

Volume 33 Issue 1 2023

Applied Language Learning

A Practice-Based Journal for Foreign Language Educators



Defense Language Institute
Foreign Language Center
Presidio of Monterey

Applied Language Learning

The mission of *Applied Language Learning* (US ISSN 1041-679X and ISSN 2164-0912) is to promote professional communication within the Defense Language Institute Foreign Language Center and academic communities on adult language learning for functional purposes. The Editor encourages the submission of pedagogy-related research and review manuscripts from areas such as: (1) instructional methods and techniques; (2) curriculum and materials development; (3) testing and evaluation; (4) implications and applications of research from related fields in linguistics, education, communication, psychology, and social sciences; and (5) assessment of needs within the profession.

Applied Language Learning, published by the Defense Language Institute Foreign Language Center and Presidio of Monterey, California, presents professional information. The views expressed herein are those of the authors, not the Department of Defense or its elements. The content does not necessarily reflect the official US Army position and does not change or supersede any information in official US Army publications. *Applied Language Learning* reserves the right to edit material.

Further reproduction is not advisable. Whenever copyrighted materials are reproduced in this publication, copyright release has ordinarily been obtained only for use in this specific issue. Requests for reprints should be directed to the authors.

Editor:

Dawn Bikowski, PhD, Associate Professor, Defense Language Institute Foreign Language Center

Editorial Support:

Natalia Barley, EdD, Director, Curriculum Support Division

Muhammad Farooq (Umer), Associate Director, Curriculum Support Division

Jeff Hansman, Copy Editor, Curriculum Support Division

To access *Applied Language Learning* online:

<http://www.dliflc.edu/resources/publications/applied-language-learning/>

To communicate with the Editor:

Send correspondence to the Editor at journals@dliflc.edu

Applied Language Learning
Volume 33, 2023

CONTENTS

ARTICLES

- 1 Assessing the Needs of Foreign Language Professionals at a UK Government
Department: A Case Study
~ *Jim Davie*
- 26 The Correlation Between Level of Learner Autonomy and Learner's Achievements
in Arabic Reading and Listening: A Quantitative, Non-experimental Correlational
Study
~ *Anas Faiz*
- 41 The Relative Effects of Motivation and Structured Input on the Acquisition of
English Past Tense Regular Forms
~ *Zilin Huang & Alessandro Benati*
- 58 Grammar as Concept: The Implementation of the PACE Model in Beginning
Spanish Classes
~ *Ewurama Okine & Gabriela C. Zapata*

COMMENTARY

- 71 Leveraging Artificial Intelligence for Assigning ILR Ratings to Authentic Content
~ *Jordan M. Eason*

GENERAL INFORMATION

- 81 Thank You Reviewers!
82 Call for Papers

ARTICLES

Assessing the Needs of Foreign Language Professionals at a UK Government Department: A Case Study

Jim Davie

Independent Researcher

Graduates who typically have L1 English, have majored in one foreign language (FL, L2) or more at university and have gone on to occupy FL posts in the UK civil service have reported mismatches between their pre-employment L2 learning and the tasks they face in the workplace. Such reported divergences in UK civil service capability have not, however, been investigated in any detail to date. Correspondingly, to provide a more informed picture of the knowledge gaps of (typically) graduate language professionals working at a UK government department, a qualitative needs analysis was undertaken involving participants who principally carry out transcription and translation tasks into English from an L2. The needs analysis identified seven key gap subjects – stylistic understanding; slang; dysphemism/swearing and euphemism; humour; language and culture; multilingualism; and discourse analysis – but also pinpointed narrower questions vis-à-vis each gap subject that participants reported to merit particular attention. Examination of both the key gaps and narrower questions suggests that more formal familiarisation with subjects such as pragmatics, sociolinguistics and other linguistic disciplines vis-à-vis FL use would give graduate FL professionals more informed analytical scaffolding for exploring the seven gap subjects. Furthermore, insofar as the gap subjects and narrower questions are identified as important needs in their own right, the study's results suggest that knowledge of linguistics and associated disciplines is occupationally relevant. While the study is small in scale, its results lend support to calls to better integrate linguistics subjects into FL learning, including within professional development and higher education.

Keywords: Needs analysis, Language for Specific Purposes, foreign language learning, curriculum design, professional development

INTRODUCTION

In recent years concerns have been voiced by language professionals in a number of UK government (hereafter HMG – ‘His Majesty’s Government’) departments that the foreign language (FL, L2) tasks assigned to them are sometimes inconsistent with the linguistic and cultural knowledge base formed during their pre-employment language learning (personal communications, p.c.). These language professionals, who typically have L1 English and have majored in one FL or more in university bachelor’s/Honours programmes, predominantly transcribe and/or translate FL material into English. They report that colleagues with little-to-no FL experience frequently assume that they possess greater authority regarding their language, culture and society of professional interest than is typically acquired during undergraduate studies (p.c.). Assumptions might be made, for instance, that subtle questions of inference, nuance and (in)directness are covered in detail when FLs are being learned when this might not be the case.

An illuminating example of such reported divergences in FL programme content and the knowledge base required in one workplace can be seen in the results of an informal survey at one HMG department in June 2019. A questionnaire asked 106 graduates of 27 UK universities whether their FL degree (major) programme had sufficiently prepared them for professional language work – typically transcription and translation from the FL into English – in the organisation.¹ Table 1 shows the results (p.c.). It is worth noting that these survey data do *not* form part of the needs analysis exercise on which the present article is based – the needs analysis described herein was conducted in 2017. Be that as it may, the 2019 survey results are instructive where reported mismatches are concerned. They suggest low levels of academic attention paid to, and a perceived lack of professional preparedness regarding, certain important subjects relevant to FL work in the given workplace. More specifically, they suggest low levels of instruction and assessment in language variation, linguistics, discourse analysis, professional translation, and language as used in occupational contexts across the relevant FL programmes as a whole. Indeed, only about 8% of teaching and 7% of testing was reported to address linguistics and discourse. If accurate, these figures, together with those for teaching language variation (5%), teaching professional approaches to translation and interpreting (8%), and the study of applied language skills in occupational contexts (0%) suggest that dedicated analysis of significant aspects of real-life language use and professional application were absent from or under-addressed in many UK FL programmes.

Table 1*Reported FL Degree Programme Coverage and Graduate Preparedness for Employment (N=106)*

Subjects	Approx. % proportion of teaching dedicated to subjects in programme	Approx. % of assessment dedicated to subjects	Approx. % of graduates who felt prepared for language work in the organisation
Speaking; understanding speech in the L2 across topics	20	19	51
Understanding written text	26	31	59
Linguistics and discourse	8	7	36
Variation in usage (e.g., register, dialect)	5	2	39
Culture (literature, film, history)	31	32	57
Professional approaches to translation and interpreting	8	9	46
Applied language in professional contexts (e.g., purchasing, law)	0	0	No figure given

Why such gaps or emphases in programme content existed (at least in the eyes of the survey participants) was not examined. There could be many reasons, and the study was more indicative than definitive. However, some considerations might well apply, singly or in combination. Some instructors might primarily associate employability with lexicogrammatical knowledge (Canning, 2009) and/or de-accentuate or overlook other competences integral to successful L2 communication (Amaya, 2008; Celce-Murcia, 2007; Correa, 2014; Roever, 2009; Sykes & Cohen, 2018). Additionally, some FL programme developers might not view linguistics (e.g., syntax, morphology, phonology, phonetics, semantics, pragmatics, as well as associated disciplines such as sociolinguistics, discourse analysis, translation and linguistic anthropology) as relevant (Correa, 2014; Hornsby, 2003) or sufficiently important. Some departments might choose not to offer linguistics within FL programmes through political (Coleman, 1988), systemic, resource or other constraints (Hyland, 2009; Miñana, 2017; Paesani, 2017). Gaps in knowledge or training might also prevent staff from adapting to evolving demands (Lomicka & Lord, 2018; Ryshina-Pankova & Byrnes, 2017). Finally, some instructors might rely heavily on more traditional textbooks (Maxim, 2004) or resist change (Lomicka & Lord, 2018; Maxim, 2004). Other factors might also obtain, of course.

Whatever the causal dynamics, the 2019 survey results pointed to perceptions that the respective university FL programmes as a whole paid insufficient attention to important questions of FL use such as linguistic variation and to related occupational practice such as professional approaches to translation, to name but two topics. These gaps in attention were correspondingly felt to have had implications for language professional preparedness for

undertaking FL tasks allocated in the workplace. This can be inferred through certain data in the last column of Table 1 – for instance, only 36% reported feeling prepared for tasks in respect of their knowledge of linguistics and discourse – although results for preparedness varied across subjects and course content was not specified beyond the topic level.²

The question that lies at the heart of the illustrative 2019 survey – of how aligned university FL programmes are with the needs of language professionals in a particular workplace – resonates strongly in this article. In particular, the article discusses the results of a small-scale needs analysis at one HMG department in 2017 to ascertain its language professionals' specific linguistic and cultural development needs (thus going into more detail than the 2019 survey, for example). It explores the reported disconnect between the skills, knowledge and competences required of the language professionals and those learned primarily in higher education, and why mismatches might occur. Finally, it considers potential implications for higher education and employers.³

THE STUDY: RESEARCH QUESTIONS

In 2017 an anonymous HMG department conducted a needs analysis to understand the professional development requirements of its language professionals. The questions posed were:

- Did pre-entry FL and cultural learning equip graduates/other newcomers sufficiently for FL work in the organisation?
- If not, what were the most important gaps that language professionals discerned in their own and others' knowledge?

These questions led to a third: How might any results be acted upon? The thinking was that if, when joining the organisation, FL majors had analytical repertoires that only *partly* matched the knowledge required in post, how might their needs be captured and acted upon? Here *acted upon* meant: (a) the potential development of in-house or combined internal/external language instruction to bridge gaps; and (b) contributing to any UK debate concerning the language and culture supply chain, particularly from higher education to employment.

METHODS, PROCEDURE, AND PARTICIPANTS

The Value of Needs Analysis

Needs encompass “learners’ goals and backgrounds, their language proficiencies, their reasons for taking the course, their teaching and learning preferences, the situations they will need to communicate in, and the genres most often employed in them ... [they] can involve what learners know, don’t know, or want to know” (Hyland, 2009, p. 204). Needs can be complex, multifaceted and emergent (Bocanegra-Valle, 2015; Coleman, 1988; Hyland, 2009) and conceptually conflicting (Liu et al., 2011).

Needs analysis is a vital component of curriculum development. It can entail several methods, including questionnaire surveys, student learning journals, interviews and participant observation (for more see Bocanegra-Valle (2015); Brown (2009); Long (2005); Serafini et al. (2015)). Once learner needs are determined, appropriate objectives can be identified and a syllabus designed. The identification of teaching and testing approaches, development of teaching materials and evaluation of a “defensible curriculum” (Brown, 2009, p. 269) – one that is valid, relevant and meets learner and instructor needs – can then follow (Brown, 2009; Hyland, 2009; Long, 2005; Trace et al., 2015).

Needs analysis can support different levels and types of provision. It can inform required language capacity at national or societal levels (Brecht & Rivers, 2005); the needs of specific employers (e.g., Lehtonen & Karjalainen, 2008; Serafini et al., 2015); or improvement in university language courses (e.g., Marina et al., 2019; Zakaria et al., 2017). English for Specific Purposes (ESP) is undergirded by a strong needs analysis literature that includes occupationally-directed learning – for example, English for business, healthcare, science and technology, education, and the hospitality sector, among others (Serafini et al., 2015). Although the corresponding literature set is weaker for languages other than English, its research base has nonetheless grown (for more see Hyland, 2009 and Trace et al., 2015).

Procedure and Participants

To determine the gaps in language professionals’ knowledge bases, 36 language professionals from a single HMG department were interviewed as part of a formal needs analysis. To work in the organisation in question, the language professionals had to possess a bachelor’s/Honours degree in a FL (or equivalent experience) and pass entry tests in transcription and translation into a high standard of English. Once employed, they translated and/or transcribed from one or more L2 into English: some individuals worked with more than one L2, though it was more common to work with one L2 only. Among the subjects addressed in work tasks were serious crime, terrorism and military topics. The department made extensive use of its FL and culture capability. Because this capability was consistently in demand, time was limited for engaging participants for research purposes. Therefore, a multistage procedure involving different contributors was undertaken involving participant observation, literature review and group interviews.

Stage One: Background research was conducted to define the professionals’ linguistic and cultural learning needs and work context. This involved (a) direct observation informed by the author’s 15+ years of performing L2 (Russian) language tasks in the department – translating, transcribing, training, and quality checking others’ language outputs (i.e., checking the accuracy of others’ transcripts and translations); (b) direct experience (5+ years) of setting tasks and assessing performance levels required of language professionals on the organisation’s behalf; and (c) reviewing documentation outlining language role descriptions, including types of language work per role and expected proficiency (e.g., some professionals translated from an L2 while others checked those outputs). The review was augmented by paperwork describing the gradated proficiency levels assigned to language professionals and how associated assessments were undertaken per criteria defined by the department’s own language community.

On the basis of this research, six broad gaps were identified as areas where language professionals lacked key conceptual scaffolding to support their L2 work: structured knowledge of discourse analysis, stylistic understanding (including non-standard varieties, stylistic appropriateness), slang (as social dialect/practice), swearing/dysphemism and euphemism, humour, and the intersection of language and culture. In each case and in combination, the assessed learner needs had strong pragmatic and sociolinguistic dimensions – for instance, what L2 communicants look to accomplish situationally through swearing or humour. The definition of “language and culture” was intentionally broad and included sociopragmatic and pragmalinguistic differences in communication. While the six gaps represented broad subject areas, it was determined that subsequent focus group examination would pinpoint more precise focal points for deliberation, should others agree about the gaps.

Stage Two: A 60-minute unstructured group interview was held with a convenience sample of five Russian language leads after all senior Russian language professionals had received an email seeking volunteers. The informants were known to the researcher and one another, and constituted ‘expert insiders’ (Long, 2005, p. 22) able to introspect about typical linguistic and cultural learning needs and their implications for the resultant product (transcripts and translations). Each had 10-30+ years of professional experience of producing transcripts and translations in the department; working with relevant L2 discourse genres; training and mentoring; and quality checking others’ work. None had mother tongue (L1) Russian; however, all approximated to Level 4 on the Interagency Round Table (ILR) framework for translation and Level 3+ or 4 in audio translation (transcription).

The researcher encouraged open and spontaneous discussion to understand what the informants saw as key knowledge gaps. Notes were taken contemporaneously and subsequently assessed. Unaware of the researcher’s initial conclusions, the participants determined that pre-entry learning brought significant gaps – the six topics identified in Stage One, less euphemism. However, they added another: understanding multilingualism. None had formally learned discourse analysis – i.e. “the close reading of *actual use of language along with other multimodal resources* for the purposes of dissecting its structures and devising its meanings” (Waring, 2018, p. 9; original emphasis) – although all agreed on its methodological value when it was raised and explained.

Stage Three: Time constraints precluded repeating Stage Two across analogous language communities. Consequently, a focus group was formed of ten lead language professionals representing different language communities across the department. The communities were: East Asian languages (Mandarin Chinese and Korean), African languages, Iranian languages,⁴ Arabic dialects and Indo-Aryan languages.⁵ These insider ‘domain experts’ (Long, 2005, p. 27) constituted a convenience sample, having responded to an email seeking senior-level volunteers.

The communities represented varied in composition. Some had small numbers of native speakers of their FLs of interest; however, most language professionals across the communities either held a bachelor’s/Honours degree in their L2(s), or – less commonly – learned them *ab initio* through private sector and/or in-house instruction. It was recognised that, where language communities had such L1-L2 diversity, the research question on whether pre-entry learning equipped language professionals for FL work in the organisation could elicit diverging insights. At the same time, however, there was also room for convergence, such as language professionals

lacking structured knowledge of disciplines such as discourse analysis, regardless of their background.

The Stage Three focus group examined the Russianists' determinations from Stage Two as part of an incremental analytical approach (Brown, 2009; see also Coleman, 1988, and Zhu & Flaitz, 2005, on sequencing focus groups and remits). Participants were free to confirm, challenge, amend, reject and/or replace the Russianists' findings in two 60-minute sessions. Discussion showed high levels of participant engagement and was facilitated by the same moderator as Stage Two (the author). The moderator's main role was to present and – where needed, explain – the Stage Two (Russianists') findings; encourage open deliberation and exchange; and probe for clarification. Notes were taken both contemporaneously and post-interview. The Stage Three group confirmed the Russianists' gaps and corresponding needs as important, valid and relevant to language communities department-wide.⁶

Stage Four: Four focus groups were arranged to elaborate the needs confirmed in Stage Three. They comprised convenience samples of representatives from four language communities – Russian (11 participants), Indo-Aryan languages (4), Iranian languages (4) and Chinese (3). The last two were selected partly due to their greater use of *ab initio* learners and the Indo-Aryan group for its mix of L1 and L2 speakers. Each informant volunteered to participate after a general call to their respective language communities, with one individual (in the Indo-Aryan group) previously involved in Stage Three. Unlike Stage Three, this phase comprised participants with greater variation in professional experience and competence, including staff that had completed university or *ab initio* instruction 1-2 years previously. It also involved two native speakers in the Indo-Aryan group. Interviewee competence spanned ILR Levels 3 to 4+ in translation and 2+/3 to 4+ in audio translation. All focus groups were set the same questions, and sessions – conducted over three months – lasted 60-90 minutes in total per group. Discussion was open and enthusiastic and led by a moderator. Notes were taken contemporaneously and *post hoc*. The more elaborated needs identified by the Stage Four groups are outlined in Results.

Merits and Demerits of the Data Collection Method

The use of participant observation and focus group interviews meant that the data collection methodology was qualitative. This presented some methodological challenges. The first concerned consistency in the definition of perceived needs. Learners and their needs are not homogenous (Coleman, 1988). Needs “mean different things to different participants, carry marked political implications” and are influenced by stakeholders' teaching and learning beliefs and agendas (Hyland, 2009, p. 204). It is not unknown for some studies to point, for example, to divergences about desired degree programme content between faculty (Marina et al., 2019; Maxim, 2004), while some students (Lenard & Pintarič, 2018) and educators (Canning, 2004) have judged professional relevance differently from employers.

Secondly, informants' views of their own capability entailed self-assessment. If uniquely subjective, self-assessment can occasion questions about data reliability. Its accuracy can be influenced by factors such as personality traits, learning style, previous learning experience, L2 proficiency and skill assessed (e.g., Birjandi & Bolghari, 2015; Blanche, 1988; Brantmeier et al., 2012).

Finally, a lack of quantitative data might limit representativeness and validity. Two of the strengths of questionnaire-based surveys, for example, are that they can enable larger cohorts to be surveyed and their uniformity can bolster reliability (Liu et al., 2011; Long, 2005). Similarly precluded were other methods such as the testing of departmental language professionals to acquire statistical data for error analysis.

Despite these challenges, however, it was decided that the limitations of a qualitative-only approach could be satisfactorily mitigated:

- Firstly, regardless of their L2(s), each participant typically undertook the same type of transcription/translation duties and was accredited a proficiency level based on the same overarching departmental assessment criteria by multiple third parties.⁷ This helped to ensure cohort consistency and validity in the identification of local and/or wider needs;
- Secondly, interviewing – sometimes accompanied by other methods such as observation – has been successfully applied in other studies of learner motivation and curricular relevance (e.g., Galishnikova, 2014; Kember et al. 2008; Talif and Noor, 2009);
- Thirdly, some research suggests that advanced L2 learners *can* produce accurate self-assessment, and if anything might tend to underestimate their capability (Blanche, 1988; Brantmeier et al., 2012);
- Fourthly, a combined data collection approach offered triangulation of sources and methods (e.g., Bocanegra-Valle, 2015; Brown, 2009; Long, 2005; Serafini et al., 2015).

No less importantly, the participation of insider experts at each stage was also integral to establishing valid and relevant needs (Long, 2005). Stage Two and Three participants in particular had typically reached ILR Level 4 or 4+ in translation, 3+, 4 or 4+ in audio translation (sometimes in multiple L2s) and enjoyed 10+ years' experience of FL work in the organisation: they knew their communities' patterns of language use (Coleman, 1988), work contexts and needs. Moreover, they had considerable experience of in-house language training and knew the professional development and other challenges facing the department's language community more broadly. Their ability to identify and situate individual and shared community needs within larger landscapes as informed participant observers in their own right was pivotal.

RESULTS

Moderators reviewed the Stage Four data independently and then jointly to identify individual and recurring themes. Focus group responses are outlined in Tables 2-7.

Stylistic Understanding

Table 2 shows the wide-ranging questions that arose when participants considered questions of stylistics, including stylistic strength, formality and register, in their respective languages.

Table 2

Perceived Needs – Stylistic Understanding

Language/ Language Family	Comments
Russian	<ul style="list-style-type: none"> What are the differences in colloquial Russian, <i>prostorechie</i> (a socially stigmatised demotic variety that indexes speakers as uneducated or uncouth) and swearing?
Mandarin Chinese	<ul style="list-style-type: none"> Differences between official vs. unofficial, formal vs. informal Mandarin: only some levels of style understood by participants Familiarity with genre (structures, registers); what is typical or missing? Colloquial language: recognising its categorisation, form and function How to understand informal language of all types to interpret interpersonal politics and relationships Perceptions of Chinese accent/dialect (non-Beijing/Shanghai) The relationship between style and power Understanding choices stylistically (e.g., implicit meaning, dropped subjects)
Iranian	<ul style="list-style-type: none"> Why educated speakers use colloquial language How to better understand differences in register and what this says about speakers and their aims How and why communicants accommodate to others stylistically
Indo-Aryan	<ul style="list-style-type: none"> Understanding the use of swearing and euphemism from a stylistic perspective Conventions regarding language and appropriateness

Needs largely concerned understanding differences between formal and informal language, and between varieties of the latter, chiefly because questions of register and variation were previously un- or under-addressed. The Chinese group understood high and formal literary registers such as those present in classical Chinese, official documents and newspapers. This meant, however, a weaker grasp of vernacular usage found online or in speech. The Russianists had difficulty understanding differences between some non-standard varieties, as did the Indo-Aryan and Iranian groups. For example, only Standard Pashto had been taught in courses, which were *ab initio* and somewhat limited in stylistic scope. The question of language ideologies was also apparent insofar as some participants associated the use of (normative) standard varieties with educated L2 speakers; non-observance of standard rules correspondingly pointed to

educational deficit. This was the case with the Iranian group, for instance, who saw a link between Standard Pashto and education. Pashto speakers who used a lower, “bastardised form of the standard” with different grammatical and syntactic characteristics were believed to lack in education.

The Iranian and Chinese contributors effectively raised questions of language variation in wanting to better comprehend colloquial usage, including by educated speakers, and accommodation to speakers of other varieties (Iranian). These, along with other observations, effectively foregrounded sociolinguistic matters in this section. Finally, the Chinese group’s need to understand how language can convey authority related to notions of power and rhetoric.

Slang

Table 3 outlines the points considered when interviewees explored the subject of slang in their language(s) of professional interest. Questions included motivations and social dimensions of use, among others.

Table 3

Perceived Needs – Slang

Language/ Language Family	Comments
Russian	<ul style="list-style-type: none"> • What slang is currently used? • How do speakers use slang in different groups/contexts? • Does slang change through time? • Is slang about snubbing the mainstream? • How to translate jargon into appropriate English, e.g., in technical translation (e.g., legalese, official registers)
Mandarin Chinese	<ul style="list-style-type: none"> • Countercultural usage: why and how is it created? • Does slang involve appropriation? • How to understand Internet Chinese
Iranian	<ul style="list-style-type: none"> • How to understand what slang is and how it is manifested • Cryptolact: how is meaning hidden?
Indo-Aryan	<ul style="list-style-type: none"> • How to understand slang and why it is used <ul style="list-style-type: none"> • Identity? • Group status? • Caste/social stratification? • Why is there so much borrowing from other languages in slang? • How to understand dialects and their specificity: social, geographic

This subject had generally not been taught in previous instruction (all) – sometimes because L1 educators generally avoided it (Russian). If anything, informants had encountered it during residence abroad (Russian, Chinese, Indo-Aryan), especially in vernacular communication.

When it *was* taught, teaching was considered outdated (Russian). Either way, participants expressed difficulty in identifying it (some confusing social and geographic dialects), in interpreting usage (Internet Chinese) and in understanding its purpose (all groups). Notably, some Iranian language professionals claimed no great difficulty in understanding the semantic meaning of slang items in context, but still lacked a finer understanding of the subject, including in pragmatic terms.

Swearing/Dysphemism and Euphemism

Table 4 includes the questions raised by the groups with regard to dysphemism (mainly described in terms of swearing) and euphemism. Factors identified included, *inter alia*, illocutionary force, motivation, appropriateness and interpretation.

Table 4
Perceived Needs – Swearing/Dysphemism and Euphemism

Language/ Language Family	Comments
Russian	<ul style="list-style-type: none"> • What is swearing? • What links are there to taboo or other socially unacceptable behaviour? • When is swearing appropriate? • What impact does it/is it designed to have, if any? • Is there an in-group/out-group dimension to its use? • How to translate swearing: the function of the word/phrase and how to convey this
Mandarin Chinese	<ul style="list-style-type: none"> • How to evaluate swearing and emphatic force • When/in what context is swearing used? To express frustration? • What/who is the target of insults? • How to gauge level of insult • Euphemism: how to identify it, especially in colloquial Chinese • Why is euphemism used? <ul style="list-style-type: none"> • To avoid conflict? • To avoid showing feelings, keep or appear composed? • To test the water to get another's point of view?
Iranian	<ul style="list-style-type: none"> • What are the functions of swearing, and why do some people use it and others not? • Non-swearing dysphemism, especially references to animals: why does it occur? • Euphemism: why is there a lack of euphemism in some interactions?
Indo-Aryan	<ul style="list-style-type: none"> • What is the emotional aspect of swearing? <ul style="list-style-type: none"> • Is it used for emphasis and intensification? • How does swearing relate to taboo in Punjabi? • What are the gender and cultural aspects of swearing?

-
- Women as targets of swearing
 - Perceptions of (young) women swearing in Punjabi
 - What cross-cultural differences are there in women swearing in Punjabi vs. mainstream British cultures?
 - Euphemism: its functions
 - Talking about sex: which euphemisms are used?
-

These subjects had similarly not featured in previous learning, mainly for sociocultural reasons (all). Teachers of Iranian languages, for example, were guarded and uncomfortable talking about expletives: classroom discussion was considered inappropriate, in part due to graphic and/or sexual dimensions. The same tensions applied to Russian, Chinese and Indo-Aryan instruction. Consequently, the participants had encountered dysphemism, including swearing, *ad hoc*: in personal encounters, in-country immersion, and watching television (Russian, Chinese). So, while it was fairly easy to identify swearing semantically, it was difficult to understand why expletives were used and what illocutionary force obtained. Cultural sensitivity gave L1 speakers of Indo-Aryan languages solid pragmatic anchoring, but for others across languages the matter was less clear-cut. What taboos and transgressions were involved (Russian, Indo-Aryan)? What did the use of swear words and phrases mean (Russian, Iranian, Indo-Aryan) and what was their impact (Russian, Chinese, Iranian)? Who was able to use them, with whom and when (all)? What differences were there in levels of insult (Chinese, Indo-Aryan)? What other forms of dysphemism were there (Iranian)? How did swearing and dysphemism compare cross-culturally (Russian, Indo-Aryan)?

Some informants stated that they were not entirely unsighted, however. Some participants working with Farsi felt able to instinctively understand from the context that someone was swearing and could determine why – this was working, admittedly, with a low level of intuition and no formal conceptual framework to guide interpretation.

Discussions also showed this subject interleaving with others. For instance, questions of multilingualism and language ideologies arose where Urdu was described as a more formal, respectful language, while Punjabi was a vernacular language of less educated, agricultural workers. As “the best language to swear in,” Punjabi enabled more emotive expression. Moreover, questions of gender came into scope in two respects: (a) where targeting women (e.g., mothers, wives) through swearing was considered especially insulting in Indo-Aryan cultural practice; and (b) the right to swear. Although swearing in Punjabi was taboo, men could do so. Such behaviour by young women was regarded, however, as not “virtuous”. Understanding speakers’ latitude to use taboo terms or other dysphemisms and when rules might be broken was important.

Euphemism was also considered relevant. Understanding its function, what topics it typically referenced (e.g., sex in Punjabi, death in Urdu) and how it was conveyed (understatement, borrowing) was salient for those working with Chinese, Iranian and Indo-Aryan languages.

Humour

Table 5 outlines questions raised by the participants in respect of humour. These included, among other things, communicant intention, interpersonal implications, and recognition of typical humour types and targets.

Table 5

Perceived Needs – Humour

Language/ Language Family	Comments
Russian	<ul style="list-style-type: none"> • What is Russian humour? • What is “different funny”? • What are communicants trying to achieve? <ul style="list-style-type: none"> • Does humour involve strategies to optimise social situations? • How is humour used to negotiate relationships? • Why and how are stereotypes used (e.g., racism)? • How to translate puns and language play
Mandarin Chinese	<ul style="list-style-type: none"> • How to understand meaning: it can seem clear that something is a joke or funny due to laughter, but what does the speaker actually mean and what makes it funny? • How to recognise and understand humour online • How to identify and understand humour through cultural allusion and the use of stereotypes
Iranian	<ul style="list-style-type: none"> • Why a speaker finds something funny • Who are the typical targets of humour/jokes? • How to understand levels of formality and relationships when humour is used • What are the interpersonal dynamics? <ul style="list-style-type: none"> • Identity? • Common values?
Indo-Aryan	<ul style="list-style-type: none"> • What kinds of humour are there? • Why is stating the obvious funny? • Role of antithesis in humour; why people often criticise one another and laugh? • How is humour used to build rapport? • What contextualisation cues are used for humour?

This subject too was not typically studied pre-employment in any dedicated way (Russian, Chinese, Iranian languages). Understanding humour is difficult for language professionals, including heritage speakers with limited contact with their background culture(s). The key needs here concerned how to understand cultural allusions; what made an utterance funny; and what

a social actor was trying to accomplish. The study of stereotypes, personal or social characteristics and typical targets was proposed by some informants as a possible gateway to better understanding (Russian, Chinese, Iranian), as was humour for social rapport.

Notably, the interpretation of humour sometimes rested on identification of laughter. Those working with Chinese and the languages of Iran felt able to discern when an interactant found something funny or amusing. However, different types of and conventions regarding laughter went unmentioned but undoubtedly require consideration.

The Intersection of Language and Culture

Table 6 shows the topics raised by the interviewees when addressing language and culture. The group identified a wide range of questions – for example, the role of religion in certain cultures, how language is used to reflect (and/or construct) culture and identity, recognising cultural practices, assumptions about the Other and rules of linguistic politeness.

Table 6
Perceived Needs – Language and Culture

Language/ Language Family	Comments
Russian	<ul style="list-style-type: none"> • What is culture? • Identifying different cultural environments • Understanding cultural narratives and perceptions • How to recognise and adopt Russian communicative cultural practices • What rules of interpretation are there in communication? <ul style="list-style-type: none"> • For instance, what does it mean if we believe Russian speakers to be shouting? • What are politeness rules? • What are the appropriate forms of address and history behind them? • What is the importance of kinship? • What can linguistic relativism tell us?
Mandarin Chinese	<ul style="list-style-type: none"> • More detail needed about written language and culture: how they correspond to social status, level of education, etc. • How, if at all, does the grammar and vocabulary of the language characterise its speakers? • How is language used to create and reflect identity? • Is language used to reflect or create imagined communities? • What links are there between language and political culture? • What resistance is there to Western practices in Chinese language and culture? • Identifying and interpreting stance in Chinese: are there cultural angles?
Iranian	<ul style="list-style-type: none"> • How to learn the cultural insights normally picked up during a visit/year abroad (for non-higher education <i>ab initio</i> students)

	<ul style="list-style-type: none"> • How/why do Pashto speakers try to avoid silence by phatic communion (e.g., greetings, even if this means repetition)? • How does language reflect the culture of individual speech communities? • What are the religious and ideological aspects of language use? • What are the gender dimensions of language use in society?
Indo-Aryan	<ul style="list-style-type: none"> • How is language used to create, present and reflect ideas of South Asian culture and ethnicity? • How is language used for social stratification? • How are identities performed, e.g., British Asian youth culture and use of slang? • Social identity: how are conformity and belonging represented linguistically by speakers of Indo-Aryan languages? • What assumptions do “host societies” have about smaller ethnic groups? • What are the use and perception of stereotypes based on? • How to understand essentialism and how language is used to reflect individual experience and heritage (the inaccurate homogenisation of “Indian culture”) • The merits of collectivist vs. individualist labels for understanding South Asian societies • The merits of Hofstede’s analysis for understanding South Asian cultures • The role of religion in South Asian cultures • Cryptolect within South Asian cultures: how and why?

No definition of culture was given to groups to avoid constraining discussion. However, certain trends are noteworthy. The first is that, in their definitions, informants often cited various facets of cultural practice encountered in pre-employment instruction. Informants often referred to *Big C* cultural practices such as the study of film, literature, music and history (Russian, Indo-Aryan), although they believed them to represent only a certain snapshot of cultural practice. Those who had learned Iranian languages *ab initio* claimed to have received no formal instruction in cultural studies. They were thus keen to capture the insights they believed to be assimilated during university immersion learning.

The Chinese group differed somewhat in their appreciation of linguistic and cultural cross-currents as they had studied how the written language conveyed social status, level of education, etc. while also constituting a form of cultural practice. They had also briefly explored how culture is manifested phraseologically in everyday Mandarin (e.g., asking how someone is – literally translated as “Have you eaten?”).

The question of gender arose. Those working with Iranian languages articulated a need to know more about gender in sociocultural practice and convention: certain societies were male-dominated and males were reported to have greater communicative latitude, including vis-à-vis swearing.

Chinese, Iranian and Indo-Aryan language participants reported a need to understand the function and role of stereotypes – in the Indo-Aryan case to explore host society assumptions

about South Asian cultures and raise awareness of the perils of essentialism. Heritage speakers in this group underlined the need to understand and challenge cultural bias in FL analysis, and to understand how native speakers situationally construct, project and negotiate multiple identities.

Comments across all seven subjects concerned culture in one respect or another, and related to questions of sociolinguistics and linguistic and cultural anthropology. Some examples are identity, agency and whether a language can characterise speakers (Indo-Aryan, Chinese); language and social stratification (Chinese, Indo-Aryan); gender (Iranian); ethnicity (Indo-Aryan); religion and ideology (Iranian, Indo-Aryan); different cultural environments or communities (Russian, Chinese, Iranian); and language, power and political culture (Chinese).

Understanding pragmatics also emerged as a need: in one case this meant knowing more about linguistic politeness (Russian); in others rhetoric and stance (Chinese) and interpreting silence (Iranian, Indo-Aryan). The participants similarly articulated a need to learn about assumptions that groups share and how these influence practice, including communication (Russian, Iranian). Finally, cultural aspects of the more conscious non-disclosure of information interested the Indo-Aryan and Iranian groups.

Multilingualism

Table 7 shows the questions identified by participants when considering multilingualism. Propositions concerning (lexical) borrowing also featured in this discussion. The motivations behind code-switching and borrowing featured especially prominently.

Table 7
Perceived Needs – Multilingualism (and borrowing)

Language/ Language Family	Comments
Russian	<ul style="list-style-type: none"> • None were identified
Mandarin Chinese	<ul style="list-style-type: none"> • Why loan words and ephemeralisms are used in Chinese • Multilingualism and code-switching: functions and forms
Iranian	<ul style="list-style-type: none"> • What are communicants trying to achieve through code-switching? <ul style="list-style-type: none"> • For example: code-switching involving Dari, Urdu and English with Pashto • What are the social statuses of different languages or dialects? • Why does borrowing occur, especially from English? • Why do communicants use synonyms from another language to clarify?
Indo-Aryan	<ul style="list-style-type: none"> • Why do communicants switch codes? <ul style="list-style-type: none"> • Identity? • Solidarity? • Specificity? • Nuance?

-
- Does a word from another language carry a specific meaning or sense that can only be conveyed by that word?
 - Why does borrowing occur?
 - To show level of education?
 - To show authority?
 - For prestige?
 - What is the role of religion in multilingual contexts?
-

This topic was sometimes conflated with borrowing. Although it did not elicit any requirements from the Stage Four Russianists (although many Russian speakers do live, work and socialise in multilingual environments and the Stage Two group had recommended the subject), it was raised by others – particularly the Iranian and Indo-Aryan groups. These participants needed a deeper appreciation of how social status and power can be signalled through variation in languages/dialects, and of the functions of code-switching and code-mixing more generally.

Lexical borrowing, particularly from English, was noted by the Iranian group, who pinpointed a need to comprehend why communicants used synonyms from other languages. The functional dimensions of borrowing were similarly mentioned by the Chinese and Indo-Aryan groups (e.g., the potential use of borrowed terms to indicate authority).

Discourse Analysis

When outlined to participants in general terms, all agreed on the need for formal study of the subject. None had a background in it. However, they appreciated the potential benefits of understanding Conversation Analysis (e.g., turn-taking, adjacency pairs, preference), Interactional Sociolinguistics (e.g., contextualisation cues) and Critical Discourse Analysis (e.g., the relationship between language and power and the maintenance of ideologies).

DISCUSSION AND IMPLICATIONS

Let's recall the study's research questions:

- Did pre-entry FL and cultural learning equip graduates/other newcomers sufficiently for FL work in the organisation?
- If not, what were the most important gaps that language professionals discerned in their own and others' knowledge?
- How might any results be acted upon?

Across Stages Two to Four the answer to the first research question was consistent: pre-entry learning provided a solid enough basis for language professionals to secure employment through passing transcription and translation tests; however, they reported feeling underprepared for many aspects of their jobs.

Concerning the second question regarding learner perceptions of key knowledge gaps, these are outlined in Tables 2-7.

Regarding acting on the results, the following aspects were examined:

- Application of the results across languages and roles;
- Implications for educators;
- Course development to address the gaps (not examined here; see Davie, 2022).

Application of the Results across Languages and Roles

Although the informants had different pre-employment learning experiences through inevitable variations in curricula and life paths, the more precise gaps identified in Stage Four showed consistencies in two respects: in the recurrence of specific questions (e.g., the role of stereotypes for understanding culture), and in the importance of comprehending:

- a. what communicants are seeking to achieve;
- b. how power and social distance are conveyed and interpreted;
- c. the relationship between language variables and social groups;
- d. how social actors build and negotiate relationships;
- e. the expression and negotiation of identities;
- f. 'non-standard' language practices: e.g., ideologies, linguistic form, social and pragmatic meaning;
- g. pragmatic force;
- h. others' sociocultural conventions and how these influence, are shaped by and are reflected in discourse;
- i. cross-cultural and -linguistic comparison and implications for translation.

Subjects a-i suggest that the language professionals would benefit from a structured introduction to disciplines such as pragmatics, sociolinguistics and discourse analysis to more profitably investigate the seven overarching topics (slang, multilingualism, etc.) and the associated narrower questions identified in Stage Four (shown in Tables 2-7). Discrete questions such as what slang is and how it might feature in social practice; when, why and with whom interactants might swear and so on can be explored as self-standing questions. However, effective examination inevitably draws discussants to fields such as pragmatics and sociolinguistics for conceptual anchoring and to better understand contexts of use. One can envisage where discussions might throw up concepts such as speech acts, speech aims, overt/covert prestige, identification and language variation, for example. A language-specific or -agnostic course on such subjects would certainly help to provide valuable scaffolding for better understanding the identified gaps.

There is also the question of whether such learning would benefit other HMG language professionals and those outside government. Insofar as the informants applied receptive skills in translating and transcribing from different L2s, their roles differed from those of HMG language professionals using receptive *and* productive skills (speaking, writing) – for example, diplomats who produce HMG messaging for embassy social media accounts, meet foreign officials, participate in media interviews, or engage social groups in an L2. Equally, the informants' job profiles diverged from those requiring productive skills in the private sector.

This distinction, however, is only one factor when comparing across professions. Investigating the seven gap topics (multilingualism, etc.), not to mention the propositions in a-i through disciplines such as sociolinguistics, can benefit other roles. Knowing how identities are manifested in discourse, for instance, is relevant when negotiating social and business relationships, as is recognising linguistic manifestations of power and social distance. Equally, knowing how to compare one's own assumptions and sociocultural norms with those of others is also essential (Lehtonen & Karjalainen, 2008), especially as grammatical or lexical errors may lead to fewer adverse consequences in intercultural communication than failures in politeness (Rathmayr, 2008; see also Celce-Murcia (2007)).

Implications for Educators

Several considerations merit attention by educators. The **first** concerns occupational relevance. Tertiary-level FL programmes look to cater for learners with different motivations (Huhtala et al., 2019) and different goals (social, professional, intellectual, etc.). In all these cases relevance is important. In their study of professional programmes, Kember et al. (2008) report that demonstrating the relevance of abstract theories to the real world helps students to understand them; lends authenticity; stimulates; and motivates. The same principle clearly applies to FL learning, both general and occupation-specific.

However, while university departments, national agencies and others in the UK have been increasingly motivated "... to respond to changes in contemporary society and the changing needs of learners, with the aim of making languages more attractive, more accessible and more relevant" (*Quality Assurance Agency for UK Higher Education Subject Benchmark Statement for Languages, Cultures and Societies* (hereafter 'QAA'), 2019, pp. 3-4), the present study suggests that FL major graduates in one organisation felt underprepared for important aspects of L2 work there.

Disconnects between higher education and the workplace are not new, and there have been calls for universities to better attune FL learning to employment needs (e.g., Lancereau-Forster, 2015; Lehtonen & Karjalainen, 2008; Marina et al., 2019) – in many cases complementing more general L2 learning. Several benefits have been advanced, including better matching skills and learning to occupational activities such as securing employment (Enkin & Correa, 2018; Marina et al., 2019), navigating informal workplace exchanges (Holmes, 2005; Kaufmann & Grünhage-Monetti, 2003), and dealing with demanding situations (Lehtonen & Karjalainen, 2008). Be that as it may, the present study suggests a need for reflection regarding how far FL departments have taken their work "... out into the world and test[ed] it against the pressing issues of our time" (Miñana, 2017, p. 421) and to what extent FL degree outcomes have become "tangible and more than vague 'global citizenship'" (Languages Forum, 2010). Many universities may insist that they already demonstrate relevance, among other outcomes, in majors programmes at bachelor's/Honours level. However, together with the informal 2019 survey (Table 1), the present needs analysis raises questions about how far initiatives to align the professional and academic have come and with what results, or at least how *explicitly* undergraduate FL programme content is linked to and driven by application in the workplace.

A **second** consideration concerns the intimation that FL majors know about and understand the structures, registers and, where appropriate, varieties of the L2, as well as the

linguistic principles necessary for L2 work (QAA, 2019, p. 19). This aim aligns to some extent with Wyburd (2011, p. 5), who suggests that UK FL graduates develop “heightened literacy, textual analysis and oracy in [their] mother tongue” by examining grammatical, syntactic and stylistic accuracy, style and register in their L2 through courses involving written language and translation, linguistics, and examination of a wide range of L2 texts (e.g., historical, literary).

While this multifaceted aim is certainly appropriate, the needs analysis suggests that the provision of more formal translation studies, stylistic analysis or linguistics options – e.g., syntax, morphology, phonology, phonetics, semantics, pragmatics, as well as associated disciplines such as sociolinguistics, stylistics, discourse analysis and linguistic anthropology – in UK undergraduate FL programmes is patchy (see also the informal 2019 survey (Table 1)). Joint degree programmes in FLs and linguistics could well be an exception, of course, depending on their content.

Similarly, graduates’ ability to “delve beneath surface meaning and to interpret intention” (Wyburd, 2011, p. 7), in line with the QAA’s (2019, p. 18) statement that students should be able to “recognise implicit meaning”, also requires closer inspection. There are undoubtedly cases where reference works, online data and knowledge of intertextuality and metaphor, for example, can aid interpretation here. However, given the HMG respondents’ stated needs, one must question how much undergraduates’ L2 metapragmatic awareness is *explicitly* developed (Amaya, 2008; Roever, 2009; Sykes & Cohen, 2018) and how far the QAA statement on implicit meaning actually reaches. The HMG interviewees sought enhanced pragmatic awareness; they saw “language as more than a compilation of the prescriptive grammatical rules that can be found in traditional grammar texts” (Correa, 2014, p. 169). However, their forays into the universe of implicature, inference and indirectness that lies behind the notion of implicit meaning to which the QAA refers appear somewhat limited. The same might be said of discourse analysis. While Wyburd (2011, p. 4) identifies the value of literary, cultural, historical and sociological courses for engendering critical textual analysis, no mention is made of discourse analysis, a centrally important discipline for textual analysis which both aids and leans on the building of metapragmatic ability and pragmatic capacity in an L2 (Amaya, 2008; Roever, 2009; Sykes & Cohen, 2018). Fostering greater awareness of these subjects would certainly help to meet the needs articulated by the given HMG language professionals.

A **final** implication concerns the purpose of universities. How does enhancing the professional relevance of FL programmes relate to the role and objectives of academia? Some might read any substantial recalibration of FL learning towards greater occupational relevance as jeopardising the intellectual and humanistic aims of academic departments or institutions (for more on this discussion, see Doyle (2019); Maxim (2004); Teichler (2016)). In particular, they may perceive a threat to literature and culture courses, and/or to critical thinking. The HMG interviewees, however, neither sought the supplanting of such aims, nor perceived the relevance of university FL programmes in zero-sum terms. If anything, they saw cultural competence as relevant to their professional endeavours. Analysing culture and semiosis in global, local and *glocal* contexts could certainly help to meet their professional needs.

However, as was intimated in the Introduction, to pursue humanistic and intellectual goals while demonstrating professional relevance requires effective communication. The UK’s University Council of Modern Languages (UCML) and employers have long emphasised the need for collaboration to better align undergraduate FL programmes with workplace demands (e.g., British Academy, 2013; CBI/Pearson Education and Skills Survey Report 2019). Options suggested

by the UCML for cultivating productive relationships with employers (and others outside the academy) include, *inter alia*, reciprocal visits to enable staff and students to better appreciate the importance of FLs *and* intercultural knowledge in occupational settings, and employer involvement in guiding curriculum development and/or review (Collis, 2011). Inasmuch as universities would benefit from such cooperation – for example, it might help some educators to recalibrate their approaches and provision as part of ongoing professional development⁸ – employers would gain by better appreciating the typical limitations of tertiary-level FL learning and, consequently, the imperative of continuous professional development. Cooperation of this kind might therefore nurture and fine-tune the linguistic and cultural agility that employers require (Cambridge Public Policy Strategic Research Initiative, 2015) and help managers/employers to develop more realistic expectations of employees, where those are poorly understood.

CONCLUSION

Using combined qualitative methods, this study sheds light on knowledge gaps and learning needs that language professionals at one HMG department reported to have arisen due to omissions in pre-entry learning – typically undergraduate FL major programmes. Filling these gaps was considered important for their professional purposes. The gaps identified were: stylistic understanding; slang; dysphemism/swearing and euphemism; humour; the intersection of language and culture; multilingualism; and discourse analysis. Deeper exploration established a more precise set of requirements for each gap subject, but also indirectly illuminated underlying sociolinguistic, pragmatic and other questions that, if investigated, would aid professionals' understanding and analysis. A conclusion to emerge from the analysis was that a fundamental knowledge of linguistics and related disciplines could help to explore competences beyond the linguistic (e.g., Celce-Murcia, 2007) and provide important scaffolding for more informed and productive analysis of the seven gap subjects.

Ultimately, the study represents a single investigation involving 36 language professionals undertaking translation and transcription tasks at an individual UK civil service department. Its results should accordingly be understood within the methodological, institutional and occupational contexts within which they were obtained. For example, in view of constraints on contact time with participants, the present study by necessity aimed to determine fairly general trends. What patterns or results would more surgically targeted investigations yield? Furthermore, would a needs analysis involving similar cohorts in other organisations record the same gaps? And would it raise any questions regarding the alignment of the professional and academic and, if so, with the same focal points (e.g., promoting more formal examination of sociolinguistics, more actively fostering explicit pragmatic awareness)?

Nonetheless, even if representing a single, limited endeavour, it is hoped that the present study provides helpful and relevant insights for colleagues insofar as it illuminates needs that a given professional community believe went unmet in pre-employment FL provision; lends a basis for potential remedies to be further researched and developed, locally or otherwise; and adds weight to calls for greater coordination between employers and instructors, particularly in higher education, to enhance preparedness for the workplace.

ACKNOWLEDGMENTS

I would like to thank all needs analysis participants, those who provided helpful feedback on earlier drafts of this article and the article's anonymous reviewers. I am also grateful to those who facilitated the study and provided consent and ethical review for the research, and (particularly) to my fellow moderator for her valuable contribution.

NOTES

1. The survey was conducted to obtain an indication of general trends only. Most respondents had a bachelor's/Honours degree. A small number (quantity unknown) had a post-graduate degree, typically a Masters in Translation and/or Interpreting.
2. For deeper discussion of curriculum content and professional relevance see, for example, Doyle (2019), Ruggiero (2014) and Teichler (2016).
3. For reasons of confidentiality, this article does not provide specific examples of the language professionals' work or of L2 materials featuring in their assignments.
4. This group of languages in the Indo-Iranian sub-family includes Farsi, Dari, Tajik, Kurdish, Balochi and Pashto.
5. This is a language group in the Indo-Iranian sub-family. It includes languages such as Hindi, Urdu, Punjabi, Bengali and Gujarati, among others.
6. The needs corresponded to Hutchinson and Waters' (1987) target needs as *necessities*, *wants* and *lacks* (cited in Bocanegra-Valle, 2015, p. 69).
7. Although this process itself was essentially qualitative, the overarching framework and associated assessments were routinely reviewed to optimise and ensure applicability across the organisation's greater language community.
8. I am grateful to one of the journal's anonymous reviewers for this helpful observation.

REFERENCES

- Amaya, L. F. (2008). Teaching culture: Is it possible to avoid pragmatic failure? *Revista Alicantina de Estudios Ingleses*, 21, 11–24.
- Birjandi, P., & Bolghari, M. S. (2015). The relationship between the accuracy of self- and peer-assessment of Iranian intermediate EFL learners and their learning styles. *Theory and Practice in Language Studies*, 5(5), 996–1006.
- Blanche, P. (1988). Self-assessment of foreign language skills: Implications for teachers and

- researchers. *RELC Journal*, 19(1), 75–93.
- Bocanegra-Valle, A. (2015). Foreign language learning needs in higher education: Reasons for convergence and accountability. *Revista de Lenguas para Fines Específicos*, 21(1), 67–87.
- Brantmeier, C., Vanderplank, R., & Strube, M. (2012). What about me? Individual self-assessment by skill and level of language instruction. *System*, 40, 144–160.
- Brecht, R. D., & Rivers, W. P. (2005). Language needs analysis at the societal level. In M. H. Long (Ed.), *Second Language Needs Analysis* (pp. 79–104). Cambridge University Press.
- British Academy. (2013). *Lost for Words: The need for languages in UK diplomacy and security*. The British Academy.
- Brown, J. D. (2009). Foreign and second language needs analysis. In M. H. Long & C. Doughty (Eds.), *The handbook of language teaching* (pp. 269–293). Blackwell.
- Cambridge Public Policy Strategic Research Initiative. (2015). *The value of languages. Ideas for a UK strategy for languages*. Cambridge University Press.
- Canning, J. (2004). Enhancing employability: A guide for teaching staff in languages, linguistics and area studies. *LLAS Centre for Languages, Linguistics and Area Studies*. Retrieved September 27, 2020 from <https://www.llas.ac.uk/resources/paper/2124.html>
- Canning, J. (2009). A skill or a discipline? An examination of employability and the study of modern foreign languages. *Journal of Employability and the Humanities*, 3, 1–12.
- CBI/Pearson Education and Skills Survey Report. (2019). *Getting Young People 'Work Ready'*.
- Celce-Murcia, M. (2007). Rethinking the role of communicative competence in language teaching. In E. A. Soler & M. P. Safont Jordà (Eds.), *Intercultural language use and language learning* (pp. 41–57). Springer.
- Coleman, H. (1988). Analysing language needs in large organizations. *English for Specific Purposes*, 7, 155–169.
- Collis, J. (2011). *Expanding Horizons: Connecting with professional organisations and external partners toolkit*. University Council of Modern Languages.
- Correa, M. (2014). Teaching (theoretical) linguistics in the second language classroom: Beyond language improvement. *Porta Linguarum*, 22, 161–171.
- Davie, J. (2022). The workplace development of language professionals after university: A United Kingdom case study. *Global Business Languages*, 22, 55–75. <https://doi.org/10.4079/gbl.v22.5>
- Doyle, M. S. (2019). Curriculum development activism (CDA): Moving forward with Spanish for the Professions and Specific Purposes (SPSP). *Hispania*, 102(4), 475–480.
- Enkin, E., & Correa, M. (2018) Evaluating learner and teacher perceptions of program outcomes in the foreign language major. *Electronic Journal of Foreign Language Teaching*, 15(1), 66–80.
- Galishnikova, E. M. (2014). Language learning motivation: A look at the additional program. *Procedia – Social and Behavioral Sciences*, 152, 1137–1142.
- Holmes, J. (2005). When small talk is a big deal: Sociolinguistic challenges in the workplace. In M. H. Long (Ed.), *Second Language Needs Analysis* (pp. 344–372). Cambridge University Press.
- Hornsby, D. (2003). Starting from scratch: French linguistics courses at Kent. *LLAS Occasional Papers*. Centre for Languages, Linguistics and Area Studies, University of Southampton. Retrieved August 24, 2020 from

- <https://www.llas.ac.uk/sites/default/files/nodes/179/papersoct03.pdf>
- Huhtala, A., Kursiša, A., & Vesalainen, M. (2019). "This language still motivates me!" Advanced language students and their L2 motivation. *Studies in Second Language Learning and Teaching*, 9(2), 287–311.
- Hyland, K. (2009). Specific purpose programs. In M. H. Long & C. Doughty (Eds.), *The Handbook of Language Teaching* (pp. 201–217). Blackwell.
- Interagency Language Round Table (2020). Retrieved September 29, 2020 from <https://www.govtilr.org/>
- Kaufmann, S., & Grünhage-Monetti, M. (2003). Communication needs analysis. In M. Grünhage-Monetti, E. Halewijn, & C. Holland (Eds.), *Odysseus – Second language at the workplace* (pp. 34–42). European Centre for Modern Languages: Council of Europe.
- Kember, D., Ho, A., & Hong, C. (2008). The importance of establishing relevance in motivating student learning. *Active Learning in Higher Education*, 9, 249–263.
- Lancereau-Forster, N. (2015). An observatory for foreign language usage in the workplace. *CLEaR2014: Conference Proceedings*. Nitra: SlovakEdu. 5–19.
- Languages Forum. (2010). First meeting at BIS, 17th March. Retrieved November 20, 2020, from <http://www.aulc.org/wp-content/uploads/2017/10/language-forum-minutes-march-2010.pdf>
- Lehtonen, T., & Karjalainen, S. (2008). University graduates' workplace language needs as perceived by employers. *System*, 36, 492–503.
- Lenard, D. B., & Pintarić, L. (2018). Comparison of employers' and students' perceptions regarding communication skills. *The Journal of Teaching English for Specific and Academic Purposes*, 6(1), 63–82.
- Liu, J., Chang, Yu., Yang, F., & Sun, Yu. (2011). Is what I need what I want? Reconceptualising college students' needs in English courses for general and specific/academic purposes. *Journal of English for Academic Purposes*, 10, 271–280.
- Lomicka, L., & Lord, G. (2018). Ten years after the MLA report: what has changed in foreign language departments? *ADFL Bulletin*, 44(2), 116–120.
- Long, M. H. (2005). Methodological issues in learner needs analysis. In M. H. Long (Ed.), *Second language needs analysis* (pp. 19–76). Cambridge University Press.
- Marina, O., Yakusheva, I., & Demchenkova, O. (2019). Examining undergraduate students' and in-service graduates' perceptions of their professionally oriented foreign language needs. *Journal of Language and Education*, 5(1), 63–84.
- Maxim, H. H. (2004). Expanding visions for collegiate advanced foreign language learning. In H. Byrnes & H. H. Maxim (Eds.), *Advanced foreign language learning: A challenge to college programs* (pp. 180–193). Heinle Thomson.
- Miñana, R. (2017). Making change happen: The new mission and location of language departments. *The Modern Language Journal* 101, 413–423.
- Paesani, K. (2017). Redesigning an introductory language program: A backward design approach. *L2 Journal*, 9(1), 1–20.
- Quality Assurance Agency for UK Higher Education. (2019). *Subject benchmark statement languages, cultures and societies*. QAA.
- Rathmayr, R. (2008). Intercultural aspects of new Russian politeness. *WU Online Papers in International Business Communication, Series One: "Intercultural communication and*

- language learning*" 4. Vienna University of Economics and Business Administration.
- Roever, C. (2009). Teaching and testing pragmatics. In M. H. Long & C. Doughty (Eds.), *The handbook of language teaching* (pp. 560–577). Blackwell.
- Ruggiero, D. M. (2014). Graduate courses in languages for specific purposes: Needs, challenges, and models. *Global Business Languages*, 19, 55–69.
- Ryshina-Pankova, M., & Byrnes, H. (2017). Embracing the language-educational challenge of FL departments: Reflections on ways forward. *The Modern Language Journal* 101, 424–427.
- Serafini, E. J., Lake, J. B., & Long, M. H. (2015). Needs analysis for specialized learner populations: Essential methodological improvements. *English for Specific Purposes*, 40, 11–26.
- Sykes, J. M., & Cohen, A. D. (2018). Strategies and interlanguage pragmatics: Explicit and comprehensive. *Studies in Second Language Learning and Teaching*, 8(2), 381–402.
- Talif, R., & Noor, R. (2009). Connecting language needs in the workplace to the learning of English at tertiary level. *Pertanika Journal of Social Sciences and Humanities*, 17(2), 65–77.
- Teichler, U. (2016). The professional relevance of study. *International Journal of Chinese Education*, 5, 135–161.
- Trace, J., Hudson, T., & Brown, J. D. (2015). An overview of language for specific purposes. In J. Trace, T. Hudson, & J. D. Brown (Eds.), *Developing courses in languages for specific purposes* (pp. 1–22). National Foreign Language Resource Center.
- Waring, H. Zh. (2018). *Discourse analysis*. Routledge.
- Wyburd, J. (2011). *The Languages Graduate*. University Council of Modern Languages.
- Zakaria, A. K. M., Arefin, Md. S., & Hassan, M. K. (2017). Needs analysis: The take off point for designing English for Specific Purposes (ESP) course for engineering learners. *International Journal of English: Literature, Language and Skills*, 5(4), 132–141.
- Zhu, W., & Flaitz, J. (2005). Using focus group methodology to understand international students' academic language needs: A comparison of perspectives. *Teaching English as a Second or Foreign Language – Electronic Journal*, 8(4). Retrieved August 13, 2020 from www.tesle.org/wordpress/issues/volume8/ej32/ej32a3/

AUTHOR

Jim Davie, Ph.D., pursued an academic career as a researcher and lecturer in Russian at the University of Portsmouth (1993-98) before moving on to work in the UK civil service. He has written his article as an independent researcher. His most recent publication is *Slang Across Societies* (Routledge, 2018).

The Correlation Between Level of Learner Autonomy and Learner's Achievements in Arabic Reading and Listening: A Quantitative, Non-experimental Correlational Study

Anas Faiz

Defense Language Institute Foreign Language Center

The shift from teacher-centered to student-centered instruction requires language teachers' roles to change from being lecturers and controllers to facilitators and advisors. This shift requires students to have a high degree of learner autonomy. However, some teachers hesitate to promote learner autonomy. If they remain uncertain about the significance of learner autonomy and its impact on student achievement, they will likely avoid promoting learner autonomy. This quantitative non-experimental correlational study investigated the relationship between students' level of learner autonomy and their achievement in Arabic listening and reading as measured by In-Course Proficiency Tests (ICPTs). The participants were 86 U.S. military service members learning Arabic and its dialects as a foreign language in the Arabic Basic Course at the Defense Language Institute Foreign Language Center in Monterey, California. A learner autonomy Questionnaire (LAQ) was used to determine the students' level of learner autonomy. Bivariate Pearson correlation and linear regression tests determined the relationship and its predictive power between the LAQ and ICPT variables. The findings indicate that high scores on the LAQ were positively and directly associated with high scores on the listening and reading portions of the ICPT. The study concluded that promoting learner autonomy in learning Arabic and its dialects is essential due to its impact on students' scores on proficiency tests. Recommendations are included and aim to enhance teachers' and learners' awareness of the importance of learner autonomy and encourage teachers to promote learner autonomy in their instruction.

Keywords: *Learner autonomy, learner achievement, language for specific purposes*

INTRODUCTION

Learner autonomy (LA) refers to learners' ability to take charge of their learning, be aware of their learning goals and strategies, and monitor and evaluate their education progress (Little, Leni, & Legenhausen, 2017; Oxford, 2015). The concept of LA in adults learning a foreign language has been at the center of academic discussions for many years, and teachers are encouraged to

promote LA. Some teachers hesitate to promote LA, which could be due to either a lack of information about its importance or a lack of awareness about the positive influence of LA on students' performance on tests (Borg & Alshumaimeri, 2017; Buenner, 2019; Szocs & Jozsef, 2017). Awareness on the part of language teachers and learners about the importance of learner autonomy, motivation, and providing a learning environment where students and teachers can practice autonomous learning skills were essential in promoting learner autonomy.

This research investigated the topic of learner autonomy through the lens of Malcolm Knowles's andragogy theory. Malcolm Knowles, known as "the father of andragogy" in the United States, claimed that adults and children learn differently and introduced six assumptions associated with adult learning (Knowles et al., 2014). The Knowles andragogical model encourages educators to understand the difference between children's and adults' learning and incorporate adult learning's core principles in their teaching (Ferreira & MacLean, 2017). The core assumptions of adult learning theory are the learner's *need to know*, *self-concept*, *prior experience*, *readiness to learn*, *orientation to learning*, and *motivation to learn* (Knowles et al., 2014). Studying the characteristics of learner autonomy and its relationship to learning a foreign language through the lens of adult learner theory can inform the importance of utilizing learning strategies and explaining the teachers' and students' roles and responsibilities in teaching and learning a foreign language.

LITERATURE REVIEW

The concept of learner autonomy in adults learning a foreign language has been at the center of many academic discussions for many years (Borg & Alshumaimeri, 2017; Oxford, 2015; Salimi & Ansari, 2015). The importance of promoting learner autonomy and finding how language teachers and learners can contribute to developing learner autonomy have become a prominent research agenda (Alonazi, 2017; Gholami, 2016; Hsieh & Hsieh, 2019; Oxford, 2015; Salimi & Ansari, 2015). Unlike traditional teaching methods in which the teacher is at the center of instruction, the current trend in education places more emphasis on putting students at the center of instruction and increasing their responsibilities in the process of learning (Hsieh & Hsieh, 2019; Rana, 2019). Engaging students in decisions, such as choosing learning materials, learning strategies, and the strategies to demonstrate their understanding, such as writing essays, explaining orally, or creating video clips, is an essential element of promoting learner autonomy (Cotterall, 2017; Mynard, 2017; Rana, 2019). Little et al. (2017) explained teachers' and students' roles in developing learner autonomy in learning a language and introduced three principles for promoting learner autonomy. These principles are engaging learners in the process of learning, providing learners with opportunities to reflect on their learning, and using the target language communicatively (cited in Mynard, 2017). Language teachers need to ensure that students understand and acknowledge their critical role in learning activities such as planning for learning, setting learning goals and objectives, assessing the effectiveness of learning strategies, and evaluating the progress of language learning (Ceylon, 2015; Little et al., 2017; Oxford, 2015; Yasmin & Sohail, 2018). Many of these researchers' works have confirmed the effectiveness of promoting learner autonomy in teaching a foreign or second language. The discussion that follows includes the findings of various studies related to learner autonomy.

Various studies have explored the topic of learner autonomy. Ayash (2018) studied the correlation between the level of learner autonomy among Yemeni postgraduate students and their achievement on an English proficiency test. The study found no significant relationship between students' degree of learner autonomy and their scores on the university's standardized English language proficiency test. However, Ayash acknowledged that there were limitations in the study and encouraged teachers to foster learner autonomy in their classroom instruction. Szocs and Jozsef (2017) conducted a study in a Hungarian secondary school to explore teachers' and learners' perspectives on learner autonomy and classroom activities that promote autonomous learning skills. Unlike Ayash (2018), Szocs and Jozsef found a direct and positive relationship between learners' level of learner autonomy, learners' motivation, and their achievements in learning a foreign language. Interestingly, their study also revealed that teachers' and learners' perceptions about learner autonomy did not match. In addition, the notes from classroom observation indicated that the learning activities teachers wanted to do were not similar to what they did in their classrooms. Similarly, Dogan and Mirici (2017) conducted mixed research with EFL teachers in a Turkish university to explore their perceptions of learner autonomy and practices that could promote learner autonomy. The findings indicated that the EFL teachers in the Turkish university perceived learner autonomy in learning the English language as a positive factor. Almost all EFL teacher participants supported integrating activities that increase students' responsibility in learning and give them authority to control their education. However, as in Ayash (2018) and Szocs and Jozsef (2017), Turkish EFL teachers in the Dogan and Mirici study expressed their concerns about the feasibility of promoting learner autonomy in their classroom and involving students in some activities such as making decisions about learning materials and monitoring and assessing learning. Some of the language teachers in the above studies demonstrated insufficient awareness about developing learner autonomy and expressed their concerns about the feasibility of promoting learner autonomy (Ayash, 2018; Dogan & Mirici, 2017; Szocs & Jozsef, 2017). However, except Ayash (2018) who found no correlation between the level of learner autonomy and students' performance on the test, the majority of the above studies confirmed that teachers play a critical role in helping students to become autonomous learners by building a close relationship with students, give them sufficient freedom in learning, and incorporating their interests and experiences in the learning process (Ayash, 2018; Dogan & Mirici, 2017; Szocs & Jozsef, 2017).

In the same vein, Buenner (2019) studied university teachers' and students' perspectives on learner autonomy in Thailand and reported that EFL teachers expressed an insufficient understanding of learner autonomy's theoretical and practical aspects and considered learner autonomy as mere self-study or independent learning. Buenner also found that teacher respondents believed autonomous learning practices are not applicable inside the classroom. Instead, they thought students might want to practice autonomous learning skills outside of school independently and in their free time. Buenner's findings in the Thai context confirmed the findings of Abdolrezapour and Fallah (2015) in the Iranian context and Khairallah et al. (2020) in the Lebanese context. Abdolrezapour and Fallah (2015) found that the way teachers practice teaching and how they view the learning process influenced their students' perspectives and views toward learning. In other words, when teachers reflect on their teaching and think about the practices they used in the classroom, their students are most likely follow the same thinking and behavior (Abdolrezapour & Fallah, 2015). Khairallah et al. (2020) studied the factors of

Lebanese students' resistance to engaging in autonomous learning activities in learning English as a foreign language. In their study, the Lebanese English-language teachers expected their students to be more independent and responsible for their learning. In contrast, the students believed that the more they rely on their teachers and get guidance and support, the better achievement they would have on the language tests. As a result, Khairallah et al. suggested that Lebanese universities need to change the top-down decision-making approach and consider university learners as capable individuals who can set learning goals and identify and select appropriate resources and learning strategies without entirely relying on their teachers' instructions.

Borg and Alshumaimeri (2017) studied teachers' beliefs about learner autonomy in the Saudi Arabian context. They found that teachers showed dependence on structured official curriculums and believed that promoting learner autonomy might not help learners succeed on standardized achievement tests. The findings showed that the teachers, due to a dependence on the official curriculum, failed to demonstrate the promotion of learner autonomy in their teaching. However, the participating teachers accepted the value of learner autonomy in learning a foreign language and acknowledged the importance of integrating learner autonomy in their teaching. Both Borg and Alshumaimeri (2017) and Szocs and Jozsef (2017) acknowledged and confirmed the critical roles of language teachers in developing learner autonomy and highlighted the importance of raising teachers' education and level of awareness about the notion of learner autonomy. Many studies concluded that without a teacher who is knowledgeable, experienced, and most importantly willing to transfer his or her authority to students, it is challenging for language learners to increase their participation and improve their autonomy in learning (Abdolrezapour & Fallah, 2015; Ayash, 2018; Borg & Alshumaimeri, 2017; Buenner, 2019; Smith et al., 2018; Szocs & Jozsef, 2017).

While most of the studies on promoting learner autonomy in learning a foreign language highlighted the importance of developing autonomous learning skills as well as the critical roles of language teachers and learning activities in promoting learner autonomy, uncertainty about the effectiveness of autonomous learning skills and lack of information to illustrate the impact of learner autonomy on a learner's results on language proficiency tests were evident in some studies. Uncertainty about the effectiveness of autonomous learning and insufficient evidence to support the positive impact of learner autonomy in learning a foreign language could prevent some language teachers and learners from promoting learner autonomy (Borg & Alshumaimeri, 2017; Ceylon, 2015; Gholami, 2016). Both teachers and students want to know more about the relationship between learner autonomy and students' achievement (Borg & Alshumaimeri, 2017; Rana, 2019). The current study addressed the association between the students' level of learner autonomy in learning Arabic and its dialects as a foreign language and their achievement in the Arabic language proficiency tests using the following research questions:

RQ1. What is the relationship, if any, between participants' level of learner autonomy in learning Arabic and its dialects as a foreign language and their scores on the Arabic In-Course Proficiency Tests?

RQ2. To what extent could participants' level of learner autonomy in learning Arabic and its dialects as a foreign language predict their achievements on the In-Course Proficiency Tests?

H1₀. The participants' level of learner autonomy in learning Arabic and its dialects as a foreign language has no statistically significant influence on their scores on the In-Course Proficiency Tests.

H1_a. Participants' level of learner autonomy in learning Arabic and its dialects as a foreign language has a statistically significant influence on their scores on the In-Course Proficiency Tests.

H2₀. Participants' level of learner autonomy in learning Arabic and its dialects as a foreign language is not a potential predictor for their achievements on the In-Course Proficiency Tests.

H2_a. Participants' level of learner autonomy in learning Arabic and its dialects as a foreign language is a potential predictor for their achievements on the In-Course Proficiency Tests.

The knowledge derived from this research revealed that students' level of learner autonomy in learning Arabic as a foreign language is positively related to their achievements on the Arabic language proficiency tests.

METHOD

The purpose of this quantitative non-experimental correlational design was to investigate the relationship between the participants' level of LA in learning Arabic and its dialects as a foreign language (as measured by the Learner Autonomy Questionnaire) and their achievement on the Arabic listening and reading (as measured in In-Course Proficiency Tests). A quantitative non-experimental correlational design was used to answer the research questions, test hypotheses, and achieve the study's purpose. A quantitative inquiry is considered to be a suitable methodology for this study because both the Learner Autonomy Questionnaire (LAQ) and In-Course Proficiency Tests (ICPT) variables generate numerical and continuous data. Additionally, a quantitative non-experimental correlational design was aligned with the purpose of the research because this study intended to investigate the relationship between the LAQ and ICPT variables as they exist naturally and without any intervention or manipulation (Apuke, 2017; Rutberg & Bouikidis, 2018). The correlational design was suitable because it reveals whether the scores generated by one continuous variable are statistically and significantly associated with the scores generated by another continuous variable.

Setting and Participants

The population for this non-experimental quantitative study was the students of the Arabic Basic Course in the Arabic schools within the Defense Language Institute Foreign Language Center (DLIFLC) located in the Presidio of Monterey, California. Data was collected after the local Scientific Review Board, Northcentral University IRB, and Human Research Protections Program at DLIFLC approved the study. At the time of the study, around 500 students studied Arabic and its dialects in three separate Arabic schools. A minimum sample size of 84 participants is necessary to measure the correlation between two variables with a power of 80%, at medium strength, and with a value of alpha of 0.05, per guidelines for correlational analysis using power analysis and sample size (PASS), developed by Bujang and

Baharum (2016). The participants were 86 U.S. military service members of different branches with a minimum of a high school graduate diploma. The participants' level of education, age, and military affiliation to a specific U.S. military branch, such as the Army, Navy, Air Force, or Marines, were not variables in this study.

A multistage sampling method was utilized to draw the sample. The multistage sampling method was a mixture of clustering and simple random sampling methods. In a clustering sampling method, researchers usually divide the population into different clusters or groups (Alvi, 2016). The students in each classroom constituted a cluster or a group of learners, and then a simple random sampling method was used within each group. All participants were informed about the study's purpose and objective, and they were given the freedom to choose whether to participate or refuse participation.

Instruments

Learner Autonomy Questionnaire (LAQ)

The data were collected online via a self-reported learner autonomy questionnaire (LAQ), which determined each participant's level of learner autonomy in learning Arabic. The LAQ was developed by Zhang and Li (2004) and hosted in a Microsoft 365 Form. Zhang and Li developed the English version of the LAQ based on the learning strategies organized by Oxford (1990), Wenden (1998), and O'Malley and Chamot (1990). The LAQ has 21 items divided into two parts. The LAQ has 21 questions divided into two parts. The first part has eleven questions on a five-point Likert scale with options of *never, seldom, occasionally, often, and always*. Numbers 1, 2, 3, 4, and 5 were equal to never, seldom, occasionally, often, and always respectively. The numbers 1, 2, 3, 4, and 5 were used to generate numerical and computable data. The second part of the LAQ has ten multiple-choice questions with choices of A, B, C, D, and E, that were represented in the study with the numbers 1, 2, 3, 4, and 5 respectively. Each participant selects the closest and most accurate answer to each question based on their beliefs, attitudes, and ideas. The survey has been reviewed by an independent Department of Defense (DoD) organization and licensed by assigning the DoD Survey License Exemption (# 9)—Exempt # 0103. The result of all 21 answers collectively determines the student's level of learner autonomy in learning a language. Zhang and Li's LAQ has demonstrated a high level of content validity and a high level of reliability for determining the degree of level of learner autonomy (Koosha, Abdollahi, & Karimi, 2016). Additionally, the LAQ has been widely used for determining the degree of learner autonomy in learning a foreign language in other countries such as Iran, Yemen, China, and Ethiopia, and has been translated into multiple languages, such as Farsi, Arabic, Amharic, and Thai (Ayash, 2018; Desta, 2020; Koosha et al., 2016).

In-Course Proficiency Tests (ICPTs)

The Arabic language ICPT is a standardized proficiency test comprised of multiple-choice tests for listening comprehension and reading comprehension. Arabic language program students take two ICPTs per semester to prepare for the final Defense Language Proficiency Test, whose scores are used to determine graduation status. Only the students in semesters two and

three (ICPT 201, 202 and 301, and 302) were permitted to access the LAQ link using their DLIFLC email accounts, and only these students accessed and completed the LAQ. The participants' most recent Arabic reading and listening ICPT scores were collected from their respective instructors.

Analysis

The Pearson correlation and simple linear regression were used for the analysis of data. The LAQ and the Arabic listening and reading ICPT results for each student were paired up in the Statistical Package for Social Sciences (SPSS) and analyzed. A bivariate Pearson correlation measure was conducted to investigate the correlation between the LAQ and listening and reading ICPT variables. The bivariate Pearson correlation indicates whether or not there is a statistically significant association between two continuous variables (Green & Salkind, 2017; Mcleod, 2019). A simple linear regression test was conducted to measure whether the LAQ result of a student could potentially predict his or her achievement on the Arabic listening and reading ICPTs. The bivariate Pearson correlation and simple linear regression statistical findings provided inferential information to answer the first and second research questions of the current study. The statistical results were used to reject or accept the predetermined hypotheses. The relationship between the LAQ and listening and reading ICPTs variables was studied, with no intervention.

RESULTS

The findings revealed a significant direct and positive relationship between the higher scores in the LAQ and the higher scores on the Arabic ICPT listening and reading. The statistical information also indicated that the LAQ scores could be a possible predictor of student performance in the ICPTs.

To examine the correlation between the LAQ and ICPT listening and reading, a bivariate Pearson correlation test was conducted for ICPT Listening Comprehension (ICPT-LC) and ICPT Reading Comprehension (ICPT-RC). Table 1 illustrates descriptive statistics of LAQ, ICPT-LC, and ICPT-RC.

Table 1*Descriptive Statistics of LAQ, ICPT-LC, and ICPT-RC*

			Statistic	Std. Error
LAQ (Learner Autonomy Question- naire)	Mean		71.44	.727
	95% Confidence	Lower Bound	70.00	
	Interval for Mean	Upper Bound	72.89	
	5% Trimmed Mean		71.42	
	Median		71.00	
	Variance		45.40	
	Std. Deviation		6.7	
	Minimum		57	
	Maximum		87	
	Range		30	
	Interquartile Range		9	
	Skewness		.110	.260
	Kurtosis		-.425	.514
ICPT_LC (In-course Proficiency Test for Listening Compre- hension)	Mean		74.90	1.484
	95% Confidence	Lower Bound	71.94	
	Interval for Mean	Upper Bound	77.85	
	5% Trimmed Mean		75.22	
	Median		75.00	
	Variance		189.48	
	Std. Deviation		13.7	
	Minimum		40	
	Maximum		98	
	Range		58	
	Interquartile Range		20	
	Skewness		-.287	.260
	Kurtosis		-.464	.514
ICPT_RC (In-course Proficiency Test for Reading Compre- hension)	Mean		84.08	1.228
	95% Confidence	Lower Bound	81.64	
	Interval for Mean	Upper Bound	86.52	
	5% Trimmed Mean		84.62	
	Median		86.50	
	Variance		129.61	
	Std. Deviation		11.4	
	Minimum		57	
	Maximum		100	
	Range		43	
	Interquartile Range		15	
	Skewness		-.688	.260
	Kurtosis		-.314	.514

The Pearson correlation results in Table 2 show a statistically significant positive relationship between LAQ and ICPT-LC scores, $r = .23$, $p = .03$. Table 2 shows that the significant value is smaller than 0.05. Therefore, the null hypothesis is rejected in favor of the alternative hypothesis. Thus, there is a direct and positive correlation between students' level of learner autonomy in learning the Arabic language and their performance on the Arabic listening comprehension ICPT.

Table 2*Correlations of LAQ and ICPT-LC*

		LAQ	ICPT_LC
LAQ	Pearson Correlation	1	.23*
	Sig. (2-tailed)		.030
	N	86	86
ICPT_LC	Pearson Correlation	.23*	1
	Sig. (2-tailed)	.030	
	N	86	86

*Correlation is significant at the 0.05 level (2-tailed)

The bivariate Pearson correlation test results in Table 3 below for the LAQ and ICPT-RC scores indicated a statistically significant positive relationship, $r = .26$, $p = .016$.

Table 3*Correlations of LAQ and ICPT-RC*

		LAQ	ICPT_RC
LAQ	Pearson Correlation	1	.26*
	Sig. (2-tailed)		.016
	N	86	86
ICPT_RC	Pearson Correlation	.26*	1
	Sig. (2-tailed)	.016	
	N	86	86

*Correlation is significant at the 0.05 level (2-tailed)

Table 3 illustrates that the significant value is smaller than 0.05. Therefore, the null hypothesis is rejected in favor of the alternative hypothesis. Thus, there is a direct and positive correlation between students' level of learner autonomy in learning the Arabic language and their performance on the Arabic reading comprehension ICPT.

A simple linear regression test was conducted two times to check if the students' LAQ scores could predict students' scores on the ICPT-LC and ICPT-RC. The first simple linear regression test evaluated the predictive power of LAQ scores on ICPT listening comprehension scores, as shown in Table 4.

Table 4*Regression Coefficient LAQ on ICPT-LC (N = 86)*

Variable	<i>B</i>	β	<i>SE</i>
Constant	40.70		15.6
Learner Autonomy (LAQ)	0.48	0.23	0.22
R^2	0.06		

Table 4 illustrates the potential influence of the student's level of learner autonomy in learning the Arabic language on the Arabic listening comprehension ICPT results. The R^2 value of .06 revealed that the predictor variable (LAQ) explained 6% of the variance in the outcome variable with $F(1, 84) = 4.9$, $p = .03$. The significant value is smaller than .05; therefore, the null hypothesis is rejected in favor of the alternative hypothesis. Thus, the findings indicated that the LAQ scores could be a possible predictor for students' performance in the ICPT-LC, but the predictive power is weak ($\beta = .23$, $p < .05$).

Similarly, the second simple linear regression test evaluated the predictive power of LAQ scores on ICPT reading comprehension scores, as shown in Table 5 below.

Table 5*Regression Coefficient LAQ on ICPT-RC (N=86)*

Variable	<i>B</i>	β	<i>SE</i>
Constant	52.90		12.80
Learner Autonomy (LAQ)	0.44	0.26	0.18
R^2	0.07		

Table 5 shows the potential influence of students' level of learner autonomy in learning the Arabic language on their Arabic reading comprehension ICPT scores. The R^2 value of .07 revealed that the predictor variable (LAQ) explained 7% of the variance in the outcome variable with $F(1, 84) = 6.03$, $p = .016$. The null hypothesis is rejected in favor of the alternative hypothesis because the significant value is smaller than .05. The findings indicated that the LAQ scores could be a possible predictor of students' performance in the ICPT-RC, but the predictive power is weak ($\beta = .26$, $p < .05$).

The outcomes produced by this study shed light on the type and direction of the relationship between the students' level of learner autonomy in learning Arabic as a foreign language and their achievements on the Arabic listening and reading Proficiency Tests. Additionally, the results of the statistical analysis indicated that the LAQ scores could be a possible predictor for students' performance in listening and reading comprehension proficiency tests. However, the predictive power of LAQ scores on students' outcomes on the Arabic listening and reading comprehension proficiency tests was found weak.

DISCUSSION

This study achieved its proposed objectives, and the findings add to the growing body of knowledge highlighting the importance of promoting LA in learning a foreign language. Even though the results confirmed a positive and direct correlation, the statistical analyses indicated a weak predictive power of learners' level of learner autonomy on students' scores on the Arabic listening and reading ICPTs. This result could be because of the time of the study. The current research took place during the COVID-19 pandemic, and the restrictions due to the COVID-19 pandemic resulted in a negative impact on overall language learning at the DLIFLC. There could also be some other explanations. First, unlike the study in other contexts, students in this research studied Arabic rather than English. Second, DLIFLC students are U.S. military service members and may have a different motivation to study Arabic than most of the student participants in other studies, who seem to have largely been university students. Finally, DLIFLC students study Arabic full-time, whereas most of the participants in other studies probably take only one English class per semester. The findings of this study are still significant not only for DLIFLC Arabic language educators, students, curriculum developers, and faculty professional development trainers but also to benefit other foreign languages and learning institutions. The recommendations are as follows:

Provide opportunities for students to reflect on their LA skills throughout the course

The first recommendation for practice is to provide opportunities for students to reflect on their use of autonomous learning skills in learning foreign languages periodically throughout the foreign language program. For example, a learner autonomy questionnaire (such as that used in this study or one that is more relevant to the teaching context) can be administered at key times throughout students' coursework. Instructors can analyze questionnaire results and keep records for each student. Identifying students with lower scores in the early stage of a course would enable teachers and support offices to intervene and help students improve their autonomous learning skills. Many studies have found that autonomous learning skills are not inherited and can be acquired by ongoing and consistent practices (Erturk, 2016; Gholami, 2016; Yuliani & Lengkanawati, 2017); therefore, providing opportunities for students to reflect on their LA skills periodically is crucial.

Raise students' awareness about the significance of LA in learning Arabic and introduce them to practices and competencies that enhance autonomous learning skills

The second recommendation for practice is developing some type of Learner Autonomy Workshop for language learners. Workshops would serve to boost students' awareness about the importance of learner autonomy in learning a foreign language and introduce them to evidence-based practices that promote autonomous learning skills. Workshops could be offered periodically and include a follow-up system, for example, a portfolio or journal, to ensure students implement autonomous learning strategies inside and outside the classroom.

Many studies have confirmed that students with a high awareness of learner autonomy and practices that promote learner autonomy achieve a high level of language proficiency (Dogan & Mirici, 2017; Erturk, 2016; Szocs & Jozsef, 2017). Similar to the current study's findings, many previous studies revealed a positive relationship between students' knowledge of learning strategies, their success in learning a foreign language, and the need for developing a sort of learning strategies awareness training for students (Bajrami, 2015; Ceylon, 2015; Ranjan & Philominraj, 2020).

Raise teachers' awareness about the importance and positive impact of LA on achieving higher levels of language proficiency

The third recommendation is to offer workshops for teachers, highlighting the importance of promoting learner autonomy from a language educator's perspective and providing evidence-based teaching methodologies that promote autonomous learning skills. Many studies have documented teachers' insufficient knowledge and their concerns about the feasibility of promoting learner autonomy (Ayash, 2018; Dogan & Mirici, 2017; Szocs & Jozsef, 2017). Moreover, many studies have confirmed that experienced teachers with adequate awareness and education about learner autonomy implemented autonomy practices more often and effectively than teachers with insufficient knowledge (Asiri & Shukri, 2018; Azin et al., 2019; Khairallah et al., 2020; Liu, 2015; Smith et al., 2018; Yagcioglu, 2018). Samples of strategies and practices that promote autonomous learning inside and outside of classrooms can be included in the workshops.

Integrate Autonomous Learning Practices in curricular and extracurricular activities

The fourth recommendation is regarding curriculum and materials development. Previous studies have acknowledged that a single training session will likely be insufficient to help students develop learner autonomy, and students likely need to practice autonomous learning skills in an ongoing manner (Gholami, 2016). Incorporating learning practices that foster learner autonomy in the foreign language curriculum and implementing them consistently in the classroom is critical (Alonazi, 2017; Yasmin & Sohail, 2018). The current study outcomes can be used by curriculum and materials developers as further evidence of the importance of incorporating curricular and extracurricular supplemental activities that promote autonomous learning skills.

CONCLUSION

The knowledge derived from this study revealed a statistically significant direct and positive relationship between students' level of learner autonomy and their scores in the Arabic listening and reading comprehension ICPTs. The findings showed the scores generated by the Learner Autonomy Questionnaire were positively and directly associated with Arabic listening and reading ICPT scores. Thus, the study concluded that promoting learner autonomy in learning a foreign language may positively impact students' scores on foreign language proficiency tests.

The study's findings were in line with previous studies investigating similar research questions. Although the current study documented a statistically significant direct and positive relationship between LAQ and ICPTs, the predictive power of learner autonomy on students' achievement in the Arabic proficiency tests was weak, which opened doors for future studies. Future researchers' aims could be to investigate the roles of teachers and language curricula in relationship to students' learner autonomy and their achievement on Arabic proficiency tests. A replication of the current study with students engaged in standard face-to-face learning is highly recommended for future researchers.

REFERENCES

- Abdolrezapour, P., & Fallah, E. (2015). An investigation into the impact of reflective teaching on EFL learners' autonomy and intrinsic motivation. *Cypriot Journal of Educational Science*, 10(4), 305–315. <https://doi.org/10.18844/cjes.v10i4.229>
- Alonazi, S. (2017). The role of teachers in promoting learner autonomy in secondary schools in Saudi Arabia. *English Language Teaching*, 10(7), 183–202. <https://doi.org/10.5539/elt.v10n7p183>
- Alvi, M. (2016). A manual for selecting sampling techniques in research. *Munich PersonalRePEc Archive*. Retrieved from https://mpira.ub.uni-muenchen.de/70218/1/MPRA_paper_70218.pdf
- Apuke, O. (2017). Quantitative research methods: A synopsis approach. *Arabian Journal of Business and Management Review*, 6(10), 40–47. <https://doi.org/10.12816/0040336>
- Asiri, J., & Shukri, N. (2018). Female teachers' perspectives of learner autonomy in the Saudi context. *Theory and Practice in Language Studies*, 8(6), 570–579. <https://doi.org/10.17507/tpls.0806.03>
- Ayash, N. (2018). The relationship between learner autonomy and English proficiency of Yemeni postgraduate English students: A correlational study in Hodeidah University. *Journal of Education and Practice*, 9(26), 80–89. Retrieved from www.iiste.org
- Azin, N., Biria, R., & Golestan, A. (2019). Iranian EFL instructors' perceptions and practices concerning learner autonomy. *Journal of Modern Research in English Language Studies* 5(3), 73–97. <https://doi.org/10.30479/JMRELS.2019.10530.1318>
- Bajrami, L. (2015). Teacher's new role in language learning and in promoting learner autonomy. *Procedia – Social and Behavioral Sciences*, 199(1). The proceedings of the 1st GlobELT conference on teaching and learning English as an additional language). 423–427. <https://doi.org/10.1016/j.sbspro.2015.07.528>
- Borg, S., & Alshumaimeri, Y. (2017). Language learner autonomy in a tertiary context: Teachers' beliefs and practices. *Language Teaching Research*, 23(1), 9–38. <https://doi.org/10.1177/1362168817725759>
- Buenner, P. (2019). University students' perception on autonomous learning: A case of private university, Thailand. *The 1st International Conference on English Studies: Innovation English Language Teaching and Learning*. Retrieved from <https://www.researchgate.net/publication/330496841>

- Bujang, M., & Baharum, N. (2016). Sample size guideline for correlation analysis. *Journal of Social Science Research*, 3(1), 37–46. Retrieved from <http://www.scholink.org/ojs/index.php/wjssr/article/view/398/468>
- Ceylon, N. (2015). Fostering learner autonomy. *Procedia-Social and Behavioral Sciences*. (199) 85–93. <https://doi.org/10.1016/j.sbspro.2015.07.491>
- Cotterall, S. (2017). The pedagogy of learner autonomy: Lessons from the classroom. *Studies in Self-Access Learning Journal*, 8(2), 102–115. Retrieved from <http://sisaljournal.org/archives/jun17/cotterall>
- DLIFLC General Catalog (2019-2020). Defense Language Institute Foreign Language Center. Retrieved from <https://www.dliflc.edu/resources/publications/general-catalog/>
- Desta, M. (2020). An investigation into the association between learning autonomy, language anxiety and thinking style: University student in focus. *Journal of Language Teaching and Research*, 11(2), 309–317. <https://doi.org/10.17507/jltr.1102.21>
- Dogan, G., & Mirici, H. (2017). EFL instructors' perception and practices on learner autonomy in some Turkish universities. *Journal of Language and Linguistic Studies*, 13(1), 166–193. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1140425.pdf>
- Erturk, N. (2016). Language learner autonomy: Is it really possible? *Procedia – Social and Behavioral Sciences* 232(1), 650–654. <https://doi.org/10.1016/j.sbspro.2016.10.089>
- Ferreira, D., & MacLean, G. (2017). Andragogy in the 21st century: Applying the assumptions of adult learning online. *Language Research Bulletin*, 1(32), 10–19. Retrieved from <http://bit.ly/ferreira2018>
- Gholami, H. (2016). Self-assessment and learner autonomy. *Theory and Practice in Language Studies*. 1(6), 46–51. <https://doi.org/10.17507/tppls.0601.06>
- Green, S., & Salkind, N. (2017). *Using SPSS for Windows and Mac analyzing and understanding the data* (8th ed.). Pearson.
- Hsieh, H., & Hsieh, H. (2019). Undergraduates' out-of-class learning: Exploring EFL students' autonomous learning behaviors and their usage of resources. *Education Sciences*, 9(159), 1–15. <https://doi.org/10.3390/educsci9030159>
- Khairallah, M., Fleonova, O., & Nicolas, M. (2020). Understanding students' resistance to autonomous learning in an L2 English language course at the university in Lebanon. *European Journal of Education*, 3(1), 14–30. <https://doi.org/10.26417/ejed.v3i1.p18-30>
- Knowles, M. S., Holton III, E. F., & Swanson, R. A. (2014). *The adult learner: The definitive classic in adult education and human resource development*. Routledge.
- Koosha, M. Abdollahi, A., & Karimi, F. (2016). The relationship among EFL learner's self-esteem, autonomy, and reading comprehension. *Theory and Practice in Language Studies*, 6(1), 68–78. <https://doi.org/10.17507/tppls.0601.09>
- Little, D., Leni, D., & Legenhausen, L. (2017). *Language learner autonomy: Theory, practice and research*. Multilingual Matters.
- Liu, H. (2015). Learner Autonomy: The role of motivation in foreign language learning. *Journal of Language Teaching & Research*, 6(6), 1165–1174. <https://doi.org/10.17507/jltr.0606.02>
- McCusker, K., & Gunaydin, S. (2015). Research using qualitative, quantitative or mixed methods and choice based on the research. *Perfusion*, 30(7), 537–542. <https://doi.org/10.1177/0267659114559116>

- McLeod, S. (2019). What does effect size tell you? *Simply Psychology*. Retrieved from <https://www.simplypsychology.org/effect-size.html>
- Mynard, J. (2017). Review of the book *Language learner autonomy: Theory, practice and research*, by D. Little, L. Dam & L. Legenhausen. *Studies in Self-Access Learning Journal*, 8(4), 362–370. <https://doi.org/10.37237/080408>
- O'Malley, M., & Chamot, A. (1990). *Learning strategies in second language learning*. Cambridge: Cambridge University Press.
- Oxford, R. (2015). Expanded perspectives on autonomous learners. *Innovation in Language Learning & Teaching*, 9(1), 58–71. <https://doi.org/10.1080/17501229.2014.995765>
- Rana, S. (2019). Exploring learner and teacher perceptions about learner centeredness. *Dialog on Language Instruction*, 29(2), 85–89.
- Ranjan, R., & Philominraj, A. (2020). Language learning strategies, motivation and gender in foreign language context. *Universal Journal of Educational Research*, 8(2), 591–604. <https://doi.org/10.13189/ujer.2020.080231>
- Rutberg, S., & Bouikidis, C. (2018). Exploring the evidence. Focusing on the fundamentals: A simplistic differentiation between qualitative and quantitative research. *Nephrology Nursing Journal*, 45(2), 209–213. Retrieved from <http://www.homeworkgain.com/wp-content/uploads/edd/2019/09/20181009143525article2.pdf>
- Salimi, A., & Ansari, N. (2015). Learner autonomy: Investigating Iranian English teachers' beliefs. *Theory and Practice in Language Studies*, 5(5), 1106–1115. <https://doi.org/10.17507/tpls.0505.28>
- Smith, R., Kuchah, K., & Lamb, M. (2018). Learner autonomy in developing countries. In A. Chik, N. Aoki, & R. Smith (Eds.), *Autonomy in language learning and teaching*, (pp. 7–27). Palgrave Pivot. https://doi.org/10.1057/978-1-137-52998-5_2
- Szocs, K., & Jozsef, E. (2017). Teachers' and learners' beliefs about language learning autonomy and its implications in the classroom. *Journal of Applied Language Studies*, 11(2), 125–145. <https://doi.org/10.17011/apples/urn.201708233542>
- Yagcioglu, O. (2018). The autonomy approach in language learning and teaching. *European Journal of Education Studies*, 5(3), 263–271. <https://doi.org/10.5281/zenodo.2539776>
- Yasmin, M., & Sohail, A. (2018). A creative alliance between learner autonomy and English language learning: Pakistani university teachers' beliefs. *Creativity Studies*, 11(1), 1–9. <https://doi.org/10.3846/23450479.2017.1406874>
- Yuliani, Y., & Lengkanawati, N. (2017). Project-based learning in promoting learner autonomy in an EFL classroom. *Indonesian Journal of Applied Linguistics*, 7(2) 285–293. <https://doi.org/10.17509/ijal.v7i2.8131>
- Zhang, L., & Li, X. (2004). A comparative study on learner autonomy between Chinese students and West European students. *Foreign Language World*, 24(4), 15–23.

AUTHOR

Anas Faiz, Ed.D., is an Associate Professor at the Defense Language Institute Foreign Language Center in Monterey, California, anas.faiz@dliflc.edu. Dr. Faiz has worked at DLIFLC in various academic and administrative capacities.

The Relative Effects of Motivation and Structured Input on the Acquisition of English Past Tense Regular Forms

Zilin Huang

Guangdong University of Education, People's Republic of China

Alessandro Benati

University College Dublin, Republic of Ireland

This study investigates the effects of motivation and structured input on the acquisition of English past tense regular forms. The role of motivation in the positive effects generated by processing instruction has been investigated in one previous study (Farhat & Benati, 2018). However, more research is needed to generalize the initial findings to a different language, language background, processing problem, and age. In addition, the present study measures the effects of the practice component of processing instruction, as no explicit information was given to the participants of this classroom experiment. The Academic Motivation Scale (AMS) questionnaire was used to capture different variables that influence motivation in order to create two different groups (highly- and nominally-motivated). In this experimental study, 43 native Chinese secondary school-age learners (L1 Mandarin) were assigned to three groups: a highly-motivated group ($n = 16$), a nominally-motivated group ($n = 14$), and a control group ($n = 13$). The two instructional groups received structured input activities, which lasted for two hours. The control group received no instruction. A one-sentence-level interpretation test was used in a pre-test and post-test design to measure instructional effects on the interpretation of sentences containing past tense forms. The results indicated that both instructional groups improved equally from the pre-test to the post-test. Both groups outperformed the control group. Structured input was shown to be the main factor for learners' improvement in performance regardless of their level of motivation.

Keywords: *Structured input, motivation, the lexical preference principle, interpretation test, input processing*

BACKGROUND

The Input Processing Theory and the Redundancy Principle

VanPatten's theory of input processing (VanPatten 1996, 2004, 2007, 2015a) in second language acquisition frames the research questions and procedures used in the current

investigation. One of the overarching principles of input processing is the so-called Primacy of Meaning Principle. The Primacy of Meaning Principle asserts that when L2 learners are engaged in communicative, meaningful interchanges, they are primarily concerned with meaning. That is, "...learners are driven to look for the message or communicative intent in the input" (VanPatten, 2004, p. 7). According to the input processing theory, second language (L2) learners attend to grammatical forms in the input to connect one form to one meaning at a time. Research has repeatedly uncovered the varying conditions under which learners successfully make meaning from the input (VanPatten, 2015a). One of these conditions, influencing correct tense assignment, is the presence or absence of temporal adverbials in the input sentences.

The Lexical Preference Principle argues that if grammatical forms express a particular meaning, that specific meaning can also be encoded lexically (making that grammatical marker redundant). If this is the case, L2 learners will prefer to initially process lexical items before grammatical items if they both encode the same semantic information (VanPatten, 2007).

A great number of grammatical features encodes some kind of semantic information (e.g., English verbal inflection *-ed* encodes the past, as in *watched*). The same semantic notion is, however, also expressed in a sentence by lexical items such as temporal adverbs (e.g., in English *yesterday*). Given, as postulated, that L2 learners are driven to process content words before anything else, they would therefore attend to lexical temporal references of pastness before verbal inflections (past tense *-ed*). In other words, L2 learners tend to mark tense lexically before they mark it morphologically.

Overall, languages have ways of encoding and thus signalling the same information in multiple ways, and we refer to this as the natural redundancy of languages. The background research in support of the Lexical Preference Principle has been focused on tense assignment (Benati, 2019; Lee et al., 1997; VanPatten & Keating, 2007). The main result generated from this research is that the factor influencing correct tense assignment is the presence or absence of temporal adverbs in the input (VanPatten, 2007).

The Preference for the Nonredundancy Principle is highly related to the Lexical Preference Principle. According to this principle, L2 learners are more likely to process nonredundant meaningful grammatical markers before they process redundant meaningful markers. A grammatical marker might well have semantic value but other sentence elements might make it redundant. Again, the presence of a lexical item would make the grammatical marker redundant and cause a delay in the ability of L2 learners to make accurate form-meaning connections.

Structured Input: Characteristics and Empirical Evidence

The pedagogical model derived from the input processing theory is called processing instruction (VanPatten, 1996). The main aim of processing instruction is to intervene in the processes L2 learners use to get data from the input and help them process input more accurately. Processing instruction consists of two main components: (1) L2 learners are given explicit information on a particular processing principle that may negatively affect their picking up of forms/structure during comprehension, and (2) L2 learners are pushed to process forms through structured input activities.

During structured input, L2 learners are pushed to process the target form or structure through activities in which the input is manipulated in particular ways to make L2 learners

dependent on the form or the structure to get meaning. Given that L2 learners are driven to process content words before anything else, they tend to rely on lexical temporal references to process tenses rather than verbal inflections, for example. L2 learners prefer to mark tense lexically before they mark it morphologically and therefore a delay in processing grammatical forms would occur. In structured input activities, lexical items are removed from the input so as to facilitate and ensure that L2 learners make appropriate form-meaning mappings.

Prior studies have measured the effects of processing instruction and its components. VanPatten and Oikennon (1996) measured the effects of the three components of processing instruction (explicit information vs. structured input activities vs. full processing instruction) on the acquisition of object pronouns in Spanish. Three groups were compared. One received only explicit information, the other structured input activities, and the third both components (full processing instruction). The main findings from this study revealed that the gains made (on both the interpretation and production sentence-level tests) by both the processing instruction group and the structured input activities group were greater than the group receiving only explicit instruction on the targeted form. The main outcomes of this study revealed that the structured input component was responsible for learners' gains.

Benati (2004a, 2004b) replicated this study by comparing the effects of processing instruction, structured input activities, and explicit information on the acquisition of the Italian future tense and Italian gender agreement. In both studies, the results confirmed that the structured input activities component was responsible for L2 learners' improved performance.

Wong (2004) found positive results for structured input activities when she compared the effects of processing instruction, structured input activities, and explicit information in the acquisition of the French negative plus indefinite article. The addition of explicit information in the processing instruction treatment did not cause improvement.

The main finding of the line of research teasing out the three components of processing instruction has confirmed that the structured input component is responsible for the changes in the learners' ability to interpret sentences containing the feature under investigation. Structured input as the main component of processing instruction is an effective pedagogical intervention that has consistently demonstrated significant improvement in L2 learners' performance on both interpretation and production tests (VanPatten, 2015b).

The effects of structured input have also been compared to other pedagogical interventions (see the full review of this line of research in Benati & Lee, 2010, 2015, and Benati, 2019) using a variety of languages (e.g., Arabic, French, German, Italian, Japanese, Spanish), and linguistic features (e.g. word order, passive and causative constructions, tense markers, aspectual markers, mood with an expression of doubt) affected by different processing principles, populations (groups and individuals from different L1s), and contexts. Overall, the findings from this line of research demonstrate that structured input activities provide L2 learners with opportunities to process forms and structures in a target language more efficiently than other pedagogical interventions (e.g., meaning-output-based instruction, textual enhancement, input flood).

More recently, the online effects of structured input have also been measured using eye-tracking (Benati, 2020) and self-paced reading tests (see the full review of these studies in Benati, 2021). This new line of inquiry is limited to three empirical studies. The use of online testing to measure the effects of structured input is very much in its infancy; however, this new branch of

investigation might offer the possibility of more fine-grained information and analysis about moment-by-moment sentence comprehension and might even be a way to better measure implicit knowledge.

Structured Input and Motivation

Motivation is one of the individual difference factors which has attracted much attention in second language acquisition from both a socio-psychological and a cognitive perspective. Motivation can be defined as the willingness to learn, and scholars in this field (Ryan, 2019) have been attempting to associate motivation with successful language acquisition.

According to Gardner (2001), we should distinguish between (1) integrative and (2) instrumental motivation. Integrative motivation refers to L2 learners' efforts and desire to be part of the target language culture. Instrumental motivation instead refers to L2 learners' desire to learn a language for educational or economic purposes.

Only a few empirical studies have investigated the question of whether or not motivation might have an effect on the results generated by structured input/processing instruction. Abbasian and Minagar (2012) carried out a study comparing the effects of processing instruction and dictogloss in developing grammar ability and motivation among EFL learners. Results from this study suggested that both instructional interventions were significantly beneficial to the development of English grammar ability. A positive correlation between highly motivated participants and the effectiveness of dictogloss instruction was found in this study. However, this was not the case for processing instruction.

Farhat and Benati (2018) provided empirical evidence of the effectiveness of processing instruction no matter the level of motivation of the learners. The main findings from the study conducted by Farhat and Benati (2018) indicated that both groups (nominally motivated vs. highly motivated) benefited equally from processing instruction when measured by interpretation and production sentence-level tests.

Motivation and Research Questions in the Present Study

Although the relationship between motivation and processing instruction has been explored in two previous empirical studies, the results are not conclusive and need to be generalised to a different population, a different linguistic feature, and a different processing principle. In view of the limitations of the existing data and intending to measure the effectiveness of structured input, this study attempts to address the following question: Do L2 learners benefit from structured input activities in equal measure, regardless of their motivation?

The current study builds upon the work of Farhat and Benati (2018) and seeks to investigate the relative effects, at the sentence level, of structured input on one linguistic item of the English verbal system (the past tense regular form *-ed*) among two groups with different motivational levels. The main contributions of the present research to the ongoing debate on the possible role of motivation in the effects generated by structured input are:

1. to investigate the effectiveness of structured input activities on the acquisition of English past tense forms no matter the level of motivation of L2 learners.

2. to generalize previous findings to a different population (secondary school learners), language background (Mandarin as L1), language feature (English regular past tense forms), and a combination of processing principles (The Lexical Preference Principle, The Preference for Nonredundancy Principle).

The study compares the immediate effects of structured input activities when delivered to a group of highly- and nominally-motivated Chinese school-age L2 learners on the acquisition of English past tense regular forms. One specific question was formulated for this study:

Do school-age L2 learners (L1 Mandarin) exposed to structured input correctly interpret sentences that contain past tense forms in English, regardless of their level of motivation?

RESEARCH DESIGN

Participants

In order to address the research question, a classroom experimental study was carried out. The participants of this study were Chinese native school-age learners of English (aged 12 and 14) studying at a junior high school in China and were a balanced mixture of male and female learners. They had little or no knowledge about the target grammatical form (the English past tense) before the beginning of the instructional treatment. The original pool of 62 participants was reduced to 43, as participants were included in the final pool according to the following criteria: (a) scored below 60% of the maximum score on the pre-test, (b) attended all the phases of the experiment, (c) scored high or low in the motivation questionnaire. All the criteria had to be met for inclusion in the final pool. Subjects who received a middle score in the motivation questionnaire were excluded from this experimental study.

The final pool of participants ($N=43$) was divided into three groups: a highly-motivated group (16 participants), a nominally-motivated group (14 participants), and a control group (13 participants) that received no instruction during normal class time. An academic motivation questionnaire (see Appendix 1) was implemented before the beginning of the instructional period and used to form the two instructional groups (highly motivated vs. nominally motivated) and the control group (highly motivated).

Procedure

The Academic Motivation Scale (AMS) questionnaire (see Appendix) was used in this experiment to measure participants' motivation to study in school in order to form the three groups. Vallerand and Blssonnette (1992) originally developed the AMS based on the so-called Self Determination Theory. This theory suggests that three elements constitute motivation: (i) intrinsic motivation, (ii) extrinsic motivation, and (iii) amotivation.

Intrinsic motivation includes the following aspects: (a) the desire to act for experiencing the pleasure or satisfaction when learning something; (b) the desire to act for experiencing the

pleasure or satisfaction from accomplishing something; and (c) the desire to act for experiencing some form of personal stimulation.

Extrinsic motivation can be defined as (a) the desire to act in order to acquire a sense of importance or personal value, (b) the desire to act to avoid pressure and guilt, and (c) the desire to accomplish acts to gain rewards. “Amotivation” refers to L2 learners who lack any motivation.

AMS consists of twenty-eight items that assess L2 learners’ motivation levels. The main reason for using this questionnaire in this study is that it is an effective tool for capturing several elements related to L2 learners’ motivation. The questions answered by this questionnaire determine L2 learners’ level of motivation.

Three days before the implementation of the instructional treatment (structured input), a pre-test was conducted to assess participants’ previous knowledge of English past tense forms. The pre-test consisted of an interpretation sentence-level test. This test required participants to interpret correctly whether the sentences heard referred to activities that happened in the past or the present. As previously noted, the original pool of participants was reduced following the established criteria. An immediate post-test was administered to the three groups at the end of the two-hour instructional period.

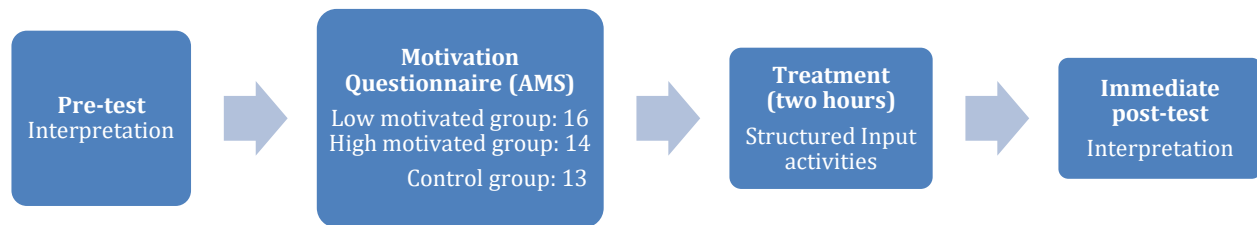
The Academic Motivation Scale is one of the most widely used instruments to measure motivational levels toward learning. The scale consisted of 28 items utilizing a seven-point Likert scale [1–7]. Based on the mean scores obtained in the AMS questionnaire, participants were assigned to three groups (Table 1). Individuals who scored between 2.00 and 3.25 were assigned to the nominally motivated instructional group. Those who scored between 5.00 and 6.25 were assigned to the highly motivated instructional group and the control group (see Table 1). Participants who scored in the middle range (3.26–4.99) were not included in the study as they could not be classified as highly or nominally motivated individuals.

Table 1
Motivation Scores (Mean range)

Mean range	Group	Number of subjects
5.00-6.25	Highly-motivated instructional group	16
2.00-3.25	Nominally-motivated instructional group	14
5.00-6.25	Control group	13

The researcher, not the regular class teacher, was the instructor during the experiment. The duration of the instructional treatment (structured input) was two hours for both groups. The pre-post-test procedure of this experiment is displayed in Figure 1.

Figure 1
Overview of the Experiment



Target Grammatical Form

The target grammatical form for this research was English past tense regular forms. This grammatical form was chosen because it is affected by a combination of processing principles: The Lexical Preference Principle and The Preference for Nonredundancy Principle (VanPatten, 2004). According to the Lexical Preference Principle (VanPatten, 2004), L2 learners are likely to process lexical items before grammatical items when both contain the same semantic information. In the sentence *Last week Daisy cleaned her room*, both the lexical item *Last week* and the *-ed* verb ending convey pastness (the idea that an activity is completed). Based on the redundancy principle (The Preference for Nonredundancy Principle), L2 learners will first rely on the temporal adverb *last week* rather than processing *-ed* to interpret the meaning of the sentence (the activity happened in the past and not in the present).

In addition, this target feature was chosen because, as Benati (2005) noted, in Mandarin, a verb has no tense form to show that something is happening, happened in the past, or will happen in the future. Mandarin uses different lexical devices such as time adverbs or the particle *le* to indicate a time frame. As a result, it is likely that L1 Mandarin learners of English would find tenses difficult (Benati, 2005; Benati & Lee, 2008). L2 learners whose first language does not mark tense grammatically (e.g., L1 Mandarin learners) would process past tense forms and their meaning later than other learners whose first language marks tense grammatically (VanPatten, 2004).

Instructional Materials

One instructional packet was developed and used for this classroom experiment. The two instructional groups received the same treatment (structured input activities), which included both reading and listening activities. Unlike Farhat and Benati (2018), the instruction in this study consisted of only structured input; no explicit information about the target form was provided. Feedback was kept to a minimum, and the instructor told participants only when they were right or wrong. No other explanation was offered. The instructor guided participants to follow the activities step by step and made sure that they were familiar with the words in these activities. High-frequency words were used to facilitate overall understanding. The language used to instruct the participants was Mandarin. The reason for this was to ensure that participants were able to understand and were clear about what they needed to do for each activity.

In the instructional treatment, structured input activities contained seventy-four tokens. More specifically, the treatment included six referential activities and four affective activities which required students to listen or read to comprehend sentences containing the target feature.

Structured Input activities (Farley, 2004; Lee & VanPatten, 2003; VanPatten, 2004), consist of referential and affective activities. Referential activities are those that have only one correct answer for each item and L2 learners must rely on what they hear or read to make a form-meaning connection. Affective activities refer to those in which L2 learners express their opinions or beliefs or produce other affective responses based on their own perception and/or experience.

An example of a referential activity used in the experiment is provided in Figure 2. In this activity, participants listened to several sentences and had to correctly interpret whether these sentences referred to Cai Xukun's past or present life. This superstar was chosen because he is now very popular among young Chinese learners.

Figure 2

Referential Activity (sample)

Referential Activity: Cai Xukun: Now or in the past?

Step 1

Listen to the sentences about the life of superstar Cai Xukun and decide whether the activity is about his past life as a student or his life now as a superstar.

	STUDENT (PAST)	SUPER STAR (NOW)
1.	<input type="checkbox"/>	<input type="checkbox"/>
2.	<input type="checkbox"/>	<input type="checkbox"/>
3.	<input type="checkbox"/>	<input type="checkbox"/>
4.	<input type="checkbox"/>	<input type="checkbox"/>
5.	<input type="checkbox"/>	<input type="checkbox"/>

Activity continues in a similar fashion

Sentences heard:

- 1) ...joins many TV programs
- 2) ...wanted to be a singer
- 3) ...used more time to practice singing
- 4) ...studied in the U.S.
- 5) ...plays basketball with friends

The affective activity in Figure 3 required participants to read sentences containing the target feature and then decide whether they did the same thing on New Year's Day of the previous year.

Figure 3*Affective Activity (sample)***Affective Activity: New Year activities!****Step 1**

Read these activities and think about whether you did the same or different things last year on New Year's Day.

	Yes	No
1. I visited my grandparents.	<input type="checkbox"/>	<input type="checkbox"/>
2. I dressed in new clothes.	<input type="checkbox"/>	<input type="checkbox"/>
3. I mailed New Year cards to friends.	<input type="checkbox"/>	<input type="checkbox"/>
4. I played with friends.	<input type="checkbox"/>	<input type="checkbox"/>
5. I enjoyed myself.	<input type="checkbox"/>	<input type="checkbox"/>

Activity continues in a similar fashion

Step 2

Compare your answers with a classmate's to find out how many things you did that were the same.

Structured input activities in the instructional treatment were designed to follow these main guidelines (VanPatten, 1996):

1. Present one thing at a time
2. Keep the meaning in focus
3. Move from sentence to discourse
4. Use both written and oral input
5. Have learners do something with the input
6. Keep learners' processing strategies in mind

Assessment and Scoring

In the present experimental research, the assessment included a sentence-level interpretation test (see sample in Figure 4). The interpretation sentence-level test was administered pre-and post-test. In the interpretation test, vocabulary consisted of high-frequency words to **minimize** participants' difficulty in processing. The interpretation assessment test consisted of 20 audio sentences, of which 10 conveyed pastness (verbs with *-ed*) and 10 were distractors (present tense verbs). These audio recordings were recorded by the same native-language instructor to make sure they were consistent and were played only once at a normal pace. Participants received 1 point for every correct response and 0 points for an incorrect answer. The maximum possible score was 10 points and the minimum score was 0

points. Distractors were not scored in the analysis (forms in the present). Participants had the option to check *Not Sure* if they were not sure of the answers (these items were not scored).

Figure 4

Interpretation Test (sample)

Interpretation test			
Listen to the sentences and decide whether the action happened yesterday or today.			
YESTERDAY	RIGHT NOW	NOT SURE	
<i>[Sentences heard by the students]</i>			
1. _____	_____	_____	I walk to school.
2. _____	_____	_____	I talked with my brother.
3. _____	_____	_____	I cooked dinner.
4. _____	_____	_____	I buy beautiful clothes.
5. _____	_____	_____	I finished my homework.

The directions for the interpretation test were created in such a way to be easily understood by the participants, thereby utilizing the terms “yesterday” for past tense and “right now” for non-past tense. Only the sentences with past tense forms were scored. The raw scores for each assessment were submitted to an analysis of variance with repeated measures where the independent variable was the instructional treatment (structured input) and the dependent variable was the assessment of that treatment (interpretation test).

RESULTS

A one-way ANOVA (Analysis of Variance) was used on the pre-test scores. This analysis showed that no significant differences were found between the highly-motivated group ($M=1.38$; $SD=0.455$), the nominally motivated group ($M=1.42$; $SD=0.655$), and the control group ($M=1.28$; $SD=0.365$) before instruction ($F(2, 43) = 1.084$, $p = .194$). Any differences found after the treatment can only be ascribed to the effects of that instructional treatment. Table 2 shows the means of the three groups in the sentence-level interpretation tests (pre-test, post-test). The means indicate an improvement for both structured input groups but not for the control group.

A repeated-measures ANOVA was conducted on the raw scores of the sentence-level interpretation test. Treatment (structured input highly-motivated group, structured input nominally-motivated group, control group) was the between-subjects factor, while Time (pre and post-tests) was the within-subjects factor. The results of the statistical analysis showed a significant main effect for Treatment ($F(2, 43) = 17.213$, $p = .001$); a significant main effect for Time ($F(2, 43) = 19.690$, $p = .001$); and significant interaction between Treatment and Time ($F(2, 43) = 13.129$, $p = .001$).

A post hoc Tukey test was carried out on the raw scores of the interpretation test to show possible contrasts among the three groups. The results of this analysis indicated that the effects of the Treatment were due to the following contrasts: the two structured input groups performed significantly better than the control group (SI highly-motivated = SI nominally-motivated > control).

group ($p = .001$). The scores of the two structured input groups were not significantly different from each other ($p = .528$).

The statistical analysis indicated that both instructional groups improved significantly and equally from pre-test to post-test no matter the level of motivation of participants. The control group, on the other hand, did not show a significant difference in pre-test and post-test scores.

Table 2

Sentence-Level Interpretation Data (Mean and SD)

Groups	<i>n</i>	Pre-test (0-10)		Post-test (0-10)	
		Mean	SD	Mean	SD
SI highly-motivated	16	1.38	0.455	6.41	1.021
SI nominally-motivated	14	1.42	0.655	6.45	1.120
Control	14	1.28	0.365	1.23	1.023

Note: SI = structured input

DISCUSSION

Main Findings

The results of this experimental study reveal that both structured input groups improved equally and obtained similar gains from pre-test to post-test after receiving their treatment. The pre-test mean scores of the three groups were similar, indicating that the groups were equal before the beginning of the instructional period (see Table 2). After the end of the treatment, the mean scores for the highly-motivated and nominally-motivated structured input groups increased to just over 50% from pre-tests to post-tests. The control group did not improve from the pre-test to the post-test. The mean scores suggest that both instructional groups improved equally from pre- to post-test. The statistical analysis ANOVA and the post hoc analysis confirmed that both structured input groups performed equally and better than the control group.

The main research question that guided this investigation was: Do school-age L2 learners (L1 Mandarin) exposed to structured input correctly interpret sentences that contain past tense forms in English no matter their level of motivation?

The statistical analysis of pre- and post-test scores showed no significant differences between groups, but significant effects in terms of instruction (structured input vs. no instruction) and time (pre- vs. post-test). Both instructional groups (highly-motivated and nominally-motivated) improved no matter their level of motivation. The control group made no improvement and motivation was not considered a key element in the results generated by this study.

The structured input treatment helped participants in both groups equally in processing and correctly interpreting new regular past tense forms. Structured input positively affects the way L2 learners process English past tense forms. In the case of this study, the target form is affected by The Lexical Preference Principle and The Preference for Nonredundancy Principle. This study shows that school-age L2 learners (L1 Mandarin) exposed to structured input can learn to interpret correctly sentences containing past tense forms in English no matter their level of motivation. In addition, the main findings from this study build on previous research in this field by demonstrating that motivation does not affect the positive results generated by structured input among a population with a different background (Mandarin L1), age (school-age secondary school learners), acquiring a different grammatical form (English past tense forms).

Implications

Overall, the main findings from this study make several theoretical and pedagogical contributions.

First of all, the main results from this study contribute directly to the discussion on the crucial role that input processing plays in second language acquisition. The scores on the pre- and post-tests further demonstrate that L2 learners rely on a combination of processing principles to process forms such as the English past tense. Subsequently, structured input activities are an effective pedagogical intervention, no matter the level of motivation of the L2 learners, to help learners to become better processors of input.

Secondly, the main findings from this study reaffirm the importance of input-based practice as a key pedagogical tool for grammar instruction. Structured input provides an effective option for teaching grammar when linguistic features such as past tense forms are affected by processing problems. Grammar instruction should not be viewed as the explanation and practice of grammatical forms but as a pedagogical intervention to facilitate the processes and strategies used by L2 learners during input processing. Structured input practice does not intend to pour knowledge into L2 learners' heads, but rather to assist certain processes which can aid the growth of learners' internal language system. L2 learners receiving structure input can make appropriate meaning connections (in the case of this study by processing English past tense forms correctly). In this respect, a subsequent pedagogical implication of this study is that acquiring grammar requires L2 learners to process meaning-bearing input first. A coherent approach to grammar instruction takes L2 learners from processing a grammatical feature in the input to accessing the feature from the internal language system to create output.

Thirdly, the findings from this study provide support to several hypotheses previously formulated within the processing instruction/structured input research framework (Benati & Lee, 2008; Benati & Lee, 2010; Benati & Lee, 2015). The results obtained in this study lend further support to the so-called Age Hypothesis (Benati & Lee, 2008). Structured input is just as effective an intervention with young learners as it is with older learners. Our results, with secondary school-age L2 learners, confirmed previous research (Benati, 2019) demonstrating the effectiveness of structured input activities with learners of different ages, no matter the level of motivation of those learners. The results obtained in this study in using native speakers of other languages than English (Mandarin L1) also support the so-called Native Language Hypothesis (Benati & Lee, 2008). Structured input is effective for instilling target language-specific processing

strategies, no matter the native language of the learners. The present study contributes to the expansion of the Native Language Hypothesis by adding Mandarin to the current list of languages in which the effectiveness of structured input practice has been observed.

Limitations and Further Research

As with all empirical research, the present study is limited in several ways. While the positive outcomes in this study have been underscored, we acknowledge certain limitations. Firstly, the lack of a delayed post-test (due to practical reasons) is a methodological limitation. Further research should measure the durability of the structured input treatment among L2 learners with different levels of motivation.

Secondly, the number of participants in this study was relatively small. Therefore more research needs to be conducted in the future with a larger population to generalize these initial findings.

Thirdly, the present study measures motivation and offline effects of structured input on the acquisition of English past forms. Future research should investigate the effects of this pedagogical intervention and motivational factors with the linguistic features of other languages that might be affected by similar processing principles, among different populations (e.g., L1s, language backgrounds) and using online measurements (e.g., eye-tracking and self-paced reading tests) to measure language processing.

Fourthly, future research should consider a longer instructional period to see whether structured input practice might have further beneficial effects in improving the accuracy of processing of forms/structures among L2 learners.

Finally, future studies measuring structured input and motivational factors should use two motivation questionnaires to make the measurement of motivation more accurate in capturing different and additional aspects of it. For example, the Attitude Motivation Test Battery (AMTB) questionnaire could be added.

CONCLUSION

The results from this study indicate that academic motivation as an individual difference does not seem to have an impact on the positive effects generated by structured input. Previous research investigating a possible correlation between structured input and motivation is limited and more research is needed within this field of inquiry. From the data collected in this study, structured input alone is beneficial for school-age learners' (L1 Mandarin) interpretation and effective processing of English past tense forms (affected by a combination of processing principles) regardless of learners' level of motivation.

ACKNOWLEDGMENTS

Special thanks should go to our colleagues at a junior high school in China who helped us recruit the participants and organize the groups. We would also like to thank the anonymous

reviewers for their comments and suggestions on the first draft of this paper. The usual caveat of responsibility (ours alone) applies.

REFERENCES

- Abbasian, G., & Minagar, N. (2012). PI-based vs. DG-oriented instruction in developing grammar ability and motivation of EFL learners. *Iranian EFL Journal*, 8, 94–112.
- Benati, A. (2004a). The effects of processing instruction and its components on the acquisition of gender agreement in Italian. *Language Awareness*, 13, 67–80.
- Benati, A. (2004b). The effects of structured input and explicit information on the acquisition of Italian future tense. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 207–255). Erlbaum.
- Benati, A. (2005). The effects of PI, TI and MOI in the acquisition of English simple past tense. *Language Teaching Research*, 9, 67–113.
- Benati, A. (2019). Classroom-oriented research: Processing instruction. *Language Teaching*, 52, 343–359.
- Benati, A. (2020). The effects of structured input and traditional instruction on the acquisition of the English causative passive forms: An eye-tracking study measuring accuracy in responses and processing patterns. *Language Teaching Research*, 24, 1–21.
- Benati, A. (2021). *Input processing and processing instruction: The acquisition of Italian and Modern Standard Arabic*. Benjamins Publishing.
- Benati, A., & Lee, J. (2008). *Grammar acquisition and processing instruction: Secondary and cumulative effects*. Multilingual Matters.
- Benati, A., & Lee, J. (2010). *Processing instruction and discourse*. Continuum.
- Benati, A., & Lee, J. (2015). Processing instruction: New insights after twenty years of theory, research and application. *International Review of Applied Linguistics*, 53, 87–90.
- Farhat, A., & Benati, A. (2018). The effects of motivation on processing instruction in the acquisition of Modern Standard Arabic gender agreement. *Instructed Second Language Acquisition*, 2, 61–82.
- Farley, A. (2004). *Structured input: Grammar instruction for the acquisition-oriented classroom*. McGraw-Hill.
- Gardner, R. C. (2001). Language learning motivation: The student, the teacher, and the researcher. *Texas Papers in Foreign Language Education*, 6, 1–18.
- Lee, J., Cadierno, T., Glass, W., VanPatten, B. (1997). The effects of lexical and grammatical cues on processing temporal reference in second language input. *Applied Language Learning*, 8, 1–24.
- Lee, J., & VanPatten, B. (2003). *Making communicative teaching happen* (2nd ed.). McGraw-Hill.
- Ryan, S. (2019). Language learner motivation: What motivates motivation researchers? In J. Schwieter & B. Benati (Eds), *The Cambridge handbook of language learning* (pp. 409–429). Cambridge University Press.
- Vallerand, R., & Blissonnette, R. (1992). Intrinsic, extrinsic, and amotivational styles as predictors of behavior: A prospective study. *Journal of Personality*, 60, 599–620.
- VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Ablex.

- VanPatten, B. (2004). *Processing instruction: Theory, research, and commentary*. Lawrence Erlbaum Associates.
- VanPatten, B. (2007). Input processing in adult second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp.115–135). Lawrence Erlbaum Associates.
- VanPatten, B. (2015a) Input processing in adult SLA. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition* (pp.113–135). Routledge.
- VanPatten, B. (2015b) Foundations of processing instruction. *International Review of Applied Linguistics*, 53, 91–109.
- VanPatten, B., & Oikkenon, S. (1996). Explanation versus structured input in processing instruction. *Studies in Second Language Acquisition*, 18, 495–510.
- VanPatten, B., & Keating, G. (2007). Getting tense: Lexical preference, L1 transfer, and native and non-native processing of temporal reference. [Unpublished manuscript].
- Wong, W. (2004). Processing instruction in French: the roles of explicit information and structured input. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 187–205). Erlbaum.

APPENDIX

Academic Motivation Scale

Your name 您的名字：

WHY DO YOU GO TO SCHOOL?

Using the scale below, indicate to what extent each of the following items presently corresponds to one of the reasons why you go to school.

你为什么上学？

使用下面的量表，表明以下每个项目在多大程度上与您上学的原因之一相对应。

1 表示“完全不符合”，4 表示“基本符合”，7 表示“完全相符”。

1 完全不符 2 挺不符合 3 比较不符合 4 基本符合 5 比较符合 6 挺符合 7 完全符合

1. Because I need at least a middle school degree in order to find a high-paying job later on.

因为我至少需要一个中学学历，这样以后才能找到一份好工作。

2. Because I experience pleasure and satisfaction while learning new things.

因为我在学习新事物的同时能体验到快乐和满足。

3. Because I think that a middle school education will help me better prepare for the career I have chosen.

因为我认为中学教育将帮助我更好地为我理想的职业做好准备。

4. Because I really like going to school.

因为我真的很喜欢上学。

5. Honestly, I don't know; I really feel that I am wasting my time in school.

老实说，我不知道；我真的觉得我在学校是在浪费时间。

6. For the pleasure I experience while surpassing myself in my studies.

为了我在学习中超越自我时所体验到的快乐。

7. To prove to myself that I am capable of completing my middle school degree.

向我自己证明我有能力完成我的中学学业。

8. In order to obtain a more prestigious job later on.

为了以后获得一份更有声望的工作。

9. For the pleasure I experience when I discover new things never seen before.

为了当我发现以前从未见过的新事物时，我所体验到的快乐。

10. Because eventually it will enable me to enter the job market in a field that I like.

因为上学最终能让我进入我喜欢的领域的就业市场。

11. Because for me, school is fun.

因为对我来说，学校很有趣。

12. I once had good reasons for going to school; however, now I wonder whether I should continue.

我曾经有充分的理由来上学；然而，现在我不知道我是否应该继续。

13. For the pleasure that I experience while I am surpassing myself in one of my personal accomplishments.

为了我在我的个人成就中超越自我时所体验的快乐。

14. Because of the fact that when I succeed in school I feel important.

因为当我在学校取得成功时，我觉得自己很重要。

15. Because I want to have "the good life" later on.

因为我希望以后拥有“美好的生活”。

16. For the pleasure that I experience in broadening my knowledge about subjects which appeal to me.

为了在拓宽我喜欢的科目的知识时，我所体验到的快乐。

17. Because this will help me make a better choice regarding my career orientation.

因为这样可以帮我更好地选择我的职业方向。

18. For the pleasure that I experience when I am taken by discussions with interesting teachers.

为了当我被有趣的老师吸引时，我所体验到的快乐。

19. I can't see why I go to school and frankly, I couldn't care less.

我不明白我为什么要上学，坦白说，我一点也不在乎。

20. For the satisfaction I feel when I am in the process of accomplishing difficult academic activities.

为了我在完成困难的学术活动的过程中所感受到的满足感。

21. To show myself that I am an intelligent person.

为了向我自己证明我是一个聪明的人。

22. In order to have a better salary later on.

为了以后有更好的薪水。

23 Because my studies allow me to continue to learn about many things that interest me.

因为我学习能让我继续学很多我感兴趣的事情。

24 Because I believe that my middle school education will improve my competence as a worker.

因为我相信我的中学教育将会提高我的工作能力。

25. For the "high" feeling that I experience while reading about various interesting subjects.

为了我在阅读各种有趣的科目时所体验到的兴奋感觉。

26. I don't know; I can't understand what I am doing in school.

我不知道; 我不明白我在学校做什么。

27. Because middle school allows me to experience a personal satisfaction in my quest for excellence in my studies.

因为中学让我体验到了在学习上追求卓越的个人满足感。

28. Because I want to show myself that I can succeed in my studies.

因为我想证明我能在学习上取得成功。

AUTHORS

Zilin Huang, M.A., *Teacher of English, Guangdong University of Education, People's Republic of China.*

Alessandro Benati, Ph.D., *Professor, School of Education, University College Dublin, Republic of Ireland. Dr. Benati is known for his research on how L2 learners process information and what internal strategies they use in language processing across romance and non-romance languages.*

Grammar as Concept: The Implementation of the PACE Model in Beginning Spanish Classes

Ewurama Okine

Texas A&M University

Gabriela C. Zapata

University of Nottingham

This study investigates the effectiveness of the PACE approach (Adair-Hauck & Donato, 2002) for grammar teaching in beginning (novice mid/high) second language (L2) Spanish classes. Even though research has reported instructional benefits for PACE in L2 French (Groenveld, 2011), Alutiiq (Branson, 2015), and ESL (Harris, 2017; Ngo, 2018), to the best of the authors' knowledge, no studies have been conducted with beginning L2 Spanish learners. Thus, this work seeks to contribute to the existing literature by focusing on the adoption of the approach in L2 Spanish university classes. The study focused on stem-changing verb forms in the present tense, and it involved the participation of 47 students in a public institution in the Southern United States. The participants were enrolled in two different sections of a first-semester class taught by a PACE-trained instructor. Data were collected through pre- and post-tests consisting of two multiple-choice tasks at the sentence and paragraph levels. Findings showed statistically significant differences between pre-and post-test results, which not only mirror previous studies, but also appeared to point to the effectiveness of the PACE approach for grammar teaching in L2 Spanish classes as well as for L2 use in all modes of communication—interpersonal, interpretive, and presentational.

Keywords: *L2 Spanish, PACE, L2 grammar teaching, modes of communication*

INTRODUCTION

The teaching of grammar in second language (L2) classrooms has been one of the most important research foci in the field of L2 pedagogy in the past decades. According to Richards and Rodgers (2001), different approaches have been proposed and investigated as far back as the 1840s, when translation was used as a tool to focus learners' attention on forms and the rules associated with said forms. With the advent of communicative language teaching (Savignon, 1983, 1987) and, more recently, literacy-based approaches to L2 teaching (Allen & Paesani, 2010; Zapata, 2022), it has become clear there is a need for instruction that will facilitate contextualized, meaningful grammar learning through students' active L2 use for communication and their guided discovery of the connections between meaning and form, instead of rote, decontextualized learning.

To address this need, in 2002, Adair-Hauck and Donato developed a pedagogical approach with a focus on form that relies on the use of texts and students' and instructors' dialogic, collaborative analysis of the relationship between grammar and meaning. The approach is known as PACE—an acronym for its four instructional stages, namely *Presentation*, *Attention*, *Co-Construction*, and *Extension*. The implementation of PACE in L2 instruction has been recommended by Glisan and Donato (2017) as one of the core L2 teaching practices “that are essential for... teachers to enact in their classrooms to support second language learning and development” (p. 1). Because of its reliance on text and learners' collaborative construction of knowledge, PACE offers opportunities for active, contextualized L2 use in all modes of communication—interpersonal (speaking), interpretive (reading and listening/viewing), and presentational (writing) (Glisan & Donato, 2017). The implication here is that the approach can not only facilitate students' guided discovery of the relationship between meaning and form, but it can also play a role in overall L2 development. Additionally, through the use of multimodal stories, L2 teachers can connect instruction to students' lifeworld, communities, and/or experiences, as well as contribute to their growth as multiliterate persons. That is, by choosing materials that are relevant for learners at both the instructional and personal levels that combine various semiotic elements (e.g., linguistic and visual/gestural), practitioners can foster belonging and can promote their students' understanding of multimodal meaning-making (Kalantzis et al., 2016, 2019; Zapata, 2022).

Even though PACE has been deemed an effective instructional tool for grammar teaching in the existing literature, and it is compatible with current literacy-based L2 approaches (Allen & Paesani, 2010; Zapata, 2022) it has not, to the best of the authors' knowledge, been widely adopted in L2 Spanish instruction. For example, the model is not included in current commercial textbooks, which, according to Fernández (2011, p. 165), appear to be “leaving out many of the newer approaches that encourage students to notice and connect form with meaning, ..., without offering significant, more innovative alternative[s] to enhance grammar learning” (see also Cubillos, 2014; Sessarego, 2016; Yoon, 2019). Additionally, no studies have yet focused on the implementation of the approach, as developed by Adair-Hauck and Donato (2002) in beginning of Spanish L2 classes.¹

The purpose of this paper is to fill this gap by investigating the application of PACE in two first-semester L2 Spanish classes in a public university in the Southern United States. The first section of the paper provides an overview of the approach and presents the existing studies on PACE in other languages. The next part introduces the study, including its research focus, participants, the institutional context in which it took place, and data collection and analysis. This is followed by the results of the investigation and their discussion. The final sections address the limitations of this work and conclude it.

¹ A recently published study by Davin & Kushki (2022) investigated grammar instruction in Spanish L2 learning from a sociocultural perspective. The work did not use the same version of PACE discussed in the present study, and originally created by Adair-Hauck and Donato (2002), nor did it focus specifically on the approach. Thus, it was not considered in the present study.

LITERATURE REVIEW

The PACE Approach and Its Application in L2 Classes

PACE was developed by Adair-Hauck and Donato in 2002 to offer the field of L2 instruction a practice that would allow for:

a focus on form... in a meaningful context [that would] make visible to the learner how a particular language form is used for the various modes of communication, how forms construct social and cultural meanings, and how they address larger communicative goals and purposes (Glisan & Donato, 2017, p. 92).

PACE consists of four pedagogical phases—*Presentation*, *Attention*, *Co-construction*, and *Extension*, which Adair-Hauck and Donato (2002) describe as follows. In the first phase, *Presentation*, the teacher foreshadows the grammar lesson by introducing students to a short, oral text, typically a narrative, that incorporates both the L2 and other semiotic elements such as images and audio (e.g., music or sound effects).² The instructor's oral narrative constitutes learners' first exposure to the structures of focus in connection to meaning. In the second phase, *Attention*, students participate in a guided discovery of the grammatical patterns tied to the lesson's forms based on the meaning conveyed by them in the chosen text. That is, in this stage, learners work with a written version of the oral text introduced in the previous phase, and the L2 teacher develops pedagogical interventions that allow for the identification of patterns and connections without much difficulty. In this phase, the students' work can also be facilitated by cooperative learning tasks. In the third phase, *Co-construction*, the L2 instructor and learners collaboratively construct the grammatical rules that guide the structures identified in the *Attention* phase, based on the form they exhibit and their connection to the message conveyed. In the fourth and final phase, *Extension*, students attempt to actively use their new forms in interpersonal, interpretive, and/or presentational communication.

Existing research on the adoption of PACE in L2 classes is quite limited; to the best of the authors' knowledge, only four studies have examined the incorporation of the approach as it was developed by Adair-Hauck and Donato (2002).³ These works have focused on four different languages in both high school and university/adult educational contexts. For example, Groeneveld (2011) conducted a study with 34 English-speaking high school students in French and Dutch language classes. The researcher was in charge of the French course, and another teacher taught the Dutch one. The objective of this work was to investigate whether the application of PACE in six different lessons could help learners identify the grammatical class or syntactic function (i.e., the focus was declarative knowledge). Data were collected through a pretest and a posttest consisting of multiple-choice questions in which the participants were

²Adair-Hauck and Donato (2002) emphasize that, when choosing a text, L2 practitioners need to consider both connections to the target culture(s), as well as the learners' lifeworld and personal/academic needs.

³There is some work (e.g., Davin & Kushki, 2022; Haight et al., 2007; Vogel et al., 2011) that has investigated the use of modified versions of PACE. Other existing publications have provided solely pedagogical information in the form of sample tasks (e.g., Li & Paul, 2019; Miguel & Solana, 2016) or recommendations for the implementation of PACE in combination with other approaches to grammar or L2 teaching (e.g., González-Bueno, 2021).

provided with a sample sentence featuring pronouns, direct objects, verb agreement, or the superlative and were asked to identify the grammatical class or syntactic function. Groeneveld found that there was an overall increase in the participants' accuracy in structure identification between the pre-and post-tests. Additionally, data collected through a final informal evaluation showed that students had enjoyed the implementation of PACE for grammar learning. Despite these positive results, there were problems with this study. For instance, the researcher did not provide much information about the instruments used for data collection (e.g., it was not specified how many items were included in the tests), nor did she clarify if any of the reported differences between the pre-and post-tests were statistically significant.

Branson (2015) carried out the second study with high school students that focused on the incorporation of PACE in L2 classes. This work took place in Alaska, and the linguistic focus was Alutiiq, one of the Native languages of the state. The researcher employed PACE to teach participants the past tense, and by resorting to stories directly connected to the Kodiak Alutiiq-speaking community in the Presentation phase, she also reinforced cultural values in traditional storytelling in this endangered language. The scholar in charge of this work mentioned the use of pre-and post-tests in her work. Nevertheless, like Groeneveld (2011), she did not provide enough information about data collection and analysis or specific results. That is, when discussing her findings, Branson reported that the participants in the study appeared to have enjoyed PACE as an instructional intervention, and they had broadened their cultural knowledge. Additionally, the researcher posited that the approach had not been as effective as expected in terms of L2 grammar learning. However, as previously mentioned, no empirical data were offered.

The third study that has explored the use of PACE for L2 teaching is Harris's (2017), which worked with six university ESL students. The structures of focus were the prepositions *in*, *through*, *on*, *along*, *at*, and *past*, and their use in presentational communication (L2 writing). The researcher created a series of PACE interventions during a three-week period and investigated the pedagogical effects of the approach through the analysis of the participants' use of the target prepositions in their writing at the end of the three-week instructional phase. The results showed that all participants were able to accurately use the prepositions *in* and *on*, but still struggled with the rest of the target structures, mostly omitting them. Harris believed that, even if the students in the study had not been able to incorporate some of the target structures into their presentational communication, different aspects of their work in the co-construction phase had revealed a better understanding of the relationship between meaning and form. That is, evidence for this understanding had been the multimodal collaborative products that the learners had created to express textually, visually, and gesturally the rules guiding these prepositions. Even though this work was more comprehensive than the two studies previously discussed, the researcher failed to provide information about data collection and analysis or empirical results.

The fourth and most recent study on PACE was carried out by Ngo (2018), who worked with 10 adult Vietnamese students enrolled in an online ESL program. The researcher employed PACE in eight virtual classes to teach participants sentence structure, subject-verb agreement, verb tenses, nouns, adjectives and adverbs, irregular adverbs, comparison, and pronouns. Data were based on the participants' impressions of PACE (recorded in class diary entries) and the results of a pre-and post-test administered, respectively, at the beginning and end of the study. These tests consisted of 20 items: 10 multiple-choice sentences and 10 incomplete statements. The qualitative analysis of the students' diaries showed that most learners felt more confident in

their L2 use in interpretive and presentational communication (writing) after their exposure to PACE. To determine the effectiveness of PACE for grammar learning, Ngo calculated the differences between the participants' correct answers in the tests completed at the beginning and end of the instructional period, reporting improvement in 7 of the 10 participants' L2 grammar knowledge. Nevertheless, no statistical analyses were offered by the researcher, and, therefore, it is not possible to determine if the differences reported were statistically significant.

The works presented in this section appeared to have had a largely instructional focus and therefore did not offer in-depth information on data sources and the analysis on which the findings were based. Nevertheless, the studies have provided some evidence of the effectiveness of PACE for grammar teaching in L2 classes. This paper seeks to contribute to the existing literature on PACE (as conceived by Adair-Hauck and Donato, 2002) by focusing on a population not thus far investigated—L2 Spanish university students. The next sections of the article introduce the present study.

Research Question

The objective of this study was to contribute to the existing literature on the adoption of PACE (Adair-Hauck & Donato, 2002) for the teaching of grammar in L2 classes. In particular, this work sought to answer the following research question:

RQ: Does the application of PACE as originally conceived by Adair-Hauck and Donato (2002) for grammar teaching in beginning L2 Spanish university classes contribute to the development of university students' declarative knowledge of stem-changing verbs in the present tense?

In the next sections of the paper, we describe the participants, instruments, and methods of data collection and analysis that were employed to investigate the research question, as well as the instructional context in which the study was carried out. This is followed by the presentation and discussion of results.

METHOD

Participants

The participants in this study were enrolled in two sections of an L2 beginning Spanish class in a public university in the southern United States during the Spring 2020 semester. The total enrolment was 47 students who had been placed in the class because they did not have any previous experience with the language. Forty-four of the 47 enrolled learners participated in all the phases of the study. Thirty-five of these participants were female and 9, male. All of these students, except for one, were between 18 and 24 years of age. Forty-two participants had received their schooling in English, 1 in Tagalog, and 1 in Chinese, and all of them characterized their Spanish proficiency as minimal. Most of the participants (89%) mentioned fulfilling the university's language requirement as the main reason for studying Spanish.

The two participating sections were taught by a female graduate teaching assistant in a Hispanic Linguistics doctoral program. Her classes were chosen for this study for a variety of reasons. First, the instructor received comprehensive (one semester) L2 methodology training, including the use of PACE. Also, classroom observations had shown that her teaching style was not only student-centered and highly innovative, but also exceeded expectations in the implementation of high-leverage teaching practices (the use of PACE is considered one of these practices) (Glisan & Donato, 2017), and the four phases of the method.

Instructional Context

The PACE lesson conducted with the study participants focused on Spanish stem-changing verbs in the present indicative tense. These verbs undergo three classes of inflectional transformations within their stem ($e \rightarrow ie$, $e \rightarrow i$, and $o \rightarrow ue$), or the syllable containing the stressed e or o vowel in the infinitive form (Bowden et al., 2011). In the *Presentation* phase, two pre-narrative collaborative activities were developed by the instructor to review key vocabulary items in the text and to activate students' schemata (both thematically and linguistically) (Carrell, 1984; Rumelhart, 1980). The first was a conversation task that introduced students to the topic of the narrative—a story connected to the popular, worldwide lodging company *Airbnb*. The participants were asked to discuss the social practices and expectations attached to the use of this service, sharing their knowledge of and personal experiences with it. In the second task, learners tried to guess the meaning of key vocabulary terms in the text before it was presented with the help of contextualized examples (i.e., words were presented in a sentence-level context). The objective of this step was to facilitate the easy recognition and comprehension of keywords and phrases during the instructor's narration of the story.

After the two pre-narrative tasks, the text was narrated by the instructor accompanied by images with the Creative Commons licenses CC BY-SA and CC BY-SA-NC. A content comprehension task, with seven questions, concluded the *Presentation* phase. In the *Attention* phase, the participants received a written copy of the text. Resorting to text enhancement to draw learners' attention to form (Doughty & Williams, 1998; Wong, 2005), the instructor asked students to identify common traits in the highlighted stem-changing verbs. In the *Co-construction* phase, the participants worked in groups of three and categorized the verbs they had identified in the previous phase based on similar stem-changing patterns (i.e., *-ie-*, *-i-*, or *-ue-*). In the next step, the instructor took the information elicited from the students, and together with them, developed the grammar rules guiding stem-changing verbs. Once this activity had been completed, the participants attempted a cloze test to practice the structures they had just analyzed. To wrap up the PACE lesson, in the *Extension* phase, the students engaged in an interpersonal and presentational task that involved the use of the newly learned verbs in a discussion and class presentation on possible scenarios to continue the narrated story.

Instruments

This study was designed to follow a pre-, post-, and delayed post-test design. However, due to disruptions to instruction brought about by the COVID-19 pandemic, the planned delayed post-test could not be administered. Thus, the findings in this study are based on the results of

the pre-and post-tests. This design was deemed appropriate by the investigators because it had been chosen in two of the four existing studies on PACE for the teaching of L2 grammar (Groeneveld, 2011; Ngo, 2018).

The first source of data in the study was a biographical and language questionnaire with six questions that focused on the participants' gender, age, language of schooling, self-reported proficiency level in Spanish, and reason for being enrolled in their Spanish class. The second set of instruments consisted of a pre-and a post-test. Both tests included two multiple-choice tasks with 15 items each and three options per blank. 10 of the 30 total items (five in each multiple-choice task) were target forms, while the remaining 20 focused on structures to which students had been exposed and used earlier in the semester. The first task in both the pre-and post-tests consisted of 15 sentences, and the second one was a short paragraph with 15 incomplete sentences. The content of both tasks in the pre-and post-tests was the same, and it incorporated vocabulary and topics with which students were familiar and had been discussed in class. Item examples are presented as follows:⁴

1. Sentence level

1a. Mi mamá _____ solamente cinco horas por día. Siempre está muy ocupada.

- a. *dorme* b. *dormir* c. *duerme*

2a. Yo _____ a las 11:00 todos los días.

- a. *almorzo* b. *almorzar* c. *almuerzo*

2. Contextualized

...También este lugar 3. _____ muchos años, más de ciento veinte, pero la mamá de María 4. _____ este tipo de casas porque 5. _____ renovarlas y expresar su creatividad. María no 6. _____ como su mamá: esta casa le da mucho miedo (fear)...

- | | | |
|----------------------|--------------------|--------------------|
| 3. a) <i>tene</i> | b) <i>tener</i> | c) <i>tiene</i> |
| 4. a) <i>prefere</i> | b) <i>prefiere</i> | c) <i>preferir</i> |
| 5. a) <i>puede</i> | b) <i>pode</i> | c) <i>poder</i> |
| 6. a) <i>opino</i> | b) <i>opina</i> | c) <i>opinar</i> |

Procedures

Data collection took place during a four-day week of instruction in the Spring 2020 semester, just before the COVID-19 shutdown, while the participating classes were still taught face-to-face. The biographical questionnaire and pre-test were administered on day 1, and the next two days were devoted to the PACE lesson (each lesson was 50 minutes long). On day 4, the participants completed the post-test, and their answers were transferred to an Excel sheet by a graduate research assistant not involved with the study. Each correct option in the pre-and post-tests was given a value of 1, and incorrect ones were recorded as 0. The results were first analyzed using descriptive statistics. However, since the outcome variable was binary (i.e., correct vs. incorrect), and the relevant distribution was not Gaussian, the resulting pre-and post-

⁴ The instructional unit and data collection instruments are available from the authors upon request.

test values were also analyzed with R version 4.0.2 (R Core Team, 2023) using a mixed-effects logistic regression (Agresti, 2002), with the lme4 package (Version 1.1-23; Bates et al., 2015) with Participant and Item as crossed random effects (the overall model is presented in Table 2 in the Results section). This analysis allowed for the determination of statistically significant differences.

RESULTS

The results of the descriptive statistical analysis (Table 1) show differences between the pre-and post-tests in both tasks, sentence-level and contextualized. The percent and average number of correct answers in both tasks were lower in the pre-test than in the post-test, which suggests learners were able to identify the correct options at a higher rate after they had worked with PACE. Also, the standard deviation values indicate less variability with respect to the average number of correct answers in both tasks in the post-test. These differences were statistically significant for both the sentence level and paragraph tasks. The results of the overall model comparing the pre-and post-tests are presented in Table 2.

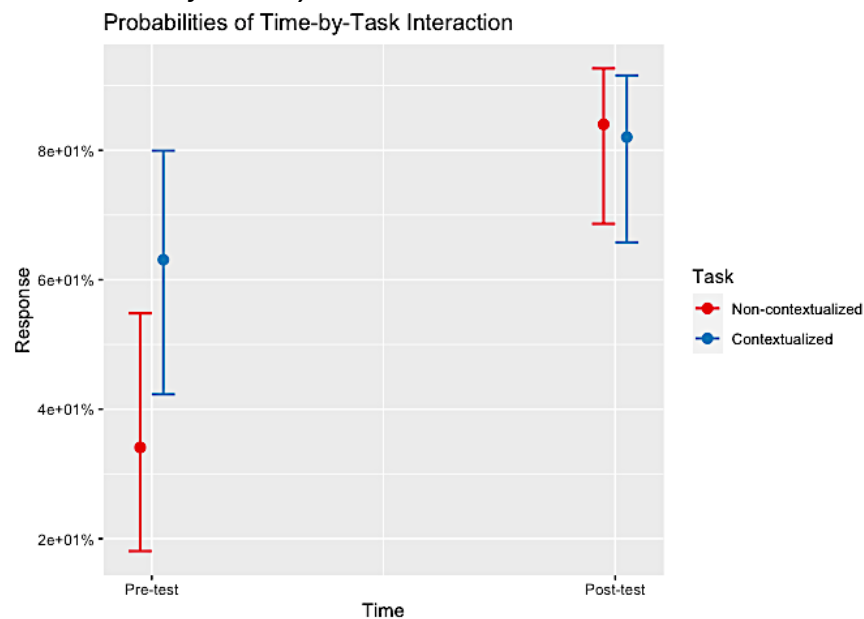
Table 1
Descriptive Statistics Results for Pre- and Post-Tests

	% Correct Answers	Mean	SD
Pre-Test (sentence level)	38%	16	10.23
Pre-Test (contextualized)	59%	26	5.64
Post-Test (sentence level)	76%	33	7.53
Post-Test (contextualized)	75%	33	2.28

Table 2*Results of Overall Model Comparing Pre- and Post-Tests*

Predictors	Response		
	Odds Ratio	CI	<i>p</i>
(Intercept)	0.52	0.22 – 1.21	0.129
Time (post-test)	10.16	6.04-17.08	<0.001
Task (contextualized)	3.30	1.13-9.66	0.029
Time [Post-test] * Task [Contextualized]	0.26	0.13 – 0.52	<0.001
Random Effects			
σ^2	3.29		
τ_{00} Participant	1.68		
τ_{00} Item	0.61		
ICC	0.41		
N Participant	44		
N Item	10		
Observations	880		
Marginal R ² / Conditional R ²	0.134 / 0.489		

Additionally, the findings suggest that even though participants performed better in the paragraph, contextualized task in the pre-test, in the post-test, there was improvement in both tasks, with students exhibiting better performance in the non-contextualized task (Figure 1).

Figure 1*Probabilities of Time-by-Task Interaction*

Analysis also revealed that some verbs might have affected the overall results. For example, *almorzar* (to have lunch; example 2a in instrument section) appeared to have performed differently than the other test items in the pre-test (34 students chose the correct option). This could have been caused by participants' previous exposure to the noun *almuerzo*, which has the same form as the verb used in the study's instruments. Nevertheless, the inclusion of this verb did not seem to have affected the overall results significantly (Table 3).

Table 3

Results of Overall Model Without the Item Almorzar

Predictors	Response		
	Odds Ratio	CI	<i>p</i>
Time (post-test)	12.42	7.05-21.87	<.001
Task (contextualized)	5.87	2.72-12.68	<.001

DISCUSSION

The results of this study suggest that PACE can have a beneficial effect on the development of L2 Spanish learners' declarative knowledge or conceptual understanding of the rules guiding stem-changing forms, which mirrors the findings reported by Groeneveld (2011) and Ngo (2018). However, the present work offers a more comprehensive statistical analysis than previous investigations, and it also appears to point to the role that context might play in L2 learning. That is, the data analysis revealed that participants performed better in the contextualized task in both the pre-and post-tests, which might point to the beneficial effects of context for L2 learning that have been reported in previous studies (e.g., Beheydt, 1987; Bolger & Zapata, 2011). This is the case when the context used is transparent enough to offer students a clear semantic environment for them to accurately infer what a word or structure might be referring to (Frantzen, 2003). In the case of this work, the passage in the pre-and post-tests was based on a story connected to the topic with which students were working (houses; a person moving to a new house) and included previously learned L2 vocabulary. Thus, thematic and linguistic familiarity might have helped the participants in the completion of the task and might have resulted in a more accurate performance in the pre-test than in the sentence-level portion of the test (Figure 1). Also, the reported standard deviation values (Table 1) suggest a more consistent performance among the participating students in the contextualized paragraph task in both pre-and post-tests, which might point to learners' reliance on context for the determination of their choices.

PACE also benefitted the participants by providing them with opportunities for active L2 use in the interpretive, interpersonal, and presentational modes of communication. For example, learners were exposed to an authentic text both in oral and written form, which was complemented by three other semiotic modes (visual, gestural, and auditory), and they were able to use the L2 for comprehension and interpretation. Students then employed the target

language (in combination with their mother tongue) to co-construct rules collaboratively with both peers and instructor and to discover connections between meaning and form. Additionally, in the extension phase, the participants worked in groups resorting to the L2 to discuss, write, and orally present possible conclusions to the story told in the presentation stage. Clearly, this approach to grammar was not based on isolated, decontextualized practices; instead, it offered beginning Spanish students chances for L2 use in controlled, but meaningful (in terms of student needs, topic, and tasks) instructional interventions. Finally, like previous work, based on the participating instructor's informal observations, the learners in this work appeared to have enjoyed the approach, which transpired in the active behavior and L2 use they exhibited in all of the PACE phases.

Limitations and Suggestions for Future Research

The most important limitation of this work lies in its sources of data. Even though the researchers planned to include a delayed post-test and interviews with participating students, this plan was thwarted by COVID-19, which not only disrupted face-to-face classes but also brought high anxiety to students and instructors. In an educational environment characterized by uncertainty and fear, it was not deemed appropriate to continue with the study as planned. Future research, however, should incorporate other sources of data such as interviews, think-aloud protocols, and the analysis of students' products to provide more comprehensive evidence in support of PACE for L2 grammar teaching. A second limitation is the lack of comparison between PACE and other approaches for grammar teaching. Although this was not what the researchers originally intended for this work, it would be interesting to investigate whether PACE can offer the same or more effective benefits than other ways of incorporating grammar into L2 classes. A final limitation is connected to some of the items that were included in the pre-test, such as *almorzar*. When developing the instruments for this study, the authors did not consider the possibility that, even though the verb had not been formally introduced in class, students might have encountered it before. Therefore, in future research, the frequency of certain verbs and learners' possible exposure to them might be considered a factor that could affect results.

CONCLUSION

This study sought to contribute to the scarce empirical literature on PACE in L2 classes by focusing on Spanish and an L2 student population not studied previously. Even though the researchers were not able to complete the study as intended due to the COVID-19 mandatory quarantine, this work still offered evidence for the effectiveness of the approach for the development of L2 beginning Spanish learners' declarative knowledge of stem-changing verbs in the present tense. Additionally, the incorporation of PACE offered the participating learners the opportunity to use the L2 in interpretive, interpersonal, and presentational communication to discuss a topic related to their lifeworld and lived experiences. Also, students' exposure to and analysis of an authentic, multimodal text might have contributed to their growth as multiliterate L2 speakers and appeared to have positively influenced their level of class participation. Therefore, despite its limitations, this study could be regarded as a first step in the much-needed

investigation of PACE in L2 Spanish instruction. Since this approach is considered a core practice in L2 teaching and its reliance on text is compatible with current literacy-based L2 approaches, it is important to continue discovering how it can contribute to L2 development and use.

REFERENCES

- Adair-Hauck, B., & Donato, R. (2002). The PACE model: A story-based approach to meaning and form for standards-based language learning. *The French Review*, 76(2), 265–276.
- Agresti, A. (2002). *Categorical data analysis* (2nd ed.). Wiley.
- Allen, H. W., & Paesani, K. (2010). Exploring the feasibility of a pedagogy of multiliteracies in introductory foreign language courses. *L2 Journal*, 2(1), 119–142. doi:10.5070/L2219064
- Beheydt, L. (1987). The semantization of vocabulary in foreign language learning. *System*, 15(1), 55–67. doi:10.1016/0346-251X(87)90048-0
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi:10.18637/jss.v067.i01
- Bolger, P. A., & Zapata, G. C. (2011). Semantic categories and context in L2 vocabulary learning. *Language Learning*, 61(2), 614–646. doi:10.1111/j.1467-9922.2010.00624.x
- Bowden, H. W., Gelfand, M. P., Sanz, C., & Ullman, M. T. (2010). Verbal inflectional morphology in L1 and L2 Spanish: A frequency effects study examining storage versus composition. *Language learning*, 60(1), 44–87. doi:10.1111/j.1467-9922.2009.00551.x
- Branson, C. (2015). *Quliangnanek Litnauwilita - Let's teach through stories* [Unpublished master's thesis]. University of Alaska Fairbanks.
- Carrell, P. L. (1984). Evidence of a formal schema in second language comprehension. *Language Learning*, 34(2), 87–112. doi:10.1111/j.1467-1770.1984.tb01005.x
- Cubillos, J. H. (2014). Spanish textbooks in the U.S.: Enduring traditions and emerging trends. *Journal of Spanish Language Teaching*, 1(2), 205–225. doi:10.1080/23247797.2014.970363
- Doughty, C., & Williams, J. (1998). *Focus on form in classroom second language acquisition*. Cambridge University Press.
- Fernández, C. (2011). Approaches to grammar instruction in teaching materials: A study in current L2 beginning-level Spanish textbooks. *Hispania*, 94(1), 155–170.
- Frantzen, D. (2003). Factors affecting how second language Spanish students derive meaning from context. *Modern Language Journal*, 87(2), 168–199.
- Glisan, E. W., & Donato, R. (2017). *Enacting the work of language instruction: High-leverage teaching practices*. ACTFL.
- González-Bueno, M. (2021). S-PACE to teach L2 Grammar: Adding structure to the PACE model. *Applied Language Learning*, 31(1–2), 29–49. <https://eric.ed.gov/?id=EJ1289689>
- Groeneveld, A. (2011). Adjective or adverb? Teaching grammar with the PACE model. *Levende Talen Tijdschrift*, 12(3), 24–32.
- Harris, E. (2017). *Bringing Twygs to life: PACE based lessons in an adult ESL classroom* [Unpublished master's thesis]. University of Alaska, Fairbanks.
- Kalantzis, M., Cope, B., Chan, E., & Dalley-Trim, L. (2016). *Literacies*. Cambridge University Press.
- Kalantzis, M., Cope, B., & Zapata, G. C. (2019). *Las alfabetizaciones múltiples: Teoría y práctica*.

Octaedro.

- Li, H., & Paul, J. (2019). A usage-based approach to L2 grammar instruction delivered through the PACE model. In C. Shei, M. E. McLellan Zikpi, & D.-L. Chao (Eds.), *The Routledge handbook of Chinese language teaching* (pp. 257–271). Routledge.
- Miguel, N. M., & Vázquez Solana, A. (2016). Implementación de la metodología PACE en el aula de alemán L2. In B. Balzer Haus & I. Szumlakowski Morodo (Eds.), *La lengua alemana vista desde dentro y desde fuera: Estudios sobre su sistema, su enseñanza y su recepción Die deutsche Sprache - intern und extern Untersuchungen zu System, Vermittlung und Rezeption* (pp. 201–210).
- Ngo, H. (2018). *The use of storytelling in online foreign language learning: A case study of a basic English grammar course provided to Vietnamese adult e-learners* [Unpublished master's thesis]. University of Lapland.
- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching*. Cambridge University Press.
- Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence and education* (pp. 33–58). Erlbaum Associates.
- Savignon, S. J. (1983). *Communicative competence: Theory and classroom practice*. Addison-Wesley.
- Savignon, S. J. (1987). Communicative language teaching. *Theory into Practice*, 26(4), 235–242. doi:10.1080/00405848709543281
- Sessarego, C. (2016). A discourse-pragmatic approach to teaching indicative/subjunctive mood selection in the intermediate Spanish language class: New information versus reformulation. *Hispania*, 99(3), 392–406.
- Vogel, S., Herron, C., Cole, S. P., & York, H. (2011). Effectiveness of a guided inductive versus a deductive approach on the learning of grammar in the intermediate-level college French classroom. *Foreign Language Annals*, 44(2), 353–380. doi:10.1111/j.1944-9720.2011.01133.x
- Wong, W. (2005). *Input enhancement: From theory and research to the classroom*. McGraw-Hill.
- Yoon, J. (2019). Teaching and learning vocabulary as L2: Approaches in Spanish textbooks. *Language Teaching and Educational Research*, 2(2), 114–131. doi:10.35207/late.647156
- Zapata, G. C. (2022). *Learning by design and second language teaching: Theory, research, and practice*. Routledge.

AUTHORS

Ewurama Okine, Doctoral student at Texas A&M University, eokine@tamu.edu.

Gabriela C. Zapata, Ph.D., University of Nottingham, Gabriela.Zapata@nottingham.ac.uk.

COMMENTARY

Leveraging Artificial Intelligence for Assigning ILR Ratings to Authentic Content

Jordan M. Eason

Doctoral Student, University of Coimbra, Portugal

When teaching Brazilian Portuguese and other languages for the Department of Defense (DoD), the decision on which authentic texts to choose can be difficult. One complicating factor is choosing an authentic text that is at the correct proficiency level for the students—not too challenging and not too easy. The DoD uses the Interagency Language Roundtable (ILR) levels to indicate the complexity of any text or listening passage. Learning how to “level” (i.e., assign a rating score) a passage according to the ILR is a skill that must be developed and takes a significant amount of time. This paper leverages experience working for the DoD as a member of a team that was assigned the task to level audios, videos, and text; during this process, differences in opinions among members occurred regarding whether a passage was ratable/unratable, or able/unable to fit in a specific level on the ILR scale. I argue that machine learning processes available in artificial intelligence (AI), specifically a natural language processing (NLP) platform such as Open AI’s ChatGPT or Google’s Bard, among other AI chatbots, offer human raters a tool that can assist them in increasing their efficiency while removing potential subjectivity from the leveling process.

Keywords: *artificial intelligence, AI chatbots, ChatGPT, leveling content, ILR scale*

INTRODUCTION

Artificial intelligence (AI) systems are producing communication responses that are increasingly similar to those produced by humans. Although they cannot yet replace humans, these systems can facilitate and help language professionals do complex tasks. One such complex task that AI can assist with is the leveling (i.e., assigning a rating or score) of language texts from authentic written sources according to the descriptions established by the Interagency Language Roundtable (ILR). The pilot study outlined in this paper provides a glimpse of AI’s potential to assist the language professional in assigning levels to texts in the DoD context. With time, AI will

likely be able to support human raters as they determine ILR reading levels for authentic texts, bringing efficiency and reducing potential subjectivity.

BACKGROUND

Machine Learning and Natural Language Processing

AI chatbots such as OpenAI's ChatGPT or Google's Bard are becoming increasingly well-known and utilized across all fields. These tools use a chat-like interface in which the user can ask the application just about anything and the application uses AI to answer the question. These AI tools are developing rapidly and will continue to improve in quality and accuracy over time. The most common method AI chatbots use patterns of language to train is via machine learning algorithms (MLAs). MLAs are used to analyze a large corpus of annotated text data. This algorithm is then trained to use these patterns to predict the sentiment of new, unannotated text. Sentiment analysis, which involves analyzing the language used in text to determine its emotional tone, is a Natural Language Processing (NLP) technique.

For some languages such as Brazilian Portuguese, where high-quality training data exists, the corpora used to train MLAs can be fine-tuned to a specific domain or context, which further improves the ability of the AI chatbot to interpret sentiments in each domain. For example, the AI chatbot may take into consideration specific terms that trigger sarcasm such as, "*Pois é*," in Brazilian Portuguese which is equivalent to the term "sure" in English. Furthermore, it can use contextual factors, such as the tone of the overall text, the relationship between parties in a conversation, and the subject matter of the text to determine the emotional undertones of a message. Although limited by their finite nature in reference to an infinite world, AI chatbots may be considered more capable than limited in their ability to "read between the lines." Considering that humans are imperfect, can be irrational, and have diverse educational backgrounds and world experiences, one may argue that an AI chatbot's ability to "read between the lines" and provide consistent analysis makes it more capable than most humans because of the AI's ability to collect and process information from an infinite repository of online information.

AI chatbots use NLP platforms to understand and respond to human language. NLP platforms have been used to tackle minor tasks for language professionals and teachers of foreign languages because of NLP's capabilities with language recognition and manipulation. NLP is capable of differentiating text types to include some intentions from authors. Academics such as Guren (2022) have demonstrated that NLP platforms can be used to simplify difficult-to-understand technical texts and make them more comprehensible. This relates to AI rating levels because it confirms that NLP platforms can identify and recognize the level of difficulty and even styles in various texts. If NLP is used to manipulate Brazilian Portuguese and other languages when an adequate amount of data is present in a specific manner, it can be used to recognize the difficulty and assign an ILR level as well. Despite achieving great success with the lexical simplification process, Guren (2022) does not claim that AI is a solution or remedy for all situations. Notably, additional data and surveys were needed to explain technical terms and broaden the fields of study. This means NLP could have equal challenges with technical terms and idiomatic expressions as well, despite the major advances in NLP since 2019. In his survey of

automated text simplification tools, Shardlow (2014) discusses how these tools can identify complex texts and generate substitutions that are determined to be simpler, focusing on lexical and syntactic simplification as well as text simplification and machine translation in languages other than English. Similarly, Bott et al. (2012) demonstrated how structural simplification operations can be accomplished for Spanish. In English, popular sites such as Grammarly and Rephrasely use NLP to make suggestions for students' papers. Again, to successfully substitute, the NLP must first recognize what is classified as a particular style or difficulty. This is what the NLP must do to rate the language level of the excerpt as well. These systems can provide summaries of a writer's tone or style and can also make suggestions on how to revise a paragraph to make it sound more "academic, human, simple, clear, elaborate, creative, formal, cohesive, confident, smart, or thoughtful" (Rephrasely.com). Although it is not possible to truly understand what an author intends, we can argue that most would agree on some specific meaning such as when trying to classify them as creative, formal, etc.

AI and NLP tools are also increasingly able to identify and analyze the inferential meaning of words in a text. Shaik et al. (2022) discuss how AI and NLP use feature extraction, feature selection, topic modeling, and text evaluation to analyze texts (in this case, student feedback from a course), including how these analyses can identify and interpret challenging language features such as sarcasm or ambiguity. Similarly, Nenkov et al. (2016) demonstrated how the integration of NLP platforms with tutoring systems can be used to improve learning management systems (LMSs) by providing language feedback via a chatbot. This confirms that NLP is capable of differentiating text types and, to some degree, interpreting authors' intentions.

Classification of complex texts is a critical step in leveling and one that can be supported by AI as well. Dzikovska et al. (2014) discuss the potential of NLP techniques for students in STEM areas such as electricity and electronics, fields that are useful for military electronics training and therefore relevant for DoD texts and related tasks. Dzikovska et al.'s research, which focused on developing an intelligent tutoring system, utilized automatic text classification and information extraction. It shows that automatic text classification can assist language professionals to classify texts, which is one of the critical decisions content developers and ILR text-level raters make daily.

Thus, the above studies provide support that NLP platforms can successfully differentiate between complex and simple texts, can simplify language, and to some degree can interpret subtleties in human communication such as sarcasm. They are also increasingly being utilized for their potential to make teaching and learning languages more efficient. In fact, the Department of Defense has created, along with MIT, an online tool for leveling texts. The Auto-ILR was developed in 2017 and automatically levels and classifies texts for DoD curricular contexts. It aids teachers and students as they assign levels to texts, making their teaching and studying processes more efficient. At this time, however, this tool does not appear to be accessible to the public. As can be seen, AI and tools such as ChatGPT or Google Bard have the potential to support DoD language professionals in terms of consistency and efficiency as they assign ILR levels to authentic texts.

Text Leveling in the Department of Defense

Text leveling in the Department of Defense is performed by human language professionals. While overall the process is accurate and consistent, there are some complications due to staffing shortages, inter-rater reliability, and other considerations. For example, some of the language professionals doing the rating are not highly proficient in the language for which they provide ILR levels. These professionals use tools to translate between English and the target language to determine the ILR level. Given the complexities and subtleties of language, this is not an ideal solution. Another challenge in rating passages objectively arises when a team working to assign a level to a text disagrees on what the appropriate rating should be, which reflects the natural complexity of authentic language. A team leader ultimately makes the final rating decision, which introduces subjectivity; depending on the individual and their training and background with the language and project, this may be significant. However rare this occurrence may be, it does highlight the potential challenges introduced by human subjectivity and the opportunity for AI to assist human language experts as they make leveling decisions.

OPPORTUNITIES FOR AI AS A TOOL TO ASSIGN ILR LEVELS

The Interagency Language Roundtable Scale

The intricacies of the ILR scale are important in this pilot study. For the sake of national security and organizational confidentiality, this paper will focus only on the information given on the ILR website (<https://www.govtilr.org/>). The ILR scale is briefly described in the table below:

Table 1

The Interagency Language Roundtable Scale

Scale Number	Description
0	No proficiency
0+	Memorized Proficiency
1	Elementary Proficiency
1+	Elementary Proficiency, Plus
2	Limited Working Proficiency
2+	Limited Working Proficiency, Plus
3	General Professional Proficiency
3+	General Professional Proficiency, Plus
4	Advanced Professional Proficiency
4+	Advanced Professional Proficiency, Plus
5	Functionally Native Proficiency

Each level has a more detailed description. To demonstrate the subtle differences between levels, we will analyze the different definitions for Levels 2 and 2+ Reading. A more complete

description of ILR level 2 Reading is found in Figure 1. (note that this reading descriptor is from the most recent ILR scale).

Figure 1

ILR Level 2 Reading Descriptor

“Sufficient comprehension to read simple, authentic written material in a form equivalent to usual printing or typescript on subjects within a familiar context. Able to read with some misunderstandings straightforward, familiar, factual material, but in general insufficiently experienced with the language to draw inferences directly from the linguistic aspects of the text. Can locate and understand the main ideas and details in material written for the general reader. However, persons who have professional knowledge of a subject may be able to summarize or perform sorting and locating tasks with written texts that are well beyond their general proficiency level. The individual can read uncomplicated, but authentic prose on familiar subjects that are normally presented in a predictable sequence which aids the reader in understanding. Texts may include descriptions and narrations in contexts such as news items describing frequently occurring events, simple biographical information, social notices, formulaic business letters, and simple technical material written for the general reader. Generally the prose that can be read by the individual is predominantly in straightforward/high-frequency sentence patterns. The individual does not have a broad active vocabulary (that is, which he/she recognizes immediately on sight), but is able to use contextual and real-world cues to understand the text. Characteristically, however, the individual is quite slow in performing such a process. Is typically able to answer factual questions about authentic texts of the types described above” (govtilr.org).

The definition has a strong presence of subjectivity. For example, several terms in the definition are ambiguous. It’s difficult to determine what exactly is meant by the following terms: “familiar context,” “some misunderstandings,” “uncomplicated, but authentic prose,” “quite slow,” or “typically able.” One may find it difficult to pinpoint what is considered “familiar” or quantify “some.” This is particularly true for the professional language community because of a diverse set of backgrounds. This reveals an advantage of using AI when rating authentic texts. AI will define such terms consistently throughout any project.

Figure 2 shows the ILR description for the 2+ Reading level.

Figure 2

ILR Level 2+ Reading Descriptor

“Sufficient comprehension to understand most factual material in non-technical prose as well as some discussions on concrete topics related to special professional interests. Is markedly more proficient at reading materials on a familiar topic. Is able to separate the main ideas and details from lesser ones and uses that distinction to advance understanding. The individual is able to use linguistic context and real-world knowledge to make sensible guesses about unfamiliar material. Has a broad active reading vocabulary. The individual is able to get the gist of main and subsidiary ideas in texts which could only be read thoroughly by persons with much higher proficiencies. Weaknesses include slowness, uncertainty, inability to discern nuance and/or intentionally disguised meaning” (govtilr.org).

Again, the 2+ Reading level uses ambiguous terms. For example, the following terms are particularly ambiguous: “non-technical prose,” “familiar topic,” “real world knowledge,” “sensible guesses,” “unfamiliar material,” “broad active reading vocabulary,” “slowness,” and “disguised meaning.” The term “slowness” may be a comprehension response time of three seconds for some readers while others may think it as a different value. The term may be defined in an entirely different fashion. Similarly, what is a “familiar topic” to one individual may be unfamiliar to another when cultural experiences and interests are taken into account. The use of AI has the potential to lessen this subjectivity, which is preferred in ILR ratings.

The definition of Reading Level 3 includes statements shown in Figure 3.

Figure 3

ILR Level 3 Reading Descriptor

“Text-types include news stories like wire service reports or international news items in major periodicals, routine correspondence, general reports, and technical material in his/her professional field; all of these may include hypothesis, argumentation and supported opinions. Misreading rare. Almost always able to interpret material correctly, relate ideas and “read between the lines,” (that is, understand the writers’ implicit intents in text of the above types).” (govtilr.org)

A fundamental question for higher levels on the ILR scale is whether the AI chatbot itself can “read between the lines.” At this point, it appears that it can, and given the nature of AI and its potential to improve, it is likely that its ability to “read between the lines” will also be more consistent than with human levelers. An important step in the process of working with AI is to determine the degree to which the systems have access to the information they need. It therefore might be prudent to also query the AI chatbot about what it knows about the ILR, perhaps at specific levels, to ensure that the system can provide accurate information.

Limitations of AI Chatbots

AI chatbots are naturally limited by the input that they receive. They are only capable of learning from the input and language patterns that they are fed, and some language patterns are more complex and difficult than others. Content developers for the DoD or those who have personally taken a Defense Language Proficiency Test (DLPT) understand the importance of being able to “read between the lines” in the target language. For example, one may be asked when a writer is being indifferent, sarcastic, or provocative among other emotions and sentiments. To explore this issue further, I used Open AI’s ChatGPT to determine if it can understand various emotions expressed in text. It returned the following answer to the given question (see Figure 4).

Figure 4***ChatGPT Example Demonstrating “Reading Between the Lines”***

<p><i>User</i></p> <p>“Given this text in Brazilian Portuguese, what is the tone and emotion that the author intends to convey in the following text: “Não interessa se foi A, B ou C, todo o episódio foi como uma facada nas minhas costas”</p>
<p><i>ChatGPT</i></p> <p>“Based on the given text in Brazilian Portuguese, the tone and emotion that the author intends to convey can be described as one of betrayal, disappointment, and possibly anger. The phrase “Não interessa se foi A, B ou C” suggests a sense of indifference or dismissal towards the specific individuals involved. However, the following statement “todo o episódio foi como uma facada nas minhas costas” (translated as “the whole episode was like a stab in the back”) conveys a strong feeling of betrayal, implying that the author feels deeply hurt and betrayed by the situation. The use of the metaphorical expression “facada nas minhas costas” emphasizes the emotional impact and the sense of personal betrayal experienced by the author. Overall, the author’s tone appears to be one of bitterness and resentment towards the events described.”</p> <p>(https://chat.openai.com/)</p>

While this particular assessment by ChatGPT of the tone and emotion in this passage is consistent with my interpretation as a native Brazilian Portuguese speaker, it is important to keep in mind that the ability of ChatGPT and AI tools to provide accurate assessments will vary depending on the language and also on the type of input the AI tool has been trained on up to this point.

Pilot Study

For this pilot study, I compared the levels produced by human raters and Open AI’s ChatGPT chatbot for nine authentic texts available in the Defense Language Institute’s (DLI) Global Language Online Support System (GLOSS, <https://gloss.dlilfc.edu/>). The texts chosen for this pilot study were written in Brazilian Portuguese and were all well under the maximum amount of text allowed in ChatGPT. GLOSS includes the official, human-rated ILR level for each of these texts. For the ChatGPT level, I entered each text into ChatGPT and asked for it to assign an ILR level using the following request:

“Based on the Interagency Language Roundtable scale, please rate the following passage in Brazilian Portuguese for the Department of Defense context: ‘(INSERT TEXT HERE)’”

Table 2 shows the title of the texts, the GLOSS human-leveled rating, the ChatGPT AI-leveled rating, and the difference between the two (i.e., Delta). Note that the GLOSS passages included in this pilot study could have been rated under a previous version of the ILR Reading scale.

Table 2*GLOSS and ChatGPT Ratings of Texts*

Title of Authentic Text (year)	GLOSS Rating (human-rated)	ChatGPT Rating (AI-rated)	Delta
Dilma Appoints New Commanders of the Armed Forces (2015)	1	2+	1.5
142 Million Voters Can Check Polling Places (2014)	1	2	1
Brazilian Airlines in the Air (2013)	1	2	1
CIA Documents about the Dictatorship (2018)	2	2	0
The Kidnapping of Charles Burke Elbrick (2015)	2	2	0
Farming in Southern Brazil (2016)	2	2	0
The Controversy of Cracolândia (2017)	3	3	0
Scientific Experiment Tests Ethical Limits (2017)	3	3–4	0.5
The Unconscious Mind (2017)	3	4	1

The average delta between human- and machine-assigned ratings in this sample is 0.56. with ChatGPT rating texts at least half a level higher than the DoD human experts. Also, out of nine texts compared, only four have the same rated level in both human and AI ratings. Based on my experience in content creation for the DOD, the deltas seem to be an accurate representation of potential variations of ratings assigned by human experts.

One of the benefits of using ChatGPT for leveling is that it explains why it chose a specific level and provides specific examples. For example, when asked why the authentic text in Brazilian Portuguese entitled “The Unconscious Mind” was rated as Level 4 on the ILR scale, ChatGPT responded with these details after demonstrating it understood what Level 4 was defined as (see Figure 5).

Figure 5*ChatGPT's Explanation of Rating The Unconscious Mind at ILR 4*

"In the given text, the interviewees discuss the theories and contributions of Sigmund Freud, including his concept of the unconscious, the interpretation of dreams, and the importance of understanding the meaning behind the symptoms and behaviors. The interview also touches on the history and evolution of psychiatry and psychoanalysis. The conversation contains several characteristics that support a Level 4 rating: 1. Abstract and complex vocabulary: The text includes advanced vocabulary related to psychology and psychiatry, such as "neuroses," "associação livre," "psicoterapêutico," "Psiquiatria dinâmica," "processos psíquicos," "sintomas," and "delírio." These terms may not be familiar to the average person and require a higher level of proficiency to understand. 2. Long and complex sentences: The text includes several long and complex sentences that require advanced grammatical structures to understand. For example, "A interpretação dos sonhos, ela vai fazer uso da associação livre, que é um dispositivo a mais que a gente tem no atendimento." 3. Use of idiomatic expressions: The text includes several idiomatic expressions, such as "chamar uma pessoa por outro nome," which require a deep understanding of the language and its nuances. 4. Discussions of complex and abstract concepts: The text delves into complex and abstract topics related to psychology, including the unconscious, the interpretation of dreams, and the meaning behind symptoms and behaviors. These topics require a high level of intellectual engagement and critical thinking to understand fully." (<https://chat.openai.com/>)

This response highlights how AI can be a useful tool for DoD content developers and foreign language instructors because of its ability to highlight the language that would qualify a text as a specific level using the ILR definitions.

ChatGPT highlighted ambiguity issues in rating for the text entitled "Scientific Experiment Tests Ethical Limits" when it returned an ILR level range of 3–4 for this particular text. One can conclude that because ChatGPT produced a range, it shows the tool's ability to interpret the definitions of the various ILR levels; this range also highlights the general difficulty of assigning a level to a text. Given that having ILR level ranges for a given text is often not permitted, one can assume that outputs with score ranges are equivalent to what human levelers would determine "unratable texts."

CONCLUSION

NLP platforms such as AI chatbots show potential in their ability to support human language professionals who assign ILR levels to texts. Because of the limitations of AI, it is not recommended that these tools replace human levelers at this time, but rather that professionals explore ways to use AI as a type of "personal assistant" when making rating decisions. AI offers the potential benefits of increased efficiency and consistency, decreased subjectivity, and a starting point for further evaluation. It also provides extensive explanations of how a machine-assigned ILR rating was determined. The small pilot comparison between ChatGPT-assigned and human-assigned ratings presented in this article can be used to inform future research and development. For example, texts in languages other than Brazilian Portuguese can be compared to gather baseline information on similarities and differences in the ratings given. Also, research can be conducted on the specific needs of human language professionals and how AI might help

meet those needs. AI developers may benefit from collaboration with DoD professionals to provide input for AI algorithms that will enhance the training of AI in this area. Tools such as chatbots have potential implications for the classroom as well. Teachers can use them to determine the ILR level of a text they want to use in class, or to find an authentic text at a certain ILR level. Research is therefore warranted on how best to engage teachers in professional development for using AI in the classroom. Over time, opportunities for AI to contribute to the teaching and learning process will increase and be limited only by our imaginations.

REFERENCES

- Bott, S., Saggion, H., & Mille, S. (2012). *Text simplification tools for Spanish*. Proceedings of the eighth international conference on language resources and evaluation (LREC'12), Istanbul, Turkey. http://www.lrec-conf.org/proceedings/lrec2012/pdf/762_Paper.pdf
- Dzikovska, M., Steinhauser, N., Farrow, E., Moore, & J., Campbell, G. (2014). BEETLE II: Deep natural language understanding and automatic feedback generation for intelligent tutoring in basic electricity and electronics. *International Journal of Artificial Intelligence in Education*, 24(3). 284–332. <https://doi.org/10.1007/s40593-014-0017-9>
- Guren, E. G. (2022). *Exploring automatic text simplification of requirements*. [Masters thesis. University of Oslo]. https://www.duo.uio.no/bitstream/handle/10852/95635/8/thesis_final.pdf
- Nenkov, N., Dimitrov, G., Dyachenko, Y. & Koeva, K. (2016). Artificial intelligence technologies for personnel learning management systems. *2016 IEEE 8th International Conference on Intelligent Systems (IS)*. Sofia, Bulgaria, 2016, pp. 189–195, <https://doi.org/10.1109/IS.2016.7737420>
- Shardlow, M. (2014). A survey of automated text simplification. *International Journal of Advanced Computer Science and Applications*, 4(1). 58–70. <https://doi.org/10.14569/SpecialIssue.2014.040109>
- Shaik, T., Tao, X., Li, Y., Dann, C., McDonald, J., Redmond, P., & Galligan, L. (2022). A review of the trends and challenges in adopting natural language processing methods for education feedback analysis. *IEEE Access*, 10, 56720–56739. <https://doi.org/10.1109/ACCESS.2022.3177752>

AUTHOR

Jordan M. Eason is a doctoral student at the University of Coimbra, Portugal. A former active duty Air Force officer and Language Enabled Airman (LEAP), Mr. Eason has worked as a Portuguese instructor contractor for ICA Languages on behalf of the Defense Language Institute, the US Air Force Academy, the University of Maryland, and the University of Michigan.

THANK YOU, REVIEWERS

Applied Language Learning relies on expert reviewers for quality of the journal. Special thanks go to the individuals listed below. The publication of *Applied Language Learning* was made possible with their generous time and support.

Natalia Barley
Jean Chun
Grazyna Dudney
Myoyoung Kim
Youngae Kim
Hye-Yeon Lim
Tatiana McCaw
Yaniv Oded
Erin O'Reilly
Thomas Stovicek
Deanna Tovar
Sabine Wolpers
Rong Yuan
Shen S. Zhu

CALL FOR PAPERS

Applied Language Learning, a refereed journal published annually by the Defense Language Institute Foreign Language Center and Presidio of Monterey, is soliciting articles for publication.

The Journal (US ISSN 1041-679X and ISSN 2164-0912 for the online version) is to provide a forum for the exchange of ideas and information on

- Instructional methods and techniques
- Curriculum and materials development
- Assessment of needs within the profession
- Testing and evaluation
- Implications and applications of research from related fields such as linguistics, education, communications, psychology, and the social sciences

The journal seeks to serve the professional interest of language teachers, administrators, and researchers concerned with the teaching of foreign languages to adult learners. We welcome articles that describe innovative and successful practice and methods and/or report educational research or experimentation.

Deadline: Submissions are welcome at any point. Send your manuscript electronically to the Editor: journals@dliflc.edu.

View the **Guide for Authors** and read the recent and past issues of *Applied Language Learning* at <http://www.dliflc.edu/resources/publications/applied-language-learning/>

Send all questions, comments, and/or submissions to the Editor: journals@dliflc.edu.