). Writing for Readability. Association for Talent Development.
[37] CSR - Corporate Social Responsibility.

findings included evidence that "the traditional readability formulas are less predictive than models of text comprehension, processing, and familiarity derived from advanced natural language processing tools" (Crossley et al., 2017, p. 340).

Thus, to answer the research questions on similarities/differences of the words 'readability' and 'chitabelnost' in American and Russian discourse over the period from 1925 to 2019, we brought together the resources that did not only help to evaluate the development of the concept in two different academic environments, but are also of great scientific value on their own. We used the Google Books Ngram Viewer to observe trends in the usage of the terms 'readability' and 'chitabelnost' as well as the stages of concept formation. The major obstacle to this endeavor was the inaccessibility of some of the resources, which caused low compatibility in the datasets of the 1920s to 1960s.

Conclusion

The study demonstrated language-specific features of the semantic values of the words 'readability' and its Russian analogue 'chitabelnost'. The American word 'readability' has had two meanings verified in the contexts of the studied period: 'the quality of being legible or decipherable' and 'the quality of being easy or enjoyable to read'. The word 'chitabelnost' has been used in Russian discourse with three meanings: two are similar to the English mentioned above, and the third, "a number of library book checkouts", was used only in the late 1920s. Russian dictionaries published between the 1920s and the 2000s register only one meaning of the noun 'chitabelnost', defining it as a derivative of the adjective 'chitabelnyi' (Russian 'readable'). Another difference between the word 'chitabelnost' and 'readability' is that it is marked in the dictionaries as colloquial, but its active functioning in Russian academic and scientific discourse over the period from the 1920s to the 2000s testifies to its belonging to the high register of communication.

The research also showed that debates over the concept of readability have involved a certain level of disagreement on the range of variables affecting it. All the existing approaches to the notion of 'readability' differ in the number of parameters they explore. American and Russian schools experienced peaks of activities in the 1980s when scientists advanced the concept to the next level of its development and added a number of new constituents to its structure. By the 2000s, the American concept of 'readability' had manifested in a wide range of scientific publications and was viewed as a function of lexical, syntactic, semantic, and pragmatic parameters of the text. As for the Russian school of text readability of that period, their records of achievements were mostly gained in the 1970s and 1980s. By the 2000s, the Russian school had accumulated extensive data on qualitative parameters of readers' cognitive and linguistic abilities, but possessed limited resources in the quantitative domain: text readability was estimated based on four variables only, i.e. word count, word length, sentence length, and noun abstractness as functions.

The features summarized in the article as 'predicting readability' provide a stable ground for further research on the metrics and parameters of different readability levels both in English and Russian.

Acknowledgements

This research was financially supported by the Russian Science Foundation, grant No 18-18-00436.

Conflict of interests

The authors declare that they have no conflict of interest.

References

Aron, E. I. (1972). *Metody issledovaniya i proyektirovaniya organizatsii truda na predpriyatii* [The methods of research and planning the work organization in the enterprise]. https://books.google.ru/books?id=K7cwAQAAIAAJ –

Baldwin, T. S., & Bailey, L. J. (1971). Readability of technical training materials presented on microfiche versus offset copy. *Journal of Applied Psychology, 55*(1), 37-41. http://psycnet.apa.org/record/1971-21933-001

Blom, M., & Blom, T. (1983). *Canada Home: Juliana Horatia Ewing's Fredericton Letters*, 1867-1869. University of British Columbia Press.

Bonsall, S. B., & Miller, B. P. (2017). The impact of narrative disclosure readability on bond ratings and the cost of debt. *Review of Accounting Studies, 22(2),* 608-643. https://doi.org/10.1007/s11142-017-9388-0

Boztas, N., Omur, D., Ozbılgın, S., Altuntas, G., Piskin, E., Ozkardesler, S., & Hanci, V. (2017). Readability of internet-sourced patient education material related to "labour analgesia". *Medicine, 96*(45), 1-5. https://clck.ru/JBxuL

Chia, E.N. (1998). Guide to readability in african languages. *Linguistic Society of America, 74*(3), 665. https://muse.jhu.edu/article/451344/pdf

Choldin, M. T. (1979). Rubakin, Nikolai Aleksandrovic. In A. Kent, H. Lancour & J. E. Daily (Ed.), *Encyclopedia of library and information science* (pp. 178-179). CRC Press.

Coleman, M., & Liau, T. L. (1975). A Computer Readability Formula Designed for Machine Scoring. *Journal of Applied Psychology*, *60*(2), 283-284. https://doi.org/10.1037/h0076540

Crossley, S. A., Skalicky, S., Dascalu, M., McNamara, D. S., & Kyle, K. (2017). Predicting text comprehension, processing, and familiarity in adult readers: New approaches to readability formulas. *Discourse Processes, 54(5-6),* 340-359. https://doi.org/10.1080/0163853X.2017.1296264

Dale, E. (1967). *Can you give the public what it wants?* World Book Encyclopedia.

Dash, N. S. (2008). Context and contextual word meaning. *SKASE Journal of Theoretical Linguistic, 5,* 21-31. http://www.skase.sk/Volumes/JTL12/pdf_doc/2.pdf

Dawkins, J. (1975). *Syntax and Readability*. International Reading Association. https://files.eric.ed.gov/fulltext/ED103814.pdf

DuBay, D. H. (2004). The principles of readability. https://ru.scribd.com/document/346200577/Dubay-principles-of-readability-pdf.

Dubois, P. & Floyd, F. (1998). Lawmaking by initiative: issues, options, and comparisons. Agathon.

Fellbaum, C. (Ed.) (1998). *WordNet: An Electronic Lexical Database*. MIT Press.

Fernbach, N. (1990). La lisibilité dans la reaction juridique au Québec [Legibility in the legal reaction in Quebec]. Le Centre de promotion de la lisibilité, Centre Canadiend'informationjuridique.

Filologicheskie nauki [Philological Sciences]. (1977). Trudy Tiflisskogo gosudarstvennogo universiteta [Proceedings of the Tbilisi State University]. Tbilisi University Press.

Filologicheskie nauki [Philological Sciences]. (1988). *Nauchnye doklady vysshej shkoly* [Scientific Reports of the Higher School]. Vysshaya shkola.

Firth, J. R. (1957). A synopsis of linguistic theory. *In Philological society: Studies in Linguistic Analysis* (special vol., pp. 1-32). Basil Blackwell.

Flesch, R. (1949 and 1974). *The art of readable writing*. Harper.

Flesch, R. (1964). *The ABC of style: A guide to plain English*. Harper.

Fries, P. (1986). Language features, textual coherence and reading. *Word, 37,* 13-29. https//doi.org/10.1080/00437956.1986.11435764

Fry, E. B. (1968). A readability formula that saves time. *Journal of Reading, 11,*513-516. http://www.jstor.org/stable/40013635

Gray, S. (2019). Readability of Patient-reported Outcome Measures for Persons with Aphasia. https://clck.ru/JAtzz

Gricenko, V.I. (Ed.) (1991). *Upravlyayushchie sistemy I mashiny* [Control systems and machines]. Nauk. Dumka.

Gunning, R. (1952). *The technique of clear writing.* McGraw-Hill.

Hai-Jew, S. (2014). *Querying google books ngram viewer's big data text corpuses to complement research.* https://www.researchgate.net

Halbert, M. G. (1944). The teaching value of illustrated books. *American school board journal, 108*(5)*,* 43-44.

Hiebert, E. H. (2012). *The common core state standards and text complexity*. Text Project & University of California.

Ivanov, V. V. (2013). K voprosu ob ispolzovanii lingvisticheskikh kharakteristik slozhnosti teksta pri issledovanii okulomotornoy aktivnosti pri chtenii u podrostkov [Back to the linguistic characteristics of the complexity of the text in the study of oculomotor activity in reading in adolescents]. *New research, 2,* 42-50. https://cyberleninka.ru/article/

Johnson, K. G. (1947). *Factors influencing the readability of science*. University of Wisconsin.

Karpov, N. V. (2014). *Identifikatsiya urovnya slozhnosti teksta i ego adaptatsiya* [Identification of the complexity level and its adaptation]. http://www.slideshare.net/karpnv/ss-31225145#14356960593761&fbinitialized

Kayam, O. (2018). The readability and simplicity of Donald Trump's language. *Political Studies Review, 16*(1), 73–88. https://doi.org/10.1177/1478929917706844
Kerlinger, F. N. (1973). *Foundations of behavioral research* (2nd ed.). Holt Rinehart and Winston.
Killam, G. D. (Ed.) (1984). *The writing of East and Central Africa*. Heinemann Educational Books Ltd.
Kim, C., Wang, K., & Zhang, L. (2019). Readability of 10-K reports and stock price crash risk. *Contemporary accounting research*, *36(2),* 1184-1216. https://doi.org/10.1111/1911-3846.12452
Kitson, H. D. (1921). *The mind of the buyer*. Macmillan.
Klare, G. R. (1952). Measures of the readability of written communication: an evaluation. *Journal of Educational Psychology, 43*(7), 385-399. https://doi.org/10.1037/h0058972
Klare, G. R. (2000). Readable computer documentation. *ACM journal of computer documentation, 24*(3), 148–168. https://doi.org/10.1145/344599.344645
Klare, G.R., & Buck, B. (1954). *Know your reader, the scientific approach to readability*. Hermitage House.
Lenin, V. I. (1917). *Polnoye sobraniye sochineniy* [Complete Works]. http://uaio.ru/vil/49.htm
Leonov, V. P., & Elepov, B.S. (1986). *Referirovanie I annotirovanie nauchno-tekhnicheskoj literatury* [Precis-writing and annotating of the scientific and the technical literature]. Nauka.
Leontyev, A. A. (1976). *Smyslovoe vospriyatie rechevogoso obshcheniya (v usloviyah massovoj kommunikacii)* [Semantic perception of speech communication (in the conditions of mass communication)]. Nauka.
Lively, B. A., &Pressey, S. L. (1923). A method for measuring the 'vocabulary burden' of textbooks. *Educational administration and supervision*, *9*, 389-398. https://doi.org/10.1080/00220671.1931.10880190
Lomperski, T. J. (1995). *Enhancing interior building sign readability for older adults: Lighting color and sign color contrast*. University of Wisconsin.
Macias-Chapula, C. A. (Ed.) (1999). *Seventh conference of the international society for scientometrics and informetrics*. Universidad de Colima
Mackovskij, M. S. (1976). *Problemy chitabel'nosti pechatnogo materiala. Smyslovoe vospriyatie rechevogo soobshcheniya v usloviyah massovoj kommunikacii* [Problems of readability of printed material. Semantic perception of speech messages in conditions of mass communication]. Nauka.
Markel, V., Vaccaro, M., & Hewett T. (1992). Effects of paragraph length on attitudes toward technical writing. *Technical Communication, 39*(3), 454-456. http://www.jstor.org/stable/43090117
McLaughlin, H. (1969). SMOG grading a new readability formula. *Journal of reading, 22*, 639-646. http://webpages.charter.net/ghal/SMOG_Readability_Formula_ G._Harry_McLaughlin_(1969).pdf.
Mikk, J. A. (1974). Metodika razrabotkif ormul ˇchitabel'nosti [Methods for developing readability formulas]. *Sovetskaja Pedagogikai Skola*, *9,* 78-163.
Mikk, Y. A. (1981). *Optimizaciya slozhnosti uchebnogo teksta: V pomoshch' avtoram i redaktoram* [Optimizing the complexity of the text: To help authors and editors]. Prosveshchenie.
Miller, W. A. (1937). Reading with and without pictures. *Elementary School Journal, 38,* 676-682. https://doi.org/10.1086/462248
Mutt, O. (1984). *Aktual'nye voprosy otbora uchebnogo materiala dlya vuzovskogo kursa inostrannogo yazyka* [Actual questions of selection of the educational material for the university course of a foreign language]. The Tartu State University.
Paterson, D. G., & Tinker, M. A. (1946). The relative readability of newsprint and book print. *Journal of Applied Psychology, 30*(5), 454-459. https://doi.org/10.1037/h0054012
Pearson, P. D. (Ed.) (1984). *Handbook of reading research*. Lawrence Erlbaum Associates.
Pechenick, E. A., Danforth, C. M., & Dodds, P. S. (2015). Characterizing the Google Books Corpus: Strong limits to inferences of socio-cultural and linguistic evolution. *PLOS ONE, 10*(10), *e0137041.* https://doi.org/10.1371/journal.pone.0137041
Peng, C. (2015). Textbook readability and student performance in online introductory corporate finance classes. *The Journal of Educators Online-JEO, 13*(2), 35-49. https://doi.org/10.9743/JEO.2015.2.6
Peterson, E. M. (1955). Aspects of readability in the social studies. *Journal of Applied Psychology 39*(2), 141-142. https://doi.org/10.1037/h0039753
Randal, R. (Ed.) (2013). *Novel and short story writer's market.* Writer's digest books.
Ravin, Y., & Leacock C. (2000). *Polysemy: Theoretical and computational approaches*. Oxford University Press.
Rezak, D. (2008). *Svyazi reshayut vse. Biznes-skazka o Carevne-lyagushke* [Communications solve everything. Business tale about the Frog Princess]. MIF.
Risi, S. (2016). *Google Ngrams: From relative frequencies to absolute counts.* http://stanford.edu/~risi/tutorials/absolute_ngram_counts.html.
Rubenstein, H., & Aborn, M. (1958). Learning, prediction, and readability. *Journal of Applied Psychology, 42*(1),

28-32. https://doi.org/10.1037/h0039808.
Schumacher, E., & Eskenazi, M. (2016). *A readability analysis of campaign speeches from the 2016 US Presidential campaign.* Carnegie Mellon University.
Sherman, L.A. (1893). *Analytics of literature: A manual for the objective study of English prose and poetry.* Ginn and Co.
Shorley, Ch. (1985). *Queneau's fiction: An introductory study.* Cambridge University Press.
Sides, C.H. (1999). *How to write and present technical information*. Cambridge University Press.
Sologub, O. P. (2002). *Usvoenie inoyazychnyh strukturnyh ehlementov v russkom yazyke* [The assimilation of foreign language structures in the Russian language]. In Proceedings of the 2002 the 3rd Scientific Conference (pp. 130-134). Nauka.
Solovyev, V. D. (2013). A frequency-based approach to language dynamics. https://events.spbu.ru/eventsContent/files/corpling/corpora2013/Solovjov.pdf.
Storino, A., Castillo-Angeles, M., Watkins, A. A., Vargas, C., Mancias, J. D., Bullock, A., Moser, A. J., Demirjian , Demirjian, A., & Kent, T. S. (2016). Assessing the accuracy and readability of online health information for patients with pancreatic cancer. *JAMA surgery*, *151*(9), 831-837. https://doi.org/10.1001/jamasurg.2016.0730
Strang, A. (1941). *A study of gains and losses in concepts as indicated by pupils' reading scores after the addition of illustrations to reading material* [Unpublished doctoral dissertation]. Temple University.
Taylor, M. C.,& Wahlstrom, M. W. (1986). Readability as applied to the assessment instrument. *International Journal for Basic Education*, *10*(3), 155–170. http://en.copian.ca/library/research/report4/rep31-35/rep34.pdf
Taylor, Z.W. (2017). A failure to communicate: are university websites readable? *International Journal of Education and Multidisciplinary Studies 8*(2), 149-163. doi: http://dx.doi.org/10.21013/jems.v8.n2.p1.
Thorndike, E. L. (1916). An improved scale for measuring ability in reading. *Teachers college record, 17,* 40-67.
Thorndike, E. L. (1921). *The teacher's word book*. Columbia University.
Thorndike, E. L. (1932). *A teacher's word book of 20,000 words*. Columbia University.
Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. Columbia University.
Thorndike, R. L. (1973-1974). Reading as reasoning. *Reading research quarterly, 9,* 135-137. Columbia University.
Tuldava, Yu. A. (1975). Ob izmerenii trudnosti tekstov [About the measurement of the difficulty of texts]. *Uchenye zapiski Tartuskogo universiteta, 345*(IV), 102-120. http://dspace.ut.ee/bitstream/handle/10062/25443/acta_827_1988.pdf.
Ulijn, J. M., & Strother, J. B. (1990). The effect of syntactic simplification on reading EST texts as L1 and L2. *Journal of research in reading, 13*(1), 38-54. https://eric.ed.gov/?id=EJ405025
Vari-Cartier, P. (1916). Development and validation of a new instrument to assess the readability of Spanish prose. *The Modern Language Journal, 65,* 141-148. https://clck.ru/HSPNr
Vieth, N. A. (1988). *Mathematics in composition: A defense of Flesch's readability formula* [Unpublished MA thesi, Iowa State University]. https://clck.ru/HSPQD
Vogel, M., & Washburne, C. (1928). An objective method of determining grade placement of children's reading material. *Elementary school journal, 28,* 373-381. https://doi.org/10.1086/456072.
Wang, Z., Hsieh, T. S., & Sarkis, J. (2018). CSR performance and the readability of CSR reports: Too good to be true? *Corporate Social Responsibility and Environmental Management, 25(1), 66-79*. https://doi.org/10.1002/csr.1440
Wierzbicka, A. (1996). Semantics: Primes and universal. Oxford University Press.
Willems, K. (2012). Intuition, introspection and observation in linguistic inquiry. *Language Sciences, 34,* 665-681. https://doi.org/10.1016/j.langsci.2012.04.008
Zakharov,V., & Masevich, A. (2014). Diachronic investigations on the base of Russian corpus of Google Books Ngram Viewer. https://www.academia.edu/8592136
Zhang, S. (2015). *The Pitfalls of using Google NGram to study language*. https://www.wired.com.
Zimnyaya, I. A., Dridzhe, T. M., & Leontyev, A. A. (1976). *Smyslovoye vospriyatiye rechevogo soobshcheniya: V usloviyakh massovoy kommunikatsii* [Semantic perception of the voice message: In the context of mass communication]. https://books.google.ru/books?id=SI8ZAAAAMAAJ.

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 1, 2020

Syed, H. (2020). Factors Preventing in-service University English Language Teachers from Becoming Action Researchers in Pakistan. *Journal of Language and Education, 6*(1), 120-133. https://doi.org/10.17323/jle.2020.9697

Factors Preventing in-service University English Language Teachers from Becoming Action Researchers in Pakistan

Hassan Syed
Sukkur IBA University

Correspondence concerning this article should be addressed to Hassan Syed, Sukkur IBA University, Airport Road, Sukkur Sindh, Pakistan. E-mail: hassan.ali@iba-suk.edu.pkd

Research suggests that TR enables classroom teachers to address their classroom-specific problems without having to rely too heavily on published research. However, despite the fact that TR narratives of language teachers have increased lately, there is still need for studies investigating the perceptions of university English language (EL) teachers about TR, and exploring the factors affecting their engagement with TR in culturally diverse contexts. The current study was designed to explore the perceptions of university EL practitioners about the factors that inhibit or encourage them to engage in TR in a university context in Pakistan. Data were obtained from fifteen EL teachers from four public sector universities through semi-structured interviews. Results show that teachers appeared to possess only simplistic knowledge of research as an activity aimed at finding something new. While a majority admitted to have done no research, even the engagement of others who claimed to have done some research seemed sporadic and less than systematic. The main factors responsible for teachers' dis-engagement from TR included academic culture, management's attitude, power relations between senior and junior faculty, workload, lack of monetary benefits, and large classes. A number of implications have also been discussed.

***Keywords*:** teacher research, english language teaching, higher education, teacher education, education policy

Introduction

Teacher research (TR) is defined as a systematic, intentional and rigorous enquiry (individual or collaborative) conducted by teachers to address their classroom-specific issues (Borg & Sanchez, 2015; Cochran-Smith & Lytle, 1993). The idea of TR emerged in reaction to the traditional knowledge transmission models of education (Burns, 2005; Cochran-Smith & Lytle, 1998), and the dependency of classroom teachers on professional research for solving classroom-specific pedagogical problems (Kasi, 2010). For over a decade there was little empirical evidence in the form of English language (EL) teachers' narratives regarding their engagement in research to believe in the transformational potential of TR in real-life classroom contexts. Lately, however, the sheer number of edited publications involving EL teachers' narratives and accounts of action research bears testimony to the increasing popularity and overwhelming reception of TR by classroom teachers in various contexts, for example, Japan, China, Turkey to name but a few (Borg & Sanchez, 2015; Etherington & Daubney, 2018; Levkina, 2018; Saglam & Dikilitas, 2019; Slimani-Rolls, 2018; Suzuki, 2018).

Studies have shown that TR has the potential to empower EL teachers by developing their research skills, and help them become autonomous decision makers, consultants, curriculum developers, analysts and activists (Atay, 2008; Barkhuizen, Burns, Dikilitaş, & Wyatt 2018; Cochran-Smith & Lytle, 1998; Farreh & Saeed, 2011; Kirkwood & Christie, 2006; Macaro, 2003; Trent, 2010; Tuyan, 2018). Research has found multifarious factors that inspire or prevent teachers from engaging in TR, such as teachers' conceptions of research, and their willingness to solve students' problems, research skills, appropriate mentoring, collegial and organizational support, and access to research material (e.g. books and journals) (Borg, 2009; Magos, 2012; McDonough, 2006; Moore, 2011; Tien, 2010; Sanchez & Borg, 2015).

In the context of Pakistan, for example, while the government emphasises its commitment to promoting innovative research in higher education institutions (HEIs) (National Education Policy, 2017)[1], anecdotal evidence suggests that most of the teachers hardly engage in TR to solve classroom-specific problems. According to a report by British Council (2018) on the research culture in the HEIs in Pakistan, some of the senior EL faculty with research skills consider research to be a solitary activity and a number game; they conduct research for their career growth, i.e. promotion and other incentives. Conversely, encumbered with teaching workload the junior EL faculty with little knowledge of TR and lack of research skills in general seem to consider research to be unrewarding because faculty promotions in higher education are subject to Mphil and PhD rather than research productivity[2] (Halai, 2011; Khan, Shah, & Khan, 2018). They are, therefore, more concerned about coping with their teaching workload and dealing with exam-oriented education system as well as improving their qualifications. In addition to that, despite 300 teacher training resource centers, 279 pre-service and in-service training institutions in the country, out of which 68 are in Sindh province alone (Coleman, 2010[3]; Halai, 2015), the obsolete pre-service training structure and…less than adequate in-service training regime with only limited or no TR training is also one of the major stumbling blocks to TR in Pakistan (National Education Policy, 2017).

Although recently a few teacher education institutions in Pakistan have introduced TR in their degree programmes, it is by no means the case in teacher training programmes in general (Halai, 2011; Shamim, 2008). What is even more worrisome is that there are only few studies reflecting upon the current state of the challenges EL teachers face in conducting classroom research in the higher education in Pakistan. The ones that are available are rather dated (Halai, 2011; Kasi, 2007; Shamim, 1996), and focus mainly on primary and/or secondary school teachers in urban areas, e.g. Karachi (Shamim, 1996). Thus, the potential of TR by teachers engaged in English language pedagogy still remains under-exploited in the context of Pakistan. Therefore, there is a need for more context-specific studies into factors affecting teachers' research engagement in higher education in order to make teaching a research-informed practice in Pakistan. The present study, therefore, aims to contribute to literature by exploring factors affecting EL teachers' engagement in TR in higher education in Pakistan.

Factors Influencing Teachers' Engagement in Classroom Research (CR)

Research suggests that a number of factors, such as professional, cognitive, attitudinal, technical, logistic and others, contribute to teachers' willingness or unwillingness to conduct teacher research (Sanchez & Borg, 2015). Professional development, promotion and employer expectations have been reported as the most important factors that motivate teachers to do research in their own classrooms (Atay, 2008; Borg, 2007; 2009; McDonough, 2006; Wyatt, 2011). For instance, MCDonough (2006) investigated the experiences of graduate teaching assistants' (TAs) experiences of conducting action research in a second and foreign language teaching context. The results showed that participants reported that they developed a better understanding of research and insights into language teaching practices inside the classroom.

More recently, Tuyan (2018) conducted a study with seven instructor mentees involved in action research in a Turkish context. Based on a survey, the study showed that participants engaged in action research experienced professional development, developed understanding of research and sensitivity to students' problems, and understood the significance of collaboration. The study argued that teachers' motivation to remain engaged in TR is associated with professional gains, a sense of enjoyment in critical self-reflection, and self-empowerment. It is also noteworthy that teachers' engagement with classroom research also depends on their expectations of types of rewards and their personal disposition regarding how they perceive their professional development (Worral, 2004; Tien, 2000). For example, Worral (2004) conducted a study into school teachers' perceptions regarding the impact of TR on their professional development. Participants in the study reported that research engagement for them proved to be unrewarding as it failed to have any significant impact on their teaching practices.

Besides, TR faces enormous challenges including teachers' conceptions of research, attitudinal factors, such as management's attitude, lack of mentoring and collegial support, technical factors, for instance, lack of research

[1] National Education Policy (2017). Government of Pakistan, Ministry of Education, Islamabad. https://planipolis.iiep.unesco.org/sites/planipolis/files/ressources/pakistan_national_education_policy_2017-2025.pdf

[2] British Council of Pakistan (2018). The university research system in Pakistan. https://www.britishcouncil.pk/sites/default/files/the_university_research_system_in_pakistan.pdf

[3] Coleman, H. (2010). Teaching and Learning in Pakistan: The Role of Language in Education. The British Council. https:// www.britishcouncil.org/pakistan-ette-english-language-report.pdf

skills, and other factors, including lack of time have been reported to have a debilitating effect on TR. The main cognitive barriers involve teachers' positivist conceptions of research, the dichotomy between teaching and research, and the myth that good teachers do not have questions (Cochran-Smith & Lytle, 1992; Macaro, 2003; McDonough, 2006; Schapper & Mayson, 2010). Research has shown that a majority of teachers subscribe to a somewhat positivistic research paradigm considering objectivity, hypothesis, statistics and generalizability as the main criteria of good quality research (Borg, 2007; 2009; Shkedi, 1998). They consider research to be a highly complicated and technical activity (McDonough, 2006). For example, Borg (2009) conducted a study into EL teachers' conceptions of research and factors influencing their research into classroom-specific issues. The data was collected from a cross section of EL teachers (n=505) of varying qualifications and language teaching experiences from across the world. Results showed that participants' definition of research included tests, experiments, peer-observations, and publications as essential elements. Thus, teachers' positivistic notions about research were the main hurdles preventing them from TR.

Furthermore, research has also shown that management's attitude, lack of research skills, and lack of time as some of the recurrent factors hindering TR. For instance, research shows that management's attitude has a strong bearing on teachers' engagement in CR (Allison & Carey, 2007; Borg & Alshumaimeri, 2012; Worral, 2004; Xie, 2015). However, the role of management varies from context to context: In some contexts management was encouraging but not supporting; in other contexts the management was encouraging and supporting teachers to do research, while in yet other contexts it was neither encouraging nor supporting it (Allison & Carey, 2007; Borg & Alshumaimeri, 2012; Everton, Galton, & Pell, 2000; Taylor, 2007; Worral, 2004). For instance, Taylor's (2007) study examined the perceptions of both teachers and management of four universities in England and Sweden. The study revealed that the management and policy making staff in all the four universities had a firm belief that teaching and research are intertwined and the nexus between teaching and research is beneficial for teachers as well as the institution. On the contrary, in a study conducted by Borg and Alshumaimeri (2012) with teacher educators in a university context in Saudi Arabia, an overwhelming majority (i.e. 79.3%) reported that while their management expected them to do research, it did not offer any material or moral support to the teachers interested in Classroom Research. in some contexts management's expectations were reported to have a low impact on TR (Borg, 2009), while a relatively high impact in other contexts.

Similarly, lack of research skills, and mentoring, also emerged as the most commonly cited constraints in TR (Anderson & Herr, 1999; Atay, 2008; Xie, 2015). Xie (2015), for instance, conducted a TR project with a purpose to enhance second language learners' motivation to attend her classes. According to her, the main hindrances she experienced while doing TR involved a lack of knowledge and research skills; lack of mentoring and support from colleagues; and pressure from exam-oriented educational system. Based on her experiences, she argued that in order for TR to thrive, the teacher-researchers should be encouraged to work together, and receive adequate moral and technical support from the management. Additionally, a lack of time, and pressure from children's parents have also been cited for teachers' disengagement from TR (Atay, 2008; Gow & Chow, 2012). For example, Atay's (2008) study into English language teachers (n=62) in a university in Turkey reported excessive workload (between 18 and 21 hours), and consequently, lack of time to be the two main reasons for teachers' disengagement from TR.

Most notably, the teachers blamed management for having no 'release time or support' to do research (Christenson et al., 2002). Some studies reported that teachers did not do research because there was no institutional pressure on them to do so (Allison & Carey, 2007; Worral, 2004). Additionally, the pressure from students' parents also affects teachers' motivation to engage in TR. For example, Magos' (2012) study into TR in Greece found students' age and stress caused by students' parents as the main factors preventing TR. The stress caused by parents' intervention in teachers' work did not allow them to engage in TR.

Studies into TR in Pakistan have also reported a host of factors that prevent affect teachers' intentions to do TR, such as lack of research skills, large classes, institutional expectations, and teachers' perceptions of research, and the pressure on teachers for the completion of the syllabus (Halai, 2011; Kasi, 2010; Sarwar, 2001; Nadeem, 2011). For instance, Halai (2011) conducted a qualitative meta-synthesis of studies into TR in a university context in Pakistan. The study showed that TR is widely regarded by practitioners as messy, tedious and time consuming. The study also highlighted management negatively affected teachers' willingness to engage in TR by forcing them to choose between the role of teacher and researcher, and invest their energy and time on getting their students to score higher marks in the board exams.

Notwithstanding its benefits, TR has faced criticism on account of lacking validity and reliability. This line of argument holds that research conducted by teachers in highly specific settings on a relatively smaller scale involving a small number of students/respondents tends to be subjective, circumstantial and pedagogically ineffectual since it fails to ensure generalisability (Burns, 2005; Cochran-Smith & Lytle, 1998). It is argued that if TR is to be regarded as research, it must be quantitative, naturalistic, value-free, and objective. In contrast to that, others adopt a constructivist and interpretivist perspective considering TR to be inherently subjective, interpretive and context-bound inquiry. According to them, TR has its own procedures of validity, that is, instead of internal/external validity, reliability and objectivity, it works with credibility, transferability, and conformability (Dornyei, 2007; Kirkwood & Christie, 2006). Thus, making positivist demands on TR, especially when TR does not subscribe to or claim to follow it, is unrealistic and unjustified (Borg & Sanchez, 2015).

The review of literature illustrates that while TR is beneficial for teachers' professional and personal development, it faces numerous challenges, including teachers' misconception regarding TR being too technical and statistical, as well as the dichotomous relationship between teaching and research (Allwright, 2005); management's discouraging attitude; lack of mentoring and research skills; and lack of time (Borg, 2009; Wyatt, 2011). What is conspicuously missing in the literature is a discussion on the role and nature of research culture prevalent in academia. Given the complexity and specificity of each classroom context, and TR being an attempt at understanding and solving classroom-specific issues, more context-specific studies are needed to reveal the challenges faced by TR in other contexts. In Borg's (2006) words, 'if teacher research is to become an integral part of teachers' professional practices, then it needs to extend beyond such settings and purposes' (p. 23). The present study, therefore, aims to explore the perceptions of English language teachers about the factors which influence TR in Pakistan. The main questions this study aims to investigate are:

Research Questions

RQ 1 What are the perceptions of in-service university English language teachers about classroom research in Sindh Pakistan?

RQ 2 What are the factors which influence in-service university English language teachers to do research in their own classrooms in Sindh Pakistan?

Method

Context and Participants

The study was conducted in four public sector universities in the Sindh province of Pakistan. Sindh offered a fertile ground for obtaining data due to a significant number of universities in the province (n=47) including 17 public and 30 private universities. Moreover, it was due to lack of time and financial resources, and access to research participants were the main reasons the present study was not conducted in any other province. Since I belong to Sindh province, moving from Sindh to any other province involved financial and contextual challenges. For instance, living in a different province for the period of data collection might involve unforeseen expenses. Secondly, due to lack of familiarity with participants, obtaining consent from institutions and the English language teachers for participating in the study could be another challenge overcoming which could take a long time and resources. However, since I had been a teacher in English in various universities of the Sindh province, it was relatively easier for me to get access to the participants personally, or through common friends and acquaintances. Thus, emails for participation in the study were sent to the English faculty of four (n=4) public sector universities of the Sindh province of Pakistan.

Most importantly, since the aim of the study was to understand the perceptions of English language teachers serving in university, participants were recruited using purposive and convenience sampling. EL teachers of four local universities were invited by email to participate in the study. However, due to low response from the population, personal and friends' contacts were used to get the interviews with teachers. Initially, I approached the teachers I had worked with in the past since they were less likely to conceal any information with regard to their engagement in research, attitude of faculty members and management. It was seemingly difficult to get the data from teachers who were not familiar with me personally. There were chances that some of the participants

might not agree to reveal their thoughts and experiences of TR. Therefore, friends' contacts and references were used to reach out to teachers from Uni-3 (Newby, 2010). Detailed information about participants is given in Appendix-1.

Fifteen EL teachers from four universities volunteered to participate in the study. Three types of faculty participated in the current study: Permanent faculty with more than 10 years of language teaching experience at a university (henceforth senior faculty); permanent faculty with less than 8 years of teaching experience at a university (henceforth junior faculty with a permanent job)' and faculty working on a contract basis (henceforth junior faculty on a contract). There were significant differences in the privileges and workload of the different types of faculty. Senior faculty received higher salary and relatively less teaching workload; while the junior permanent faculty received better salary compared to those on a contract, their amount of workload was not much different from contractual faculty; lastly, the junior faculty on contract were given relatively higher workload but paid significantly lower amount for their services than the permanent junior faculty.

Six of the participants (i.e. T1, T2, T3, T4, T5, T6) were from a public sector university (henceforth Uni-1) in the Khairpur region of the province. Uni-1 offers undergraduate and postgraduate courses in arts, sciences, and business. The English department of this university offers courses in both literature and linguistics to undergraduate and postgraduate students. The department had a total of fifteen faculty members, including five permanent teachers, including four senior faculty and one junior faculty, while others were on contract teachers whose contract was renewed every six months. Two of the senior faculty possessed PhD in linguistics, seven had a master of arts (MA) in literature, while the rest held an MA in linguistics. The linguistics faculty taught courses, including introduction to linguistics, introduction to second language acquisition, psycholinguistics, and sociolinguistics. The teachers in English literature taught courses, such as English drama, classical poetry, English prose, etc. Most notably, the permanent faculty, including both senior and junior, enjoyed job security, got regular promotions, and received government salaries with perks and benefits, while the contractual teachers received only a quarter of what the former were taking in salary, with no job security or promotions, and no perks and benefits.

Five of the participants (i.e. T7, T8, T9, T10, T11) belonged to another public university (henceforth Uni-2), located in the Sukkur region of Sindh province, offered both undergraduate and postgraduate courses in business, education, engineering, computer science and maths. The faculty comprised of seven teachers of different age, qualifications and experiences of which only five consented to interview. Since the university did not offer courses in linguistics or literature, the various courses that the English faculty taught, included Functional English, Creative writing, Business communication, across all the departments. All the teachers of Uni-2 were permanent faculty members. This was a well resourced university with sufficient public and private sector funding to provide students with scholarships and running various talent hunt programs for candidates aspiring to study at the university. The English faculty taught regular courses in the morning and extra classes in the evening for the purpose of increasing their income.

Three participants (i.e. T12, T13, T14) were associated with a university (henceforth Uni-3) located in Hyderabad region of Sindh. The university was one of the largest universities of the province offering a variety of courses in a wide range of disciplines, including the sciences, arts, business, computers and engineering. English faculty of Uni-3 is comprised of more than twenty teachers of varying age groups, qualifications and experience. EL teachers at Uni-3 taught a list of similar courses as the teachers in Uni-1. One of the participants (T15) was also from an Engineering university (Uni-4) in the Nawabshah region of the Sindh province. The university offered undergraduate and postgraduate courses in disciplines, including Civil, Mechanical, and Electrical Engineering. The list of English courses that the faculty taught was very similar to the list of courses in Uni-2.

Data Collection: Interviews

Semi-structured interviews were conducted in order to elicit participants' responses regarding the factors affecting their TR. The interview guide comprised of open-ended questions and was divided into three main sections: The first section consisted of questions related to the personal information of the participants/ interviewees; i.e., their age, qualifications and experience; second section comprised of some general questions about their perceptions of research and whether they read published research and conduct any TR; third section involved questions about the factors that encouraged or discouraged them to engage in TR. Interviewees were

free to answer in as much detail as they like. Some of the participants (n=12) agreed to a face-to-face interview, while others (T13, T14, T15) preferred to send their responses via email. Their responses were followed up through the phone later. Participants were given the choice to opt for mother tongue (Sindhi or Urdu) or English to respond to interview questions. Ten participants agreed to do the interview in their native language (Sindhi), while the rest of them (T1, T7, T12, T13, T14) voluntarily responded to questions using English. Interviews in Sindhi were translated into English and transcribed with the help of a lecturer in English and a colleague.

Data Analysis

The data obtained in the form of transcripts and e-mails were constantly read and analysed during data collection. The type of analysis of the data was interpretivist: based on the examination of the participants' responses. The data from each participant was read iteratively to examine teachers' perceptions and analysed under the given themes, i.e. teachers' conceptions and practices of research, and secondly, the factors influencing TR.

The categories adopted lied on a continuum between deductive and inductive categories. Deductive categories, such as teachers' positivist conception of research, binary between the roles of teacher and researcher; lack of research skills; lack of time; and management's' attitude, were derived from literature, while inductive categories, such as attitude of senior faculty and large classes, emerged through content analysis and constant comparison approach. Part of the data was independently coded and compared by a colleague in the same institution. Coding agreement (90%) was achieved using: *number of agreement/total number of agreements-disagreements*. Findings of the study are presented in the next section.

Results

Teachers' Conception of Research

The data show that the majority of participants considered research to be a process of 'finding' something, while others described research as something which explores something new'; yet others said, it is a kind of discovery which finds something 'that already exists'. For instance, T1 said, 'when we explore something, when we put all our efforts to prove or disprove something'.

Some indicated a lack of confidence in defining research. T8, for instance, stated 'to find out something, *right* [my italics]!? Something that already exists but we research it, find it out'. The word in italics (i.e. right) in the response of T8 indicates her lack of confidence in her definition of research as she seemed to be trying to confirm definition from the interviewer. Some of the participants gave somewhat elaborate and specific answers which gives an impression of their familiarity with research. For example, T2 described research as something which 'explores a problem or a question in a particular area and tries to find out answers, which can be used for practical purposes and benefits. Or at least finds basis for future explorations'.

This definition not only defines what research is but also suggests the practical purpose it is carried out for. T6 gave an even more elaborate definition describing it as an organised activity which serves a variety of purposes besides discovering facts, i.e. proving or disproving a hypothesis, devising a line of action for practitioners in the light of facts. In T6's own words, 'research is an organized study or investigation into a subject in order to discover facts, to establish or revise a theory, or to develop a plan of action based on the facts discovered'. T10 called it a 'conscious inquiry' for solving [classroom-specific] problems. In defining research, T3 made an attempt to illustrate the difference between 'natural science research' and 'social science research'. He said:

> Scientific research requires a lot of physical things, like if somebody is a patient of tuberculosis so there has to be some virus or some bacteria that needs to be identified. In social sciences we don't have such a luxury. In social sciences it is more cooperative and behavioural. Otherwise in scientific research that is happening in chemistry u need to be finding some matters and putting the things together and finally getting some results. But we cannot in social sciences get things like that (T3).

According to this statement, research in natural science and social science diverges: the former deals with something concrete (i.e. tuberculosis, bacteria), finds out facts and puts them together, whereas, the social

science deals with abstract issues related to behaviours and opinions of people. The complexity of this definition is to some extent emblematic of teachers' understanding of research, and the complexity teachers unconsciously attribute to the activity of research.

Teachers' Practice of TR

Furthermore, when they were asked whether they themselves do any research in their own classrooms, two participants (T10, T14) did not respond to the questions, eight participants (T2, T3, T5, T6, T7, T10, T11,T13) said they do not do it, while five (T1, T4, T8, T9, T12) said they do only 'sometimes'. For example, T1 said while he 'keeps thinking' about finding innovative ways to solve students' problems, he does not engage in classroom research. T8 also said he keeps observing her students and sincerely tries to solve students' problems regarding language learning. T9 reported that he did research to improve students' oral communication in the past wherein he tried to develop strategies to develop students' fluency.

> I interviewed some students once in order to learn what I could do to improve their L2 fluency. But that was a year ago. (T9)

T12, however, seemed to be relatively more skilled in research. He claimed to have conducted research into motivation for reading in L2 using Vygotskian sociocultural theory. It suggests that T12 was familiar with research skills; however, his research was conducted as an obligation for passing his master degree rather than for solving classroom-specific issues. T12 stated, 'Yes, I have conducted research on students' motivation, and on devising strategies to enhance their zone of proximal development in improving reading'.

Factors Affecting Teachers' TR

Various factors were found to have a bearing on teachers' research activity, of which two factors emerged to be the main ones: teachers' concern about solving students' problems, and their professional development as teachers. For example, T1 said he does TR because it would bring him as well as his institution 'a good name'. T4 and T6 said they do it because they are concerned about their own and their students' improvement. T5 said he has been doing it for improving his students' 'fluency level', as well as for saving himself from stagnation as a teacher. T7 does it for motivating his students to give 'oral presentations'. T8 said she does it solely for her students' better performance. T9 gave various reasons for doing research, i.e. 'solving students' problems', 'conveying the message properly', 'becoming a successful teacher' and 'increasing [his] confidence'. T10 said he investigates the problems of his students because he is 'also a part of teaching and learning processes'. Following excerpts summarise the responses of the participants:

I started to dig out the truth as to why my students feel hesitation to give oral presentation? Is dice [rostrum] a snake or a scorpion? Why is it that only a few students turn up on the day presentations are scheduled? (T7)

> …it is just because of my commitment to my students, with my job. This *extra work* [research] on the part of the teacher will give him and his institution a good name. (T1)

Academic Culture

Majority of the participants reported that the main reason, amongst many others, for their not involving in research was due to academic culture. By academic culture they meant the routine practices and discourse within the department. Participants reported that teachers mostly discuss the tangible advantages for any activity beyond their routine workload. As T1, himself a senior faculty member, reported,

> We [teachers] talk about monetary benefits and packages that we will get for doing research. The intellectual benefits of research are not discussed.

Some of the teachers' comments were a bit blunt. For instance, T2 said, 'academic culture discourages not only from doing research but also from conducting classes'. T7 reported that, 'the culture is such that you come to university, conduct class, talk drivel, kill time and leave'. Similarly, T12 and T13 reported that research is not even discussed in their university; what is discussed is 'politics'. For instance, T13 stated that most of the teachers are busy 'lobbying, playing foul against the opposite groups'. Discussing the academic culture in his

institution, T6 indicated towards lack of research activity even amongst the senior faculty as a discouraging factor. He stated, 'we don't do research because we don't see even the senior faculty do it'.

Attitude of the Senior Faculty

Some of the participants also referred to the non-supportive attitude of senior faculty with more work experience, research skills and qualifications, as the main stumbling block to TR. Participants reported that the lack of cooperation and help on the part of senior and highly qualified, i.e. PhD, faculty also discouraged the junior faculty to do research. For instance, T5 reported,

> Young teachers are full of spirit and enthusiasm for performing their jobs even if they try to go into the depth of the problem but the attitude of seniors with juniors is not supporting and encouraging, they just treat the juniors like slaves.

Responses of some of the senior faculty also bear witness to the non-cooperative attitude of their peers. For example, T1 reported that senior faculty does not support junior faculty in TR because they are afraid of losing their authority that the seniors have by virtue of possessing research skills,

> I have always found them in a state of fear, if they shared their research with juniors, they would lose their position in the department. And the authority senior teachers have been enjoying for many years so they don't want to share their authority (T1).

The comprehensive response above indicates the insecurities of the qualified and skilled senior teachers and the power relations between senior and junior faculty.

Management's Attitude

Participants also reported that management's attitude also exerted a strong impact on their engagement with research. Management expected teachers to be engaged in classroom teaching, assessment and administrative jobs rather than research. Also, management maintained a strict distinction between teaching and research. For example, T13 reported, 'management is the problem; it discourages you; it forces you to quit the idea of research'. And T5, replied as, 'the role of the heads of the department regarding this (research) is not supportive'. T15 told a similar story with regard to research in his institution, 'management is concerned about syllabus, exams and tests in our universities rather than research'.

The responses of some of the participants were inconsistent regarding management's attitude. For example, T7 and T9 said that the senior faculty is supportive, and the management facilitates research. However, T11 stated that management 'is not pro-research'. Most interestingly, while T8 also admitted that management was not research friendly, he held the teachers responsible for indifference towards TR. In his words, 'the management does not favour research but somehow we are also responsible for not doing it'. He also admitted that teachers do not find time to do research because on the one hand they have to do excessive teaching, and on the other hand, the management involves them in administrative jobs, such as coordinating programmes, students' hostel wardenship. Some of the teachers (i.e. T2, T11, T14) reported that encouragement on the part of management is enough to get involved in TR. For instance, T2 reported, 'I think just appreciation, recognition, encouragement, respect are sufficient to make teachers do TR'. Similarly, T11 stated, 'management's result-centric approach does not let us engage in research; we are supposed to engage in teaching and assessment only'. Finally, T14 complained about 'management's unsupportive attitude with respect to providing facilities, i.e. electronic resources and training, enabling teachers to conduct research'.

Monetary and Logistic Factors

In addition to that, a number of participants (i.e. T3, T4, T5, T6, T7, T14) expressed monetary benefits as the main incentive for them to engage in TR. For instance, T7 candidly admitted to have never done TR because of his desire to earn extra money by conducting extra classes. His answer is worth-quoting as, 'we are bound to do extra classes in order to get extra remuneration'. The quote explicitly shows that his enormous workload was due to his desire to earn extra remuneration for extra classes. Thus, in response to what incentives he needs to

conduct research, T7 replied, 'management should offer packages [financial benefits] to those who want to do research'.

Also, the staggering difference between the senior permanent faculty and junior ad-hoc/contractual faculty also affected teachers' motivation to engage in research. For instance, T6 expressed his sense of lower monetary gains than the senior faculty in these word: 'the permanent faculty gets between 70 and 80 thousands (PKR=Pakistani rupees), whereas we get a meagre 12 thousands (PKR). They have job security, whereas we are on a contract and can be dismissed whenever the management decides as such'. Similarly, T4 specifically mentioned financial benefits as an important incentive. He noted, 'first of all, the salary of us contract teachers should be increased'. T5 also regarded 'money as well as specific equipment, and time liberty' as necessary incentive for him to conduct CR.

Participants also indicated a lack of logistics, such as proper offices, laptops and electronic resources as a major impediment to TR. They expressed their desire for logistic support on the part of authorities For instance, T3 reported 'we should be provided offices, where we can work on research and discuss problems with the students'; similarly, T14 mentioned that teachers 'should be provided with laptops as well as internet facility to make use of modern technology to keep ourselves updated about the state-of-the-art research in language teaching'. Most interestingly, for T13, material incentives were not a strong incentive for classroom research. According to her, 'a teacher ought to be self-motivated to engage in CR'.

Workload

Heavy workload also emerged as a strong factor impeding teachers' research engagement. T6 reported that junior faculty was assigned too many courses in a semester. According to him, he had 21 contact-hours/teaching hours a week during the current semester. T7, however, reported an incredibly higher number of contact-hours (32 hours) as the reason he could not engage in research. Similarly, T9 and T13 considered reduction of workload as an incentive for them to do research.

> Language teacher's workload is a major issue, so this may be reduced…(T9)
> First of all, there should be minimum workload. (T13)

Large Classes

Finally, some of the participants, i.e. T4, T8 and T13, mentioned large classes to be an impediment to teachers' CR. For instance, T4 stated, 'it is difficult to conduct research in a class of 40-50 students'. T14's response is worth quoting,

> I do have four classes with four different subjects along with 150+ (One fifty plus) students per class. We get limited opportunities in large classes to identify problems of the students.

Discussion

The current study aimed to explore the factors that motivate or prevent teachers to engage in TR in four public sector universities in the Sindh province of Pakistan. The study showed that a majority of participants regarded research to be an activity aimed at finding something new, while others considered research to be analogous to scientific inquiry conducted in a lab. Most notably, while participants displayed awareness of and interest in solving students' problems, their engagement in TR seemed to be limited as well as occasional. For example, while some participants claimed to have engaged in research *whenever* [my italics] a problem arose in the class, a majority seemed to have conducted it sometime in the past, recent or distant. Furthermore, their research also lacked systematicity and rigor. Thus, the conceptions as well as the research activity did not strictly fit into the definition of TR as a 'systematic self-study that seeks positive impact and is made public' (Sanchez & Borg, 2015, p. 185). When asked what methods they use to investigate and solve classroom-specific problems, most of the participants mentioned only two methods, such as observations and reflections. Whether they were aware of other methods, such as questionnaires, interviews/individual meetings with students, and etc., was not reflected in their responses. This reflects lack of knowledge of research and research skills among participants.

These findings are in partial consonance with previous research in the sense that while participants expressed willingness to address classroom related issues through observations and reflections, however, they did not use the data to solve the issue (Alison & Carey, 2007; Borg, 2007; Borg & Sanchez, 2015).

Most interestingly, while a few participants were intrinsically motivation to conduct research in order to solve students' problems and improve their own teaching skills, majority of them showed disinterest in TR and held exogenous factors, including academic culture, management's attitude, senior faculty's attitude, excessive workload, lack of incentives, and large classes, as responsible for it. The role of academic culture appeared as a recurrent theme in the findings. It indicates a deep and subtle impact it had on teachers' classroom practices as well as on them as teachers. Most of the participants unequivocally reported that research is not discussed in their respective institutions. Contractual teachers reported that permanent faculty mostly indulge in leg-pulling, lobbying and petty politics. They also reported that the attitude of senior faculty towards their juniors was discouraging and disparaging. Although the senior faculty possessed research skills, they were less willing to guide the juniors in research. What is strange is that some of the participants who belonged to senior faculty also admitted to the unequal power relations between senior and junior faculty. Thus, the present study contributes to literature by highlighting the imperious and paranoiac attitude of senior faculty as a barrier in a Pakistani context.

Furthermore, teachers expressed anomalous views about the role of management in supporting or impeding TR. Some of the teachers said that the management does not favour or facilitate teachers' research; rather, it over-burdens them with workload, involving teaching and administrative jobs, while others regarded the management to be supportive, and instead, held themselves responsible for not engaging in research. This ambivalence on the part of participants reveals, among other things, the dualistic attitude of university authorities/management towards the idea of TR. Also, these incongruous responses also suggest that management pays lip-service to TR but not encourage it practically, for instance, by offering them incentives or lessening their workload. It is also noteworthy that teachers' performance in all the four institutions was measured by the number of chapters covered and the score of their students in exams. The inconsistent attitude of management is also congruent with previous research (Borg & Alshumaimeri, 2012; Nadeem, 2011).

Also, workload or syllabus completion acts like a spectre haunting the EL teachers all the time. While in previous research teachers were reported to have 18-20 hours of workload per week (Atay, 2008), teachers in the present study reported on an even higher workload (32 hours/week). Higher workload, in turn, was also found to be responsible for lack of time. This finding is consistent with previous studies in which the lack of time was cited as the most common factor (Atay, 2008; Borg, 2009; Worral, 2004). Besides workload, large classes with a disproportionate number of students was also highlighted by a few students. This seems to have been corroborated by previous studies in the context of Pakistan (Bughio, 2013; Shamim, 1996). Teachers demanded a few incentives in order for them to do TR, including rise in salaries, and availability of resources, such as books, magazines, offices, laptops and offices. The lack of these incentives explain the lack of motivation and, consequently, a moderate engagement of teachers in TR. This information reinforces the need for providing incentives to enhance teachers' motivation to engage in TR. Some of the findings of the current study, such as dubious role of management, lack of research skills, time and resources, are consistent with previous research (Borg & Alshumaimeri, 2012; McDonough, 2006; Moore, 2011); however, a number of findings, such as unprofessional attitude of teachers, imperious attitude of senior faculty, excessive teaching hours and large classes, were unique to the context of the current study.

Conclusion

This study was inspired by the increasing popularity of TR and specifically by the dearth of studies into English language (EL) teachers' TR in universities in the Sindh province of Pakistan. The aim of this study was to explore the factors which affect EL teachers' research into classroom-specific problems. Based on semi-structured interviews with fifteen EL practitioners in four public sector universities, the study showed that a majority of EL teachers were only modestly familiar with and interested in conducting TR on their own. The main factors that affected TR included lack of familiarity with TR, academic culture, management's attitude, lack of support from senior faculty, lack of incentives, lack of time and logistics, and large classes.

The findings of this study show that TR cannot become a classroom reality until major changes are made at i) macro-level, i.e. in educational policy, ii) micro-level, i.e. institutional/university-level, and iii) individual level. At the policy-level, i) action research ought to be made an integral part of teacher education programs whereby teachers are trained into conducting TR. ii) At a university-level, management's dichotomous conceptions regarding the role of teacher and researcher need to change by an top-down intervention the policy level. However, merely a change in policy might not convince management at a university level to provide teachers with more autonomy with regard to addressing classroom-specific issues by intervening through TR. Awareness campaigns and workshops on TR are required to inform the management of the potential implications of TR for the improvement of teachers' professional growth, and enrichment of students' learning experiences.

In addition, unsupportive or indifferent attitude of research-trained faculty towards their novice colleagues can be addressed by management's encouragement and incentivising mentoring faculty interested in TR. While changes in policy and management's attitude are necessary, what is even more essential is the teachers' motivation and willingness to engage in TR. Therefore, iii) at an individual level, teachers' concerns regarding judicious workload, large classes and monetary incentives ought to be addressed by the management and government (Halai, 2011; Kasi, 2007; Nadeem, 2011). In other words, there needs to be a balance between top-down institutional support and a bottom-up teacher motivation for TR to be a sustainable classroom practice on the part of language teachers (Tuyan, 2018). In a Pakistani context, the findings of this study can help HEC Pakistan in accomplishing its Vision-2025[4]; that is, promoting constructive research culture in academia and increase the production of research aimed at solving local issues.

The results of the current study cannot be generalised due to a small sample size (nearly 25% of the total workforce) of EL teachers of the four universities in the region. However, the aim of the study was not to generalise the findings but to analyse teachers' perceptions of research, and explore the factors preventing teachers from engaging in research in order to contribute to research aiming it exploring the feasibility of TR in a university context, and secondly to inform teacher educators and policy-makers. Besides, while this study could examine the perceptions of EL teachers from four public sector universities, future studies may explore the factors in other public and private sector universities to get a better understanding of facilitating and debilitating factors affecting TR in Pakistan. Also, future studies must also focus on exploring strategies for integrating research with teachers' classroom practices.

References

Allison, D., & Carey, J. (2007). What do university language teachers say about language teaching research? *TESL Canada Journal*, *24*(2), 61-81. https://doi.org/10.18806/tesl.v24i2.139

Allwright, D. (2005). Developing principles for practitioner research: The case of exploratory practice. *The Modern Language Journal*, *89*(3), 353-366. https://doi.org/10.1111/j.1540-4781.2005.00310.x

Atay, D. (2008). Teacher research for professional development. *ELT Journal*, *62*(2), 139-147. https://doi.org/10.1093/elt/ccl053

Barkhuizen, B., Burns, A., Dikilitaş, K., & Wyatt, M. (2018). *Empowering teacher researchers, empowering learners*. IATEFL

Borg, S. (2006). Conditions for teacher research. *English Language Forum*, *6*, 22-27.

Borg, S. (2007). Research engagement in English language teaching. *Teaching and Teacher Education*, *23*, 731–747. https://doi.org/10.1017/S0261444810000170

Borg, S. (2009). English language teachers' conceptions of research. *Applied Linguistics*, *30*(3), 358-388. https://doi.org/10.1093/applin/amp007

Borg, S., & Alshumaimeri, Y. (2012). University teacher educators' research engagement: Perspectives from Saudi Arabia. *Teaching and Teacher Education*, *28*, 347-356. https://doi.org/10.1016/j.tate.2011.10.011

Borg, S., & Sanchez, H. S. (2015). Key issues in doing and supporting language teacher research. In S. Borg, & H. S. Sanchez, *International Perspectives on Teacher Research* (pp. 1-12). Palgrave Macmillan.

Bughio, F. A. (2013). *Improving English language teaching in large classes at University level in Pakistan* [Unpublished doctoral dissertationб University of Sussex]. http://sro.sussex.

[4] Higher Education Commission of Pakistan (2017). Higher Education Commission of Pakistan Vision-2025. https://www.hec.gov.pk/english/HECAnnouncements/Documents/Announcement/HEC-Vision-2025.pdf

ac.uk/45170/
Burns, A. (2005). Action research: An evolving paradigm? *Language Teaching*, *38*(2), 57-74. https://doi.org/10.1017/S0261444805002661
Christenson, M., Slutsky, R., Bendau, S., Covert, J., Dyer, J., Risko, G., & Johnston, M. (2002). The rocky road of teachers becoming action researchers. *Teaching and Teacher Education*, *18*, 259–272. https://doi.org/10.1016/S0742-051X(01)00068-3
Cochran-Smith, M., & Lytle, S. L. (1992). Communities for teacher research: Fringe or forefront? *American Journal of Education*, *100*(3), 298-324. https://www.jstor.org/stable/1085491
Cochran-Smith, M., & Lytle, S. L., (1993). *Inside/Outside: Teacher research and knowledge*. Teachers College Press
Cochran-Smith, M., & Lytle, S.L. (1998). Teacher research: The question that persists. *International Journal of Leadership in Education: Theory and Practice*, *1*(1), 19-36.
Creswell, J. W. (2014). Research design: qualitative, quantitative and mixed methods approaches (4th ed.). *Sage*.
Everton, T., Galton, M., & Pell, T. (2000). Teachers' Perspectives on Educational Research: Knowledge and context. *Journal of Education for Teaching: International research and pedagogy*, *26*(2), 167-182. https://doi.org/10.1080/02607470050127081
Gao, X., & Chow, A. (2012). Primary school English teachers' research engagement. *ELT Journal*, *66*(2), 224-232. https://doi.org/10.1093/elt/ccr046
Halai, N. (2011). How teachers become action researchers in Pakistan: Emerging patterns from a qualitative metasynthesis. *Educational Action Research*, *19*(2), 201–214.
Halai, A. (2015). Teacher education continuum: A case for continuing professional development. In A. Rashid, & I. Muzaffar's '*Educational Crisis and Reform: Perspectives from South Asia'* (pp. 347-381). Oxford University Press.
Mackay, J., Birello, M., & Xerri, D. (2018). *ELT research in action: Bridging the gap between research and classroom practice*. IATEFL
Kasi, F. (2010). Collaborative action research: An alternative model for EFL for teacher professional development in Pakistan. *The Asian EFL Journal Quarterly*, *12*(3), 98-117.
Khan, S., Shah, S. M., & Khan, T. (2018). An investigation of attitudes towards the research activities of university teachers. *Bulletin of Educational Research*, *40*(1), 215-230.
Khurram, B. A. (2018). Promoting learner engagement in a large university-Level ESL class in Pakistan. In K. Kuchah & F. Shamim's (Eds.), *International perspectives on teaching English in difficult circumstances: Contexts, challenges and possibilities*' (pp. 73-88). Palgrave Macmillan
Kirkwood, M., & Christie, D. (2006). The role of teacher research in continuing professional development. *British Journal of Educational Studies*, *54*(4), 429-448. https://doi.org/10.1111/j.1467-8527.2006.00355.x
Levkina, M. (2018). Teaching pronunciation through tasks: Myth or reality? In J. Mackay, M. Birello, & D. Xerri,'s (Eds.), *ELT in Action: Bridging the gap between research and classroom practice* (pp. 51-54). IATEFL
Macaro, E. (2003). Second language teachers as a second language classroom researchers. *Language Learning Journal*, *27*, 43-51. https://doi.org/10.1080/09571730385200071
Magos, K. (2012). But I cannot do research: action-research and early childhood teachers. A case study from Greece. *Teachers and Teaching: Theory and Practice*, *18*(3), 331-343. https://doi.org/10.1080/13540602.2012.629839
McDonough, K. (2006). Action research and the professional development of graduate teaching assistants. *The Modern Language Journa*l, *90*(1), 33-47. https://doi.org/10.1111/j.1540-4781.2006.00383.x
Moore, S. (2011). Cambodian English teachers' conceptions of and engagement with research. http://scholar.google.co.uk/scholar?q=Cambodian+English+Teachers%E2%80%99+Conceptions+of+and+Engagement+with&btnG=&hl=en&as_sdt=0%2C5
Nadeem, M. (2011). Re-searching research culture at higher education. *Journal of Research and Reflections in Education*, *5*(1), 41 -52.
Newby, P. (2010). *Research Methods for Education*. Pearson Education Limited.
Saglam, A. L. G., & Dikilitas, K. (2019). *Stories by teacher researchers in an online research community*. IATEFL.
Sanchez, H., & Borg, S. (2015). Teacher research: Looking back and moving forward. In S. Borg & H. S. Sanchez (Eds.), *International perspectives on teacher research* (pp. 185-193). Palgrave Macmillan.
Sarwar, Z. (2001). Innovations in large classes in Pakistan. *TESOL Quarterly*, *35*(3), 497-500. https://doi.org/10.2307/3588033
Schapper, J., & Mayson, S. E. (2010). Research-led teaching: Moving from a fractured engagement to a marriage of convenience. *Higher Education Research & Development*, *29*(6), 641-651. https://doi.org/10.1080/07294360.2010.489236

Shkedi, A. (1998). Teachers' attitudes towards research: A challenge for qualitative researchers. *International Journal of Qualitative Studies in Education*, 11(4), 559-577. https://doi.org/10.1080/095183998236467

Slimani-Rolls, A. (2018). Using exploratory practice to develop teacher research. In S. Etherington & M. Daubney's (Eds.), *Developing as EFL researcher: Stories from the Field* (pp. 29-34). IATEFL

Suzuki, N. (2018). How do teacher questions affect students' L2 task performance in EFL classrooms? In J. Mackay, M. Birello & D. Xerri,'s (Eds.), *ELT in Action: Bridging the gap between research and classroom practice* (pp. 55-60). IATEFL

Taylor, J. (2007). The teaching-research nexus: A model for institutional management. *Higher Education*, *54*(6), 867-884. https://doi.org/10.1007/s10734-006-9029-1

Tien, F. F. (2000). To what degree does the desire for promotion motivate faculty to perform research? Testing the expectancy theory. *Research in Higher Education*, *41*(6), 723-752. https://doi.org/10.1023/A:1007020721531

Tuyan, S. E. (2018). On the way to achieve sustainability: An evaluative look at a three-year action research program in an ELT context. In G. Barkhuizen, A. Burns, K. Dikilitaş & M. Wyatt's (Eds.), *Empowering teacher researchers, empowering learners* (pp. 19-28). IATEFL

Xie, J. (2015). Learning to do teacher research independently: Challenges and solutions. In S. Borg & H. S. Sanchez's (Eds.), *International Perspectives on Teacher Research* (pp. 47-56). Palgrave Macmillan.

Worrall, N. (2004). Trying to build a research culture in a school: Trying to find the right questions to ask. *Teacher Development*, *8*(2-3), 137-148. https://doi.org/10.1080/13664530400200020228

Wyatt, M. (2011). Teachers researching their own practice. *ELT Journal*, *64*(4), 417-425. https://doi.org/10.1093/elt/ccq074

Appendix 1

Participants' information

No	Name/ Code	Age (in years)	Qualifications (B.ED/MA/M.ED/M. Phil/PhD)	Experience of ELT at a university level	Experience of ELT at a college level	Status	Monthly income
Uni-1	T1	45	MA (English literature)	21	Not applicable (NA)	Senior faculty with permanent job	160,000/-
	T2	44	M.Phil (Linguistics) and MA (Literature)	16	NA	Senior faculty with permanent job	140,000/-
	T3	31	MA (English literature)/ B.Ed	5	NA	Junior faculty with permanent Job	80,000/-
	T4	31	Masters in English literature/M.Ed	4	NA	Junior faculty on contract	12000/-
	T5	27	Masters in English literature	3	NA	Junior faculty on contract	12000/-
	T6	26	Masters in Linguistics	4	NA	Junior faculty on contract	12000/-
Uni-2	T7	42	Masters in English literature	7	12	Senior faculty with permanent job	80,000/-
	T8	40	Masters in English literature/M.Ed	7	11	Senior faculty with permanent job	80,000/-
	T9	32	Masters in English literature	5	4	Senior faculty with permanent job	80,000/-
	T10	33	Masters in Linguistics	3	NA	Junior faculty with permanent job	80,000/-
	T11	29	Masters in English literature	5	NA	Junior faculty with permanent job	80,000/-
Uni-3	T12	28	Masters in English literature	4	NA	Junior faculty on contract	80,000/-
	T13	29	Masters in English literature	4	NA	Junior faculty on contract	80,000/-
	T14	28	Masters in English literature	4	NA	Junior faculty on contract	80,000/-
Uni-4	T15	30	Masters in English literature	4	4	Junior faculty with permanent job	80,000/-

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 1, 2020

Wariyo, L.G. (2020). Instructional Goal Structure, Gender, and Second Language Motivation Affecting English Language Achievement. *Journal of Language and Education, 6*(1), 134-156. https://doi.org/10.17323/jle.2020.7766

Instructional Goal Structure, Gender, and Second Language Motivation Affecting English Language Achievement

Lemecha Geleto Wariyo
Addis Ababa University

Correspondence concerning this article should be addressed to Lemecha Geleto Wariyo, Department of Curriculum and Instruction, Addis Ababa University , NBH1, 6killo King George VI St, Addis Ababa, Ethiopia. E-mail: lemechageleto@yahoo.com

The study opted to: 1) Investigate differences between cooperative and competitive learning modes in the extent to which they affect English language achievement; 2) Find gender, intra-gender, and inter-gender differences in English language achievement within and across the cooperative, competitive, and control group learning conditions; and 3) Study the relationship between students' motivation to learn the English language and English language achievement. The cooperative and competitive learning groups were used as treatment groups while the control group was the individualized learning group. An English language test was administered to 120 secondary school grade nine students. The 12-item mini-Attitude/Motivation Test Battery (Mini-AMTB) was administered to the students. Different parametric tests were used in the pre-test and post-test data analysis. Post-test data analysis results revealed that cooperative learners significantly outperformed both competitive learners and the control group, but the control group significantly outperformed competitive learners. The gender difference in English language achievement was not significant. The cooperative learning (CL) mode was favored by both male and female students but more favored by males. The five aggregate measures of the mini-AMTB (Integrativeness, Motivational Intensity, Attitudes towards the Learning Situation, Instrumental Orientation, and Parental Encouragement) produced significant positive correlations with English language achievement; however, Language Anxiety negatively and significantly correlated with English language achievement. The CL mode, with its effective CL technique, was recommended to be researched and applied by trained teachers to improve student achievement. Its implications for teacher training were also given. Treatment of second language motivation was recommended as an important issue in second language learning.

Keywords: cooperative learning, competitive learning, instructional goal structure, second language motivation

Introduction

Cooperative learning (CL) results in highest student achievement and good pro-social behavior, as well as attitudes toward learning and higher-order reasoning (Alghamdy, 2019; Gillies, 2016; Johnson, Johnson, Roseth, & Shin, 2014; Peldon & Chalermnirundorn, 2018; Sharan, 1980; Yager, Johnson, & Johnson, 1985; Zarei, 2012; Zhang, 2010). Kyndt and et.al (2013) also reported that CL positively affects achievement and attitudes. As a modern pedagogy, CL is a method that has a significant positive effect on achievement, motivation to learn, group cohesion, critical thinking, and problem-solving (Baloche, & Brody, 2017). CL is effective because it uses an active learning approach between learners of heterogeneous abilities and backgrounds (Azizinezhad, Hashemi, & Darvishi, 2013, Ghodbane & El Achachi, 2019).

When it comes to language learning, current language teaching methodologies emphasize both communication and the learner (Lopez & Rua, 2006; Oxford, 1997). The CL methodology has attracted practitioners of communicative language teaching because of the innovative methodologies it applies in the language classroom (McCafferty, Jacobs, & Iddings, 2006).

In Ethiopian schools, a student-centered teaching approach was introduced decades ago. The Institute for Curriculum Development and Research, Ministry of Education, has prepared new English language textbooks for high schools (Gezahegn, 2003). Gezahegn also stated that the course is more student-centered and communicative than its predecessors. Even though this teaching approach has been applied in schools, problems related to its implementation have been reported (Lucha & Bongase, 2015; Kumar, 2017; Moges, 2019; Reda & Hagos, 2015; Serbessa, 2006; Yayo, 2013). Tirussew, Amare, Jeilu, Tassew, Aklilu, and Berhannu (2018) found that the implementation of CL through the Education Development Army has been found to be poor, contrary to its positive effects on student learning as confirmed by studies from around the world.

In addition to the language teaching-learning approach, gender difference has become a hot topic of discussion in language studies. The empirical gender research in the field of second language learning underlies gender differences and the role of gender in second language learning (Feery, 2008). Second language education and the field of gender has a long history, and it has been a burning issue for language researchers and teachers since the outbreak of the women's movement (Sunderland, 2010). Gender is an issue that has significant implications in second language learning. Gender difference theories state that females use learning strategies more effectively (Tercanlioglu[1], 2004). In the Ethiopian context, the gender differences in English language achievement have been inadequately researched. To identify gender-related factors that contribute to under-achievement, local empirical studies on gender differences in language achievement are needed.

Even nowadays, research findings on gender differences in second language achievement show that female performance dominates male performance. Van Der Slik, van Hout, and Schepens (2015) studied gender differences across countries and language families in the world, using a dataset consisting of 27,119 adult Dutch learners as a second language. In this study, female students repeatedly outperformed male students in writing and speaking proficiency. Mady and Seiling (2017) also found that being female was associated with increases in French achievement. Females in some parts of the world may have better opportunities and psychosocial differences than females from other parts of the world. As a result, the natural advantages for females in learning may be biased due to these factors. For this reason, a critical approach to this topic is paramount. However, taking a critical and contextual approach to gender differences in language achievement is a recent phenomenon. Supporting this, Schmenk (2004) stated that gender difference positions in previous research to understand gender and language learning were culture blind because they regarded gender as an unchanging, contextually independent variable. Studies stereotype language performance as a female's domain and mathematics performance as a male's domain. The cultural, economic, and other local factors that may affect females' achievement ought to be controlled to arrive at more valid conclusions regarding subject area difference between sexes.

Furthermore, gender differences in the modes of language learning are another issue that has been discussed in recently conducted studies around the world (Alsheikh & Elhoweris, 2011; Esiobu, 2011; Hossain, Tarmizi, Aziz and Nordin, 2013; Mohammadjani, & Tonkaboni , 2015;). Studying intra-gender differences in English language achievement across the three learning conditions was another objective of this study. In other words, the study opted to analyze the differences between same-sex learners in English language achievement across the three learning conditions. By doing this, the study aimed at analyzing a learning mode that results in significantly higher achievement for both a female same-sex group of learners and a male same-sex group of learners that are compared to their peer same sexes (e.g., assume that both female and male cooperative leaners outperform their same-sex groups in other learning conditions). This implies that cooperative learning is favoured by both sexes because it works best for both sexes when compared to their own peer same-sex groups. Previous studies on learning modes did not consider these issues in the Ethiopian context. The researcher believes this issue can help educators obtain more information on how males and females learn.

Moreover, another variable that received attention in previous studies for affecting students' second language achievement is second language motivation. In a meta-analysis that explained seventy-five different samples consisting of 10,489 participants, the findings revealed that the correlations between motivation and achievement were consistently and similarly higher than the correlations of the other variables, such as integrativeness, instrumental orientation, and attitudes toward the learning situation with student achievement (Masgoret & Gardner, 2003). The Socio-Educational Model (Gardner, 2005) proposed that motivation and situational anxiety

[1] L., T. (n.d.). IIER 14: Tercanlioglu - gender effect on adult foreign language learning [Web logpost]. http://www.iier.org.au/iier14/tercanlioglu.html

are more important in second language learning. Female students' integrative motivation and attitudes towards the English language are higher than male students (Al Harthy, 2017). A large scale meta-analytic study by Teimouri, Goetze, and Plonsky (2019) reported a strong significant association between language anxiety and second language achievement.

In Ethiopia, although there is a shortage of studies on the associations between second language motivation and second language performance, two studies by Gezahegn (2003) and Kumar (2017) reported that students' low motivation, language anxiety, and poor proficiency in the English language were identified as impediments to the implementation of group work in English language teaching classes. This report runs counter to previous studies that argued that when teachers use CL effectively in language study classrooms, the learners' motivation to learn the language is heightened (Johnson, Johnson, Roseth, & Shin, 2014; Zareian & Jodaei, 2015).

Statement of the Problem

Contrary to a myriad of findings that support the positive effects of CL on students' performance, the findings in Ethiopian schools on the effectiveness of CL are still controversial. A study by Gezahegn (2003) revealed that teachers' resistance to the new instructional procedure is most likely to be attributed to their lack of adequate training and skills necessary to devise and manage group-work activities. Despite the arguments against the teacher-centered methodology in education, the teaching-learning process in most schools in Ethiopia has continued to be teacher-centered (Desta, Chalchisa, Mulat, Berihum, & Tesera, 2009). The major factors impacting the implementation of active learning in Ethiopia are teachers' shortage of skills, an unsuitable curriculum and resources for CL, learners' shortage of prior experience to actively cooperate in the learning process, inappropriate classroom materials, a shortage of language proficiency, shyness, and the lack of skill to use various CL methods (Ayele & Olamo, 2019; Lucha & Bongase, 2015; Serbessa, 2006; Yalew, 2004; Yayo, 2013).

Generally, studies report that the responsibility for the effective application of CL is shared by students and teachers (Ferguson-Patrick, 2016). According to Gillies (2016), teachers are the leaders in organizing and facilitating CL activities in classrooms. Learners' attitudes toward the cooperative method has caused problems in applying this approach with students (Çelik, Aytın, & Bayram, 2013). Although the effectiveness of CL was well recorded, applying it in classrooms is a challenge for many teachers (Gillies, Ashman, & Terwel, 2007; Jolliffe, 2019). Commonly used group-work activities with language teaching are dissimilar to CL. Although group work, such as role-playing and problem-solving activities are prerequisites to CL and consist of certain CL principles, CL tasks in second language classrooms usually underutilize CL principles (Siciliano, 2001).

Because language is a key to academic success, the underutilization of the student-centered teaching methods can result in low language proficiency and this can be a factor in academic failures. Concerning this issue, a recent study published in Ethiopia by Reda and Hagos (2015) reported that there is a shortage of skills and knowledge in using CL principles and structure in elementary schools in Ethiopia. These researchers pointed to the problem. However, the study warrants focusing on how the existing implementation problems can be solved. Therefore, conducting empirical studies on how to utilize appropriate student-centered techniques (i.e., CL techniques and principles in this study) and ways to motivate teachers and students toward the method is very important.

Another gap in language research in Ethiopia is a shortage of studies on gender differences in language achievement, although it has received a great deal of attention in many studies around the world. In the Ethiopian educational context, a critical study on gender differences in second or foreign language learning is also crutial. This may help in dealing with gender-related learning problems. Observing favorable learning modes for males and females is also an important issue for designing effective instructional methods.

Therefore, this study is primarily opted to: 1) Find the differences between the three learning conditions (cooperative, competitive, and individualistic learning) in English language achievement; 2) Analyze gender differences in English language achievement; 3) Find the correlations between the aggregate measures of the mini-AMTB and the English language achievement; 4) Analyze gender differences in English language achievement within a group of students who learn using similar learning modes; and 5) Find the learning mode that is most favorable for both males and females.

The following research questions were formulated to achieve the objectives of this study: A) Do groups of students who learn using cooperative and competitive learning modes significantly differ in English language achievement? B) Is the gender difference in English language achievement significant? C) Is there a significant gender difference in English language achievement within a group of students who learn using a similar learning mode? D) Which learning mode is significantly favored by both males and females? E) Is there a significant correlation between the aggregate measures of the mini-AMTB and English language achievement?

Theoretical Framework

CL is the use of small groups in which the learners study together to enhance their own and their peers' learning (Boling & Robinson, 1999; Johnson and Johnson, 1974; Johnson, Johnson, & Holubec, 2013; W. Johnson, & T. Johnson, 2018). Goal structure (i.e., learning mode) is an interdependence that exists among students and how the students will interact with one other and with their teacher in studying toward the achievement of learning objectives (Johnson, & Johnson, 2013; W. Johnson, & T. Johnson, 2018; Owens & Straton, 1980). These authors listed three different conditions of interdependence: 1) A *cooperative goal structure* exists in a learning situation in which students can achieve their own individual learning goals only through studying with their group as they achieve their learning goals; 2) *Competitive goal structures* exist in a learning situation in which learners are able to achieve their own individual learning objectives only through causing others fail to achieve their own learning goals; and 3) An *individualistic learning goal structure* happens in a learning situation in which learners work to achieve their own individual learning goals regardless of what other students choose to work on to achieve their individual goals. CL can be defined as a learning method in which learners work in small groups to enhance their team's learning (Slavin, 2014). Generally, the definitions view CL as a learning approach in which all group members work toward achieving an individual goal through achieving a group goal by supporting one another. On the other hand, they defined competitive learning as a learning approach in which individual students compete to outperform others to achieve their own individual goals. In individualistic learning, students are neither inclined to support one another nor compete to outperform others to achieve their own goal. It is assumed that their achievement depends on their own efforts.

According to Deutsch's cooperative-competitive theory (1949, 2011, & 2015), a competitive social situation is a social situation in which the objectives of various participants are associated to the point that they result in a negative association between their individual accomplishments. Deutsch pointed out that the student seeks not only to succeed but also to cause other students to fail in a competitive situation; the student emphasizes the outcome that is most useful for them and most problematic to their peers. According to Deutsch, an individualistic situation is one where the goals of the individuals are unrelated. Deutsch suggested that the instructional method applied ought to be in the form of a CL method that takes the form of the conflictual and controvercial interactions in the classroom in order to facilitate the development of conflict resolution skills.

Corresponding to cooperation-competition theory- Teams-Games Tournament (TGT) is a structured blend of educational techniques such as small groups, educational games, and tournaments that apply encouragement for successful academic performance.

The TGT consists of both cooperation and competition in a way that enhances peer-group rewards for academic performance. This happens through changing the social organization of the class as follows: 1) Creating interdependence between learners; and 2) Making it possible for all learners to have an equal probability of being successful in a project regardless of various learning rates (Slavin, 1995, 2014). Another study by Slavin (2013) revealed that well-structured CL methods bring about more positive effect sizes than the use of innovative curriculum textbooks or the use of technology in maths. The carefully structured CL techniques, with which are easier for both experienced and inexperienced teachers to apply the principles of CL in the classroom, have to be carefully researched before being included in the Ethiopian curriculum.

On the other hand, similar to its application in sports, cooperation-competition has an effective application in learning classrooms. It is possible to replace interpersonal (individualistic) competition in the classroom with a competition between groups, accompanied by a competition within a group. By incorporating educational materials into intellectual games, learning can be like a sport. The students' interest and motivation toward learning can be heightened in this way (Slavin, 1991; Slavin, 1995, Slavin, 2014).

Nevertheless, making students work in groups can be quite beneficial or it can hold little value. Its effectiveness depends on how it is implemented in classroom settings. Two kinds of CL techniques were identified: 1) *Structured Team Learning*, which is characterized by individual accountability. Individual accountability implies that the success of a team depends on individual learning and rewards to the groups are dependent on the learning progress of their members; and 2) *Informal Group CL approaches* are more focused on social dynamics and discussion (Pedersen & Digby, 2014; Slavin, 1991). In order to conform to 'the how' issue of implementing CL in the classroom, in this study, a structured CL model (Teams-Games-Tournament) was implemented.

In addition to Slavin's work, Johnson and Johnson (2013) listed four CL types: informal, formal, cooperative base groups and constructive controversy. When CL types are categorized by their objectives, *formal CL* groups are often used to teach specific content; *informal CL* groups are used to accelerate the active processing of cognition while teaching students; *cooperative base groups* are often used to deliver long-term support for academic progress; and *constructive controversy* can also be used as a method of involving intellectual conflicts in the learning process.

Cooperative Learning Principles according to Slavin

Effective CL should encompass six principles: 1) *Positive Interdependence* implies that an increase in learning for one person is correlated with an increase in learning for other group members.This situation of learning is created through structuring the goals, materials, and rewards; 2) In assuring *accountability*, each student becomes accountable through testing and grading. Through the group grade, the group becomes accountable; 3) *Team Formation* is accomplished randomly by learner interest and by using criteria such as homogeneity, gender, and aptitude; 4) *Team Size*: A team works best when it is smaller than seven members; 5) *Cognitive Development*: It is the main objective of CL; and 6) The next CL principle is *Social Development* in which the development of social skills, such as turn-taking and active listening, are achieved (Slavin, 1991; Slavin, 1994; Slavin, 1995).

Slavin seemed to be concerned about the difficulties and problems implementing CL by teachers and students around the world. As a result, in his recent work, Slavin (2014) stated that teachers can gain the highest benefit from CL by using these five strategies and they can ensure that collaboration: 1) Enhances the development of communication and problem-solving skills; 2) Creates interdependent teams; 3) Sets group goals; 4) Ensures individual accountability; and 5) Integrates CL with other structures of learning.

Cooperative Learning Principles According to Johnson and Johnson

These authors pointed out that in order to enhance the effectiveness of the CL, groups have to be formed so that the five elements of the effective cooperative method are included in their structure: 1) *Creating Positive Interdependence* in the learning procedure, thereby each of the students in a group understands that they are associated in such a way that individual success depends on the group success; 2) *Promotive Interaction*: It is about encouraging others to accomplish each other's tasks to help the team achieve its objective. This interaction is enhanced when students support each other when needed; when members share resources, give each other effective feedback on specific activities, critically analyze one another's conclusions and reasoning, and work in cooperation to achieve common objectives; 3) These authors describe *Individual Accountability as* a learner's responsibility in ensuring that they accomplish their part of work while ensuring that one's peers accomplish theirs. Accountability is promoted by i) Creating positive interdependence between each member, thereby making them responsible for leading peer activities; ii) Making group members individually accountable for working on their part in the activity and making sure that their activities can be clearly recognized; 4) *Negotiation with Older Students and Teaching the Younger Students the Important Interpersonal Skills* to manage conflict between individual group members is highly crucial; 5) In the *Group Processing*, students reflect on their own performance and their own work interactions by asking what they achieved, by asking each other what they haven't achieved, and what they need to do?" (Gillies, 2016; Johnson, & Johnson, 2017).

Teams-Games-Tournament (TGT)

In TGT, the teacher assigns students to four- or five-member groups. A cross-section of academic ability levels in the class has to be shown in the group composition. This should also be applied to representatives of ethnic groups and sex groups. The objective of this grouping is to make its members ready through peer tutoring and

practice for the following day's tournament by working on the material presented by the teacher. The teaching activity is delivered in any method a teacher is accustomed to using. A tournament usually takes forty minutes, is held once per week, and consists of short-answer questions from the material (Slavin, 1980; Slavin, 1991; Slavin, 1995; Slavin, & Project, 1994).

Three students of homogeneous academic performance are assigned to tournament tables based on the previous performance in the same subject. After the game has ended, the three competitors are ranked and given marks or points (the highest performer in each group earns six points, the middle performer earns four points, and the lowest performer earns two points). The group scores (i.e., the score of the original practice group or team of four to five players) are found by adding the scores earned by each member at the contest tables then creating "reward interdependence" within the practice group: the more the team members support or help each other, the more likely they are to earn points in the contest. At the end of the week, a class newsletter announces each groups' standing in the class. In addition, some comments about the winners at each table and other issues concerning the contest are discussed (Slavin, 1980; Slavin, 1991; Slavin, 1995; Slavin, & Project, 1994).

Theoretical Backgrounds and Perspectives of Cooperative Learing

The four major theoretical backgrounds and perspectives of CL have been briefly discussed in this section. In addition to Deutsch's cooperation-competition theory discussed above, CL has its assumptions rooted in major theories. These perspectives are motivational theory's perspectives, social cohesion theory's perspectives, cognitive theory's perspectives, and developmental theory's perspectives.

Motivational Perspectives on Cooperative Learning. According to this perspective, motivation toward a task is a crucial part of the learning process. This assumes that other things are also driven by motivation. For this reason, this perspective mainly emphasizes the reward and goal structures through which students perform (Hänze & Berger, 2007; Li, 2017; Slavin, 1987, 1995). The motivationalist perspective assumes that incentive structures in CL create a condition in which the team will be successful only when group members can achieve their own individual objectives. For this reason, to achieve their individual objectives, team members need to help each other to do whatever is helpful for group to be successful. This is realized by helping their group members exert maximum potential. Rewarding teams depending on group achievement forms an interpersonal reward structure in which team members share social reinforcers such as praise and encouragement in response to peers' task-related activities (Hänze, & Berger, 2007; Panitz, 1999; Slavin, 2014).

Social Cohesion Perspective in Cooperative Learning. This perspective states that the effects of CL on academic performance are mediated by the a bond created within a team. It believes that group cohesion determines the success of the group's interactions. The secret of working in a group is that the students identify with each other and want each other to succeed. For the motivationalist perspective, the impetus that puts students together to work in a group is the students' motivatation learn together. On the other hand, for the social cohensive perspective the students work in groups because they care for one another or they want each other to succeed; they are interested in working together in a group because they like and respect each other and this helps them benefit from group work. The motivational and cohensive theoretical perspectives are similar in the fact that they highlight motivational rather than cognitive reasons for effectiveness of CL. The social cohesion theory perspective in CL emphasizes team-building activities in preparation for CL (Elliot et.al., 2016; Johnson & Johnson, 2009; Slavin, 2014).

Cognitive Perspectives in Cooperative Learning. Unlike motivationalist and social cohesiveness theories that depend on motivation, the cognitive theory asserts that interaction among learners by itself enhances learner achievement because it promotes the cognitive processing of information. Goup goals are the backbones of CL developed by motivationalist theories whereas building group cohensiveness is the essence of CL methods developed by social cohensive theories. However, CL methods by cognitive theorists lack the formulation of group goals and building group cohesiveness, although they have many unique techniques for them (Allal, Chanquoy & Largy, 2012; Slavin, 2014).

Developmental Perspectives on Cooperative Learning. Developmental perspectives assume that the interaction between students during a task that is appropriate to their developmental stage enhances their mastery of critical concepts (Slavin, 1987; Slavin, 2014). The Zone of Proximal Development (ZPD) is the gap

between the real developmental stage and the level of potential development. This gap is filled by the guidance of adults and cooperation with team members (Vygotsky, 1978). According to Vygotsky, cooperative activities among children of similar age levels enhance development because these children operate within one another's ZPDs. According to Piaget (1926), a developmental theorist, through collaborations with peers and adults, social-arbitrary knowledge of language, values, rules, morality, and symbol systems are learned.

A nother theory related to second language learning is Gardner's Socio-educational model of second language acquisition. According to Gardner (2005), an individual's motivation for learning a second language is related to and affected by classes of variables: 1) Attitudes towards the learning situation (teachers, curriculum, school environment, etc.); 2) Integrativeness (motivation to integrate with others' cultures); 3) Instrumental orientation (learning a language for practical purposes, such as to get a job); and 4) Language anxiety and parental encouragement. Motivation and gender largely affect second language learning (Fontecha, 2010), thus second language motivation is an influencing factor in second language learning and it is very important to include second language motivation as a covariate in the data analysis (Iwaniec, 2014; Tremblay and Gardner, 1995).

Method

General Background of the Research

Initially, this study identified the underutilization of the CL principles in the application of the student-centered approach in the study area as a major problem. The general question the study was motivated to answer was, "What is the effect of CL on English language achievement if the CL approach is applied appropriately using techniques that incorporate all CL principles?" Specifically and, firstly, the study was aimed at investigating the differences between the two instructional goal structures (i.e., cooperative and competitive learning) in affecting English language achievement. Secondly, it aimed at analyzing gender, intra-gender, and inter-gender differences in English language achievement within and across the cooperative, competitive, and control group learning conditions. And thirdly, it was aimed at determining the relationship between students' motivation to learn the English language and English language achievement.

Participants

The target population of the study was Arsi Negelle Shalla Secondary School grade 9 students of the year 2011 G.C. The school is located in Ethiopia, Oromia region, West Arsi Zone, Arsi Negelle town. The total population of the study was 328 students: 167 male and 161 female. The total sample size was 120 students: 61 male and 59 female. The minimum age of the participants was 15 and the maximum was 18. A one-way analysis of variance was conducted to identify age difference between CL and competitive learning, and the control group revealed that the age difference between the three groups was not significant, $F(2, 117) = 2.24$, $p > .05$.

During the sampling process of the participants for the study, the probability sampling design with a complex random sampling procedure followed by a stratified sampling method was used. Once the number of males and females in each stratum were identified, the sample size among strata was allocated using the following formula:

$$n_k = (n/N)N_k$$

Note: nk = the sample size for kth strata; Nk = the total population of kth strata; N = the total population size, and n = the total sample size

After determining how many members should be taken from each stratum, they were selected using simple random sampling.

Sampling and Sampling Procedures

After the pilot study, 180 students were selected using the stratified sampling formula above: 92 male and 88 female. Then the English language pre-test (43 items) was administered to these 180 students. After administering the pre-test, the papers were marked and the 180 students were categorized into high, medium, and low achieving ability levels; 1.5 standard deviations below and above the mean pre-test English language

score were considered low and high achievers respectively and the rest were medium achievers. After the categorizing the 180 students into the three ability levels, the number of students from the three ability levels who would be included in the 120 students (25% low achievers, 50% medium achievers, and 25% high achievers) for the main study were calculated using the stratified sampling formula. Then, the calculated number of 120 students were randomly selected from these three ability groups and remained in the study; 60 students were expelled from the study. The final step of sampling was the random assignment of the 120 students to the three ability levels. The importance of categorizing students into three ability levels is to make it suitable for the researcher to randomly select an unbiased and matching sample from the three ability levels for the experimental research.

Table 1
Assignment of Participants to the Three Groups by Gender and Ability Level

Groups	Ability	Gender		Total
		Male	Female	
Cooperative Learning	H	5	5	10
	M	10	10	20
	L	5	5	10
	Total	20	20	40
Competitive Learning	H	5	5	10
	M	10	10	20
	L	5	5	10
	Total	20	20	40
Control Group	H	5	5	10
	M	11	9	20
	L	5	5	10
	Total	21	19	40

Note. H = high; M = medium; L = low

Materials

Two types of instruments were used in this study: a 43-item English language test and the 12-item mini-AMTB. All English language test items were multiple-choice items and were constructed by the author of this study. The test was prepared from the grade 9 English language textbook (unit 1 up to unit 3) - the units the students learned or covered during the semester. The reason why the English language test was constructed from the units learned in the semester was to ensure that the treatment was not using the same material that the classroom teacher was covering at the time. The test consisted of reading, grammar, dialogue, and punctuation mark items.

The 12-item mini-AMTB uses a seven-point scale. It is based on the socio-educational model (Gardner, 2005) and is used to measure the participants' degree of motivation to learn the English language. The mini-AMTB (see Gardner, 2004) is made up of 12 items that fail into six dimensions of motivational constructs: Integrativeness (item 1 up to 3), Attitude towards the Learning Situation (items 6 and 9), Motivational Intensity (items 4, 5, and 11), Instrumental Orientation (item 7), Language Anxiety (item 8 and 10), and Parental Encouragement (item 12).

Both instruments were administered in the presence of the researcher and the English language teachers of the school. Each of the English language test items was worth one point. Therefore, the highest score was 43. For the mini-AMTB, the maximum score on the scale was 84 points and the minimum was 12. A high score for the five constructs, except language anxiety, indicates a positive measure of the constructs; a high score for language anxiety measures high language anxiety.

Cronbach alpha internal consistency reliability analysis was conducted for both of the instruments. Since the Cronbach alpha reliability analysis does not work for single items, the test-retest reliability analysis was used for item 7 and item 12. The Cronbach alpha reliabilities of the English language sub-tests were: reading (.76),

grammar (.87), dialogue (.65), and punctuation (.70). The total English language test's Cronbach alpha reliability was .91. The Cronbach alpha reliability or test-retest reliability (for single items) for the mini-AMTB sub-scales ranged from .68 (Attitude towards the Learning Situation), .74 (Integrativeness), .80 (Motivational Intensity), .90 (Instrumental Orientation), .93 (Parental Encouragement), and .93 (Language Anxiety). Gardner (1985) conducted item analysis for sub-scales of the AMTB in the French language.

The Item Discrimination Power and Item Difficulty Index (Appendix A) for the English language test were calculated. A point-biserial correlation was used to find the item discrimination power after entering the dichotomous data of the test into SPSS software. Poor distractors that were chosen by only a few students were discarded and replaced. According to Educational Data Systems[2], the items with problems will always produce low point-biserial correlations. However, the accompanying p-value can be low or high. For this reason, the use of point-biserial correlation is recommended to assess item quality whereas p-values are recommended to assess item difficulty. In this study, the point-biserial correlation was used to select quality English language test items. A point-biserial value of .15 is recommended although experience shows that good items have a point-biserial value greater than or equal to .25 (Educational Data Systems, n.d). In the English language test, except item number 22, which had a point-biserial value of .18, all items had a point-biserial value of .25 and above.

The teaching materials used in this study were the grade nine English language textbook, plasma television teacher guide, and numbered cards.

Research Design

The study is a pure quantitative research and follows the positivist paradigm. The true experimental and correlational research designs were used in this study. The pre-test post-test control group experimental design using two treatment groups and one control group was applied. Cooperative and competitive groups were treatment groups whereas the individualized learning group was a control group. The experimental design used instructional goal structure (i.e., with categories of cooperative, competitive, and control group conditions) and gender (i.e., with categories of male and female) as independent variables, and the English language pre-test and the language motivation as covariates to analyze the effect of these factors on English language achievement (i.e., the dependent variable) controlling for the effect of the covariates. The reason that this design is preferred over others is that it is believed to be the best design to analyze the effect of one variable on the other controlling for the confounding variables.

On the other hand, a correlational research design, in which the aggregate measures of the mini-AMTB (i.e., Integrativeness, Motivational Intensity, Attitudes towards the Learning Situation, Instrumental Orientation, Language Anxiety, and Parental Encouragement) were correlated with English language achievement, was also used. This is to observe to what extent these measures correlate with English language learning.

Data Analysis

SPSS software version 20 was used for data analysis. To analyze the inter-gender (e.g., males and females in the CL condition and other two learning conditions) differences in English language achievement before and after treatment, a dependent sample t-test was conducted. Dependent sample t-test was used because the data were related since the students learn in the same class. The second data analysis was on the intra-gender (e.g., same-sex groups in three conditions) difference in English language achievement in the pre-test and post-test. In this analysis, the same-sex groups in the three conditions were taken as an intra-gender factor (i.e., independent variable) with categories (for example, males in CL, males in competitive learning, and males in control group) and English language achievement (i.e., the dependent variable). These variables were analyzed using a one-way analysis of variance (ANOVA) to analyze the effect of one on the other. Using the instructional goal structure and gender as independent factors, and the English language pre-test and language motivation as covariates, Analysis of Covariance (ANCOVA) was used to analyze the effect of instructional goal structure and gender on English language achievement, controlling for the covariates. Another data analysis method, the Pearson Product Moment Correlation, was used to analyze the relationship between the aggregate measures of the mini-AMTB and English language achievement.

[2] [Web log post]. (n.d.). http://www.eddata.com

Procedure

Three weeks before the main study, the researcher visited the school, and contacted the administrators, teachers, and students. He then constructed a test from the English language course. The researcher also took the list of grade 9 students from the school office and identified the total number of males and females. During these weeks, he took a sample for the pilot study. After the pilot study, papers were marked, an item analysis was conducted, and the best English language test items were selected.

Weeks later, the researcher went to the school again and selected 180 students for the administration of the pre-test. He administered the English language test (43 items) and the mini-AMTB (12 items). He scored the tests and categorized the 180 students into three ability groups.

Next, the researcher used the stratified formula and calculated the number of students from the three ability groups who would be included in the sample of 120 students for the main study. The total sample of 120 students was randomly selected from the three ability groups. Finally, the 120 students were randomly assigned to the cooperative, competitive, and control group learning conditions.

Then, the researcher taught using the three learning conditions each for one month. During the teaching period, after each presentation of a lesson to students, the researcher gave individual and group work. At the end of a week, tournaments were held. Group scores were posted for cooperative learners. Individual scores were posted for the competitive learners after each test and tournament. Tests were administered for competitive and individualized learners. Finally, the post-test was administered to the three groups and the final data analysis was conducted.

Teams-Games-Tournament in This Study

In the CL condition, students were grouped into 10 groups of four students. Each group consisted of one high achiever, two medium achievers, and one low achiever. After grouping was completed, the students were told that they were playing academic games and each student had to earn points for his/her group and the winning teams would be awarded at the end of the tournament. Generally, an orientation was given on how to work in groups at the beginning of each class.

The students were then taught according to the curriculum in the classroom and each team of students was told to study together, practice, and quiz each other after every lesson that the researcher taught. The researcher taught the day's content and gave similar tasks to be accomplished in groups after each class. The researcher prepared the questions and informed students that they should prepare for that week's tournament. At the end of each week, 10 homogenous-ability students assembled for a tournament at three tables (see Appendix B). At the start of a tournament, a student took a numbered card and the researcher asked a question. If the student answered the question correctly, they earned a point for their group. The questions the students were asked at the table were similar to the test items administered to the competitive and control groups. Each of the students was asked two questions at a table. A participant at a table could win as many cards (points) as they could win from their group. A participant got a chance to win additional cards when other participants failed to answer a question and they could manage to answer questions others failed to answer. For example, student number 1 (Appendix B) earned one additional card for his group in addition to the two cards he could win by answering his questions.

Competitive teaching mode

In this learning condition, the students were taught according to the curriculum but the researcher used a competitive teaching mode – rewarding those students who were high performing in the class and giving less attention to low achievers. At the end of a week, this class was tested and the top ten high-scoring students were announced.

Control group

In this learning condition, the students were taught according to the curriculum and the class was tested using the same test as the test the competitive group was given. However, in this condition, the students were taught

via the individualistic learning mode in which the students were left alone to work for themselves. They were neither rewarded in groups nor individually. The researcher only told them their test results and continued teaching.

Results

The Relation of Instructional Goal Structures and Gender to Students' English Language Achievement

Pre-test analysis

In this section, the inter-gender difference is analyzed. The objective of pre-test dependent sample t-test analysis was to make sure that gender groups in a learning condition are matching in language achievement before treatment and to reasonably expect the effect of the treatment on language learning by gender in a learning condition in the post-test analysis. The dependent sample pair-wise t-test was conducted on pre-test data to analyze gender differences in English language achievement in a learning condition (i.e., cooperative, competitive, and control group independently) and the results are presented as follows: 1) The dependent sample t-test for pre-test data analysis revealed that the difference between males (M = 22.60, SD = 8.75) and females (M = 23.45, SD = 8.49) in the CL group in English language achievement was not significant, t (19) = .31, p > .05, two tailed; 2) The dependent sample t-test for pre-test data analysis also indicated that the difference between males (M = 22.15, SD = 8.52) and females (M = 20.75, SD = 8.64) in competitive learning group was not significant, t(19) = -.47, p > .05, two tailed; 3) Similarly, the pre-test dependent sample t-test analysis revealed that the difference between males (M = 21.37, SD = 8.43) and females (M = 23.16, SD = 8.78) in the control group condition in English language achievement was not significant, t(18) = .82, p > .05, two tailed.

The one-way ANOVA pre-test intra-gender (i.e., same-sex groups in the three conditions) analysis revealed that the difference between groups of males in the three learning conditions was not significant, F (2, 58) = .02, p > .05. Similarly, the result was also non-significant for groups of females, F (2, 56) = .21, p > .05.

The post-test analysis

In order to test the differences in English language achievement as a result of the treatment effect in the three learning conditions, an Analysis of Covariance (ANCOVA) was conducted. The analysis of covariance, when the goal structure and gender were used as fixed factors, and the pre-test English language score and language motivation were used as covariates, revealed that the difference between the three groups in English language achievement was significant, F (2, 112) = 166.70, p < .05. However, the gender difference was not significant again in the post-test analysis of covariance, F (1, 112) = .19, p > .05. From post-hoc pair-wise comparison of the Least Significant Difference analysis, it can be seen that the CL group (M = 32.45, SD = 7.18) significantly outperformed both the competitive learning group (M = 25.00, SD = 8.60) and the control group (M = 26.53, SD = 8.42), p < .05. However, the control group significantly outperformed the competitive group in this analysis.

The one-way ANOVA post-test intra-gender (i.e., same-sex groups in the three conditions) analysis revealed that the difference between groups of males in the three learning conditions in English language achievement was significant, F (2, 58) = 4.34, p < .05. The post-hoc Tukey HSD pair-wise comparison revealed that males in the CL group (M = 32.20, SD = 7.56, n = 20) significantly outperformed both males in the competitive learning group (M = 25.50, SD = 8.36, n = 20) and males in the control group (M = 25.90, SD = 8.32, n = 21). However, the difference between males in the competitive learning group and males in the control group was not significant.

The one-way ANOVA post-test intra-gender analysis revealed that the difference between groups of females in the three learning conditions in English language achievement was significant, F (2, 56) = 5.11, p < .05. The post-hoc Tukey HSD pair-wise comparison revealed that females in the CL group (M = 32.70, SD = 6.76, n = 20) significantly outperformed both females in the competitive learning group (M = 24.50, SD = 9.01, n = 20) and females in the control group. The difference between female competitive learners and female individualistic learners was also not significant.

Table 2
Summary of Pair-wise t-test for Pre-test and Post-test Data

		Cooperative Learning				Competitive Learning				Control Group			
Dependent Measure	Sex	M	SD	N	t	M	SD	N	t	M	SD	N	t
Pre-test	M	22.60	8.75	20	.31	22.15	8.52	20	-.47	22.24	8.53	21	.82
	F	23.40	8.46	20		21.75	8.86	20		23.16	8.78	19	
	Total	23.00	8.50	40		21.96	8.58	40		22.67	8.48	40	
Post-test	M	32.20	7.56	20	.21	25.50	8.36	20	-.28	25.90	8.32	21	1.02
	F	32.70	6.96	20		24.50	9.01	20		27.21	8.70	19	
	Total	32.45	7.18	40		25.00	8.60	40		26.53	8.42	40	

Note. N refers to the number of males or females in a learning condition.
*p < .05

The dependent samples pair-wise t-test was conducted on post-test data to analyze the gender difference in English language achievement as a result of the treatment effect and the results are presented as follows: 1) The pair-wise dependent sample t-test for post-test data analysis revealed that the difference between males (M = 32.20, SD = 7.56, N = 20) and females (M = 32.70, SD = 6.96, N = 20) in English language achievement in CL as a result of the treatment effect was not significant, t(19) = .21, p > .05, two tailed; 2) The pair-wise t-test for post-test data analysis also indicated that the difference between males (M = 25.50, SD = 8.36, N = 20) and females (M = 24.75, SD = 8.86, N = 20) in English language achievement in the competitive learning group as a result of the treatment effect was not significant, t(19) = -.28, p > .05, two tailed; and 3) The post-test pair-wise t-test analysis revealed that the difference between males (M = 25.90, SD = 8.32, N = 21) and females (M = 27.21, SD = 8.70, N = 19,) in English language achievement in the control group condition was not significant, t(18) = 1.02, p > .05, two tailed.

The Relationship between English Language Learning Motivation and English Language Achievement

The correlation between the aggregate measures of the mini-AMTB and English language achievement was found and summarized in the table below.

Table 3
Correlations between the Aggregate Measures of the Mini-AMTB English Language Achievement

	1	2	3	4	5	6	7
INT	-	.70**	.61**	.35**	-.60**	.47**	.77**
MI		-	.85**	.44**	-.67**	.50**	.83**
ALS				.48**	-.59**	.45**	.77**
IO					-.38**	.29**	.56**
LA						-.49**	-.77**
PE							.57**
Pre-test score							-

Note. INT = Integrativeness; MI = Motivational Intensity; ALS = Attitudes towards the Learning Situation; IO = Instrumental Orientation; LA = Language Anxiety; PE = Parental Encouragement
**p < .01, two tailed; N = 120

The analysis revealed that all of the correlations between the aggregate measures of the mini-AMTB and English language achievement were significant at .01 levels. The correlation for the five dimensions of the mini-AMTB [i.e., Integrativeness, r(118) = .77, p < .01; Motivational Intensity, r(118) = .83. p < .01; Attitudes towards the Learning Situation, r(118) = .76, p < .01; Instrumental Orientation, r(118) = .56, p < .01; and Parental Encouragement, r(118) = .57, p < .01] were positive; however, the correlation was negative for Language Anxiety,

$r(118) = -.77$, $p < .01$. The highest positive correlation was observed for Motivational Intensity and the lowest positive correlation was observed for Instrumental Orientation. Studies have reported different findings on the role of the specific aggregate measures of the AMTB on second language learning.

Regarding foreign language learning motivation, Gholami, Allahyar, and Rafik-Galea (2012) reported that the dominant motivational construct among high school students was instrumental orientation. They also reported that high-achieving students were more integratively motivated than low achievers. Gardner (2012) also reported that the correlations of the integrative motivation sub-items score with English grades showed that the aggregate score was a more consistent correlate across samples than the integrative motivation sub-items scores themselves. Similar to this study, Gardner also found that scores in the English language at the end of the academic year were significantly and positively correlated with motivation and integrativeness, and negatively and significantly correlated with language anxiety. Iwaniec (2014) also reported that the role of parental encouragement and language learning anxiety was limited. More interestingly, a recently published book by Lamb (2019) discussed the main theoretical approaches to language learning motivation and presented ways in which motivation theory has been applied in practice,. It showcased examples of motivation research in particular contexts and with particular types of language learners.

Discussion

Consistent with previous studies, CL had significantly and positively affected students' English language achievement. Compared to the other two learning modes, the positive effects of CL dominated the literature. Numerous studies on CL state that CL promotes high achievement; retention and transfer of knowledge gained through learning; the long-lasting implementation of knowledge; the internalization of values, attitudes, and behavior; advancing teachers' professional identity; and participation and inclusion in the community of practice (Johnson & Johnson, 2017). A meta-analysis by Johnson, Johnson, Roseth, and Shin (2014) assessed the degree to which achievement was positively correlated with motivation and revealed that those factors related with positive interdependence resulted in higher motivation and performance than those related to negative or no interdependence conditions. A more recent study in Ethiopia by Molla and Muche (2018) also found a positive impact of CL on student learning. CL had positively and significantly affected reading and vocabulary learning (Zarei, 2012). A more recent study by Alghamdy (2019) reported that CL in language classrooms enhanced students' English language skills, promoted positive relationships among learners, helped learners play different roles, promoted oral presentation skills, and built students' self-confidence.

On the other hand, some studies reported the effect of the TGT on English language achievement and motivation towards the language. The TGT significantly and positively affected maths achievement more than a traditional method (Salam, Hossain, & Rahman, 2015). Nasution (2018) also found a significant effect of the TGT on the vocabulary learning of grade 10 students. Learning motivation and efficiency can be enhanced through educational games (Liu & Chen, 2013). CL positively affects both the achievement and motivation of students towards learning (Johnson, Johnson, Roseth, & Shin, 2014). More recent studies by Fauzi, Buhun, and Purwadi (2019); Silitonga and Wu (2019); and Artha, Syam, and Priambodo (2020) reported the significant positive effect of TGT CL on both foreign learning achievement and motivation.

Additionally, the finding of this study showed the absence of gender difference in English language achievement, contrary to previous studies (Chan, 2018; Nyikos, 1990; Rua; 2006; Schmenk, 2004; Tong, Irby, Lara-Alecio, Yoon, & Mathes, 2010; Van der Slik, Van Hout, and Schepens; 2015). The gender difference was not significant both within and between the conditions of learning in this study. Although few studies have been conducted on the gender difference in language performance by the learning mode conditions, Gillies and Ashman (1995) asserted that, from learning in gender-balanced, mixed-ability CL groups, high, medium, and low-ability students all benefited academically. On the other hand, Nawaz, Hussain, and Javed (2011) found that the CL method was found to be more effective than the lecture method in enhancing the academic achievement and academic self-efficacy of the students. Another study by Ghanbari and Samar (2016) found that the heterogeneity of the second language learning groups did not bring about a significant effect between the groups' language achievement whereas gender difference in the language achievement existed between groups.

Moreover, 'Which learning mode is significantly favored by males and females?' was one of the research questions in this study. An intra-gender comparison revealed the presence of differences between groups of females in the three learning conditions. A group of females in the CL condition significantly outperformed a group of females in the competitive learning condition, showing no significant difference with a group of females in the control group. The group of female competitive learners also did not show a significant difference with the control group in English language achievement. In the case of males, the group of males in the CL condition significantly outperformed both the group of males in the competitive learning condition and the group of males in the control group. However, there was no significant difference between the competitive learner males and the control group males. This finding reveals that the difference in English language achievement between groups of students of the same sex occurred as a result of the learning mode differences. In other words, the intra-gender and inter-gender language achievement comparison revealed that the difference did not occur as result of the gender difference but occurred as a result of learning mode differences. In the case of inter-gender comparison, no difference was observed; however, in the case of intra-gender comparison, the difference between groups of students of the same sex across the three learning conditions was significant. Surprisingly, a myriad of research findings support the effectiveness of CL, especially for females (Hossain, Tarmizi, Aziz, & Nordin, 2013; Rodger, Murray, & Cummings, 2007; Wang, 2012) and for both sexes (Bilen & Tavil 2015; Esiobu, 2011).

Generally, the study revealed that the CL mode is more favored (i.e., as its impact on the achievement of both sexes showed) by both sexes than competitive learning mode. However, contrary to the previous research findings cited in the previous paragraph, the study revealed that males preferred the CL mode than females as the males in the CL condition outperformed males in both competitive and control group conditions, whereas cooperative learner females outperformed on only female competitive learners.

The Researcher's Own Experiences on the Implementation of Cooperative Learning: Challenges

The researcher implemented the Teams-Games-Tournament and Student Teams Achievement Divisions in his teaching. The challenges came from students and local classroom conditions such as class size and time constraints as well as the curriculum.

For instance, high-achieving students resisted accepting the low achieving students' contributions (individual accountability) to the group task. A study by Alghamdy (2019) also found this problem. This problem happens when, in Student Teams Achievement Divisions for example, all group members sit for a test after studying in groups for days and the instructor marks the individual tests to sum up the individual contributions to the group. At this time, the high-achieving students argue that their total group result does not represent them and is not fair. For this reason, most of the time, high-achieving students want to do all the group assignments by themselves without including the contributions of all the group members.

The large class sizes and unsuitable classrooms are also problems for managing group discussions as a part of structured CL. Also, where there are no clear guidelines and time allocation for implementation of the structured CL in the curriculum, it took extra time for the researcher to adjust his methods and procedures to fit the situation.

The next issue of this study was the association between language motivation and English language achievement. The correlations of the five constructs of the mini-AMTB with English language achievement were positive and significant. However, the correlation for Language Anxiety was negative and significant. This finding is similar to the previous study conducted by Gardner (2005) in Croatia, Poland, Romania, and Spain that found positive correlations between these five constructs of the AMTB and student grades and a negative correlation with Language Anxiety in most cases. Both language achievement and attitudes toward the language improved in CL (Bilen & Tavil, 2015). Azizinezhad, Hashemi, and Darvishi (2013) and Obinna-Akakuru, Onah, and Opara (2015) found that CL enhanced oral communicative competence and English as a foreign language learning motivation. Other studies discussed gender-based motivation towards the second language. For instance, Park and French's (2013) study results indicated that females reported higher language anxiety levels compared to males. Chan's (2018) study reported that female students were more positively oriented towards native speaker pronunciation and tended to adopt it as their teaching model and learning target. Chan also stated that female students had greater confidence than males in their ability to understand English pronunciation. Wang (2012)

reported that CL classes improved achievement motivation among female college students. Female students' integrative motivation and attitudes towards the English language are higher than male students (Al Harthy, 2017). Generally, these findings report that CL enhances females' language learning motivation than males' language learning motivation.

The teaching period used to observe the effects of implementing the teaching methods mentioned lasted only one month because of time and financial constraints and this can be one limitation. Future studies might be invited to take a longer time to obtain stronger results.

Implications

The findings of this study can make contributions to the language curriculum. When strong language performance based on the utilization of CL is desired, the utilization of CL principles and appropriately using effective CL techniques (e.g., TGT in this study) are more fruitful. Especially, in Ethiopian schools' language classrooms where the teachers show a lack of knowledge in how to use CL methods, the introduction of different CL guidelines and techniques are more important. It is believed that when CL techniques and principles are applied by trained teachers in a CL classroom, the maximum benefit from the method is expected. Misuse of the method is the problem in Ethiopian classrooms and this study will help to improve the problem.

Gender-based individual differences in the preferred learning approach is an important issue that language curriculum specialists should pay due attention to and this study opens the way for future empirical research. Firstly, the study of gender differences in language learning locally should be compared with findings from another part of the world. This helps to identify what local factors are pulling the sexes behind in learning. Secondly, when designing a non-gender-biased curriculum, it is important to study how both sexes are similar and different in the ways they learn.

The issue of language motivation raised in this study is also something new that was missing in Ethiopian language research. It is believed that language motivation serves as fuel for learning a language and this study opens the gate for future research in the Ethiopian context. In addition, the idea that the methods of teaching and learning themselves ought to be those that can trigger motivation towards the learning (i.e., the role of structured CL) has been discussed in this study and the researcher believes that this study shows directions for curriculum designers and teachers in the subject area. The careful incorporation of the CL models in the Ethiopia English language curriculum is also recommended.

Its implication for teacher education is: 1) Teacher training has to include these methods as well as procedures that the teachers can easily apply; and 2) The mechanisms by which teachers teach the groups of students how to work in groups, assure individual accountability, and apply other CL principles should be included in teacher training.

Finally, the researcher encourages future local and international researchers to conduct further studies on other CL models such as Student Teams Achievement Divisions, Learning Together, Academic Controversy, Group Investigation, Teams Assisted Individualization, and Cooperated Integrated Reading Composition.

Conclusion

Based on the findings of this study, the following conclusions can be drawn: 1) CL mode, when it is applied in English language learning classrooms with effective CL techniques, can result in higher student English language achievement than the competitive learning mode; 2) The significant gender difference in English language achievement was not supported by the findings of this study when the two sexes are compared after teaching in classes of mixed ability and mixed gender composition using cooperative and competitive learning condition treatment conditions and using individualistic learners as a control group; 3) The significant difference in English language achievement between the two sexes within a learning condition group was not supported by the findings of this study; 4) CL mode was found to be effective for both sexes, but found to be more

effective for males; and 5) English language achievement was significantly and positively correlated with the five dimensions of the mini-AMTB, while it produced a significant negative correlation with Language Anxiety.

Acknowledgments

I would like to thank Addis Ababa University for funding this work. My gratitude also goes to Dr. Girma Lemma for advising me throughout this study. I thank Arsi Negelle Shalla high school administrators, teachers, and students for their active cooperation during this study.

References

Alghamdy, R. Z. (2019). EFL learners' reflections on cooperative learning: Issues of implementation. *Theory and Practice in Language Studies*, *9*(3), 271-277. https://doi.org/10.17507/tpls.0903.03

Al Harthy, S. R. (2017). English language motivation between gender and cultures. *Journal of Educational and Social Research*, *7*(2), 123-132. https://doi.org/10.5901/jesr.2017.v7n2p123

Allal, L., Chanquoy, L., & Largy, P. (2012). *Revision Cognitive and Instructional Processes: Cognitive and Instructional Processes.* Springer Science & Business Media.

Alsheikh, N. & Elhoweris, H. (2011). United Arab Emirates (UAE) high school students' motivation to read in English as a foreign language. *International Journal of Language Studies (IJLS)*, *5*(4), 53-68.

Artha, A. A., Syam T, A. R., & Priambodo, A. (2020). The application of teams games tournament (TGT) and teaching game for understanding (TGfU) learning models on learning motivation and Volley Ball passing skills. *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal*, *3*(1), 46-53. https://doi.org/10.33258/birle.v3i1.751

Ayele, Z.A. & Olamo, T.G. (2019). The practices, challenges and opportunities of cooperative learning at Hawassa University – Ethiopia. *IOSR Journal of Humanities And Social Science* (IOSR-JHSS), 24 (3), 15-27.

Azizinezhad, M., Hashemi, M., & Darvishi, S. (2013). Application of cooperative learning in EFL classes to enhance the students' language learning. *Procedia - Social and Behavioral Sciences*, *93*, 138-141. https://doi.org/10.1016/j.sbspro.2013.09.166

Baloche, L., & Brody, C. M. (2017). Cooperative learning: exploring challenges, crafting innovations. *Journal of Education for Teaching, 43*(3), 274-283. https://doi.org/10.1080/02607476.2017.1319513

Bilen, D., & Tavil, Z. M. (2015). The effects of cooperative learning strategies on vocabulary skills of 4th grade students. *Journal of Education and Training Studies*, *3*(6), 151-165. https://doi.org/10.11114/jets.v3i6.1062

Boling, N. C., & Robinson, D. H. (1999). Individual study, interactive multimedia, or cooperative learning: Which activity best supplements lecture-based distance education? *Journal of Educational Psychology*, *91*(1), 169-174. https://doi.org/10.1037//0022-0663.91.1.169

Çelik, S., Aytın, K., & Bayram, E. (2013). Implementing cooperative learning in the language classroom: Opinions of Turkish teachers of English. *Procedia - Social and Behavioral Sciences, 70*, 1852-1859. https://doi.org/10.1016/j.sbspro.2013.01.263

Chan, J.Y.H (2018). Gender and attitudes towards English varieties: Implications for teaching English as a global language. *System*, *76*, 62-79.

Desta, D., Chalchisa, D., Mulat, Y., Berihum, A. & Tesera, A. (2009). Enhancing active learning through self- and-peer reflections: the case of selected schools in Ethiopia. *Journal of International Cooperation in Education, 12*(1), 71-87.

Deutsch M. (2015). Cooperation, Competition, and Conflict. In M. Deutsch (Ed.), A *pioneer in developing peace Psychology* (pp. 47-70). Springer. https://doi.org/10.1007/978-3-319-15440-4_3

Deutsch, M. (2011). Cooperation and Competition. *Conflict, Interdependence, and Justice*, 23–40. https://doi.org/10.1007/978-1-4419-9994-8_2

Deutsch, M. (1949). A Theory of co-operation and competition. *Human Relations*, *2*(2), 129-152. https://doi.org/10.1177/001872674900200204

Elliot, A. J., Aldhobaiban, N., Kobeisy, A., Murayama, K., Gocłowska, M. A., Lichtenfeld, S., & Khayat, A. (2016). Linking social interdependence preferences to achievement goal adoption. *Learning and Individual Differences, 50*, 291–295. https://doi.org/10.1016/j.lindif.2016.08.020

Esiobu, G.O. (2011). Research achieving gender equity in science class shift from competition to cooperative learning. *Multicultural Education & Technology Journal*, *5*(4), 244-257. https://doi.org/10.1108/17504971111185081

Fauzi, M. F., Buhun, M. F., & Purwadi, A. (2019).The influence of teams games tournament (TGT) toward students' interest in Arabic language learning. *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature*, *2*(2),135-148. https://doi.org/10.22219/jiz.v2i2.9986

Feery, K. (2008).Current perspectives on the role of gender in second language acquisition (SLA) research.*The ITB Journal, 9*(1), 21-51.

Ferguson-Patrick, K. (2016). The importance of teacher role in cooperative learning: The effects of high-stakes testing on pedagogical approaches of early career teachers in primary schools. *Education 3-13, 46*(1), 89–101. https://doi.org/10.1080/03004279.2016.1189946

Fontecha, A. F. (2010). Gender and motivation in EFL vocabulary production. In R. M. J. Catalán (Ed.), *Gender perspectives on vocabulary in foreign and second languages* (pp. 93-116). Palgrave Macmillan.https://doi.org/10.1057/9780230274938_5

Gardner, R. C. (1985). *Social Psychology and second language learning: The role of attitudes and motivation*. Hodder Arnold.

Gardner, R.C. (1985). *The attitude/Motivation Test Battery: Technical report*. University of Western Ontario, Canada. http://publish.uwo.ca/~gardner/docs/AMTBmanual.pdf

Gardner, R. C. (2004). *Attitude / Motivation Test Battery*. International AMTB research project (English version). The University of Western Ontario, Canada. http://publish.uwo.ca/~Gardner/docs/englishamtb.pdf

Gardner, R. C. (2005). *Integrative motivation and second language acquisition*. Canadian Association of Applied Linguistics: Association of Joint Plenary Talk. Social Sciences and Humanities Research Council of Canada.

Gardner, R. C. (2012). Integrative motivation and global language (English) acquisition in Poland. *Studies in Second Language Learning and Teaching*, *2*(2), 215-226. https://doi.org/10.14746/ssllt.2012.2.2.5

Gezahagn, G. (2003). In-service English language teachers' perceptions of the factors that influence the implementation of group work and activities in ELT classes. *Ethiopian Journal of Education*, *23*, 103-126.

Ghanbari, N., &Samar, R.G. (2016).Grouping strategies and writing achievement in cooperative learning. *Journal of Applied Linguistics and Language Research*, *3*(7), 85-97.

Ghodbane, T., & Hamzaoui El Achachi, H. (2019). Narrowing the achievement gap between EFL students in oral expression through cooperative learning. *Arab World English Journal*, *10*(2), 365-378. https://doi.org/10.24093/awej/vol10no2.28

Gholami, Z., Allahyar and Rafik-Galea (2012). Integrative motivation as an essential determinant of achievement: A case of EFL high school students. *World Applied Sciences Journal, 17*(11), 1416-1424.

Gillies, R. M. (2016). Cooperative learning: Review of research and practice. *Australian Journal of Teacher Education, 41*(3), 39-54. http://dx.doi.org/10.14221/ajte.2016v41n3.3

Gillies, R. M., & Ashman, A. F. (1995).The effects of gender and ability on students' behaviours and interactions in classroom-based work groups. *British Journal of Educational Psychology*, *65*(2), 211-225. https://doi.org/10.1111/j.2044-8279.1995.tb01143.x

Gillies, R. M., Ashman, A., & Terwel, J. (2007). *The Teacher's role in implementing cooperative learning in the classroom*. Springer Science & Business Media.

Hänze, M. E., & Berger, R. (2007). Cooperative learning, motivational effects, and student characteristics: An experimental study comparing cooperative learning and direct instruction in 12th grade physics classes. *Learning and Instruction, 17*(1), 29-41. https://doi.org/10.1016/j.learninstruc.2006.11.004

Hossain, A., Tarmizi, R.A., Aziz, Z., & Nordin, N. (2013). Group learning effects and gender differences in mathematical performance. *Croatian Journal of Education*, *15*(2), 41-67.

Iwaniec, J. (2014). Motivation of pupils from southern Poland to learn English. *System*, *45*, 67-78. https://doi.org/10.1016/j.system.2014.05.003

Johnson, D. W., & Johnson, F. (2013). *Joining together: Group theory and group skills* (11th ed.). Allyn & Bacon

Johnson, D. W., Johnson, R. T., & Holubec, E. J. (2013). *Cooperation in the classroom* (9th ed.). Interaction Book Company.

Johnson, D. W., & Et al. (1981). Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychological Bulletin*, *89*(1), 47-62. https://doi.org/10.1037//0033-2909.89.1.47

Johnson, D. W., & Johnson, R. T. (1974). Instructional goal structure: Cooperative, competitive, or individualistic. *Review of Educational Research*, *44*(2), 213-240. https://doi.org/10.2307/1170165

Johnson, D. W., & Johnson, R. T. (1974). Instructional goal structure: Cooperative, competitive, or individualistic. *Review of Educational Research*, *44*(2), 213-240. https://doi.org/10.2307/1170165

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38*(5), 365-379. https://doi.org/10.3102/0013189x09339057

Johnson, D. W., & Johnson, R. T. (2017). The use of cooperative procedures in teacher education and professional development. *Journal of Education for Teaching, 43*(3), 284-295. https://doi.org/10.1080/02607476.2017.1328023

Johnson, D. W., Johnson, R. T., Johnson, J., & Anderson, D. (1976). Effects of cooperative versus individualized instruction on student prosocial behavior, attitudes toward learning, and achievement. *Journal of Educational Psychology*, *68*(4), 446-452. https://doi.org/10.1037//0022-0663.68.4.446

Johnson, D. W., Johnson, R. T., Roseth, C., & Shin, T. S. (2014). The relationship between motivation and achievement in interdependent situations. *Journal of Applied Social Psychology*, *44*(9), 622-633. https://doi.org/10.1111/jasp.12280

Jolliffe, W. (2015). Learning to learn together: Cooperation, theory and practice. *Education 3-13*, *43*(1), 1-4. https://doi.org/10.1080/03004279.2015.992617

Kumar, M.K. (2017). Challenges of implementing student-centered strategies in classrooms. *International Research Journal of Engineering and Technology*, *3*(12), 1224-1227

Kyndt, E., Raes, E., Lismont, B., Timmers, F., Cascallar, E., & Dochy, F. (2013). A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? *Educational Research Review*, *10*, 133-149. https://doi.org/10.1016/j.edurev.2013.02.002

Lamb, M. (2019). *The Palgrave handbook of motivation for language learning*. Springer Nature.

Li, K. (2017). Motivational regulation and L2 Learning. In Motivational Regulation in Foreign Language Learning. Palgrave Macmillan. https://doi.org/10.1057/978-1-349-93118-7_3

Liu, E. Z., & Chen, P. (2013). The effect of Game-Based Learning on students' learning performance in science learning – A case of "Conveyance Go". *Procedia - Social and Behavioral Sciences*, *103*, 1044-1051. https://doi.org/10.1016/j.sbspro.2013.10.430

Lopez-Rua, P (2006). *The sex variable in foreign language learning: An integrative approach. PORTA LINGUARUM*, *6*, 99-114.

Lucha, Z.T., & Bongase, T.M. (2015).A Study on the practices of cooperative language learning method: Enango high school grade 9 in focus.*Science, Technology and Arts Research Journal*, *4*(2), 278-287. https://doi.org/http://dx.doi.org/10.4314/star.v4i2.39

Mady, C., & Seiling, A. (2017). The coupling of second language learning motivation and achievement according to gender. *Theory and Practice in Language Studies*, *7*(12), 1149. https://doi.org/10.17507/tpls.0712.01

Masgoret, A., & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning*, *53*(1), 123-163. https://doi.org/10.1111/1467-9922.00212

McCafferty, S. G., Jacobs, G. M., & Iddings, A. C. (2006). *Cooperative learning and second language teaching*. Cambridge University Press.

Moges, B. (2019). Practices and challenges of cooperative learning in selected college of Arsi University: As a ootivational factor on enhancing students' learning. *Universal Journal of Psychology 7* (1), 1 - 17. https://doi.org/ 10.13189/ujp.2019.070101.

Mohammadjani , F. & Tonkaboni , F. (2015). A comparison between the effect of cooperative learning teaching method and lecture teaching method on students' learning and satisfaction level. *International Education Studies*, *8*(9), 107-112. https://doi.org/ doi:10.5539/ies.v8n9p107

Molla, E., & Muche, M. (2018). Impact of cooperative learning approaches on students' academic achievement and laboratory proficiency in Biology subject in selected rural schools, Ethiopia. *Education Research International*, *2018*, 1-9. https://doi.org/10.1155/2018/6202484

Nasution, M. (2018). The effect of Team Games Tournament (TGT) method toward vocabulary mastery at the tenth grade students of high school 1 Padangsidimpuan. *IOP Conference Series: Earth and Environmental Science*, *175*, 1-5. https://doi.org/10.1088/1755-1315/175/1/012087

Nawaz, Q., Hussain, L. & Javed, M. (2011). A gender based comparative study of cooperative learning and lecture demonstration method on the academic achievement and academic self-concept at elementary school level. *Pakistan Journal of Education*, *28*(2), 47-60.

Nyikos, M. (1990). Sex-related differences in adult language learning: Socialization and memory factors. *The Modern Language Journal*, *74*(3), 273-287. https://doi.org/10.2307/327624

Obinna-Akakuru, A.U., Onah, T.A., & Opara, D.C. (2015). Cooperative learning and student's academic achievement in English language in Imo State, Nigeria. *IOSR Journal of Research & Method in Education, 5*(3)*, 26-29*

Owens, L., & Straton, R. G. (1980).The development of a co-operative, competitive, and individualised learning preference scale for students. *British Journal of Educational Psychology*, *50*(2), 147-161. https://doi.org/10.1111/j.2044-8279.1980.tb02440.x

Owens, L., & Straton, R. G. (1980). The development of a co-operative, competitive, and individualised learning preference scale for students. *British Journal of Educational Psychology*, *50*(2), 147-161. https://doi.org/10.1111/j.2044-8279.1980.tb02440.x

Oxford, R. L. (1997). Cooperative learning, collaborative learning, and interaction: Three communicative strands in the language classroom. *The Modern Language Journal*, *81*(4), 443-456. https://doi.org/10.1111/j.1540-4781.1997.tb05510.x

Panitz, T. (1999). The motivational benefits of cooperative learning. *New Directions for Teaching and Learning*, *1999*(78), 59-67. https://doi.org/10.1002/tl.7806

Park, G., & French, B. F. (2013). Gender differences in the foreign language classroom anxiety scale. *System*, *41*, 462-471.

Pedersen, J. E., & Digby, A. D. (2014). *Secondary schools and cooperative learning*: *Theories, models, and strategies*. Routledge.

Peldon, D., & Chalermnirundorn, N. (2018). The effects of cooperative learning strategies on sixth grade Bhutanese students' learning achievement and satisfaction in social studies class. *St.Theresa Journal of Humanities and Social Sciences*, *4*(2), 1-29.

Piaget, J. (1926). *The language and thought of the child*. Harcourt Brace.

Reda, W. N., & Hagos, G. T. (2015). The practices of student network as cooperative learning in Ethiopia. *Africa Education Review*, *12*(4),696-712. https://doi.org/10.1080/18146627.2015.1112162

Rodger, S., Murray, H.G., & Cummings, A.L. (2007). Gender differences in cooperative learning with university students. *Alberta Journal of Educational Research*, *53*(2), 157-173.

Salam, A., Hossain, A., & Rahman, S. (2015). The effect of using Teams Games Tournaments (TGT) cooperative technique for learning mathematics in secondary schools of Bangladesh. *REDIMAT*, *4*(3), 271-287. https://doi.org/10.17583/redimat.2015.1519

Schmenk, B. (2004). Language learning: A feminine domain? The role of stereotyping in constructing gendered learner identities. *TESOL Quarterly*, *38*(3), 514-524. https://doi.org/10.2307/3588352

Serbessa, D.D. (2006). Tension between traditional and modern teaching-learning approaches in Ethiopian primary schools. *Journal of International Cooperation in Education*, *9*(1), 123-140.

Sharan, S. (1980). Cooperative learning in small groups: recent methods and effects on achievement, attitudes, and ethnic relations. *Review of Educational Research*, *50*(2), 241 - 271. https://doi.org/10.2307/1170146

Siciliano, J. I. (2001). How to incorporate cooperative learning principles in the classroom: It's more than just putting students in teams. *Journal of Management Education*, *25*(1), 8-20. https://doi.org/10.1177/105256290102500103

Silitonga L.M., Wu TT. (2019). Increasing students' interest and learning achievement using cooperative learning (Students Team Achievement Division) through Edmodo. In L. Rønningsbakk, T. Wu, F. Sandnes & Y. Huang (Eds.), *Innovative technologies andlLearning*. *Lecture notes in Computer Science* (vol. 11937, pp. 3-13). Springer. https://doi.org/10.1007/978-3-030-35343-8_1

Slavin, R. E. (1991). *Student team learning*: *A practical guide to cooperative learning*. NEA.

Slavin, R. E. (1994). *Using student team learning (2nd ed.).* Johns Hopkins University.

Slavin, R. E. (1995). *Cooperative learning*: *Theory, research, and practice* (2nded.). Pearson College.

Slavin, R. E. (2014). Making cooperative learning powerful. *Educational Leadership*, *72*(2), 22-26.

Slavin, S. (1980). Cooperative learning. *Review of educational research,50*(2), 315-342. https://doi.org/10.2307/1170149

Slavin, R. E. (1987). Developmental and motivational perspectives on cooperative learning: A reconciliation. *Child Development*, *58*(5), 1161-1167. https://doi.org/10.1111/j.1467-8624.1987.tb01449.x

Slavin, R. E. (2013). Effective programmes in reading and mathematics: Lessons from the Best Evidence Encyclopaedia1. *School Effectiveness and School Improvement*, *24*(4), 383-391. https://doi.org/10.1080/09243453.2013.797913

Sunderland J. (2010). Theorizing gender perspectives in foreign and second language learning. In R. M. J. Catalán. (Ed.),*Gender perspectives on vocabulary in foreign and second languages* (pp. 1-19). Palgrave Macmillan. https://doi.org/10.1057/9780230274938_1

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement. *Studies in Second Language Acquisition*, *41*(2), 363-387. https://doi.org/10.1017/s0272263118000311

Tercanlioglu, L. (2004). Achievement goal theory: A perspective on foreign-language-learners' motivation. *TESL*

Canada Journal, *21*(2),34-49. https://doi.org/10.18806/tesl.v21i2.173

Tirussew, T., Amare, A., Jeilu, O. Tassew, W., Aklilu, D., & Berhannu, A. (2018). Ethiopian education development roadmap (2018-30): An integrated executive summary: Draft for discussion. https://planipolis.iiep.unesco.org/en/2018/ethiopian-education-development-roadmap-2018-30-integrated-executive-summary-draft-discussion

Tong, F., Irby, B., Lara-Alecio, R., Yoon, M., & Mathes, P. (2010). Hispanic English learners' responses to longitudinal English instructional intervention and the effect of gender: A multilevel analysis. *The Elementary School Journal*, *110*(4), 542-566. https://doi.org/10.1086/651195

Tremblay, P. F., & Gardner, R. C. (1995). Expanding the motivation construct in language learning. *The Modern Language Journal*, *79*(4), 505-518. https://doi.org/10.2307/330002

van der Slik, F.W.P., van Hout R.W.N.M., & Schepens , J.J. (2015). The gender gap in second language acquisition: Gender differences in the acquisition of Dutch among immigrants from 88 countries with 49 mother tongues. *PLoS ONE*, *10*(11), 1-22. https://doi.org/10.1371/journal.pone.0142056

Vygotsky, L. S. (1978). *Mind in society*. Harvard University Press.

Johnson, W. D., & Johnson, T. R. (2018). Cooperative learning: The foundation for active learning. In M.B. Silvo (Ed.), Book *active learning*. IntechOpen. https://doi.org/10.5772/intechopen.81086

Wang, M. (2012). Effects of cooperative learning on achievement motivation of female university students. *Asian Social Science*, *8*(15), 108-114. https://doi.org/10.5539/ass.v8n15p108

Yager, S., Johnson, D. W., & Johnson, R. T. (1985). Oral discussion, group-to-individual transfer, and achievement in cooperative learning groups. *Journal of Educational Psychology*, *77*(1), 60-66. https://doi.org/10.1037//0022-0663.77.1.60

Yalew, E. (2004). Teachers' beliefs, knowledge, and practice of learner-centered approach in schools of Ethiopia. *The Ethiopian Journal of Education*, *24*, 17-41. http://ejol.aau.edu.et/index.php/EJE/article/view/97

Yayo, B. (2013). Teachers' knowledge, attitude and practice of cooperative learning the case of Alamata Town administration second cycle primary schools [Unpublished Master's Thesis]. Addis Ababa University.

Zarei, A. A. (2012). The effects of STAD and CIRC on L2 reading comprehension and vocabulary learning. *Frontiers of Language and Teaching*, *3*, 161-173.

Zareian, G. and Jodaei, H. (2015). Motivation in second language acquisition: A State of the art article. *International J. Soc. Sci. & Education*,*5*(2), 295-308

Zhang, Y. (2010). Cooperative language learning and foreign language learning and teaching. *Journal of Language Teaching and Research*, 1(1), 81-83. https://doi.org/10.4304/jltr.1.1.81-83

Appendix A

Point-biserial and p-values of the English Language Test

Reading (1-8)								
Items	It1	It2	It3	It4	It5	It6	It7	It8
Point-biserial	.47	.25	.52	.25	.44	.68	.51	.57
p-values	.62	.59	.38	.35	.35	.29	.21	.35

Grammar (9-34)																
Items	**It9**	**It10**	**It11**	**It12**	**It13**	**It14**	**It15**	**It16**	**It17**	**It18**	**It19**	**It20**	**It21**	**It22**	**It23**	**It24**
Point-biserial	.43	.45	.36	.34	.47	.29	.37	.49	.28	.58	.48	.26	.36	.18	.54	.37
p-values	.59	.65	.74	.62	.65	.35	.35	.18	.62	.47	.35	.29	.24	.41	.53	.35
Items	It25	It26	It27	It28	It29	It30	It31	It32	It33	It34						
Point-biserial	.46	.45	.63	.57	.36	.32	.24	.67	.28	.57						
p-values	.44	.26	.45	.45	.58	.81	.71	.29	.74	.35						

	Dialogue (35-40)					**Punctuation (41-43)**			
Items	It35	It36	It37	It38	It39	It40	It41	It42	It43
Point-biserial	.41	.39	.36	.63	.27	.49	.49	.50	.47
p-values	.52	.68	.52	.42	.61	.52	.58	.65	.45

Appendix B

Summary of the Tournament Procedure

Week	Student	Table	Team Number	Ability Level	Questions Asked	Number of cards		Points earned
						Won	lost	
1	1	1	5	High	2	3	0	6
	2		7	High	2	1	1	2
	3	2	2	High	2	2	0	4
	4		4	High	2	3	0	6
	5		10	High	2	2	0	4
	6	3	3	High	2	1	1	2
	7		6	High	2	2	0	4
	8		9	High	2	2	0	4
	9		1	High	2	3	0	6
	10		8	High	2	1	1	2
2	11	4	5	Low	2	1	1	2
	12		8	Low	2	2	0	4
	13	5	1	Low	2	3	0	6
	14		10	Low	2	4	0	6
	15		6	Low	2	1	1	4
	16	6	3	Low	2	1	1	4
	17		4	Low	2	0	2	2
	18		7	Low	2	1	1	2
	19		2	Low	2	2	0	4
	20		9	Low	2	4	0	6
3	21	7	7	Medium	2	1	1	4
	22		3	Medium	2	0	2	2
	23	8	1	Medium	2	3	0	6
	24		9	Medium	2	2	0	4
	25		4	Medium	2	1	1	2
	26	9	8	Medium	2	3	0	6
	27		6	Medium	2	0	2	2
	28		10	Medium	2	2	0	4
	29		2	Medium	2	1	1	2
	30		5	Medium	2	4	0	6
4	31	10	2	Medium	2	3	0	6
	32		7	Medium	2	1	1	2
	33	11	4	Medium	2	2	0	4
	34		1	Medium	2	2	1	4
	35		9	Medium	2	0	2	2
	36	12	5	Medium	2	3	0	6
	37		6	Medium	2	2	0	4
	38		3	Medium	2	1	1	2
	39		10	Medium	2	4	0	6
	40		8	Medium	2	1	1	2

Appendix C

Rank of Practice Group According to the Total Scores Earned by Individual Group Members

Practice Team Group Number	Total Points Earned
1	22
5	20
10	20
9	16
2	16
6	14
8	14
4	14
3	10
7	10

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 1, 2020

Delvand, S. A., & Heidar, D. M. (2020). Computerized Group Dynamic Assessment and Listening Comprehension Ability: Does Self-Efficacy Matter? *Journal of Language and Education, 6*(1), 157-172. https://doi.org/10.17323/jle.2020.9834

Computerized Group Dynamic Assessment and Listening Comprehension Ability: Does Self-Efficacy Matter?

Shahin Abassy Delvand, Davood Mashhadi Heidar
Islamic Azad University

Correspondence concerning this article should be addressed to Davood Mashhadi Heidar, Department of English, Tonekabon Branch, Islamic Azad University, Tonekabon, Iran.
E-mail: davoodm_tarbiatmodares@yahoo.com

The present study investigated the effect of group dynamic assessment (DA) through software on Iranian intermediate EFL learners' listening comprehension ability. The main question of the study was whether dynamic assessment via CoolSpeech software had any effect on the listening comprehension ability of learners with high and low self-efficacy. To find the answer, 80 Iranian intermediate learners were selected from among a population of 120, based on their scores on a placement test. A self-efficacy questionnaire was then used to assign selected participants into two experimental groups as low self-efficacious experimental group (n=20) and high self-efficacious experimental group (n=20), as well as two control groups, each containing 20 participants. Next, a pretest of listening comprehension ability was administered to all groups, and no significant difference between their mean scores was observed. After a period of two months, during which the experimental groups received treatment of dynamic assessment through CoolSpeech software and the control groups received a placebo, a posttest of listening comprehension was administered to all groups. The data analysis results revealed that the participants in high self-efficacious experimental group achieved significantly better scores than the other groups. However, in the second experimental group, no significant change was observed, and participants in the second experimental group did not significantly outperform the control group. It was concluded that the group dynamic assessment method via software could have a significant effect on the listening comprehension ability of EFL learners with high self-efficacy.

Keywords: computer-assisted language learning (CALL), group dynamic assessment, listening comprehension ability, self-efficacy, zone of proximal development (ZPD)

Introduction

Computers play an increasingly significant role in our business, recreational, and educational activities. It is quite clear that in our daily lives computers are here to stay. Schools across Iran are helping students become computer literate at earlier ages than ever before (Mohammadi & Mirdehghan, 2014). Computer programs may best serve students who need an additional challenge apart from the classroom. Computers may also be helpful to students who are behind their peers academically (Nutta, 2013). They can serve as an economical tutor in order to bring these students up to levels of mastery in line with other study skills. In fact, using computers in the classroom create an amusing atmosphere for the learners to interactively take part in classroom discussions with their peers and the teacher as well (Rahimi & Hosseini, 2011). Computer-based instruction and its most popular terminology, i.e., CALL, has had a long historical background since they were first introduced (Kulik, Bangert, & Williams, 1983). Much of the research on the effectiveness of CALL occurred over the past twenty years, highlighting the need for both teachers and learners to be computer literate. Early supporters of CALL programs (Gerard, 1967; Oliver, 1999; Warschauer & Healey, 1998) believed that these educational devices might pave the way for both teachers and learners to benefit from them in their language classroom, and create variety in the learning atmosphere, which probably leads to a more cooperative learning environment. Some of the expected benefits included self-paced instruction for students, which would result in faster learning, the availability of

richer and more sophisticated materials, expert system-based instruction, and dynamic assessment (Corbeil, 2007; Wang, 2016; Zhao, 2003). The benefits for teachers were found to be the ease of modifying instructional materials and better time management, which allowed them to allocate more time to assist individual learners who required additional contact (Apperson, Laws, & Scepansky, 2006; Coleman, 2005). The effectiveness of CALL lies in the fact that, at the very least, it seems to be more effective and productive than traditional methods of instruction (Alodail, 2014; Lau, Yen, Li, & Wah, 2014), paving the way for language learners to develop their language skills more efficiently; one of the most important of which is listening comprehension.

Listening comprehension is of great importance in second and foreign language research, which should be acquired at "the early stages of L2 learning" (Nation & Newton, 2009, p. 37). It is through listening comprehension that the learners' input can be shaped, facilitating the learning of other language skills as well (Vandergrift & Goh, 2012). As stated by Rubin (1995), listening refers to "an active process in which listeners select and interpret information which comes from auditory and visual clues in order to define what is going on and what the speakers are trying to express" (p. 7). It has also been considered "the least understood and most overlooked of the four skills (Listening, Speaking, Reading, and Writing)" (Nation & Newton, 2009, p. 37). It is worth mentioning that listening comprehension seems to be overlooked in foreign language context as several Iranian researchers such as Razmjoo and Riazi (2006), Hosseini (2007), and Jahangard (2007) argued that little attention is paid to aural/oral skills in the Iranian EFL curriculum. These skills are not appropriately instructed and evaluated (Razmjoo & Riazi, 2006). In fact, Razmjoo and Riazi argued that listening tasks have received insufficient time and emphasis. Thus, innovative methodologies such as CALL aligned with assessment can be taken into account when teaching listening comprehension.

Assessment and language teaching have been found to be inter-related as Poehner (2005, 2009) believes that dynamic assessment (DA) can pave the way for language teachers to create meaningful interactions by providing them with appropriate feedback, which does not stop the flow of interaction. As Poehner and Lantolf (2005) stated, DA "is a procedure for simultaneously assessing and promoting development that takes account of the individual's (or group's) zone of proximal development" (p. 240). They also stated that group DA takes place when students actively interact and try to encourage their peers to take part in the learning atmosphere, while the teacher is monitoring their interaction. Inspired from CALL, computerized DA can be an effective tool in teaching language skills. Even though the capabilities of computers do not excite everyone, their potential for innovative computerized assessment can be of interest for multimedia test designers, instructional designers, and courseware production developers (Drasgow & Olson-Buchanan, 1999).Because of the potential of computerized assessment, numerous scholars have recommended the use of computers for evaluation (Hambleton, 1996). In addition, learners' characteristics, such as their self-efficacy, can also influence their performance.

According to Bernhardt (1997), self-efficacy can be represented as the learners' perception of their own capabilities when carrying out a task. According to Pajares (2000), it can be formed when the learners begin to judge their competence. Ehrman (1996) described self-efficacy as "the degree to which the learners think they have the capacity to cope with the learning challenge" (cited in Arnold & Brown, 1999, p.16). According to Bernhardt (1997), higher levels of self-efficacy can pave the way for language learners to be more successful in their learning, since their learning behaviors are goal-oriented. Alternatively, less confident learners are those who possess lower levels of self-efficacy, thus they assume failure from the beginning (Bernhardt, 1997, cited in Rahimi & Abedini, 2009). Therefore, the degree of learners' self-efficacy can directly or indirectly influence the learners' language learning progress (Gorban Doordinejad & Afshar, 2014).

In recent years, computer-based and web-based materials have been broadly applied by EFL instructors to compensate for the deficiencies that exist in traditional listening classes. Digital technology can contribute to listening instruction, by providing learners and teachers with intelligible input and output, and by providing opportunities for the negotiation of meaning (Chapelle, 1999; cited in Puakpong, 2008). Using digital technology in listening instruction has significant constructive effects on EFL learners' listening skills (Bingham and Larson (2006); cited in Puakpong, 2008). Hence, computerized DA can be applied as an effectual methodology to help EFL learners develop their listening comprehension ability through CoolSpeech software.

Statement of the Problem

Educators are constantly seeking new and better ways to improve instruction (Kennedy, 2005). Discovering a method of instruction that meets or exceeds conventional methods of instruction would be an appropriate

justification for its implementation. Computer-assisted instruction may be the answer to problems that have attracted educators' attention for years.

Computers have been proved to be influential in the classroom, and studies (e.g., Coiro & Dobler, 2007) related to the use of computers vary in their findings. Some studies (e.g. Gersten, Fuchs, Williams, & Baker, 2001) show the benefits of computers as supplements over varying lengths of time with various populations of students with different levels of proficiency. Many dependent measures have also been used to measure the quality of the teaching techniques. Some studies (e.g., Kennedy, 2005) have taken into account statistical analyses to probe the effectiveness of computer programs in language learning. However, the abundance of research on CALL might not necessarily look into teaching listening comprehension. Although the literature might provide some evidence regarding the role of computerized instruction in vocabulary learning (e.g., Wang, 2016) or grammar (Corbeil, 2007; Yusof & Saadon, 2012), studying computer software in teaching listening comprehension seems not to be sufficiently examined, particularly in the context of Iran and being integrated with group DA.

Apart from computer software, the nature of traditional types of assessment, including paper-and-pencil tests, was time-consuming and not having concurrent assessment might be problematic for examiners as well as the examinees. In fact, the time of administration is not a concern when utilizing computerized assessment (Thiagarajan, 1999). Moreover, the scoring procedure is immediately conducted when computerized tests occur.

Last but not least, the inclusion of DA into computerized assessment might not have been well recognized in the literature, since each feature has been taken into account individually and the advantage of each has been presented separately. However, it seems that the provision of computerized dynamic assessment through instructional software can facilitate the teaching and learning of listening comprehension, allowing both teachers and learners to benefit from meaningful interactions while working on the listening tasks. Therefore, the present study takes into account the effect of computerized DA on Iranian EFL learners' development of listening comprehension while also considering their self-efficacy.

In sum, it seems that DA and its application in technology-assisted language learning environments have been of interest to language scholars. However, the learners' performance through exposure to CALL integrated with DA while considering the learners' high and low self-efficacy might not have been examined deeply, particularly in a foreign language context, such as Iran. Hence, the current study addressed the following research questions:

Research Questions

RQ1. Does dynamic assessment via CoolSpeech software have any statistically significant effect on the listening comprehension ability of learners with high and low self-efficacy?
RQ2. Is there any significant difference in the listening comprehension ability of learners with high and low self-efficacy?

Literature Review

Theoretical Background: DA and its Components

DA has been found to be an effective methodology for evaluating the learners' performance while they are interacting with their peers. Poehner (2005) argued that teachers play the role of mediators since they attempt to take control of the learners' communication by providing the most beneficial type of feedback to keep track of the learners' engagement. In this regard, the notion of the Zone of Proximal Development (ZPD) is highlighted. The ZPD is considered to be the distance between two levels of assistance; one in which the learners need to be provided with educational support or scaffolding, and the other in which they can manage to carry out tasks autonomously (Vygotsky, 1978). It is assumed that DA should be directed to the learners' ZPD in order to help language learners feel more comfortable in their learning environment and enjoy peer communication monitored by the teacher (Poehner, 2009).

As stated by Vygotsky (1998), while traditional assessment only measures fully matured abilities, dynamic assessment measures both fully matured abilities, and abilities that are still in the process of growing; therefore,

dynamic assessment can reveal much more about the process of acquiring that information. Traditional psychological assessments are descriptive, and do not explain developmental processes (Shabani et al., 2010). Vygotsky (1984) argued that by putting a learner's ZDP at the center of the assessment procedure, the teacher is able to monitor the learners' gradual development and examine their potential to initiate an interaction with the peers (Minick, 1987).

As Sternberg and Grigorenko (2002) argued, in the context of DA, teacher-learner relationship is flexible as the tester mediates throughout the evaluation process. In the DA context, an ambience of teaching and helping replaces the traditional neutrality attitude of the traditional assessment context (Shabani et al., 2010). As claimed by Vygotsky (1998), independent problem solving shows only a small part of the cognitive ability of learners, that is, the actual level of cognitive development. Vygotsky believed that non-dynamic evaluation only determines a small fraction of the overall image of development. He also argued that responding to assistance is a very central feature for understanding a learner's cognitive ability, since it can give instructors good insight into the future progress of the individual. In the DA approach, the learners' future performance can be envisaged through their development process (Yildirim, 2008).

Vygotsky (1998) argued that teaching and evaluation should not be separated from each other in DA research. Thus, the real focus ought to be on the learners' learning, which can be shaped when learners are involved with activities, fostering peer interactions and the teachers' mediations. It can be pointed out that with the help of mediators the learners' independence in doing the tasks can be achieved (Yildirim, 2008). Dynamic testing is essentially represented as a procedure that construes personal characteristics and their consequences for education, and takes into account the mediation within the evaluation procedure. In these procedures, the learning process is accentuated, and not the learning outcomes.

As stated by Sternberg and Grigorenko (2002), in traditional evaluation, questions are provided by an examiner, and are all expected to be answered by an examinee consecutively, without any kind of feedback or intervention, unlike DA which considers teaching and evaluation inseparable, and as two sides of the same coin. DA is fundamentally based on Vygotsky's revolutionary insights which believed that teaching leads to development, in the ZPD. Before Vygotsky, it was generally believed that the only reliable marker of mental function was independent problem solving, but Vygotsky opposed this idea and proposed that independent problem solving could merely reveal a fraction of the learners' cognitive ability (Yildirim, 2008). Vygotsky claimed that the importance of receptiveness to assistance is the same as the actual developmental level, and since it offers some insight into the future development of an individual, it is considered to be an essential feature for understanding cognitive abilities (Yildirim, 2008).

It is worth mentioning that there exist two models of DA, the interventionist and interactionist models. The former is concerned with the evaluation of the learners' performance before the intervention or the instructional program, after which the target instruction takes place. This is followed by the second assessment in order to evaluate the effectiveness of the intervention (Poehner, 2009; Poehner & Lantolf, 2005). As the name suggests, the interventionist approach does not take into account the how of the learning; and only the learners' performance at the beginning and at the end of the instruction is of importance. However, the learners' process of learning and the quality of their participation are taken into account in the interactionist approach. Poehner (2009) argued that the interactionist model justifies the need for ZPD-directed feedback in which a knowledgeable peer provides dynamic assessment in order to assist the learner independently solve a task. It is here that the teacher tries to mediate the learners' interactions with their peers in order to help them benefit from each interaction and improve their autonomy (Poehner & Lantolf, 2005). Hence, group DA can be accommodated within the notion of the interactionist model since the learners' interactions are mediated by the teacher, who tries to facilitate the learners' participation in a problem-solving task (Poehner, 2005). Group DA can be integrated with other instructional methods, such as CALL, in which learners have more opportunities to interact in a more attractive learning environment.

Previous Related Studies

Researchers have been interested in applying DA to improve the learners' development of language skills. Hill and Sabet (2009) conducted a research study on four possible dynamic speaking assessment approaches in a classroom setting. They included the mediated assistance, transfer of learning, ZPD, and collaborative

engagement. Mediated assistance was done in the form of teacher-learner interaction in order to check the learners' speaking ability. The learner's ZPD was provided for a problem-solving group of students whose concentration was directed at the social and cultural dimensions of ZPD, called group-ZPD, since comparisons revealed that DA provided for individuals was not as significant as the group DA, in line with cooperative activity. The last DA approach was collaborative engagement that was done for the purpose of identifying common speaking problems as well as difficulties in assessing the learners' speaking performance. The study involved four speaking assessments for freshman Japanese university students. Their study revealed that explicit and implicit prompts were found to be effective in conditions where learners might face difficulties in comprehending or sometimes doing the target tasks. However, the nature of the prompts and the activities that stimulated the prompts remained unclear since the researchers recommended more research in this area. The study did not provide additional consideration for manipulating classroom activities and the participation of adult university students. The students' linguistic level for pairing students was not thoroughly justified by the researchers. Moreover, the study was limited to pairing as the only grouping technique and this left the reader to feel doubtful concerning the suitability of the grouping approach for dynamic speaking assessment.

Similarly, in the context of foreign language, in a study conducted by Shabani, Alavi, and Kaivanpanah (2012), it was aimed to probe whether identification of mediation strategies can be carried out through group DA as an instructional treatment, which is guided by a mediator of the learners' interactions in the listening classroom. Furthermore, the research attempted to highlight the impact of G-DA on knowledge construction among L2 learners. The researchers formulated a list of mediational tactics, embracing various forms of implicit and explicit feedback. They also revealed how beneficial the G-DA interaction could be in establishing a community of practice in which learners could greatly benefit from their classmates and instructors through their cooperative scaffolding.

As far as listening comprehension can be affected by DA, Alavi and Taheri (2014) also looked into the impacts of applying DA on the learners' ability to boost their listening comprehension. Findings demonstrated the learners' significant improvement of listening comprehension when DA was implemented in the classroom. Findings contributed to the teachers' awareness of applying DA an instructional approach to foster more peer interactions, which can pave the way for language learners to improve their listening comprehension through communication.

Not only language skills, but also the learners' pragmatic awareness can be affected by group DA. Hashemi Shahraki, Ketabi, and Barati (2015) investigated whether group DA could be effective in evaluating EFL learners' pragmatic awareness of conversational implicatures, while detecting the mediational tactics that could contribute to improving this knowledge. The results showed that through enhancing EFL learners' pragmatic comprehension of conversational implicatures, G-DA could dramatically improve the learners' listening comprehension ability. Their results supported G-DA and its usefulness for L2 listening comprehension and pragmatic instruction.

Implementing DA into a technology-mediated learning environment appears to have rarely been considered. In a recent study, Mashhadi Heidar and Afghari (2015) examined the learners' listening comprehension ability when they were exposed to dynamic assessment in Synchronous Computer-Mediated Communication via Talking and Writing technologies of Web 2.0 and Skype. In this study, the socio-cognitive progression of EFL learners was studied through DA that sought Vygotsky's willingness for instructional scaffolding in ZPD. The results showed that DA in synchronous computer-mediated communication through interaction in the ZPD enabled the educators to explore both the actual and the potential level of the learners' listening ability.

By integrating DA into computer software, Mashhadi Heidar (2016) studied the role of DA in enhancing the listening skill of EFL students via Web 2.0. His findings proved that online DA via Web 2.0 improves listening comprehension ability in Iranian EFL learners. Findings also revealed that when DA is applied in an online learning platform, it can be more flexibly ZPD-directed, which can help language learners to feel more independent in doing the related tasks.

Finally, Ashraf, Motallebzadeh, and Ghazizadeh (2016) attempted to identify whether electronic-based DA affected the listening comprehension ability of L2 learners. The experimental group received the treatment in which online application of DA was practiced in order to establish a rather different learning atmosphere for

EFL learners. The results revealed that electronic-based DA significantly improved the listening skill of the EFL learners. In fact, technology appears to be a helpful solution to benefit from DA more efficiently.

In the aforementioned studies, the impact of self-efficacy on EFL learners' achievement in conventional methods and the effects of computerized and non-computerized DA on L2 learners' listening skill were investigated, but none of the previous studies investigated the impact of EFL learners' self-efficacy on computerized group dynamic assessment. This study managed to build on previous investigations on the said variables regarding the effects of self-efficacy on computerized G-DA, and intended to focus on dimensions that could not be dealt with previously.

Method

Design

The present study adopted a quasi-experimental design in which the homogenous participation of language learners was fulfilled. Then, they were divided into four groups, including two experimental and two control groups. The study employed a pretest, treatment, and posttest design.

Participants

The study was conducted with 80 Iranian EFL learners at the intermediate level from Kish Language Institute. The participants were all adult female learners attending the institute, because the researcher aimed at keeping age and gender as fixed variables. Teenage participants needed different treatment approaches. The same problem existed for middle-aged participants. That was why the age range of the participants was kept limited. In order to make sure of homogeneity, subjects were chosen from among 120 students, based on their scores on a PET (Preliminary English Test). Having calculated the mean and the SD, the participants with the score of one SD above and below the mean were selected to take part in the study. Eighty students with an intermediate proficiency level were selected from among 120 potential participants using the PET placement test. Afterwards, the selected participants were given a Self-Efficacy Questionnaire (MSEQ), and low and high self-efficacious participants were defined. Then, 20 high self-efficacious participants were assigned to the first experimental group, 20 low self-efficacious participants were assigned to the second experimental group, and 20 participants were assigned to each control group.

Instruments and Materials

CoolSpeech Software (version 5.0)

CoolSpeech software, produced by ByteCool Software Inc. (2001), is a text-to-speech software that is compatible with Microsoft Speech API to get and read aloud texts from various sources, for the purpose of empowering learners to listen to online news from any URL, converting texts into spoken wave files (.Wav), reading text files and HTML files aloud, listening to new messages from email accounts, listening to any sentences that they have typed anywhere in the Microsoft Windows, listening to texts that they have copied to the Microsoft Windows clipboard, and scheduling files, URLs, and emails to be read aloud. Such capabilities were used in treating Iranian EFL learners on their English listening comprehension skill.[1]

Proficiency Test

To ensure the homogeneity of the participants, a PET was administered. It was used in order to select intermediate language learners. The score obtained from the test showed the level of proficiency so that learners who passed the exam with scores higher than 65 were considered to be suitable for intermediate level. It is necessary to mention that between those subjects who passed the exam the ones who could obtain the score one SD above and below mean were selected for the study.

[1] "Purchase CoolSpeech 5.0", (2019, April 29). Retrieved from http://www.bytecool.com/company.htm. Retrieved 2019-04-29.

Listening Comprehension Test

A listening comprehension test was designed and developed by the researcher. This test was designed for an intermediate level of proficiency. In order to make sure that the listening texts in the tests were of the right level, they were selected from different listening tasks of the textbook that the participants studied at Kish Language Institute. The reliability of the test was examined in a pilot study. In this phase, the researcher designed the listening test and gave it to 15 participants who were representative of the participants of the study. Afterward, the data collected from the piloting were analyzed through Cronbach's alpha. The test reliability was estimated at .71. As this test benefited from an acceptable level of reliability, it was utilized as both pretest and posttest. When two different tests are designed, the risk of different levels of difficulty in different versions of the test goes up; to avoid this problem, one test was used. As the treatment took almost two months, the same test could be used as the pretest and posttest, and there was almost no chance of remembering the questions.

Table 1
Reliability of the Listening and Reading Sections

	N of Items	Cronbach's Alpha
Listening	25	.709
Reading	35	.771

As Table 1 shows, the reliability of the listening and reading sections of the PET of the study was .709 and .771, which is acceptable.

In order to check the writing and speaking sections, Pearson correlation was administered to check the inter-rater reliability.

Table 2
Pearson correlation result for writing and speaking sections inter-rater reliability

Groups		Writing Rater 1	Writing Rater 2
Writing Rater 1	Pearson Correlation	1	.923
	Sig. (2-tailed)		.000
Writing Rater 2	N	120	120
	Pearson Correlation	.923	1
	Sig. (2-tailed)	.000	
	N	120	120
Groups		Speaking Rater 1	Speaking Rater 2
Speaking Rater 1	Pearson Correlation	1	.949
	Sig. (2-tailed)		.000
Speaking Rater 2	N	120	120
	Pearson Correlation	.949	1
	Sig. (2-tailed)	.000	
	N	120	120

The result of Pearson correlation in Table 2 revealed that the reliability level for writing and speaking was at .923 and .949, which was very high.

Self-Efficacy Questionnaire

In the present study, the MSEQ (Memory Self-Efficacy Questionnaire) was used. The MSEQ is a self-rating questionnaire that evaluates individuals' self-efficacy in two phases: first, Self-efficacy Level, which evaluates individuals' memory ability; second, Self-efficacy Strength, which evaluates individuals' confidence (Berry, West, & Dennehey, 1989). As the MSEQ enjoys high reliability and validity, it can be utilized for studies on memory self-efficacy tests (Berry et al., 1989).The questionnaire was translated into Persian to be more comprehensible to the students.

Data Collection Procedure

To collect the data of the current study, first, PET was administered to the participants. This test included 20 multiple-choice items in the form of paper-and-pencil. The learners were asked to answer the questions in 15 minutes on the provided answer sheet. The pretest and the posttest (a retest of the pretest) of the study included four paragraphs with five multiple-choice questions for each. The paragraphs were played back to the participants three times, and they were asked to answer the questions on the provided answer sheet.

In both experimental groups, the participants received the same form of treatment, based on their level of proficiency (B1). The participants in both experimental groups in this study were taught using the group dynamic assessment method through CoolSpeech software. Each session lasted twenty minutes. In these groups, the teacher selected level-appropriate listening items from the listening tasks students listened to in the class. The learners were asked to work in pairs or groups of three. They used CoolSpeech software to listen to some extracts, and then they were asked to talk about it with their peers. The teacher monitored them and helped them correct each other.

In group dynamic assessment, the teacher's role is more like a mediator who facilitates learning. In the current research, the teacher applied the following mediational strategies: first, confirming learners' correct responses that, they were not sure about; second, replaying the tape, either the total passage, or some parts of it. Whenever it was necessary, the teacher allowed learners to listen to the passage again. Third, the teacher helped learners put the words together. When learners could not comprehend an utterance after replaying it several times, the teacher tried to divide the utterance into smaller and more comprehensible parts. Fourth, whenever students guessed a sentence erroneously, the mediator repeated it with a questioning tone. This helped learners find their mistakes and correct themselves. Fifth, the mediator provided students with contextual clues. The contextual clues entailed learners' world knowledge, topical knowledge and situational awareness. Sixth, using metalinguistic clues, the teacher tried to scaffold learners. These metalinguistic clues were grammatical or lexical. Seventh, whenever learners did not know a word, and they could not guess it, they were allowed to use a dictionary. Eighth, the teacher explained the correct response when other mediational strategies did not work well.

After two months of treatment on the experimental groups, which took 20 sessions, twenty minutes each session, the participants in all groups took part in a listening posttest, and the results were compared and contrasted to check the hypotheses of the study.

Results

The statistical indexes analyzed were the mean and standard deviation of the pretest and posttest of students' listening comprehension in the groups. Having calculated the reliability of the values of the test, and established that the values were acceptable, the researcher analyzed the data through SPSS software (Version 22.0). The two research questions of the study are taken into account below.

RQ1: Does dynamic assessment via CoolSpeech software have any statistically significant effect on the listening comprehension ability of learners with high and low self-efficacy?

The first research question of the study examined the effect of dynamic assessment via CoolSpeech software on the listening comprehension ability of learners with high and low self-efficacy. In doing so, quantitative analysis

of the pretest and posttest scores was done. Initially, a normal distribution of data had to be checked. A Shapiro-Wilk test of normality revealed that the *p*-values of the pretests and posttests of the experimental and control groups are more than .05, indicating a normal distribution of data (pretest p-values: EH=.114, EL=.105, CH=.091, CL=.085; posttest p-values: EH=.118, EL=.112, CH=.091, CL=.095). Table 3 shows the descriptive statistics for the high self-efficacious learners' listening comprehension.

Table 3
Descriptive Statistics for the Pretest and Posttest of Learners with High Self-Efficacy in the Experimental Group

	N	Mean	Std. Deviation	Std. Error Mean
Pretest	20	21.04	3.05	.472
Posttest	20	22.86	3.05	.472

Table 3 reveals a little improvement in the learners' listening comprehension from the pretest (M=21.04, SD=3.05) to the posttest (M=22.86, SD=3.05), which demonstrates that dynamic assessment via CoolSpeech software could help learners gain mastery over listening comprehension tasks.

Also, paired samples t-test for the high self-efficacy experimental group showed that the level of significance was less than .05 (p= .000), highlighting the significant improvement in the listening comprehension ability of the learners with high self-efficacy. Therefore, it can be deduced that DA via CoolSpeech software was found to be effective in paving the way for high self-efficacious learners to develop their listening comprehension.

Table 4
Descriptive Statistics for the Pretest and Posttest of Learners with High Self-Efficacy in the Control Group

	N	Mean	Std. Deviation	Std. Error Mean
Pretest	20	20.54	3.23	.514
Posttest	20	20.92	3.39	.534

Table 4 indicates no noticeable improvement in the learner's listening comprehension ability from the pretest (M= 20.54, SD=3.23) to the posttest (M=20.92, SD=3.39) as a result of exposure to teaching of listening comprehension through conventional means.

Further, paired samples t-test between the pretest and posttest of high self-efficacy control group indicated that the significance level was more than .05 (p= .488), which meant that no improvement was made in the listening comprehension ability of the high self-efficacious learners in the control group (mean difference= -.38, SD= 3.16, std. error mean= .481, t= 11.23, df= 19). In other words, the control group, which was taught using conventional instructional means, could not significantly improve their listening comprehension ability.

The second part of research question looked into the effect of applying DA via CoolSpeech software on the listening comprehension ability of learners with low self-efficacy. Descriptive statistics are presented in Table 5.

Table 5
Descriptive Statistics for the Pretest and Posttest of Learners with Low Self-Efficacy in the Experimental Group

	N	Mean	Std. Deviation	Std. Error Mean
Pretest	20	21.08	2.51	.472
Posttest	20	22.02	2.40	.472

Table 5 shows a very small increase from the pretest (M=21.08, SD=2.51) to the posttest (M=22.02, SD=2.40) of the learners' listening comprehension, which reveals that low self-efficacious learners' listening comprehension was not greatly affected by dynamic assessment via CoolSpeech software. In order to inferentially take into account the learners' performance on the pretest and posttest of listening comprehension, paired samples t-test was conducted.

Paired samples t-test for low self-efficacy experimental group indicated that the significance level was more than .05 (p= .112), from which it can be inferred that no significant difference can be observed between the

low self-efficacious learners' listening comprehension on the pretest and posttest (mean difference=-.94, SD= 2.91, std error mean= .422, t=13.46, df= 19). Therefore, the findings demonstrated that dynamic assessment via CoolSpeech software resulted in no significant improvement in the listening comprehension of the learners with low self-efficacy.

Table 6
Descriptive Statistics for the Pretest and Posttest of Learners with Low Self-Efficacy in the Control Group

	N	Mean	Std. Deviation	Std. Error Mean
Pretest	20	20.85	3.16	.521
Posttest	20	21.72	3.33	.519

Table 6 shows that only a very small increase took place in the control group's pretest (M= 20.85, SD=3.16) of listening comprehension and their posttest (M= 21.72, SD=3.33). In order to inferentially take into account the learners' performance on the pretest and posttest of listening comprehension, paired samples t-test was conducted.

Paired samples t-test for the low self-efficacy control group revealed that no significant difference could be observed concerning the low-efficacious learners' listening comprehension ability in the control group since the significance level was more than.05 (p= 310). In fact, the learners in the control group did not significantly benefit from the conventional instruction for listening comprehension (mean difference= -.87, SD= 2.78, Std Error Mean= .493, t=10.98, DF= 19).

RQ2: Is there any significant difference in the listening comprehension ability of learners with high and low self-efficacy?

The second aim of the study was to explore the difference between the high and low self-efficacious learners' listening comprehension ability affected by dynamic assessment via CoolSpeech software. To do so, descriptive as well as inferential analyses were conducted. Descriptive statistics for the experimental and control groups are presented in Table 7.

Table 7
Descriptive Statistics for the Pretest of Learners with High and Low Self-Efficacy in the Experimental and Control Groups

	N	Mean	Std. Deviation	Std. Error
EH	20	21.04	3.05	.472
EL	20	21.08	2.51	.472
CH	20	20.54	3.23	.514
CL	20	20.85	3.16	.521

Table 7 reveals that the four groups acted similarly in the beginning, from which it can be inferred that there exists a small difference among the four groups' mean scores of the listening comprehension pretest. As for the comparison of the four groups' mean scores of the pretest, two-way analysis of variance (two-way ANOVA) was run.

Table 8
Two-Way ANOVA Statistics for the Pretest of Learners with High and Low Self-Efficacy in the Experimental and Control Groups

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	142.431	3	71.215	2.986	.205
Within Groups	1241.120	76	23.315		
Total	1383.551	79			

Table 8 shows no significant difference among the mean scores of the four groups' listening comprehension (F3, 76= 2.98, *p*= .205), which reveals the similarity in the intermediate learners' listening comprehension before the treatment sessions. The students were all at the intermediate level (B1).Table 13below provides descriptive data for the posttest scores of the learners' listening comprehension in the experimental and control groups.

Table 9
Descriptive Statistics for the Posttest of Learners with High and Low Self-Efficacy in the Experimental and Control Groups

	N	Mean	Std. Deviation	Std. Error
EH	20	22.86	3.05	.472
EL	20	22.02	2.40	.472
CH	20	20.92	3.39	.534
CL	20	21.72	3.33	.519

Table 9 indicates that there were differences in the four groups' listening comprehension ability (EH, M=22.86, SD=3.05; EL, M=22.02, SD=2.40; CH, M=20.92, SD=3.39; CL, M=21.72, SD=3.33). In order to compare the mean scores of the learners' listening comprehension in the posttest, a two-way ANOVA was run.

Table 10
Two-Way ANOVA Statistics for the Posttest of Learners with High and Low Self-Efficacy in the Experimental and Control Groups

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	421.198	3	210.599	6.811	.001
Within Groups	1318.413	76	26.619		
Total	1739.611	79			

The two-way ANOVA results between the posttest scores of the high and low self-efficacy in the experimental and control groups in Table 14 showed a significant difference among the posttest of four groups (F3, 76= 6.81, *p*= .001< .05). Thus, it can be inferred that the four groups were different in their listening comprehension ability.

To highlight the difference among the groups, a post-hoc Tukey HSD test was used (See Appendix). The results of Tukey HSD showed that a significant difference could be observed between the two experimental groups (*p*= .014<.05) as high self-efficacious learners outperformed the low self-efficacious ones in terms of their listening comprehension ability. Similarly, the learners with high self-efficacy in the experimental group performed better than those of the control group (*p*= .003<.05). However, the learners with low self-efficacy in the experimental group did not significantly outperform the control group (*p*= .112>.05). Hence, it can be concluded that high self-efficacious learners were significantly affected by dynamic assessment via CoolSpeech software concerning their listening comprehension ability. Moreover, they improved more than the learners with low-self-efficacy in their listening comprehension. Finally, the listening comprehension of the learners with low self-efficacy was not significantly affected by dynamic assessment using CoolSpeech software.

Discussion

The present study aimed to find out the potential effects of the group DA through CoolSpeech software on the listening comprehension ability of EFL learners with high and low self-efficacy. In this regard, four groups (two experimental and two control groups) formed the sample of the study. The findings revealed that high self-efficacious learners improved more than the other groups in terms of their listening comprehension ability. Regarding the first research question, which investigated the effect of dynamic assessment via CoolSpeech software on the listening comprehension ability of learners with high and low self-efficacy, it was found that DA via CoolSpeech software could help high self-efficacious learners to gain mastery over listening comprehension tasks. However, it was found that low self-efficacious learners' listening comprehension was not greatly

affected by such assessment. Considering the second research question, which explored the difference between the high and low self-efficacious learners' listening comprehension ability affected by dynamic assessment via CoolSpeech software, the results of the two-way ANOVA revealed differences in the four groups' listening comprehension ability. There was a significant difference between the two experimental groups as high self-efficacious learners outperformed the low self-efficacious ones regarding their listening comprehension ability. Likewise, the learners with high self-efficacy in the experimental group performed better than those of the control group. However, learners with low self-efficacy in the experimental group did not significantly outperform the control group.

The results of the current investigation can be justified on the basis that group DA method through CoolSpeech software could influence listening comprehension of the learners with high self-efficacy, but it could not significantly affect the listening comprehension of the students with lower self-efficacy. Self-efficacy, as Bernhardt (1997) suggest, is a set of various self-beliefs related to diverse areas of performance. These beliefs have both behavioral and emotional aspects. They influence the decision about whether to engage in a certain task, the effort and power an individual exerts in completing the task, and the degree of avoidance and persistence in performing it (Pajares, 2000). Thus, for inefficacious learners, tasks are perceived to be more difficult than they actually are which ultimately leads to a decrease in persistence and effort (Ehrman, 1996; Gorban Doordinejad & Afshar, 2014), thus, the less self-efficacious students demonstrated lower perceived ability, fewer amounts of invested mental effort, and lower performance as highlighted by Pajares (2000).

The findings of the current research is in harmony with Gorban Doordinejad and Afshar (2014) that revealed a significant relationship between foreign language learners' self-efficacy and their English achievement. They suggested that learners, who benefited higher foreign language self-efficacy, were more likely to attain higher English scores. This study is also congruent with Rahimi and Abedini (2009); who revealed there was a significant relationship between EFL learners' self-efficacy and their achievements in listening skill.

The results of the study are also in alignment with those of other DA studies which revealed the positive effects of DA and group DA on learners' listening comprehension ability. The present study is in harmony with Shabani, Alavi, and Kaivanpanah (2012); Alavi and Taheri (2014); and Hashemi Shahraki, Ketabi, and Barati (2015) who proved that group DA could significantly improve students' listening comprehension ability, compared to non-dynamic assessment methods. Furthermore, this study confirms the studies conducted by Mashhadi Heidar and Afghari (2015) and Mashhadi Heidar (2016) who revealed that online DA via Web 2.0 significantly improved listening comprehension ability in Iranian EFL learners. Finally, the present study is in accordance with Ashraf, Motallebzadeh, and Ghazizadeh (2016) who proved that electronic-based DA method could significantly improve the listening skill of the EFL learners.

Conclusion

By quantitatively gathering and scrutinizing the data, it was concluded that the G-DA method of teaching via software could have a significant effect on the listening comprehension skill of learners with high self-efficacy, although this result was not observed in learners with lower levels of self-efficacy. The pedagogical significance of this study is multifaceted and can be examined both at micro and macro levels. Regarding the usefulness of the results at the macro level, it can be said that more areas of inquiry were identified to help curriculum designers reauthorize the remarkable changes in learning environments and the influence on learning, teaching, and testing pedagogy. The findings of this research may assist policy makers in emphasizing the significance of the use of different approaches to skills evaluation. Moreover, it seems that students, teachers, and researchers can also benefit from the outcomes of the present study.

Learners are the first beneficiary of the study findings. Many learners appear to be worried about their listening ability in the process of language learning and are usually concerned with their listening skill as well as their grades on listening exams. By being assessed through a dynamic method of assessment, learners can overcome listening difficulties since they are consciously involved in the assessment procedure. In fact, when learners are aware of their listening skill, they can take necessary action to solve possible deficiencies in listening as well as strengthen their listening ability. Teachers, who are always concerned with teaching language skills,

can benefit from various types of assessment approaches toward listening. Different types of DA through CALL can be applied as tasks and helpful techniques that can be used in English classrooms. Based on the results of this study, both teachers and learners can apply the best assessment tools in the classroom to address any possible difficulties they may face during the listening class. Finally, the researcher hopes that this study will have far-reaching conclusions that can be practical and helpful for the researchers who are interested in DA approaches because it provides them with current literature on the topic. The researcher believes that this study can contribute to putting English language courses and objectives more in line with modern approaches to language assessment, particularly for the Iranian context in which only traditional methods of instructions are made use of and the real-world needs of the students are almost completely neglected.

The findings of the present investigation should be generalized with care as the sample and context are not representative of the whole population of Iranian English learners in different settings. Thus, further research can be carried out to explore other variables, such as different learning environments, participants' gender, different levels of proficiency, and personality types.

References

Alavi, S. M., & Taheri, P. (2014). Examining the role of dynamic assessment in the development and assessment of listening comprehension. *English Language Teaching (ELT), 1*(2), 23-40.

Alodail, A. (2014). Impact of technology (power point) on students' learning. *International Interdisciplinary Journal of Education, 3*(4), 200-206.

Apperson, J. M., Laws, E. L., & Scepansky, J. A. (2006). An assessment of student performances for power point presentation structure in undergraduate courses. *Computers and Education, 47*(1), 1-6. https://doi.org/10.1016/j.compedu.2006.04.003

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

Ashraf, H., Motallebzadeh, K., & Ghazizadeh, F. (2016). The impact of electronic-based dynamic assessment on the listening skill of Iranian EFL learners. *International Journal of Language Testing, 6*(1), 24-32. https://doi.org/10.14744/alrj.2018.36844

Bernhardt, S. (1997). Self-efficacy and second language learning. *The NCLRC Language Resource, 1*(5), 1-13.

Berry, J. M., West, R. L., & Dennehey, D. M. (1989). Reliability and validity of the Memory Self-Efficacy Questionnaire. *Developmental Psychology, 25*(5), 701-713. https://doi.org/10.1037/0012-1649.25.5.701

Bingham, S., & Larson, E. (2006). Using CALL as the major element of study for a university English class in Japan. *The JALT Journal, 2*(3), 39-52.

Brookhart, S. M., & DeVoge, J. G. (1999). Testing a theory about the role of classroom assessment in student motivation and achievement. *Applied Measurement in Education*, 12(4), 409-425. https://doi.org/10.1207/S15324818AME1204_5

Chapelle, C. A. (1999). Technology and language teaching for the 21st century. *Paper presented at the Proceedings of the Eighth International Symposium on English Teaching*. The Crane.

Coiro, J., & Dobler, E. (2007). Exploring the online reading comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet. *Reading Research Quarterly*, *42(2),* 214-257. https://doi.org/10.1598/RRQ.42.2.2

Coleman, J. A. (2005). CALL from the margins: Effective dissemination of CALL research and good practices. *ReCALL: The Journal of EUROCALL, 17*(1), 18-31. https://doi.org/10.1017/S0958344005000315

Corbeil, G. (2007). Can power point presentations effectively replace textbooks and blackboards for teaching grammar? Do students find them an effective learning tool? *CALICO Journal, 24*(3), 631-656.

Doordinejad, F.G., & Afshar, H. (2014). On the relationship between self-efficacy and English achievement among Iranian third grade high school students. *International Journal of Language Learning and Applied Linguistics World, 6*(4), 461-470.

Drasgow, F., & Olson-Buchanan, J.B. (1999). *Innovations in computerized assessment.* Lawrence Erlbaum Associates.

Ehrman, M. E. (1996). An exploration of adult language learning motivation, self-efficacy and anxiety. In R. L. Oxford (Ed.), *Language learning motivation: pathways to the new century* (pp. 23-46). The University of Hawaii Second Language Teaching and Curriculum Centre.

Gerard, R. W. (Ed.). (1967). *Computers and education.* McGraw Hill.
Gersten, R., Fuchs, L. S., Williams, J. P., & Baker, S. (2001). Teaching reading comprehension strategies to students with learning disabilities: A review of research. *Review of Educational Research*, 71,279–320. https://doi.org/10.3102/00346543071002279
Gorban Doordinejad, F., & Afshar, H. (2014). On the relationship between self-efficacy and English achievement among Iranian third grade high school students. *International Journal of Language Learning and Applied Linguistics World,* 6(4), 461-470.
Hambleton, R. K. (1996). Advances in assessment models, methods, and practices. In D. C. Berliner, & R.C. Calfee (Eds.), *Handbook of educational psychology* (pp. 889-925). American Council on Education/Macmillan.
Hashemi Shahraki, S., Ketabi, S., & Barati, H. (2015). Group dynamic assessment of EFL listening comprehension: Conversational implicatures in focus. *International Journal of Research Studies in Language Learning, 4*(3), 73-89. https://doi.org/10.5861/ijrsll.2014.955
Hill, K., & Sabet, M. (2009). Dynamic speaking assessments. *TESOL Quarterly, 43*, 537-545. https://www.jstor.org/stable/27785036
Hosseini, S. M. H. (2007). ELT in higher education in Iran and India: A critical view. *Language in India, 7*, 1-11. http://www.languageinindia.com/dec2007/eltinindiaandiran.pdf
Jahangard, A. (2007). Evaluation of EFL materials taught at Iranian public high schools. *The Asian EFL Journal, 9*(2), 130-150. https://www.asian-efl-journal.com/main-journals/evaluation-of-efl-materials-taught-at-iranian-public-high-schools/
Kennedy, M. M. (2005). Inside teaching: How classroom life under-mines reform. Harvard University Press.
Koura, A. A., & Al-Hebaishi, S. M. (2014). The Relationship between multiple intelligences, self-efficacy and academic achievement of Saudi gifted and regular intermediate students. *Educational Research International, 3*(1), 48-70.
Kulik, J. A., Bangert, R. L., & Williams, G. W. (1983). Effects of computer-based teaching on secondary school students. *Journal of Educational Psychology, 75*(1), 19–26. https://doi.org/10.1037/0022-0663.75.1.19
Lau, R. W. H., Yen, N. Y., Li, F., & Wah, B. (2014). Recent development in multimedia e-learning technologies. *World Wide Web*,17, 189-198. https://doi.org/10.1007/s11280-013-0206-8
Lave J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge University Press.
Lidz, C. S., & Gindis, B. (2003). Dynamic assessment of the evolving cognitive functions in children. In A. Kozulin, B. Gindis, V.S.Ageyev & S. M. Miller (Eds.), *Vygotsky's educational theory in cultural context* (pp. 99-116). Cambridge University Press.
Mashhadi Heidar, D., & Afghari, A. (2015). The effect of dynamic assessment in synchronous computer-mediated communication on Iranian EFL learners' listening comprehension ability at upper-intermediate level. *English Language Teaching, 8*(4), 14-23. https://doi.org/10.5539/elt.v8n4p14
Mashhadi Heidar, D. (2016). ZPD-assisted introduction via web 2.0 and listening comprehension ability. *English for Specific Purposes World, 49*(17).
Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: Self-efficacy, anxiety, and their relation to reading and listening proficiency. *Foreign Language Annals, 39*(2), 276-294.
Minick, N. (1987). The development of Vygotsky's thought: An introduction. In R.W. Reiber & A. S. Carton (Eds.), *The collected works of L. S. Vygotsky: Problems of general psychology* (pp. 17-36). Plenum Press.
Mohammadi, E., & Mirdehghan, S. S. (2014). A CMC approach to teaching phrasal-verbs to Iranian EFL senior high school students: The case of blended learning. *Procedia - Social and Behavioral Sciences, 98*, 1174–1178. https://doi.org/10.1016/j.sbspro.2014.03.531
Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL speaking and listening.* Routledge.
Nutta, J. (2013). Is computer-based grammar instruction as effective as teacher directed grammar instruction for teaching L2 structures? *CALICO Journal*, *16*(1), 49-62.
Oliver, R. (1999). Exploring strategies for online teaching and learning. *Distance Education, 20*(2), 240-254.
Pajares, F. (2000). *Self-efficacy beliefs and current directions in self-efficacy research*. http://www.emory.edu/EDUCATION/mfp/effpage.html
Poehner, M. E. (2005). *Dynamic assessment of oral proficiency among advanced L2 learners of French* [Unpublished doctoral thesis]. The Pennsylvania State University.
Poehner, M. E. (2008). *Dynamic assessment. A Vygotskian approach to understanding and promoting L2 development.* Springer International Publishing.
Poehner, M. E. (2009). Dynamic assessment as a dialectical framework for classroom activity: Evidence from second language (l2) learners. *Journal of Cognitive Education and Psychology, 8*(3), 252-268. https://doi.org/10.1891/1945-8959.8.3.252

Poehner, M. E., & Lantolf, J. P. (2005). Dynamic assessment in the language classroom. *Language Teaching Research, 9*, 233-265. https://doi.org/10.1191/1362168805lr166oa

Puakpong, N. (2008). An evaluation of a listening comprehension program. In F. Zhang & B. Barber (Eds.), *Handbook of research on computer-enhanced Language acquisition and learning*. IGI Global.

Rahimi, A., & Abedini, A. (2009). The interface between EFL learners' self-efficacy concerning listening comprehension and listening proficiency. *Novitas-Royal, 3*(1), 14-28.

Rahimi, M., & Hosseini, F. K. (2011). The impact of computer-based activities on Iranian high-school students' attitudes towards computer-assisted language learning. *Procedia Computer Science, 3*, 183-190. https://doi.org/10.1016/j.procs.2010.12.031

Razmjoo, S. A., & Riazi, M. (2006). Is communicative language teaching practical in the expanding circle? A case study of teachers of Shiraz high schools and institutes. *Journal of Language and Learning, 4*(2), 144-171.

Reeves, T. C. (2000). Alternative assessment approaches for online learning environments in higher education. *Journal of Educational Computing Research, 23*(1), 101–111. https://doi.org/10.2190/GYMQ-78FA-WMTX-J06C

Rubin, J. (1995). An overview to a guide for the teaching of second language listening. In D. J. Mendelsohn, & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp. 7-11). Dominie Press, Inc.

Shabani, K., Alavi, S. M., & Kaivanpanah, S. (2012). Group dynamic assessment: An inventory of mediational strategies for teaching listening. *The Journal of Teaching Language Skills (JTLS), 30*(4), 27-58.

Shabani, K., Khatib, M., & Ebadi, S. (2010). Vygotsky's Zone of Proximal Development instructional implications and teachers' professional development. *English Language Teaching Journal, 3*(4), 237-248. https://doi.org/10.5539/elt.v3n4p237

Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing. The nature and measurement of learning potential.* Cambridge University Press.

Thiagarajan, S. (1999). Formative evaluation in performance technology. *Performance Improvement Quarterly, 4*(2), 22-24

Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in Action*. Taylor & Francis.

Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Vygotsky, L. S. 1984a. *Sobranie Sochinenii, Tom chetvertyi: Detskaya psikhilogii*[collected works (vol. 4: Child psychology). Izdatel'stvo Pedagogika.

Vygotsky, L. S. (1998). Infancy. In R. W. Rieber (Ed.), *The collected works of L.S. Vygotsky* (pp. 207-241). Plenum Press.

Wang, Y. H. (2016). Promoting contextual vocabulary learning through an adaptive computer-assisted EFL reading system. *Journal of Computer Assisted Learning, 32,* 291-303. https://doi.org/10.1111/jcal.12132

Warschauer, M., & Healey, D. (1998). Computers and language learning: An overview. *Language Teaching, 31*, 57-71.

Warschauer, M., & Healey, D. (1998). *Computers and Language Learning*: *An overview* Published online by Cambridge University Press. https://doi.org/10.1017/S0261444800012970

Yildirim, A. G. O. (2008). Vygotsky's sociocultural theory and dynamic assessment in language learning. *Anadolu University Journal of Social Sciences, 8*(1), 301-308.

Yusof, N. A., & Saadon, N. (2012). The Effects of Web-based Language Learning on University Students' Grammar Proficiency. *Procedia-Social and Behavioral Sciences*, 67, 402-408.

Zhao, Y. (2003). Recent developments in technology and language learning: A literature review and meta-analysis. *CALICO Journal, 21*(1), 7-27. https://www.jstor.org/stable/24149478

Appendix

Table15
Multiple Comparisons: Tukey HSD Test for the Posttest of Learners with High and Low Self-Efficacy in the Experimental and Control Groups

(I) Groups	(J) Groups	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
EH	EL	.84*	.693	.014	-4.7543	1.9258
	CH	1.94*	1.213	.003	.5831	6.9870
	CL	1.14*	1.214	.011	-.0011	7.0709
EL	EH	-.84*	.693	.014	-1.9258	4.7543
	CH	1.10*	1.201	.012	2.0565	8.3410
	CL	.30	1.309	.112	1.4681	8.4319
CH	EH	-1.94*	1.213	.003	-6.9870	-.5831
	EL	-1.10*	1.201	.012	-8.3410	-2.0565
	CL	-.80	1.253	.051	-3.5980	3.0971
CL	EH	-1.14*	1.214	.011	-7.0709	.0011
	EL	.30	1.309	.112	-8.4319	-1.4681
	CH	.80	1.263	.051	-3.0971	3.5980

*. The mean difference is significant at the 0.05 level.

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 1, 2020

Roohani, A., & Iravani, M. (2020). The Relationship Between Burnout and Self-Efficacy among Iranian Male and Female EFL Teachers. *Journal of Language and Education, 6*(1), 173-188. https://doi.org/10.17323/jle.2020.9793

The Relationship Between Burnout and Self-Efficacy among Iranian Male and Female EFL Teachers

Ali Roohani, Mehdi Iravani
Shahrekord University

Correspondence concerning this article should be addressed to Ali Roohani, English Department, Faculty of Letters & Humanities, Shahrekord University, the 2nd Kilometer of Saman Road, Shahrekord, Iran. Zip Code: 88186/34141, P.O. Box: 115. E-mail: roohani.ali@gmail.com

Teaching is a job marked by high levels of burnout. Teacher burnout has been extinsively studied in connection with other important psychological qualities, such as perceived self-efficacy. However, little research has examined this relationship among teachers in the English as a foreign language (EFL) context. In this light, this mixed-method study was intended to a) investigate the relationship between the degree of perceived self-efficacy by Iranian EFL teachers and their professional burnout level, and b) see whether gender could make a significant difference in the teachers' burnout level. To these ends, 80 male and female Iranian EFL teachers from several high schools in Isfahan, selected through convenience sampling, participated in the study and responded to the Maslach Burnout Inventory-Educator's Survey (MBI-ES) and a modified version of Teacher Sense of Efficacy Scale (TSES). To triangulate the data, a semi-structured interview was also conducted with 15 teachers. Correlational and *t*-test data analysis showed that there was a strong significant correlation between the participants' perceived self-efficacy and their burnout level in a negative direction. The gender variable also had a modifying effect on the teachers' burnout. Female teachers, in fact, had a lower burnout level. The follow-up interview further confirmed the relationship and revealed the three main themes of mental fatigue, contact avoidance, and stress in explaining teacher burnout. The implications for school administrators and teacher educators are discussed.

Keywords: professional burnout, perceived self-efficacy, EFL teachers

Introduction

Professional or job burnout is caused by long-term interpersonal stress in the working environment (Maslach, 1982). The condition develops when there is no match between occupational requirements and an employee's perceived capabilities (Brown, 2012; Maslach, Schaufeli, & Leiter, 2001; Smith & Bourke, 1992). According to Maslach (2015), burnout is "a psychological syndrome that develops in response to chronic emotional and interpersonal job stressors" (p. 1415). This phenomenon is so common among human service providers that it is often considered an index of mental and physical health (Jennett, Harris, & Mesibov, 2003).

All too often, burnout can be traced back to externally driven factors such as long working hours or unpleasant working conditions (Lim, Kim, Kim, Yang, & Lee, 2010), reduced levels of infrastructure and social support (Kay-Eccles, 2012), increased levels of accountability (Dorman, 2003), and reduced management skills (Blase, 1984). Yet, it is equally important to consider individual and self-regulatory qualities that might yield important clues as to the possible genesis of the phenomenon. According to Alarcon, Eschelman, and Bowling (2009), perceived self-efficacy, i.e. beliefs in one's abilities, is one relevant factor. In this light, self-efficacy beliefs are held to shield employers against the adverse consequences of occupational stress (Brown, 2012).

Viewed from this perspective, the relationship between professional burnout and perceived self-efficacy finds great significance in the educational arena, particularly in relation to language teachers. An investigation into this relationship is particularly important for English as a foreign language (EFL) teachers who play a major role in today's globally interconnected world and amid the great demand for English language education in this world where English is the global lingua franca in many fields. As Jedynak (2011) states, the profession of a foreign language teacher can be over-stressed. EFL teachers constantly need to meet new challenges, obstacles,

and difficulties in their teaching lives (Reeve & Su, 2014). As Hong (2012) points out, state schools can become a tense site for EFL teachers with students who often lack enough motivation to learn English. They can gradually discover a conflict between their job expectations and the demands they actually experience; hence, they fail to meet their responsibilities properly. Thus, it is important to identify the culture-specific variables that can anticipate EFL teachers' burnout in different settings, such as public high schools where the pressure to satisfy professional requirements seems to be greater than private language schools (Rostami, Ghanizadeh, & Ghapanchi, 2015).

Perceived self-efficacy, as an individual protective factor, might hold the key to the prediction, prevention, and understanding of professional burnout among language teachers. Previous research, using Bandura's (1994) concept of self-efficacy, has indictaed that teachers with a higher level of self-efficacy often see classroom demands and problems more as challenges than hindrances (Brouwers & Tomic, 2000; Yang & Farn, 2005). However, an examination of the related literature suggests that the relationship between self-efficacy and burnout has not received enough attention within the EFL context, particularly in public schools where EFL teachers face new challenges. The paucity of research in this area and the important role of EFL teachers in today's world where English is the dominant language in many areas could warrant further in-depth investigation into the possible relationship between EFL teachers' self-efficacy and their burnout. Investigating such affective and personality factors, as possible determinants of burnout, can contribute to preventing teacher burnout.

The antecedents of teacher burnout are commonly thought to have multifaceted factors (Chan, 2009). One group of burnout studies (e.g., Friedman & Farber, 1992; Greenglass & Burke, 1988) identify such antecedents as individual factors, which include demographic variables such as age, gender, and years of teaching experience. Among these variables, gender has been observed to be a likely factor both in the rate and the frequency of burnout symptoms among teachers (Klassen & Chiu, 2010). However, other studies on gender and burnout (e.g., Lau, Yuen, & Chan, 2005; Sari, 2004) did not focus on EFL teachers in public schools. Rather, they were concerned with the teachers whose majors were not teaching English in secondary or high schools. Inconsistencies and gaps can also be seen, depending on the occupation and culture-specific context of the studies (Purvanova & Muros, 2010). It is thus assumed that gender-related studies on burnout among EFL teachers can enrich the related literature and shed some light on professional foreign language teachers' quality and expectations. In this light, this study was intended to investigate the possible relationship between the degree of perceived self-efficacy by Iranian EFL teachers and their level of professional burnout, and further examine whether gender could make a difference in the perceived components of burnout among Iranian EFL teachers.

Teacher Burnout

The term burnout found its way into the occupational literature in the early 1970s. Introduced by Freudenberger (1974), burnout was described as "a state of mental and physical exhaustion caused by one's professional life" (p. 160). In fact, burnout was considered a condition in which one is struck by constantly feeling stressed and, consequently, worn out due to unfulfilling working conditions.

Later on, however, Maslach (1982) offered a more scientific definition of the phenomenon, known as the state definition as the result of extensive research with her colleague, Mrs. Susan Jackson, on individuals working in

the human services sector. She posited a three-component approach to the concept of burnout, that is, *emotional exhaustion*, *depersonalization*, and *reduced personal accomplishment*. Accordingly, burnout was characterized as "a syndrome of emotional exhaustion, depersonalization, and reduced personal accomplishment that can occur among individuals who work with people in some capacity" (Maslach, et al., 1996, p. 4).

Emotional exhaustion is observed when a human services worker is bereft of any energy and motivation, represented by a sense of weariness (Maslach & Jackson, 1986). Depersonalization (or cynicism) is the adoption of a cynic, indifferent, and impersonal attitude towards the service recipients (i.e., clients, patients, or students). The last component is reduced personal accomplishment. It represents the negative self-image one harbors about one's achievements, job performance, and ability to be successful at work or in life (Maslach, Schaufeli, & Leiter, 2001).

The literature points extensively to the widespread existence of burnout among social and human services professionals (Skaalvik & Skaalvik, 2010). Among these, schoolteachers are particularly more vulnerable than

those working in other non-educational occupations, perhaps due to the effects of external social forces and internal critical agencies (Day & Gu, 2010; Hanushek, 2007; McCarthy, Lambert, O'Donnell, & Melendres, 2009; Van Tonder & Williams, 2009). According to Hanushek (2007) and McCarthy et al. (2009), burnout among teachers is related to tiredness stemming from demands in the educational context (emotional exhaustion), a pessimistic viewpoint and lack of interest towards students (cynicism), and feelings of incompetency as a teacher (reduced personal accomplishment). Additionally, Van Tonder and Williams (2009) state that one of the main reasons for focusing on teachers in burnout studies is that burnout is more common in pedagogical settings than any other professional context. This is supported by Farber's (2000) estimation. He claims that 5 to 20 percent of American teachers suffer from burnout.

Teacher Self-Efficacy

The concept of self-efficacy is mainly attributed to Bandura's (1986) social cognitive theory. According to the theory, individuals are capable of human agency, that is, they intentionally pursue future courses of action. This assumes that what individuals believe about their own abilities will determine future behavior. The factors determining future behavior are behaviors, environmental influences, and internal individual factors such as affective, cognitive, and biological processes (Bandura, 1997). Bandura (1994) states that self-efficacy is "people's beliefs about their capabilities to produce designated levels of performance that exercise influence over events that affect their lives" (p. 71). When humans believe they can do something successfully, they enjoy self-efficacy (Woolfolk, 2004). The concept of perceived self-efficacy has also been described as "beliefs in one's capabilities to organize and execute the courses of action required producing given attainment" (Bandura, 1997, p. 3). It accounts for energy, effort, and persistence one is willing to invest in a task.

Self-efficacy is usually interpreted as a domain-specific concept. Teacher self-efficacy, often viewed as a future-oriented construct, is one such domain. It refers to "the teacher's belief in his or her capability to organize and execute courses of action required to successfully accomplishing a specific teaching task in a particular context" (Tschannen-Moran, Woolfolk Hoy, & Hoy, 1998, p. 233). It is teachers' perceived competence in performing designated functions to achieve a set of educational goals, leading to students' learning facilitation and achievement (Skaalvik & Skaalvik, 2010). The concept of self-efficacy, as applied in the area of education, has also inspired a number of researchers (e.g., Tschannen-Moran & Woolfolk Hoy, 1998) to investigate how teachers' self-efficacy beliefs could influence their future behaviours and outcomes. According to social cognitive theory, advocated by Bandura (1994), teachers who do not expect success with their students are likely to invest less effort in the instruction, and are bound to give up more easily at the first hurdle. Therefore, the beliefs teachers have about their own pedagogic capabilities and knowledge, professional practice, and classroom management strategies could potentially impact their teaching productivity. In a similar vein, Collie, Shapka, and Perry (2012) emphasize that highly self-efficacious teachers firmly believe in their personal capacity to have a major impact on the educational processes, including students' learning. They are able to successfully meet school challenges, manifesting themselves mainly in issues such as students' disruptive behaviors, learning problems, communication with parents and colleagues, and school administration.

The Self-Efficacy-Burnout Theoretical Link

Burnout could have an adverse impact on professionals, particularly teachers. However, it is not judicious to claim that many teachers will suffer from this condition, and it is important to discover the factors that could cushion its negative effects. Social cognitive theory postulates that self-efficacy could be an aid to recovery from job strain (Unsworth & Mason, 2012) and a facilitator of adaptation to change in the organization (Bandura, 1994, 1995, 1997). Efficaciousness, or self-efficacy beliefs, consciously or unconsciously, can affect one's choices, efforts, persistence, and emotions in the face of adversity. Extending this to educational contexts, teachers who enjoy high self-efficacy are more confident in their capabilities to get the desired results. According to Bandura (1997), efficaciousness, consciously or unconsciously, can affect one's choices, efforts, persistence, and emotions in the face of adversity. In educational contexts, self-efficacy beliefs are believed to act between knowledge and behavior. Self-efficacy beliefs influence behavior through four mediating processes, that is, cognitive, affective, motivational, and selective processes. These mediating processes can greatly affect the performance of teachers in schools and negatively affect thought patterns, self-regulation of affective states, the selection of activities, and regulations of motivation in teachers who harbor self-doubts and inefficacy (Evers, Brouwers, & Tomic, 2002), which can be a mediator of burnout (Bandura, 1997).

More recently, as Brown (2012) states, foreign/second language (L2) researchers have begun to show interest in teacher self-efficacy and review its relationship with other affective variables, including those related to the dimensions of burnout. The common thread among their arguments is that perceived self-efficacy beliefs may be negatively correlated with feelings of being overextended and losing an emotional or cognitive involvement with work. However, it can be positively related to feelings of competence and productivity at work (Hong, 2012).

In a systematic review, Brown (2012) claims an association between self-efficacy and burnout among teachers, with the strongest association between self-efficacy with depersonalization and emotional exhaustion. Brouwers and Tomic (2000) have also reported that self-efficacy has a longitudinal effect on depersonalization and a synchronous impact on personal accomplishment, whereas the direction can be reverse for emotional exhaustion. In a similar vein, Aloe, Amo, and Shanahan (2014) have suggested that higher levels of self-efficacy in classroom management make teachers less vulnerable to feelings of burnout. Betoret (2009) also concludes that job stressors, stemming from low instructional self-efficacy, can be a major factor in burnout.

Chwalisz, Altmaier, and Russell (1992) related self-efficacy beliefs to the realm of dealing with the most stressful situations that teachers experienced at their job. In a correlational study, they examined the relationships among causal attributions, self-efficacy cognitions, and coping through data collected from 316 public school teachers in the Netherlands. They reported that teacher burnout could be moderately linked to teacher self-efficacy. They also found that teachers who were unsure about their ability to effectively cope with stressful events showed a higher level of burnout than their colleagues who demonstrated more self-confidence in this area. Evers, Brouwers, and Tomic (2002) also studied burnout among 490 teachers who were employed in the upper grades of schools of higher general secondary education in the Netherlands. Similarly, the results indicated that teachers' self-efficacy beliefs were associated with their burnout level in a moderate way. However, Skaalvik and Skaalvik (2007) reported a strong relationship between teacher self-efficacy and teacher burnout among elementary Norwegian teachers. It seems there is no clear consensus regarding the level of the relationship between teacher self-efficacy and teacher burnout in different contexts.

Despite the potential theoretical link between the teacher burnout and self-efficacy, more investigations on the issue, especially in different educational systems, are in order. As Perrewé et al. (2002) state, these concepts could go well beyond cultural and national boundaries to (in)validate the theoretical foundations and previous findings. Perhaps, the different characteristics of educational systems can result in different conceptualizations and findings. Moreover, no systematic studies have explored the association between teacher burnout and self-efficacy among EFL teachers in the context of Iranian public high schools both in a quantitative and qualitative manner. The existing literature has provided some insight for the association between teacher burnout and self-efficacy. However, the prior studies were conducted in other contexts such as in the Turkish context (e.g., Ozkara, 2019) or in private language institutes (e.g., Motallebzadeh, Ashraf, & Tabatabaee-Yazdi, 2014). Additionally, an investigation into the association between teacher burnout and self-efficacy with a mixed-methods approach can be more conducive for learning about culture-specific factors involved in burnout experience among EFL teachers in the context of Iranian public schools.

Moreover, evidence suggests that the demographic variable of gender is a likely candidate in burnout research (Biron & Link, 2014). However, controversies surround the role played by gender in the qualitative and quantitative development of the burnout phenomenon (Schwarzer & Hallum, 2008). Some research studies point to higher degrees of burnout among females (e.g., Maslach, 1982; Poulin & Walter, 1993). For instance, Sari (2004) reported higher levels of emotional exhaustion and personal accomplishment among female teachers in Turkish high schools. In contrast, Küçüksüleymanoğlu (2011) concluded that male teachers in special education schools in Turkey showed more symptoms of emotional exhaustion and depersonalization, with no significant difference for personal accomplishment. In a cross-cultural research study, Schwarzer and Hallum (2008) found no major gender effect for burnout components except for depersonalization for male Syrian and German teachers. In the EFL context, Motallebzadeh et al. (2014) and Mardani, Baghelani, and Azizi (2015) came up with different mixed results on the role of gender in teacher burnout. Additionally, Mede (2009) found gender to be an insignificant factor in the burnout experience. Given the inconclusiveness surrounding the role played by gender, and considering the fact that females constitute a large portion of EFL teachers in Iran, particularly in public high schools, understanding gender-related differences among teachers in the EFL context of Iran, where demands for expanding new knowledge and sharing it with the world outside are rapidly increasing, it is necessary to investigate the possible role of gender in the burnout experience.

The study, therefore, sought answers to the following questions:

1. Is there any significant relationship between perceived self-efficacy and professional burnout among Iranian EFL teachers?
2. Is there any significant difference between male and female Iranian EFL teachers in professional burnout?
3. Does the qualitative data help explain the results from the initial quantitative phase of the study?

Accordingly, the following null hypotheses were formulated to address the two first research questions:

H_{01}: There is no statistically significant relationship between perceived self-efficacy and professional burnout among Iranian EFL teachers.

H_{02}: There is no statistically significant difference between male and female Iranian EFL teachers regarding professional burnout.

Method

Participants

The participants were 80 EFL teachers from Isfahan, Iran. They were all native speakers of Persian. The participants consisted of 44 males and 36 females, working full time in nine public high schools. Their teaching experience ranged from 12 to 25 years. The selection of the research site was driven by the assumption that teachers working in secondary education, particularly in public schools, are vulnerable to burnout risks (Otero López et al., 2008). The selection was based on convenience sampling and the sample size was considered satisfactory for the purpose of the study. The participants' age ranged from 35 to 56. The teachers had majored in Teaching of English as a Foreign Language (TEFL, n = 36), English Translation (n = 20), English Literature (n = 14), and Linguistics (n = 10). They had bachelor's (n = 49) and master's (n = 31) degrees.

Assessments and Measures

Two questionnaires were used to collect the data on the professional burnout and perceived self-efficacy from the participants. The first was the Maslach Burnout Inventory-Educator's Survey (MBI-ES) (Maslach & Jackson, 1986). The inventory includes 22 items with a seven-point scale ranging from 0 (*never*) to 6 (*every day*) with higher scores indicating a higher frequency of occurrence of feelings described for each item. The inventory captures the three dimensions of burnout, namely, emotional exhaustion (EE) (e.g., "I feel emotionally drained from my work"), depersonalization (DP) (e.g., "I don't really care what happens to some students"), and (reduced) efficacy or personal accomplishment (PA) (e.g., "I have accomplished many worthwhile things in this job"). The participants were allotted 25 minutes to complete the questionnaire. Several researchers (e.g., Akbari, Ghafar Samar, Kiani, & Eghtesadi, 2011; Byrne, 1993; Schwarzer, Schmitz, & Tang, 2000) have established the validity of the MBI-ES. Early on, Maslach and Jackson (1986) demonstrated that the three dimensions had good psychometric properties. They confirmed the convergent validity of the questionnaire by correlating MBI-ES scores with: 1) indexes of various outcomes postulated to be associated with burnout, such as job dissatisfaction; 2) job characteristics expected to contribute to burnout, such as heavy workloads; and 3) behavioral ratings by those knowing the individuals (e.g., colleagues). Kokkinos (2007) also confirmed the three factor structures of burnout employing exploratory and confirmatory factor analysis. Regarding the internal consistency of the scale, Cronbach's alpha coefficients for all the subscales were above .70 in the current study, with EE (α = .78), DP (α = .74), and PA (α = .83).

The second questionnaire was a modified version of Teacher Sense of Efficacy Scale (TSES, Tschannen-Moran & Woolfolk Hoy, 2001), developed and validated by Eslami and Fatahi (2008) in the context of Iran. Recently, using the Rasch model, Chang and Engelhard (2016) reconfirmed the validity of the original questionnaire with data from 554 teachers in a US Midwestern state. The inventory includes 12 items with a five-point Likert scale, ranging from 1 (*not at all*) to 5 (*a great deal*). It has three subscales: (a) *efficacy for instructional strategies* (IS, α = .81 in the current study) (e.g., "To what extent can you provide an explanation or example when students are confused?"), (b) *efficacy for classroom management* (CM, α = .83 in the current study) (e.g., "How much can you do to control disruptive behavior in the classroom?"), and (c) *efficacy for student engagement* (SE, α = .80 in the current study) (e.g., "How much can you do to motivate students who show low interest in school work?"). The participants were allotted 15 minutes to complete the questionnaire. Meanwhile, the participants had to

provide their demographic information, such as gender, years of teaching experience, highest academic degree, and academic major at the beginning of the two questionnaires employed in the study.

A semi-structured interview was carried out as a follow-up step to the quantitative phase for the qualitative part of the study. Several questions were carefully prepared to help contextualize and explain the quantitative data. The questions were reviewed by five experienced principals working in other school districts to provide feedback regarding their clarity, appropriateness, and ability to elicit relevant information. The questions included yes/no and open-ended questions (see Appendix). Moreover, if necessary, more probing questions were raised.

Procedure

To gather the data, the two questionnaires were distributed among the participants who were asked to fill them out in their schools. It took about one weak to have all the copies collected. The data were analyzed for descriptive and inferential statistics using SPSS.

Additionally, in order to invest more confidence in the validity of the scores and to get a more comprehensive understanding of the possible relationship between the participants' level of self-efficacy and burnout, 15 teachers sat to be interviewed. They consisted of volunteer teachers who were selected from the sample in the quantitative phase. They included the teachers with both high and low levels of self-efficacy. They comprised eight teachers with high scores on the TSES questionnaire and seven teachers with low scores on the TSES questionnaire. The interviews were conducted face-to-face in their schools, each lasting about 25 minutes. The respondents were assured that their personal information would be confidential. All the interviews were digitally recorded and transcribed. Afterwards, following Mackey and Gass (2005), the transcripts were read several times to identify salient features and organize the raw data into categories for further analysis. They were coded and themes were identified by the present researchers, who moved back and forth between the gathered data and codes to identify the emerging themes. To ensure the credibility of the interview data, peer-debriefing and member-checking were utilized. For the member-checking, the participants were invited to review the transcripts and emerging themes to assess the accuracy of the interpretations. The peer-debriefing involved an external check by a PhD student, an expert in qualitative data analysis, who was provided with the data and our interpretations to make sure the interpretation of data was credible.

Results

Quantitative Phase

Table 1 reports the descriptive statistics of the three dimensions of the MBI-ES, namely, emotional exhaustion, depersonalization, and personal accomplishment.

Table 1
Descriptive Statistics of the MBI-ES Dimensions

Dimension	N of Items	Min	Max	Total Mean	SD	Scale Mean	SSkewness	Kurtosis
EE	9	14	35	25.98	6.15	2.89	-.15	-1.25
DP	5	8	19	13.13	2.92	2.63	-.11	1.02
PA	8	11	31	20.68	7.49	2.59	.26	-1.76

Note. EE = Emotional Exhaustion; DP = Depersonalization; PA = Personal Accomplishment

According to guidelines suggested by Maslach, Jackson, and Leiter (1996), scores on the three dimensions of MBI-ES can be low, moderate, or high if the mean scores fall in the lower, middle, and upper third categories, respectively. Following Maslach et al.'s (1999) guidelines, the mean score for EE was considered moderate to high as it fell in the middle third range of score category (17-26). The mean score for DP fell in the upper third ($\geqslant 13$), which was high. The mean score for PA also fell in the upper third ($\leqslant 31$), indicating a low sense of PA and high level of reduced PA. In sum, the EFL teachers reported a relatively high level of burnout, in general.

Table 2 reports the descriptive statistics of three dimensions of the TSES, namely, student engagement, classroom management, and instructional strategies.

Table 2
Descriptive Statistics of the TSES Dimensions

Dimesnsion	N of Items	Min	Max	Total Mean	SD	Scale mean	Skewness	Kurtosis
SE	4	5	20	11.91	3.79	2.98	.167	-.94
CM	4	4	20	11.96	3.62	2.99	.166	-.66
IS	4	5	20	13.18	3.61	3.29	.026	-.08
Total	12	15	58	37.05	10.54	3.09	.025	-.76

Note. SE = Student Engagement; CM = Classroom Management; IS = Instructional Strategies

As observed in Table 2, the total self-efficacy mean was 37.05. Given that the possible range of scores for self-efficacy was 12-60 with 12 Likert-type items each with five options, the scale mean was 3.09, which was obtained through dividing the mean score by the number of items. This number was close to the third option (*Some Degree*) on 5-point Likert items. This result does not reflect a very high level of perceived self-efficacy among the teachers. This is especially the case for the subcomponents SE (2.98) and CM (2.99).

To examine the first null hypothesis stating that there is no statistically significant relationship between perceived self-efficacy and professional burnout among Iranian EFL teachers, a Pearson's product correlation was run between these two variables. The results are summarized in Table 3.

Table 3
Correlations between Teacher Self-Efficacy and Burnout Dimensions

	EE	DP	PA
Teacher Self-Efficacy	-.66** (r^2 = .44)	-.57** (r^2 = .32)	.79** (r^2 = .62)

Note. **p < .01

As displayed in Table 3, there is a statistically significant and strong negative correlation (p =.000) between EE and teacher self-efficacy, with higher levels of teacher self-efficacy associated with lower levels of EE. Additionally, following Larson-Hall's (2010) guidelines, there is a moderate to large negative correlation between DP and teacher self-efficacy, with higher levels of teacher self-efficacy significantly associated with lower levels of DP. As for PA, the correlation was also statistically significant with a large effect size, but in a positive direction (p =.000). As the teachers' self-efficacy increased, so did their PA.

Moreover, to address the second null hypothesis stating that there is no statistically significant difference between male and female Iranian EFL teachers regarding professional burnout, independent *t-test*s were run for each burnout dimension (see Table 4).

Table 4
Independent t-Tests for MBI-ES Dimension Scores for Both Males and Females

	Groups									
	Male			Female						
Dimensions	**M**	**SD**	**n**	**M**	**SD**	**n**	**t**	**df**	**Sig**	**Eta Squared**
EE	30.75	4.25	44	21.20	3.52	36	11.03	78	.000	.43
DP	15.35	1.87	44	10.90	1.91	36	10.57	78	.000	.41
PA	16.10	4.68	44	25.25	6.98	36	-6.89	78	.000	.30

As demonstrated in Table 4, there were statistically significant differences between the male and female teachers in all three burnout dimensions. The results indicate that males had significantly higher scores for EE (p =.000), with a medium effect size. The female teachers, however, received significantly higher scores for PA (p =.000), with a medium effect size.

Qualitative Phase

The present study employed a sequential mixed-method design to answer the third research question. The collection and analysis of data from the quantitative survey was followed by the semi-structured interviews, which helped as a secondary source to validate the quantitative results (Creswell, 2009). As with the quantitative part, the qualitative analysis focused on the teacher burnout and perceived self-efficacy variables. Coding and analysis of the interview data generated themes or concepts related to these variables, which are summarized in Table 6 and explained below.

Table 6
Emergent Themes from the Interview Data

Themes	Number of Teachers
Mental fatigue	8
Contact avoidance	6
Stress	7

Mental fatigue

The first theme that emerged from the interviewees' responses was mental fatigue. The low self-efficacious teacher participants felt more mentally fatigued. They were unable to recover their energy after a hard day's work at school. Almost all of them reported enduring fatigue. For example, one of the male EFL teachers with 20 years of teaching experience in several public schools commented, "at the end of the day, I feel so tired …, er, so motionless when I get up in the morning, facing a new day. My wife becomes worried about my health." As to the self-efficacy side of the story, being a low self-efficacious teacher, he also said:

> As I speak or … teach in the class, I'm repeatedly interrupted by noisy and mischievous students and, … er …well …, don't know how to restore order to the class. I will be really at a loss for words and will have no idea of how to react properly…I am tired.

However, the teachers with a higher level of self-efficacy, pointed to their energy when dealing with the students and their problems. One highly energetic female teacher said, "My energy and activity can keep me going the whole day…I like my job." Although some of the self-efficacious teachers reported transient feelings of tiredness and frustration, they were relieved by subsequent routine recovery periods. Therefore, these transient experiences of fatigue were not regarded as representative of burnout themes.

Contact avoidance

Another theme was related to avoiding too much personal and interpersonal relationship with EFL students. The low self-efficacious teachers were inclined to keep a distance from their students, especially outside the classroom. They reported keeping a low profile most of the time. Their main reason seemed to be harboring a deep resentment of disruptive students, violence, discipline problems (e.g., shouting in the class), and students' inattention and lack of interest in learning. One male teacher with a bachelor's degree in the Teaching of English said:

> I am really put off at the sight of some of my students. They don't know how to behave themselves. That's intolerable …. They're … awful. I wonder why just …. why they, … er, waste their time here. It is not good to be a teacher here.

Another male teacher participant explained, "I don't really care if my students understand the material or not, I do not want to see them in the office at all." On the other hand, teachers with high levels of self-efficacy reported a tendency to have more contact with their students and highlighted the need to build mutual trust in and out of the classroom. One of the female interviewees with 16 years of teaching experience explained, "I love my job and students; having close relationships with students could be more of a help. It can change their negative attitudes." In sum, the majority of high self-efficacious teachers strongly believed in the importance of paying undivided attention to the interactional and interpersonal relationships with their students.

Stress

The last emerging theme was related to the teachers' stress and anxiety in their job due to a lack of skills, great expectation and challenges, and long hours spent in the classroom. Stress manifested itself when the low self-efficacious teachers were bereft of the necessary skills and resources to deal with the diverse demands and expectations on the part of students and cope with unexpected challenges in the schools. One male teacher with a low level of efficacy, who had 12 years of teaching experience, explained, "Teaching English is a stressful job and all eyes are fixed on you. It puts pressure on you and you cannot motivate the students enough." Another male teacher disclosed:

> When a student raises his hand to ask a question, I become so keyed up. I wonder if I'll be on the right path. It's very very, er, dreadful. The problem gets worse when a student tries to pull your leg with a grammatical point or word you don't have the foggiest idea about.

Another source of stress was related to the expectations on the part of visiting parents. A young female teacher, who was thinking about changing her job, reported, "… the problem is that if anything goes wrong, everyone will hold the teacher responsible." Another dissatisfied male teacher claimed, "If a student gets a poor mark on an exam, his or her parent will visit the school and hold the researcher responsible."

These remarks, however, were in sharp contrast with the feelings reported by teachers with higher levels of self-efficacy. The teachers were found to be less affected by feelings of chronic stress and anxiety. One of them explained:

> I like the challenge in my job so much and I like different students attending my class with different background knowledge . . ., the variety in the job is very satisfying.

In sum, the teachers who were rated as high in self-efficacy were judged to suffer less from the symptoms of professional burnout while those who were ranked as low in self-efficacy seemed to show more signs of job burnout.

Discussion

In response to some suggestions in the literature for greater insight into the interaction between personal attributes that could play a part in teacher burnout, the present study delved into teacher burnout and self-efficacy and their relationship in the EFL high school context in Iran. Based on the descriptive statistics, the high school EFL teachers in the study displayed a risk profile on burnout when their scores were compared with the guidelines offered by Maslach et al. (1996). Considering the moderate-to-high level of EE, high level of DP, and reduced PA, one can infer that the participants did not enjoy a satisfactory level of individual achievement and self-efficacy. The descriptive statistics on the self-efficacy measure lent support to this line of interpretation. The self-efficacy scores, in general, did not depict a high sense of perceived self-efficacy among the EFL teachers. This was also observed with respect to the self-efficacy subcomponents, too. These results indicate that the high school English teachers did not feel very competent doing their job and many had feelings of burnout in their high school teaching jobs. As Friedman and Farber (1992) argued, teachers with a high level of inefficacy and a low sense of classroom efficacy feel stressed and pessimistic about their work and students' ability to improve.

Furthermore, the correlational results provided evidence to reject the first null hypothesis of the study and indicated a meaningful relationship between the teachers' self-efficacy and the three dimensions of job burnout in the EFL context of Iran. The results from correlational analysis revealed that the EFL high school teachers' depersonalization and emotional exhaustion had negative and significant relationships with their self-efficacy, that is, the beliefs they had about their own abilities to organize and implement the courses of action in teaching English. As Schwarzer and Hallum (2008) pointed out, the self-efficacy construct implies a kind of protective tool when coping with unfavorable working conditions. Enjoying a strong sense of belief in one's occupational capabilities and the self-confidence to effectively deal with work challenges will boost the motivation to engage in strategies of coping. This being the case, the highly self-efficacious teacher participants interpret

their daily instructional challenges in the classroom as being less menacing than teachers who harbored self-doubts about their job performance and achievement. It is very likely that successful adaptation to stressful expectations and demands in the EFL classroom would stop or delay job burnout. Therefore, self-efficacy can make a difference in how EFL teachers think, feel, and perform in the classroom. Furthermore, the relatively high negative correlations between depersonalization and emotional exhaustion with teacher self-efficacy, on the one hand, and the large positive correlation between personal accomplishment and teacher self-efficacy, on the other hand, support the findings reported by Mardani et al. (2015) and Motallebzadeh et al. (2014) who came up with the reverse correlations between teacher burnout and self-efficacy variables among English language teachers in different contexts. Specifically, they found that instructional efficacy was the most important factor in predicting emotional exhaustion, depersonalization, and personal accomplishment. Additionally, the results of the study by Ozkara (2019), conducted among Turkish EFL state teachers, count as a piece of corroborative evidence for this kind of relationship between self-efficacy and burnout components. More likely, the significant role of self-efficacy in burnout experience cuts across EFL boundaries and different geographical and cultural contexts. This is particularly evident in the findings of the study by Skaalvik and Skaalvik (2007), who reported that a low level of self-efficacy among Norwegian elementary teachers could increase the risk of occupational stress, which might, in turn, increase the level of emotional exhaustion and depersonalization. Therefore, the antecedents underlying teacher burnout and the mitigating effect of self-efficacy on this phenomenon seem to be not confined to a specific educational and cultural milieu.

The main themes of mental fatigue, contact avoidance, and stress emerged from the in-depth qualitative analysis also demonstrated that EFL teachers who were rated as low in self-efficacy were more vulnerable to the symptoms of professional burnout. According to the data, the first dominant symptom in the majority of the EFL teachers with low self-efficacy and who were suffering from temporal or chronic burnout was the feeling of mental fatigue. Being unable to successfully overcome some of the daily obstacles in teaching English in a high school and dealing with their students' noncompliant behavior in the classroom, the low self-efficacious teachers were gradually overcome by feelings of lassitude and decided to change their job. As Hong (2012) points out, such a situation over time drains the teacher's energy and interest to the point of deciding to quit the job. The theme of contact avoidance, which was mostly observed in the low self-efficacious male teachers, can also be interpreted as the inability of such teachers to develop effective communicative, instructional, and class management strategies for confronting new challenges in foreign language teaching. More likely, those participants who were identified as suffering from the depersonalization dimension had resorted to a kind of professional seclusion. This relational avoidance had, in turn, resulted in their adoption of an air of indifference towards their professional improvement and their students' insignificant L2 achievement. The theme of stress was also related to such issues as the students' diverse needs and expectations in L2 learning, incongruence between the teachers' expectations and job demands, class management problems, as well as great expectations on the part of other stakeholders. The high self-efficacious teachers, particularly the female ones, were reportedly less affected by this negative emotional quality. These EFL teachers, perhaps, were equipped with internal coping tools and resources like resilience to better manage stress in their profession over the long term.

Furthermore, considering the role of gender in EFL teachers' burnout, the results showed that gender, to some extent, could be considered a modifying variable in explaining the difference in burnout experiences among the EFL teachers. The male teachers reported perceived symptoms of emotional exhaustion and depersonalization more frequently than the female teachers. As a possible reason, it can be argued that a sense of social responsibility and occupational commitment was stronger among the female EFL teachers, who had more tendencies to socialize at school, build good relations with their students or colleagues, and use their coping strategies in the face of problems. As Maslach and Jackson (1985) asserted, the female role in the workplace is often characterized with a nurturing and caring attitude when dealing with different people and their problems. This can possibly protect female EFL high school teachers against the negative consequences of job burnout. However, unfavorable job-related factors such as low payment might contribute to higher burnout levels among male EFL teachers in the public high school context in Iran. The above results on gender differences support the results of studies by Küçüksüleymanoğlu (2011) and Motallebzadeh et al. (2014) who reported that female teachers would not be affected as much as male teachers by the unpleasant conditions of the teaching profession. Küçüksüleymanoğlu (2011) found out that male teachers in special education schools in Turkey felt more emotionally exhausted than female teachers. Motallebzadeh et al. (2014) also reported that male teachers' level of burnout was significantly higher than that of the female teachers in private language institutes in Khorasan Razvi, Iran. Along the same lines, while exploring the cross-sectional associations between teacher

burnout and several putative precursors, Schwarzer and Hallum (2008) found that male German and Syrian teachers experienced more depersonalization than female German and Syrian teachers. Nonetheless, these findings are not consistent with the findings of the studies by Mardani et al. (2015), who concluded that EFL female teachers experienced more burnout symptoms than male teachers, and Mede (2009), who did not find gender as a significant predictive factor in the three burnout dimensions among Turkish EFL teachers. It seems that the context of teaching and educational environment of foreign language teaching in schools can pose different types of challenges for female and male teachers, leading to mixed gender-related differences in reported burnout. Thus, the role of the gender effect in burnout studies should be treated with caution.

Conclusion

What this research study has brought home is the importance of perceived self-efficacy as an internal protective factor against burnout in pedagogical settings. The present study has confirmed the significant correlation between teacher self-efficacy and professional teacher burnout, a phenomenon affecting both teacher turnover rates and student outcomes. The findings revealed that higher burnout levels correlated meaningfully with lower levels of self-efficacy among EFL teachers. Additionally, female teachers demonstrated lower burnout levels than the male ones, highlighting the modifying function of the gender variable in burnout studies. The follow-up semi-structured interviews also supported these quantitative findings on the negative association between teacher self-efficacy and professional teacher burnout. The three main themes of mental fatigue, contact avoidance, and stress emerged in explaining the burnout experience of EFL teachers.

The negative consequences of burnout call for teacher educators in the EFL context to understand the conditions leading to this phenomenon and the factors fostering teacher self-efficacy. After all, if EFL teachers hope to function as a torch for EFL students and provide them with opportunities for personal growth and the motivation to learn, as the results suggest, they need to develop persistence and resilience when they face pedagogical challenges in teaching a foreign language. Therefore, L2 teacher educators should seize the opportunity to inculcate these essential qualities in EFL teachers. The relationship between teacher self-efficacy and professional teacher burnout also implies that L2 administrators should develop programs to train teachers with strong positive beliefs in their capabilities early on in their L2 teaching career, and set up a congenial environment at schools to boost self-efficacy among EFL teachers.

At the same time, L2 teachers should be assisted in expanding their skills for managing the feelings of incompetency, stress, and the resulting negative effects in L2 classes. The findings highlight the important role of L2 school administrators in discussing the antecedents of teacher burnout and providing a flexible and supportive environment to reduce EFL teachers' sense of exhaustion and stress. Additionally, L2 educators are recommended to pay attention to teacher self-efficacy as a pedagogical quality when recruiting foreign language teachers.

Last but not least, the findings should be interpreted in light of some limitations. First, some results were based on self-reported data. As Gay, Mills, and Airasian (2009) point out, participants in a survey might select the responses that are more desirable even if they do not actually believe them. Thus, other types of data collection procedures such as diaries, as recommend by Clarkson and Hodgkinson (2007), can be employed to make stronger generalization about the association between self-efficacy and professional burnout among EFL teachers. Future L2 burnout research might also benefit from data provided by students on EFL teachers' burnout experience and self-efficacy. This can be complementary to the primary data, and hence reduce any possible bias. Second, the current study used semi-structured interviews for the data triangulation; however, the codes that evolved from the qualitative interview data were limited to some EFL teachers who presented their reflections according to their personal views. Other codes could, therefore, be generated by conducting interviews at a different time with other foreign language teachers. Third, the degree of motivation by EFL students might have affected the teachers' burnout levels. As research (e.g., Anderson & Iwanicki, 1984; Roohani & Dayeri, 2019) suggests, student motivation could be a contributing factor in the burnout experience. Therefore, future research can take student motivation into consideration in burnout studies. Fourth, the present study did not include a control group of schoolteachers in other disciplines in dealing with the relationship between self-efficacy and the burnout experience among language teachers. Future research, however, can include a control

group of schoolteachers, such as mathematics teachers, who may suffer from similar challenges when they are engaged in teaching (Cui et al., 2018). In this way, one might gain more insight into the phenomenon of burnout and its possible cross-disciplinary differences. Fifth, this study was conducted with a relatively small sample of EFL teachers in Iran. Future research can be carried out with a larger sample and other types of sampling because research shows that a larger sample is more representative of the population (Robson, 1993). Finally, the findings, particularly those on the gender effect, must be treated with caution if they are generalized to EFL teachers in the context of decentralized educational systems like private language institutes where the context is different from a centralized educational system like public high schools in Iran. Future research can also replicate the same study among EFL teachers in other contexts.

References

Akbari, R., Ghafar Samar, R., Kiani, Gh., & Eghtesadi, A. (2011). Factorial validity and psychometric properties of Maslach burnout inventory-the Persian. *Knowledge and Health*, *6*(3), 1-8.

Alarcon, G., Eschleman, K. J., & Bowing, N. A. (2009). Relationship between personality variables and burnout: A meta-analysis. *Work & Stress, 23,* 244-263.https://doi.org/10.1080/02678370903282600

Aloe, A. M., Amo, L. C., & Shanahan, M. E. (2014). Classroom management self-efficacy and burnout: A multivariate meta-analysis. *Educational Psychology Review, 26*(1), 101-126. https://doi.org/10.17323/jle.2019.v5.i410.1007/s10648-013-9244-0

Anderson, B. G., & Iwanicki, E. F. (1984). Teacher motivation and its relationship to burnout. *Educational Administration Quarterly, 20,* 109-132. https://doi.org/10.1177/0013161X84020002007

Bandura, A. (1994). Self-efficacy. In V. S. Ramachandran (Ed.), *Encyclopedia of human behavior* (vol. 4, pp. 71-81). Academic Press.

Bandura, A. (1995). *Self-efficacy in changing societies.* Cambridge University Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

Betoret, F. D. (2009). Self-efficacy, school resources, job stressors, and burnout among Spanish primary and secondary school teachers: A structural equation approach. *Educational Psychology, 29*(1), 45-68. https://doi.org/10.1080/01443410802459234

Biron, M., & Link, S. (2014). Stress, appraisal, and work routine in wartime: Do men and women differ? *Anxiety, Stress, & Coping, 27,* 229-240. https://doi.org/10.1080/10615806.2013.836506

Blase, J. J. (1984). School principals and teacher stress: A qualitative analysis. *National Forum of Educational Administration and Supervision, 1*(2), 35-43.

Brouwers, A., & Tomic, W. (2000). A longitudinal study of teacher burnout and perceived self-efficacy in classroom management. *Teaching and Teacher Education, 16,* 239-253. https://doi.org/10.1016/S0742-051X(99)00057-8

Brown, C. G. (2012). A systematic review of the relationship between self-efficacy and burnout in teachers. *Educational & Child Psychology, 29,* 47-63.

Burke, R. J., & Richardsen, A. M. (1996). Stress, burnout, and health. In C. L. Cooper (Eds.), *Handbook of stress, medicine, and health* (pp. 101-117). CRC Press.

Byrne, B. M. (1993). The Maslach Burnout Inventory: Testing for factorial validity and invariance across elementary, intermediate, and secondary teachers. *Journal of Occupational and Organizational Psychology, 66*, 197-212. https://doi.org/10.1111/j.2044-8325.1993.tb00532.x

Chan, D. W. (2008). Teacher self-efficacy and successful intelligence among Chinese secondary school teachers in Hong Kong. *Educational Psychology, 28*(7), 735-746. https://doi.org/10..1080/01443410802259246

Chan, M. (2009). An appraisal perspective of teacher burnout: Examining the emotional work of teachers. *Educational Psychology Review, 21,* 193-218. https://doi.org/10.1007/s10648-009-9106-y

Chang, M. L., & Engelhard, G. (2016). Examining the teachers' sense of efficacy scale at the item level with Rasch measurement model. *Journal of Psychoeducational Assessment, 34*(2), 177-191. https://doi.org/10.1177/0734282915593835

Chwalisz, K., Altmaier, E. M., & Russell, D. W. (1992). Causal attributions, self-efficacy cognitions, and coping with stress. *Journal of Social and Clinical Psychology, 11,* 377-400. https://doi.org/10.1521/jscp.1992.11.4.377

Clarkson, G. P., & Hodgkinson, G. P. (2007). What can occupational stress diaries achieve that questionnaires can't? *Personnel Review, 36*(5), 684-700. https://doi.org/10.1108/00483480710773990

Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social-emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy. *Journal of Educational Psychology, 104*(4), 1189-1204. 1204.

https://doi.org/10.17323/jle.2019.v5.i4 10.1037/a0029356
Coulter, M. A., & Abney, P. C. (2009). A study of burnout in international and country of origin teachers. *International Review of Education*, *55*, 105-121.
Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage.
Cui, Q., Chao, Q., Han, J., Zhang, X., Ren, Y., & Shi, J. (2018). Job stress, burnout and the relationship among the science and mathematics teachers in basic education schools. *Eurasia Journal of Mathematics, Science and Technology Education, 14*(7), 3235-44. https://doi.org/10.29333/ejmste/85957
Day, C., & Gu, Q. (2010). *The new lives of teachers*. Routledge.
Dicke, T., Parker, P. D., Marsh, H. W., Kunter, M., Schmeck, A., & Leutner, D. (2014). Self-efficacy in classroom management, classroom disturbances, and emotional exhaustion: A moderated mediation analysis of teacher candidates. *Journal of Educational Psychology, 106*(2), 569-583. https://doi.org/10..1037/a0035504
Dorman, J. P. (2003). Relationship between school and classroom environment and teacher burnout: A LISREL analysis. *Social Psychology of Education, 6,* 107-127. https://doi.org/110.1023/A:1023296126723
Eslami, Z. R., & Fatahi, A. (2008). Teachers' sense of self-efficacy, English proficiency, and instructional strategies: A study of nonnative EFL teachers in Iran. *TESL-EJ, 11*(4), 1-19.
Evers, W. J. G., Brouwers, A., & Tomic, W. (2002). Burnout and self-Efficacy: A study on teachers' beliefs when implementing an innovative educational system in the Netherlands. *British Journal of Educational Psychology, 72,* 227-45. https://doi.org/10.1348/000709902158865
Farber, B. A. (2000). Treatment strategies for different types of teacher burnout. *Journal of Clinical Psychology, 56*(5), 675-689.
Freudenberger, H. J. (1974). Staff burnout. *Journal of Social Issues, 30,* 159-165. https://doi.org/10.1111/j.1540-4560.1974.tb00706.x
Friedman, I. A., & Farber, B. A. (1992). Professional self-concept as a predictor of teacher burnout. *Journal of Educational Research, 86,* 28-35. https://doi.org/10.1080/00220671.1992.9941824
Gay, L. R., Mills, G. E., & Airasian, P. (2009). *Educational research: Competencies for analysis and applications* (9th ed.). Pearson Education.
Greenglass, E. R., & Burke, R. J. (1988). Career orientations and psychological burnout in teachers. *Psychological Reports, 63,* 107-116. https://doi.org/10.2466/pr0.1988.63.1.107
Hanushek, E. A. (2007). The single salary schedule and other issues of teacher pay. *Peabody Journal of Education, 82*(4), 574-586. https://doi.org/10.1080/01619560701602975
Hong, J. Y. (2012). Why do some beginning teachers leave the school, and others stay? Understanding teacher resilience through psychological lenses. *Teachers and Teaching: Theory and Practice, 18*(4), 417-440. https://doi.org/10.1080/13540602.2012.696044
Jedynak, M. (2011). Teaching experience and its role in foreign language teachers' anxiety. *Acta Neophilologica, 13*, 59-72.
Jennett, H. K., Harris, S. L., & Mesibov, G. B. (2003). Commitment to philosophy, teacher efficacy, and burnout among teachers of children with autism. *Journal of Autism and Developmental Disorders, 33,* 583-593. https://doi.org/10.1023/B:JADD.0000005996.19417.57
Kahn, J. H., Schneider, K. T., Jenkins-Henkelman, T. M., & Moyle, L. L. (2006). Emotional social support and job burnout among high-school teachers: Is it all due to dispositional affectivity? *Journal of Organizational Behavior, 27,* 793-807. https://doi.org/10.1002/job.397
Kay-Eccles, R. (2012). Meta-analysis of the relationship between coworker social support and burnout using a two level hierarchical linear model. *Western Journal of Nursing Research, 34,* 1062-1063. https://doi.org/10.1177/0193945912453684
Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology, 102,* 41-756. https://doi.org/10.1037/a0019237
Klassen, R. M., Tze, V. M. C., Betts, S. M., & Gordon, K. A. (2011). Teacher efficacy research 1998-2009: Signs of progress or unfulfilled promise? *Educational Psychology Review, 23*(1), 21-43. https://doi.org/10.1007/s10648-010-9141-8
Klusmann, U., Kunter, M., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Engagement and emotional exhaustion in teachers: Does the school context make a difference? *Applied Psychology: An International Review*, *57,* 127-151. https://doi.org/10.1111/j.1464-0597.2008.00358.x
Kokkinos, C. M. (2007). Job stressors, personality, and burnout in primary school teachers. *British Journal of Educational Psychology*, *77*(1), 229-243.https://doi.org/10.1348/000709905X90344
Küçüksüleymanoğlu, R. (2011). Burnout syndrome levels of teachers in special education schools in Turkey.

International Journal of Special Education, 26(1), 53-63.
Larson-Hall, J. (2010). *A guide to doing statistics in second language research using SPSS*. Routledge.
Lau, P. S., Yuen, M. T., & Chan, R. M. (2005). Do demographic characteristics make a difference among Hong Kong secondary school teachers? *Social Indicators Research*, *71,* 491-516.
Lim, N., Kim, E. K., Kim, H., Yang, E., & Lee, S. (2010). Individual and work-related factors influencing burnout of mental health professionals: A meta-analysis. *Journal of Employment Counseling, 47,* 86-96. https://doi.org/10.1002/j.2161-1920.2010.tb00093.x
Mackey, A., & Gass, S. (2005). *Second language research: Methodology and design*. Lawrence Erlbaum Associates Publishers.
Mardani, N., Baghelani, E., & Azizi, R. (2015). Exploring the relationship between self-efficacy and burnout: The Case of Iranian EFL teachers. *Cumhuriyet* Üniversitesi *Fen Edebiyat Fakültesi Fen Bilimleri Dergisi, 36*(3), 3538-3548.
Maslach, C. (1982). *Burnout: The cost of caring*. Prentice Hall.
Maslach, C. (2015). Psychology of burnout. *International Encyclopedia of the Social and Behavioral Sciences*, *2,* 929-932.
Maslach, C., & Jackson, S. E. (1985). The role of sex and family variables in burnout. *Sex Roles, 12,* 837-851.
Maslach, C., & Jackson, S. E. (1986). *Maslach burnout inventory manual* (2nd ed.). Consulting Psychologist Press.
Maslach, C., Jackson, S. E, & Leiter, M. P. (1996). *MBI: The Maslach Burnout Inventory Manual* (3rd ed.). Consulting Psychologists Press.
Maslach, C., Jackson, S. E., & Schwab, R. L. (1986). Maslach Burnout Inventory - Educators Survey (MBI-ES). In C. Maslach, S. E. Jackson, & M. P. Leiter. (Eds.), *MBI Manual* (3rd ed., pp. 191-218). Consulting Psychologist Press.
Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology, 52,* 397-422. https://doi.org/10.1146/annurev.psych.52.1.397
McCarthy, C., Lambert, R., O'Donnell, M., & Melendres, L. (2009). The relation of elementary teachers' experience, stress, and coping resources to burnout symptoms. *The Elementary School Journal, 109,* 1-19. https://doi.org/10.1086/592308
Mede, E. (2009). An analysis of relations among personal variables, perceived self-efficacy and social support on burnout among Turkish EFL teachers. *Inonu University Journal of the Faculty of Education, 10*(2), 39-52.
Motallebzadeh, K., Ashraf, H., & Tabatabaee-Yazdi, M. (2014). On the relationship between Iranian EFL teachers' burnout and self-efficacy. *Procedia-Social and Behavioral Sciences, 98,* 1255-1262. https://doi.org/10.1016/j.sbspro.2014.03.541
Ortero Lopez, J. M., Castro Bolaño, C. C., Santiago Mariño, M. J, & Villardefrancos Pol, E. (2010). Exploring stress, burnout, and job satisfaction in secondary school teachers. *International Journal of Psychology and Psychological Therapy, 10*(1), 107-123.
Ozkara, B. (2019). An investigation into the relationship between Turkish EFL teachers' self-efficacy and burnout level. *Journal of Family Counseling and Education, 4*(1), 12-24. dhttps://doi.org/10.32568/jfce.504499
Perrewé, P. L., Hochwarter, W. A., Rossi, A. M., Wallace, A., Maignan, I., Castro, S. L., et al. (2002). Are work stress relationships universal? A nine-region examination of role stressors, general self-efficacy, and burnout. *Journal of International Management, 8,* 163-187. https://doi.org/ 10.1016/S1075-4253(02)00052-2
Poulin, J., & Walter, C. (1993). Social workers' burnout: A longitudinal study. *Social Work Research & Abstracts, 29*(4), 5-11. https://doi.org/10.1093/swra/29.4.5
Purvanova, R. K., & Muros, J. P. (2010). Gender differences in burnout: A meta-analysis. *Journal of Vocational Behavior, 77,* 168-185.
Reeve, J., & Su, Y-L. (2014). Teacher motivation. In M. Gagné (Ed.), *The Oxford handbook of work engagement, motivation, and self-determination theory* (pp. 349-362). Oxford University Press.
Robson, C. (1993). *Real world research: A resource for social scientists and practitioner-researchers*. Blackwell.
Roohani, A., & Dayeri, Kh. (2019). On the relationship between Iranian EFL teachers' burnout and motivation: A mixed methods study. *Iranian Journal of Language Teaching Research (IJLTR) 7,*1,77-99.
Sari, H. (2004). An analysis of burnout and job satisfaction among Turkish special school head teachers and teachers, and the factors affecting their burnout and job satisfaction. *Educational Studies, 30*(3), 291-306. https://doi.org/10.1080/0305569042000224233
Schwarzer, R., & Hallum, S. (2008). Perceived teacher self-efficacy as a predictor of job stress and burnout: Mediation analyses. *Applied Psychology, 57,* 152-171. https://doi.org/10.1111/j.1464-0597.2008.00359x
Schwarzer, R., Schmitz, G. S., & Tang, C. (2000). Teacher burnout in Hong Kong and Germany: A cross-cultural validation of the Maslach Burnout Inventory. *Anxiety, Stress, and Coping, 13 (3)*, 309-326. https://doi.

org/10.1080/10615800008415730
Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy, and teacher burnout. *Journal of Educational Psychology, 99,* 611-625. https://doi.org/10.1037/0022-0663.99.3.611
Skaalvik, E. M., & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education, 26,* 1059-1069. https://doi.org/10.1016/j.tate.2009.11.001
Smith, M., & Bourke, S. (1992). Teacher stress: Examining a model based on context, work load, and satisfaction. *Teaching and Teacher Education 8*(1), 31-46. https://doi.org/10.1016/0742-051X(92)90038-5
Tschannen-Moran, M., & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and Teacher Education, 17,* 783-805. https://doi.org/10.1016/S0742-051X(01)00036-1
Tschannen-Moran, M., Woolfolk Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*, 202-248.
Unsworth, K. L., & Mason, C. M. (2012). Help yourself: The mechanisms through which a self- leadership intervention influences strain. *Journal of Occupational Health Psychology, 17,* 235-245. https://doi.org/10.1037/a0026857
Van Tonder, C. L., & Williams, C. (2009). Exploring the origins of burnout among secondary educators. *SA Journal of Industrial Psychology, 35*(1), 204-218.
Woolfolk, A. (2004). *Educational Psychology* (9th ed.). Allyn and Bacon.
Yang, H., & Farn, C. K. (2005). An investigation of the factors MIS student burnout in technical- vocational college. *Computer Human Behavior, 21,* 917-932. https://doi.org/10.1016/j.chb.2004.03.001

Appendix

Semi-Structured Interview Questions

1. Do you feel tired from your work? If so, why?
2. Have you ever thought of changing your job? If so, why?
3. Have you ever felt tired of teaching at school? If so, how often?
4. Can you talk about some of your achievements as an EFL teacher?
5. Do you think you are successful in your job, managing the classroom and behaviors of your students, controlling disruptive behavior…? If so, how and how much?
6. Is it important for you when your students learn/do not English? Why and how much?
7. Can you get your students to believe that they can do well in learning English at school?
8. Can you change your students' attitudes/views toward learning English? Why and how much?

National Research University Higher School of Economics
Journal of Language & Education Volume 6, Issue 1, 2020

Pahlevansadegh, V., Vasheghani Farahani, M.V. (2020). Corpus Linguistics for Vocabulary: A guide for Research by Paweł Szudarski. Routledge Publications 2018. 239 pp. ISBN: 978-1-138-18721-4. *Journal of Language and Education, 6*(1), 189-192. https://doi.org/10.17323/jle.2020.10554

Corpus Linguistics for Vocabulary: A guide for Research by Paweł Szudarski. Routledge Publications 2018. 239 pp. ISBN: 978-1-138-18721-4

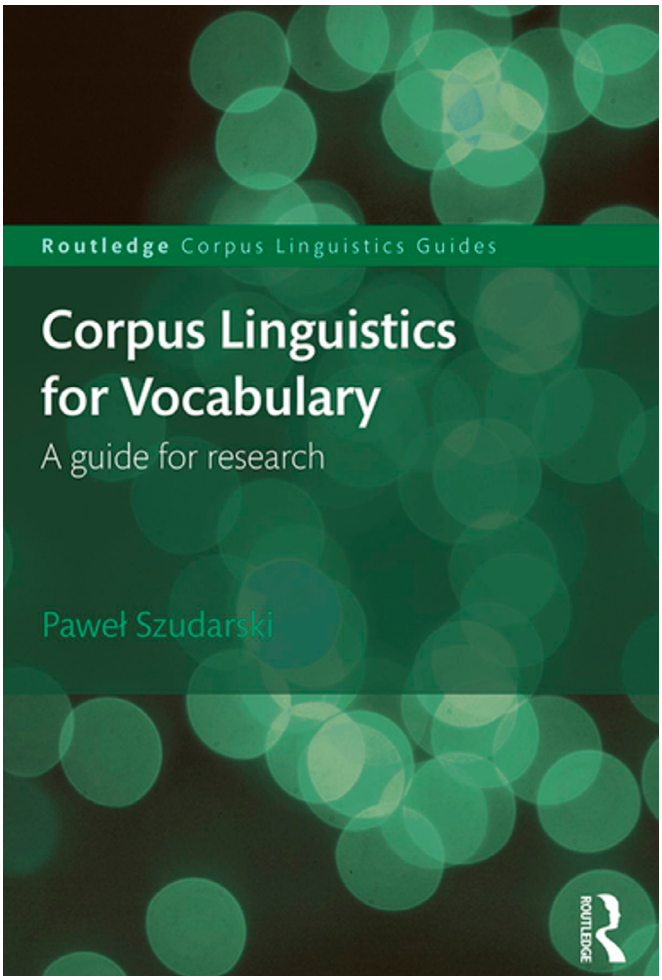
Vahid Pahlevansadegh[1], Mehrdad Vasheghani Farahani[2]

[1] Georgia University

[2] Universitat Leipzig

Correspondence concerning this article should be addressed to Vahid Pahlevansadegh, Georgia University, 77a, M. Kostava str. Tbilisi, 0171, Georgia. E-mail: vahid_pahlevansadegh@yahoo.com

Since its advent, corpora have been used in various aspects of language-related issues including teaching English (Timmis, 2015). This extensive usage of corpora in language studies has brought about the creation of many resources, among which is "Corpus Linguistics for Vocabulary: A Guide for Research" by Paweł Szudarski published by Routledge in 2018. This book is composed of 10 chapters. This book provides the competent readership with the systematic and comprehensive clarification of fundamental concepts and terms in corpus linguistics and its applications in vocabulary teaching.



The first chapter on the definition of corpus linguistics comprises five sub-sections. In the first chapter, the author defines corpus linguistics. As the author provides the reader with a very practical and fathomable definition of corpus linguistics, he refers to its main two approaches as corpus-based and corpus-driven; referring to the former as a methodology and the latter as a theory. The second sub-subsection elaborates on the criteria by which a corpus is designed. Having discussed at length and in detail the principles of corpus design and creation, in the third sub-section the author concentrates on the advantages of corpus linguistics by listing the benefits we can take from corpora. In the fourth sub-section, the author sheds light on the limitations of embarking on corpus analysis; pointing to the constraints that corpora can have. In the last sub-section, a variety of internet-based corpus types are provided, especially free databases.

Chapter 2 has three sub-sections and is about tools and statistics in corpus analysis. The first sub-section focuses on tools and types of analysis. First, the author introduces several web-based corpora. Then he uses concordance as an important tool for conducting different kinds of corpus-informed research with some examples. Next, some very important corpus analyses are discussed, such as frequency analysis and concordance, wordlists, cluster (n-gram) analysis, and keyword analysis. In the second sub-section, statistical tests in carrying out corpus-based research are elaborated on. First, log-likelihood is introduced

for conducting comparative analyses on the frequency values between diverse collections of data. Next, tests like t-score and mutual information (MI) test are introduced to determine collocations. Type-token ratio and lexical richness tests are the last tests elucidated on, which are used to examine lexical diversity. The last sub-section explains how to integrate quantitative and qualitative analyses in corpus-informed research.

Chapter 3 has five sub-sections. It elaborates on vocabulary and introduces (as the first sub-section) vocabulary as a core constituent in language use. The second sub-section provides useful terminologies in literature like words, vocabulary, and lexis. This part presents corpus-related terminologies like word form, lexeme, lemma, word family, and lexical item. In next part, the author provides different perspectives by which vocabulary as a multifaceted construct can be approached. The fourth sub-section discusses vocabulary-related notions that can be studied using corpus-based analysis, which consists of the polysemy of words, synonyms, metaphoricity, and register variation. The author, in the last sub-section, provides some research questions pertinent to the issue.

Chapter 4, entitled Frequency and Vocabulary, includes five sub-sections and discusses the importance of frequency in corpus-based linguistics and different types of frequency-based categorization of vocabulary. The first sub-section focuses on the significance of frequency, which is a rudimentary function of corpora. It also elaborates on different types of frequency in vocabulary, such as the frequency of spoken and written words and that of content and function words. Then, the author introduces Zipf's law (1935), which refers to the diverse relationship between the frequency of an occurring word and its rank in the frequency list. Next, the author focuses on the term 'frequency' itself and enumerates different types of word frequency as low-, mid-, and high-frequency words and explains each type. Lexical coverage and bands of vocabulary are two other terms that are covered in this part. In the second sub-section, the author concentrates on frequency-based research on vocabulary load. The third sub-section is about useful vocabulary lists and different examples of such lists and their specifications. In next sub-section, the author considers the impact of corpus-based research on vocabulary and frequency in teaching. Different pedagogical aspects are addressed which are based on frequency and quantitative analysis. In the last sub-section, a study on mid-frequency readers is discussed.

Chapter 5 comprises seven sub-sections and brings the readers' attention to corpora, phraseology, and formulaic language. In the first sub-section, the author discusses how meaning is conveyed beyond single words and through multiword units. It also introduces the theory of lexical priming (Hoey, 2005), which connects corpus-based research with phraseology through a psycholinguistic perspective. In the next sub-section, the author concentrates on formulaic language and defines it as an umbrella term including all types of phraseological units. In the third sub-section, the open-choice principle, the idiom choice principle, and lexico-grammar are elaborated on. In the fourth sub-section, the author describes various types of multiword units like collocation, colligation, semantic preference, and semantic prosody. The fifth sub-section concentrates on another facet of phraseological research, which is register variation referring to diverse types of language used in social contexts. In the sixth sub-section, the author discusses corpus-based lists of phraseology and research that has been carried out in this field in order to determine the most useful phrases and their usage in language. It also focuses on the process of finding useful vocabulary and issues influencing the process directly. A study related to vocabulary lists is discussed in the last sub-section.

Chapter 6 discusses corpora and teaching vocabulary and includes five sub-sections. In the first sub-section, the author introduces two approaches in which corpora are used in language teaching as direct and indirect applications. The second sub-section is concerned with corpora in indirect application and elaborates on how they influence the development of different materials. The next sub-section is concerned with pedagogical corpora and its benefits in the language teaching process. The fourth sub-section deals with teaching oriented corpora, in which the author names the benefits of corpora in language teaching. Finally, in the last sub-section a study about the use of corpora for teaching purpose is provided.

Chapter 7 includes five sub-sections and discusses the significance of leaner corpora when analyzing learner vocabulary. The first sub-section emphasizes the importance of learner language, which is regarded as all forms of language produced by second language (L2) learners represented by a number of peculiar traits. In the second sub-section, characteristics of learner corpora are elaborated on. The third sub-section deals with using learner corpora in language teaching. The author also elucidates on different examples of corpora characterized by the use of English in EFL contexts. At the end, some ways for carrying out research on learner vocabulary

improvement are put forward, such as research based on longitudinal data. The fourth sub-section starts by explaining the ways learner corpora contributes toward having a more valid and reliable assessments of learners' proficiency. There are, according to the book, some challenges and difficulties in using learner corpora for the purpose of language testing, which have been addressed thoroughly in the book. The last sub-section introduces a study on the application of learner corpora for pedagogical purposes.

Chapter 8 elucidates 'specialized corpora' and has six sub-sections. First, the author introduces this type of corpora as an approach to answer some criticism about the use of decontextualized data in corpora. The second sub-section deals with register and genre analysis. It is stated that as different registers exists in different contexts, they possess different syntactical, lexical, and discourse characteristics. Therefore, a huge part of corpus analysis is devoted to the study of register variations by using register-focused corpora. In the next sub-section, the role played by corpora in investigating English for specific purpose (ESP) is highlighted. It is stated that ESP researchers are now interested in investigating different linguistic features in specialized language applications. Next, the author brings attention to the process of identifying ESP vocabulary using corpus-informed approaches. In the fourth sub-section, the profound impact of corpora in investigating English for academic purposes (EAP) is accentuated. Different corpus-informed research focusing on academic vocabulary is presented, including academic word lists as well as phraseology and academic vocabulary lists. In the last sub-section, the author places emphasis on the application of specialized corpora in areas such as translation studies and literacy texts. .

Chapter 9 focuses mainly on studying the use of words through a corpus-informed approach and discourse analysis. It comprises five sub-sections. The first sub-section defines the terms 'discourse' and 'discourse analysis.' Regarding the use of both methods together (corpus-based approach and discourse analysis), the author mentions that although there are basic differences between these two methods, both of them share a concentration on linguistic-related data. In the second sub-section, the author elucidates the interaction between these two approaches, stating that discourse analysis and corpus linguistics are based on qualitative and quantitative analyses. The focus then moves on to lexical features in the third sub-section. In this part, the importance of vocabulary in forming the basic framework of the text and the way through which discourse analysis and corpus-based approach can highlight the significance of vocabulary on the discourse are discussed. Another feature is 'relexicalization', which is regarded as a way in which speakers use one another's vocabulary to paraphrase conversation. 'Intertextuality' is another term in discourse analysis which that investigates how a discourse refers to a prior or future discourse. In this sub-section, a huge number of authentic examples of conversation are presented and researchers can use them in their studies. Corpora and pragmatics is the topic of discussion in the fourth sub-section. First, pragmatics is defined as the study of how meaning is formed and perceived in certain context. Then, its relation to the role of corpora is elaborated on. At the end, some corpus-based pragmatic research is provided, such as the pragmatic analysis of vocabulary as well as semantic prosody and speech acts.

Chapter 10 is a summary and repopulation of the book, which includes some exercises for readers.

Szudarski's seminal work has significant contribution to the field of vocabulary learning and teaching in second language. In other words, this book provides the competent readership with the systematic and comprehensive clarification of fundamental concepts and terms in corpus linguistics and its applications in vocabulary teaching. Reading each chapter of this impressive book will lay the foundation for further and detailed research in this area of inquiry. This book is useful in that the examples are taken from BNC (British National Corpus) and COCA (Corpus of Contemporary American English) as two main monitor corpora. The so many authentic examples taken from BNC and COCA brings the reader with tangible examples which is one of the strongest points of this book. Moreover, the author has taken direct phots from these two corpora interfaces in order to familiarize the readers with the lay out of the corpora as well as the real examples. The very concise summaries provided at the end of each chapter bring the readers with a recapitulation of the concepts. Another reliable source of information is the list of jargons and terminologies listed at the end of book. By referring to the list, the author will have the chance to read the practical definitions. Moreover, the very positive point of this book is that it gives lots of concrete examples from the two corpora. These examples help the reader to understand the concepts in a practical way.

References

Hoey, M. (2005). Lexical priming: A new theory of words and language. Routledge.
Timmis, I. (2015). Corpus Linguistics for ELT: Research and Practice. Routledge.

Contents