DEAL 2023

MULTIPLE PERSPECTIVES IN ENGLISH APPLIED LINGUISTICS

EDITED BY ● **ATTILA M. WIND & BRIGITTA DÓCZI**

DEAL 2023:
Multiple Perspectives
in English Applied Linguistics

Edited by
Attila M. Wind & Brigitta Dóczi

DEAL 2023:
Multiple Perspectives
in English Applied Linguistics

An edited volume of the Department of English Applied Linguistics,
School of English and American Studies,
Faculty of Humanities,
Eötvös Loránd University, Budapest

Edited by

Attila M. Wind & Brigitta Dóczi

Budapest, 2023

Editors:
Attila M. Wind
Brigitta Dóczi

ELTE | EÖTVÖS UNIVERSITY PRESS

ELTE | FACULTY OF HUMANITIES

eltebook.hu

Table of Contents

Contributors

Ágnes Albert (PhD, ORCID ID 0000-0001-8339-7119) holds a PhD in Language Pedagogy. Her research interests include task-based language learning and individual differences in foreign language learning, with specific focus on learner creativity and the positive emotions associated with the learning of foreign languages. Recently, she was involved in a research project investigating secondary school learners' motivation, autonomy, and positive and negative emotions in connection with their English classes. She is currently a member of the MTA-ELTE Foreign Language Teaching Research Group.

Dina Al-Madanat (ORCID ID 0009-0001-9266-6285) holds an MA in Applied Linguistics and is currently a PhD student in Language Pedagogy and English Applied Linguistics. She has experience teaching English to both primary and secondary school students. As a bilingual herself, Dina's personal experience have sparked her interest in researching the relationship between bilingualism and emotions. She has conducted studies comparing emotional perception and expression in first and second languages.

Kholoud Al-Madanat is a PhD student at Eötvös Loránd University in the Language Pedagogy and English Applied Linguistics program. She has previously taught an Academic Skills course for undergraduate students at the university. Her primary research interest is English as a lingua franca. In her MA thesis, she investigated its use among international students in Hungary. Her PhD focus has shifted to exploring the implications of English as a lingua franca for language teaching and learning.

Duha Alsayed-Ahmad (ORCID ID 0000-0002-3492-4637) holds an MA in English Studies and is a PhD student in Language Pedagogy and English Applied Linguistics. Her research interests include translation studies and positive psychology, in particular language learners' experiences of flow. She was involved in a research project studying the flow experiences of EFL learners in writing tasks from different genres. She taught a Skills Devel-

opment course at ELTE, and she is currently a member of the MTA-ELTE Foreign Language Teaching Research Group.

Zsuzsanna Andréka (ORCID ID 0000-0001-6357-8895) is a PhD student in the Language Pedagogy and English Applied Linguistics PhD program at Eötvös Loránd University, Hungary. Her main research area is language testing. She is also a PhD student researcher member in an assessment project funded by the Hungarian National Research, Development, and Innovation Office.

Wijdene Ayed (ORCID ID 0000-0002-6411-2883) is a PhD student in the Language Pedagogy and English Applied Linguistics program at Eötvös Loránd University, Hungary. Her research interests include language policies and implementation and language-in-education development. She taught skills development and presentation skills courses at ELTE. She has been a leading mentor at the Faculty of Humanities of ELTE, PhD students' representative, and ELTE international students' ambassador. She has also been the Tunisian PhD students' ambassador in Hungary with the Association of Hungarian PhD and DLA candidates.

Kata Csizér (ORCID ID 0000-0003-1755-8142) is a Full Professor at the Department of English Applied Linguistics at Eötvös Loránd University, Budapest. Her main field of research interest is the social psychological aspects of L2 learning and teaching, as well as foreign language motivation. She has published over 100 academic papers and has co-authored several books, including the recent *Palgrave Macmillan handbook of motivation for language learning* with Martin Lamb, Alastair Henry and Stephen Ryan. She is editor of the *Journal of Second Language Learning and Teaching*.

Brigitta Dóczi (PhD, ORCID ID 0000-0002-1581-0079) has been teaching various courses at the Department of English Applied Linguistics since 1999. She obtained her PhD degree in 2012. Her main fields of interest are second language vocabulary development as well as task-based teaching and materials design. She has extensive experience in teacher training, while still actively teaching English as a foreign language. Apart from

offering courses related to vocabulary research, she also teaches research methodology to PhD students at the Faculty of Social Sciences. Currently she is working on a monograph about research in vocabulary learning and development in the 21st century.

Árpád Farkas (ORCID ID 0000-0002-2504-7046) teaches skills and English language modules at the International Business School in Budapest and has also worked in the Department of English Applied Linguistics at Eötvös Loránd University. In addition to teaching, Árpád has been pursuing a PhD in the Language Pedagogy Programme of Eötvös Loránd University. Árpád's professional interests include pragmatics and English as a lingua franca, and his current research focuses on the use of English in international classroom contexts.

Thet Oo Khaing (ORCID ID 0009-0001-7529-8978) is a PhD student in the Language Pedagogy and English Applied Linguistics program at Eötvös Loránd University, Hungary. She taught *Introduction to Applied Linguistics* and *Academic Skills* courses at ELTE voluntarily as an external lecturer. Her research interests include language pedagogy, second language vocabulary development, Dynamic Systems Theory, and metacognition. She is currently conducting research on the dynamics of L2 receptive and productive academic vocabulary development and vocabulary learning strategies of EFL learners. She has presented her research papers at local and international ELT and Applied Linguistics conferences.

Dávid Smid (PhD, ORCID ID 0000-0003-4484-8591) is a Postdoctoral Researcher at the Department of ELT Research and Methodology at the University of Graz in Austria. His research interests include language learner and language teacher psychology, language teacher education, and motivation. His publications include a monograph entitled *Toward an understanding of pre-service English teachers' motivation: The case of Hungary* published by Akadémiai Kiadó (2022). He has taught various courses at the tertiary level in Hungary, Japan, and Austria. Currently, he serves as the Associate Editor of the *Journal for the Psychology of Language Learning*.

Gyula Tankó (PhD, ORCID ID 0000-0002-0721-8171) is an associate professor at the Department of English Applied Linguistics at Eötvös Loránd University, Budapest, and has also taught in upper-primary, secondary, and higher education institutions in Italy, Germany, the UK, and Romania. He teaches and researches Language Assessment, English for Academic Purposes, and Mediation Studies at BA, MA, and PhD level. He has published books and articles on language assessment and academic writing, and is the principal researcher in a large-scale assessment project funded by the Hungarian National Research, Development and Innovation Office (Hungarian Scientific Research Fund, OTKA; Project number: 142536).

Attila M. Wind (PhD, ORCID ID 0000-0002-2702-5211) is an assistant professor at the Department of English Applied Linguistics at Eötvös Loránd University, Budapest. He completed his PhD at Lancaster University (UK) in 2019. His research interests are in the area of second language development, more broadly, with a special focus on second language writing development. He is also interested in the application of Complex Dynamic Systems Theory to the study of second language development and corpus linguistics. He is currently a member of the MTA-ELTE Foreign Language Teaching Research Group.

Editorial

We are proud to present the next reviewed volume in the applied linguistics series of the Department of English Applied Linguistics, Eötvös Loránd University, the fourth in line. The initiative was proposed by our late Head of Department, Enikő Öveges five years ago and we would like to dedicate this volume to her memory. We miss her very much! The aims of the current collection are based on the original concepts: to further our understanding of the interdisciplinary research work done in our department, to share our research results with our students and fellow researchers, as well as to develop our personal communication and review skills. We hope that we have managed to meet these aims and the articles of the book will reflect this variety.

This year's volume is based on a variety of theoretical and empirical research carried out by members of the department, as well as in collaboration with their PhD students, and some of our PhD students have submitted their independent work. We present very diverse strands in applied linguistics through which we try to provide further insights into our field: individual differences, the skill of listening, translation, ELF, testing, and corpus linguistics.

We start the collection with Smid and Csizér's work. In their article, they provide a critical synthesis of research on extracurricular language use both in the home environment as well as in study abroad contexts. Focusing on different individual variables, their analysis attempts to outline possible future research directions within Hungarian contexts to aid language learning.

In an exploratory study, D. Al-Madanat and Albert investigate how bilinguals perceive the emotional significance of their languages, using swearwords and taboo words. Based on their findings, they highlight the importance of incorporating emotionally resonant materials in teaching.

In another qualitative study, K. Al-Madanat sets out to explore the dynamic relationship between two emotions, anxiety and enjoyment, in two languages. Using the idiodynamic method, which allows momentary changes to be tracked, her two participants' self-rating is analyzed. Her findings reveal the fluctuations between the two emotions in both English

and French and she argues that several internal and external factors contribute to this change in emotions.

Using quantitative research methods to investigate individual differences, Ayed explores the interrelationship between willingness to communicate, motivation and the extroversion-introversion aspects of personality traits. Her findings demonstrate that the extroverted learners in her study have higher levels of willingness to communicate. The implications of the research point to the necessity of increasing learners' motivation to boost their willingness to communicate.

Khaing explores the developmental trajectories of an EFL learner's, using a mixed-methods Complex Dynamic Systems Theory approach. In her exploratory study, she compares one learner's listening performance with their metacognitive listening strategy use in an attempt to investigate the dynamic interaction between the two variables. Her findings point to non-linear developmental patterns and dynamic correlations between performance and strategy use.

Alsayed Ahmad's small-scale study aims to uncover the translatability of Arabic jokes into English by revealing the linguistic problems that translators encounter when translating them. The analysis also investigates the strategies the translators in the study used to overcome these issues of translatability.

In an attempt to provide a deeper understanding of pragmatic phenomena in English as a lingua franca communication, Farkas analyses the interactions and language use in an international university context. In his study, the researcher's etic analysis as well as the participants' emic views on a speech event are presented. He argues that this combination of etic and emic perspectives might contribute to a more complex view of such pragmatic phenomena.

Tankó's theoretical paper reviews the relevant literature on the teaching and assessment of listening with the ultimate aim of presenting a comprehensive model of listening behavior types. The proposed model is especially important for listening research related to assessment development, test evaluation, and item writer training purposes.

In her interview study, Andréka looks into Hungarian high school EFL teachers' attitudes towards the objectively scored components of the

advanced level English school-leaving examination. The teachers partici-pating in the study seem to have expressed mostly positive views regard-ing the reading and listening parts of the exam. The insights are useful for further improving the English school-leaving exam in Hungary.

Finally, Wind tackles the issue of syntactic complexity in L2 writing and sets out to observe variation in syntactic complexity between natural and social science research articles based on two corpora. His corpus-based findings reveal considerable differences, the implications of which might especially be useful in teaching academic writing at university level.

We hope you enjoy reading the articles!

Attila M. Wind & Brigitta Dóczi

1

Empirical Research on Extracurricular Language Use: The Role of Individual Differences

Dávid Smid
david.smid@uni-graz.at

Kata Csizér
wein.kata@btk.elte.hu

Abstract

The aim of our paper is to provide a critical synthesis of research on extracurricular language use both in the home environment and in study abroad contexts, with the goal of outlining possible future research directions within Hungarian contexts. To achieve this aim, we first present a critical review of studies pertaining to out-of-class language learning and use in both home and study abroad contexts, concentrating on the various individual difference variables included in earlier studies. In the final section of our article, we provide new research directions for various Hungarian contexts, taking into account the individual variables covered in the analysis.

Keywords: extracurricular language use, individual differences, study abroad contexts

Empirical Research on Extracurricular Language Use:
The Role of Individual Differences

The aim of our paper is to provide a critical synthesis of research on extracurricular language use both in the home environment and in study abroad (SA) contexts in order to outline possible future directions for research within Hungarian contexts. The importance of overviewing this subfield in applied linguistics is twofold. First, it is a well-known fact that Hungary, although decreasingly markedly, has been lagging in regard to second language (L2) proficiency when compared to other European Union (EU) states. Data show that while 57.6% of the total adult (i.e., between 25 and 64 years of age) and 40.4% of the young adult (i.e., between 25 and 34 years) population of Hungary did not know any L2s in 2016, the total EU statistics (27 countries included) were only 31.8% and 21.3%, respectively (European Statistics, 2021). For this reason, beyond considering quality issues in public education (Öveges & Csizér, 2018), it is also crucial to investigate how students use English outside the classroom and in what ways their extracurricular use of English can enhance learning processes within the classroom. Second, the field of individual difference (ID) research holds significant value in applied linguistics; as investigations show, ID variables can increase the speed of language learning processes (Dörnyei, 2005; Dörnyei & Ryan, 2015; Dörnyei & Ushioda, 2013). However, mainstream theories do not seem to consider the role of IDs in learning and using language outside the classroom with the notable exception of Illés's (2012) work on autonomy. Empirical studies tend to address issues either in the home environment (Dörnyei & Ryan, 2015) or in various SA contexts (Pérez-Vidal & Sanz, 2023), leaving a research gap concerning the role of IDs in extracurricular activities within the home environment.

Given the aim of our paper, we have organized it into three main sections. First, we present the results of our review of studies focusing on the home environment, followed by those in various SA contexts. In the final section, we outline future research directions, addressing both theoretical and empirical studies in various Hungarian contexts. Throughout the paper, we use the terms "out-of-class L2 learning," "extracurricular L2 learn-

ing," and "informal L2 learning" interchangeably to refer to an individual's deliberate engagement in L2 use in a non-instructed setting (Benson, 2011).

Out-of-Class Language Learning/Use in the Home Environment

In regard to out-of-class L2 learning in the home environment, one key ID variable emerging from the reviewed studies is L2 motivation, which can be defined as an individual's intention to invest effort in learning an L2 (Dörnyei & Ushioda, 2013). The reviewed studies paint a mixed picture of Hungarian students' out-of-class L2 motivation in their home environment. On the positive side, Lajtai (2020) found that primary and high school students tended to invest more effort into learning English in out-of-school contexts compared to the formal learning context. Similarly, Fajt's (2021) results showed that Hungarian 9th graders' engagement in English use in out-of-school contexts was significant. In contrast, studies conducted by Albert et al. (2018a, 2018b) revealed only average levels for both primary as well as high school pupils' out-of-school engagement in English (i.e., listening to music, watching movies/series, gaming, browsing the Internet, and attending cultural events such as theater performances or concerts). This was confirmed by the L2 teachers in Illés and Csizér's (2018) study, whose general opinion was that many students enrolled in Hungarian public education were not engaged with the English language outside class. Given that Albert et al.'s and Illés and Csizér's quantitative inquiries employed much larger samples, it is possible that the out-of-school L2 motivation of pupils in Hungary is indeed not as high as it appears in Fajt's and Lajtai's research, both of which were based on smaller samples. In any case, the results of these large-scale studies are in line with those of Illés and Csizér (2010), suggesting that there has been no visible progress regarding the extent to which Hungarian pupils have been engaging in contact with the English language in out-of-school contexts throughout the past decade.

Empirical research has revealed that L2 autonomy is also an essential aspect for L2 learners/users engaged in out-of-class L2 learning in their home context. L2 autonomy is typically understood as an individual's capacity to manage their own learning and thus seek opportunities to use

and learn the L2 (Littlewood, 1996). Alongside L2 learning motivation, L2 autonomy has also been extensively investigated in relation to instructed L2 learning, which is not surprising considering that the two concepts are interrelated: In Dörnyei's (2009) theoretical model, L2 autonomy is linked to the L2 learning experience, that is, the current dimension of the L2 motivational self-system. Lajtai's (2020) study is unique in that it appears to be the first to provide empirical evidence for the dependence of L2 autonomy on both in-school and out-of-school L2 learning motivation in the case of primary and high school pupils in Hungary. The study's results indicate that high levels of autonomy correlate with L2 motivation across various L2 learning contexts, which, in turn, suggests that fostering pupils' L2 autonomy should be a focal point of pedagogical practices (Illés, 2012; Littlewood, 1996).

Beliefs have been identified as another ID variable that accompanies informal L2 learning. As L2 learner beliefs are also a key concept in research focusing on classroom L2 learning, they can be understood as primarily cognitive events characterized by reflection and replete with personal meanings (Kalaja et al., 2015). Lajtai (2020, p. 137) found that Hungarian primary and high school "students believed that the outside-the-school context was significantly more conducive to language learning than school itself." This result can best be interpreted in light of the pupils' responses in Albert et al.'s (2018a, 2018b) studies, which revealed that frontal teaching remains the predominant teaching method in L2 classes in Hungarian public education. As a result, L2 teachers often fail to provide opportunities for their learners to strengthen their out-of-school contact with the L2 culture (Albert et al., 2018a, 2018b), the importance of which has been underscored by other research (e.g., Leona et al., 2021; Sayer & Ban, 2014). In fact, the absence of an L2 teaching approach that centers around authentic language and its real-life use does not seem to be unique to the Hungarian context (see, e.g., Badwan, 2017; Bai & Wang, 2020; Henry, 2014; Whitehead, 2021). One potential risk of holding the belief that informal L2 learning environments are more beneficial than formal ones is that it can negatively influence in-school L2 learning motivation, as suggested by Henry and Cliffordson's (2017) study. Although Lajtai did not find evidence for such

a debilitating effect in his Hungarian sample, additional research into this issue is definitely warranted.

Finally, our review of research targeting English learners' IDs in out-of-school L2 learning in their home context highlights three key concepts: L2 willingness to communicate (WTC), L2 linguistic self-confidence, and L2 anxiety. Given that these IDs have often been shown to be interlinked in L2 classroom research (see, e.g., MacIntyre et al., 1998), we have decided to examine them together here. L2 WTC can be defined as the "readiness to enter into discourse at a particular time with a specific person or persons, using a L2" (MacIntyre et al., 1998, p. 547). As such, it is regarded as the intentional aspect of communicative behavior and is assumed to be dependent on multiple conditions such as motivation, self-confidence, and anxiety (MacIntyre et al., 1998). L2 anxiety is understood as a negative emotional reaction that is triggered by negative self-perceptions, beliefs, and fears related to the use of the L2 (Horwitz et al., 1986). L2 self-confidence, in turn, entails "both a lack of anxiety and positive self-ratings of L2 proficiency" (Sampasivam & Clément, 2014, p. 25). Studies involving university students have recently found that L2 self-confidence does not only positively influence L2 WTC in the classroom context, but also outside the classroom and in digital settings as well. It was also uncovered that L2 anxiety was a negative predictor of L2 WTC both inside and outside the classroom (Lee & Hsieh, 2019, Lee & Lee, 2019). These results further underline the need for L2 teachers to build on their pupils' L2 contact experiences in non-traditional settings to facilitate their overall L2 achievement and promote further research into L2 teachers' own perceptions of out-of-school L2 learning.

The Study Abroad (SA) Context

With respect to the SA context, students' L2 learning motivation has also been a recurrent topic of research. A central question across such studies is whether the particularities of the SA context, and thus, the time spent abroad, have any influence on L2 learning motivation. Several studies have shown the positive effect of SA on the salience and potency of students' future L2 selves following their time abroad (e.g., Du, 2019; Fryer & Roger,

2018). In reference to Dörnyei's (2009) L2 motivational self-system model, this also indicates an intensification of their effort investment (Pawlak et al., 2020). Direct contact with other English speakers has also been identified as a key variable motivating students during their SA (Cai et al., 2013; Hessel, 2019), although the amount of contact with native speakers (NSs) tended to be lower than with non-native speakers (NNSs) (Kormos et al., 2014), pointing to the debilitating effect of L2 anxiety. These results emphasize the importance of raising students' awareness of the importance of maximizing their L2 contact during SA (Cai et al., 2013; Hessel, 2019), and suggest that both home and host institutions alike should promote opportunities for their students to use the L2 outside of class with both NSs and NNSs (Du, 2019; Hessel, 2019).

Research focusing on L2 learning in the SA context has also illuminated the interrelated roles of autonomy and self-regulation. Although both concepts involve the learner's management of their L2 learning, self-regulation differs from L2 autonomy in that it is more limited in its domain of L2 learning and is typically associated with the use of learning strategies (Kormos & Csizér, 2014; Tseng et al., 2006). Regarding L2 autonomy, Durbidge (2019) found that the L2 proficiency levels of Japanese high school students prior to their SA stay correlated with their use of technology in the L2 during their SA period. This result does not only point to the potentially significant role of technology use during SA, but also to the importance of making students aware of it. Regarding self-regulatory learning strategies, multiple studies have confirmed their positive effect on L2 learning motivation, particularly the use of strategies initiated by students before SA and those initiated by teachers during SA (Pawlak & Csizér, 2022; Pawlak et al., 2020). Again, these findings offer important pedagogical implications, suggesting that students could benefit from receiving training in the use of the various kinds of self-regulatory learning strategies prior to, during, and following their SA experiences (Pawlak et al., 2020).

Similarly to the home context, there has also been empirical evidence for the role of beliefs in students' L2 learning and use in the SA context. One category of beliefs emerging from the reviewed literature relates to the L2 and its speakers. Zaykovskaya et al. (2017) found that a L2 learner's pre-existing beliefs had a positive effect on their SA experiences as well

as on their L2 learning motivation. Although their study focused on a L2 learner of Russian, according to research on beliefs in instructed L2 learning contexts (see, Kalaja et al., 2015), we can assume that a similar relationship exists in the case of L2 learners of English during SA. Another belief identified by Smid (2018) influencing SA experiences is NS idealism, namely, the belief that NSs should be emulated when speaking a L2. More specifically, the author found that German Erasmus+ students had ascribed a role to being surrounded by NNSs during their SA stay when perceiving a lack of development regarding their linguistic accuracy in English.

The other type of belief highlighted by research is L2 self-efficacy, which refers to an individual's self-evaluation of their L2-related performance (Bandura, 1995). While this concept is similar to the above-mentioned L2 self-confidence in that both relate to an individual's self-perceived competence, L2 self-efficacy is characterized as cognitively constructed, while L2 self-confidence is considered socially constructed (Dörnyei, 2005). One of the reasons why L2 self-efficacy can be regarded as an essential aspect of the SA context is that it can provide increased opportunities for students to engage in L2 contact experiences. As Hessel's (2019) research suggests, a L2 learner's perception of their own as well as others' successful experiences with the L2 (i.e., mastery and vicarious experiences) can develop their L2 self-efficacy. The role of L2 self-efficacy in SA contexts is also evidenced by Pawlak et al. (2020), who came to the conclusion that L2 self-efficacy is likely interlinked with L2 learning motivation and self-regulatory learning strategies at all stages of SA. This complex web of interrelationships further underlines the need for pedagogical support, and, as such, the need to understand L2 teachers' perspectives of out-of-class L2 learning.

Future Research Directions

Based on the review of the literature presented above, we have identified a number of fruitful directions for future research. The need for comparative and longitudinal studies should be highlighted, as well as the importance of understanding how ultimate success is conceptualized in the learning process. In relation to success, the role of teachers and their beliefs

could provide important insights into the ways that ID variables shape the learning processes. In addition, our current understanding of the workings of ID variables calls for an integrated approach towards examining these variables as interrelated factors rather than in isolation (Ryan, 2019). As we have shown, only a limited number of ID variables have received systematic attention in Hungary. Generally speaking, the field would benefit from a broadening of the research scope to include more ID variables, especially cognitive variables contributing to the learning processes. Furthermore, the investigation of L2 WTC, L2 anxiety, and linguistic self-confidence might be worth being included in future studies (see also Albert et al., 2023, for further variables). In what follows, we summarize research ideas pertaining to the ID variables covered in this article.

As studies regarding the role of motivation to use English outside the classroom have yielded somewhat contradictory results, comparative studies are needed to map students' L2 motivation both in and outside of formal educational settings. Before carrying out empirical studies, it would be important to explore the role of various L2 motivation theories that describe the learning processes inside and outside the classroom. This theoretical understanding could then inform the conceptualization and operationalization of relevant variables. Apart from investigating conceptual and empirical differences in students' L2 motivation between classroom and out-of-classroom contexts, it would also be useful to examine interactions between the two contexts. Changes in the use of English outside the classroom would require longitudinal quantitative studies, retrospective interviews, or carefully designed case studies, which could shed light on the changing role of L2 motivation to use English outside the classroom and how this use shapes students' learning processes. In connection to L2 motivation, it is important to understand the long- and short-term goals that students have for language learning and what they consider short-term and long-term success in in this regard.

The case of autonomy seems to be different from that of L2 motivation, as Illés (2012) has already defined the concept in terms of English use outside the classroom. However, empirical studies that map the development of the autonomous use of English outside the classroom and examine how

this development can support learning processes within the classroom are lacking. Utilizing the benefits of the extracurricular use of English can also be addressed through teacher-focused studies, especially through juxtaposing exam-centered teaching practices (Csizér et al., in preparation; Öveges & Csizér, 2018). Finally, as we think that language learning through language use should be a life-long enterprise, future studies should also aim at uncovering the ways in which teachers can prepare students to be autonomous language learners throughout their professional careers.

In considering the investigation of language learning beliefs, we propose that, as language learning processes are strongly shaped by students' milieu (e.g., Csizér & Galántai, 2012), the research scope in this area should extend beyond students' and teachers' beliefs to include those of the parents of school-aged children. Insights into the parents' language learning beliefs could shed light on how students' beliefs might be shaped through their milieu, and taking teachers' views into account could further provide a complex picture of the impact of beliefs on students' efforts and achievements. Another issue that needs to be highlighted here is that language learning beliefs measured in applied linguistic studies are not neutral statements but present issues that can aid or thwart the language learning process; therefore, language learning beliefs should be investigated from this perspective as well. In light of Henry and Cliffordon's (2017) and Lajtai's (2020) studies, it goes without saying that language learning beliefs should be operationalized in future studies, taking into account classroom, out-of-classroom, and SA contexts.

When it comes to SA research in Hungary, there is a paucity of studies targeting students participating in SA programs. Considering the available small-scale pilot studies (e.g., Stamenkovska et al., 2022; Szentpáli Ujlaki, 2007), it can be argued that the field would clearly benefit from more empirical research. This could include both large-scale investigations into the experiences of outgoing students as well as those of incoming students, which could allow for generalizable results. Such studies could also involve students of different origin and with different SA destinations so as to shed potential light on contextual particularities specific to the Hungarian context (see, e.g., Pawlak & Csizér, 2022).

Conclusion

It is important to point out that it was beyond the scope of our paper to cover all relevant ID variables with regard to extracurricular language use in home and SA contexts. The appearance of new individual characteristics seems to be the rule in the field of ID research, not the exception [see, e.g., L2 learner well-being (Mercer et al., 2018), L2 grit (Teimouri et al., 2022), and L2 boredom (Pawlak et al., 2022)]. Given the challenges faced by the Hungarian education system, we think it is important to continuously take stock of the most important ID variables, such as motivation, autonomy, and beliefs, while also expanding the research scope to include new ID variables.

Having summarized what we think are the most important research implications of our literature review, what remains to be considered are the further implications that this research presents for teachers, students, and other stakeholders. Given the growing presence of technology and social media as well as increasing international mobility, out-of-school opportunities for L2 learning and use have dramatically expanded. We think that these social trends cannot go unnoticed in formal L2 instruction. L2 teachers should be aware of their L2 learners' out-of-class experiences and build on them in the classroom, considering the proposed link between formal and informal L2 learning (e.g., Henry & Cliffordson, 2017; Lajtai, 2020; Sayer & Ban, 2014). In addition, as research has indicated (e.g., Pawlak et al., 2020), students could also benefit from support in regard to effective L2 use both in out-of-school contexts in the home environment and during SA to make the most of their experiences. Finally, it should be noted that the implementation of these pedagogical practices should also be facilitated through the recognition of the benefits of extracurricular L2 learning in the national and local curricula, as well as through continued professional development.

References

Albert, Á., Bowden, H. W., & Csizér, K. (2023). Individual variables in study abroad contexts: Concepts and measurements. In C. Pérez-Vidal & C. Sanz (Eds.), *Methods in study abroad research: Past, present, and future* (pp. 319–344). John Benjamins Publishing Company. https://doi.org/10.1075/rmal.4.13alb

Albert, Á., Tankó, G., & Piniel, K. (2018a). A tanulók válaszai a 7. évfolyamon [Grade 7 pupils' responses]. In E. Öveges & K. Csizér (Eds.), *Vizsgálat a köznevelésben folyó idegennyelv-oktatás kereteiről* és *hatékonyságáról: Kutatási jelentés* (pp. 52–89). Hungarian Educational Authority.

Albert, Á., Tankó, G., & Piniel, K. (2018b). A tanulók válaszai a 11. évfolyamon [Grade 11 pupils' responses]. In E. Öveges & K. Csizér (Eds.), *Vizsgálat a köznevelésben folyó idegennyelv-oktatás kereteiről* és *hatékonyságáról: Kutatási jelentés* (pp. 90–160). Hungarian Educational Authority.

Badwan, K. M. (2017). Did we learn English or what? A study abroad student in the UK carrying and crossing boundaries in out-of-class communication. *Studies in Second Language Learning and Teaching, 7*(2), 193–210. https://doi.org/10.14746/ssllt.2017.7.2.2

Bai, L., & Wang, Y. X. (2020). Pre-departure English language preparation of students on joint 2+2 programs. *System, 90*, 102219. https://doi.org/10.1016/j.system.2020.102219

Bandura, A. (1995). Exercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 1–45). Cambridge University Press. https://doi.org/10.1017/CBO9780511527692.003

Cai, W., Li, X., & Liu, M. (2013). Academic and linguistic gains during a semester-long study abroad: A cohort case study. *Studies in Second Language Learning and Teaching, 3*(4), 505–522. https://doi.org/10.14746/ssllt.2013.3.4.4

Csizér, K., Albert, Á., Smid, D., & Zólyomi, A. (in preparation). *Motivation, autonomy, and emotions in foreign language learning: A multi-perspective investigation in Hungary*. Multilingual Matters.

Csizér, K., & Galántai, D. (2012). A tanári és szülői szerepek alakulása a középiskolások idegen nyelvi motivációjának alakításában. Egy strukturális modell tanulságai [The role of parents and teaching shaping secondary school students' L2 motivation: The results of structural equation modeling]. In A. Németh (Ed.), *A neveléstudományi doktori iskola programjai: Tudományos arculat, kutatási eredmények* (pp. 171–178). Eötvös Kiadó.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Lawrence Erlbaum.

Dörnyei, Z. (2009). The L2 motivational self-system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Multilingual Matters. https://doi.org/10.21832/9781847691293-003

Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited.* Lawrence Erlbaum. https://doi.org/10.4324/9781315779553

Dörnyei, Z., & Ushioda, E. (2013). *Teaching and researching motivation* (2nd ed.). Routledge. https://doi.org/10.4324/9781315833750

Du, X. (2019). The impact of semester-abroad experiences on post-sojourn L2 motivation. *Studies in Second Language Learning and Teaching, 9*(1), 117–155. https://doi.org/10.14746/ssllt.2019.9.1.6

Durbidge, L. (2019). Technology and L2 engagement in study abroad: Enabler or immersion breaker? *System, 80*, 224–234. https://doi.org/10.1016/j.system.2018.12.004

European Statistics (2021). *Number of foreign languages known (self-reported) by age*. Retrieved December 28, 2021, from https://ec.europa.eu/eurostat/databrowser/view/edat_aes_l22/default/table?lang=en

Fajt, B. (2021). Hungarian secondary school students' extramural English interests: The development and validation of a questionnaire. *Working Papers in Language Pedagogy, 16*, 36–53. http://langped.elte.hu/WoPaLParticles/W16Fajt.pdf

Fryer, M., & Roger, P. (2018). Transformations in the L2 self: Changing motivation in a study abroad context. *System, 78*, 159–172. https://doi.org/10.1016/j.system.2018.08.005

Henry, A. (2014). Swedish students' beliefs about learning English in and outside of school. In D. Lasagabaster, A. Doiz, & J. M. Sierra (Eds.), *Motivation and foreign language learning: From theory to practice* (pp. 93–116). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.40.05hen

Henry, A., & Cliffordson, C. (2017). The impact of out-of-school factors on motivation to learn English: Self-discrepancies, beliefs, and experiences of self-authenticity. *Applied Linguistics, 38*(5), 713–736. https://doi.org/10.1093/applin/amv060

Hessel, G. (2019). The role of international student interactions in English as a lingua franca in L2 acquisition, L2 motivational development and intercultural learning during study abroad. *Studies in Second Language Learning and Teaching, 9*(3), 495–517. https://doi.org/10.14746/ssllt.2019.9.3.4

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125–132. https://doi.org/10.1111/j.1540-4781.1986.tb05256.x

Illés, É. (2012). Learner autonomy revisited. *ELT Journal, 66*(4), 505–513. https://doi.org/10.1093/elt/ccs044

Illés, É., & Csizér, K. (2010). Secondary school students' contact experiences and dispositions towards English as an international language – A pilot study. *Working Papers in Language Pedagogy, 4*, 1–22. Retrieved July 19, 2022 from http://langped.elte.hu/WoPaLParticles/W4IllesCsizer.pdf

Illés, É., & Csizér, K. (2018). A nyelvtanárok válaszai [The language teachers' responses]. In E. Öveges & K. Csizér (Eds.), *Vizsgálat a köznevelésben folyó idegennyelv-oktatás kereteiről* és *hatékonyságáról: Kutatási jelentés* (pp. 161–188). Hungarian Educational Authority.

Kalaja, P., Barcelos, A. M. F., Aro, M., & Ruohotie-Lyhty, M. (2015). *Beliefs, agency and identity in foreign language learning and teaching*. Palgrave Macmillan. https://doi.org/10.1057/9781137425959

Kormos, J., & Csizér, K. (2014). The interaction of motivation, self-regulatory strategies, and autonomous learning behavior in different learner groups. *TESOL Quarterly, 48*(2), 275–299. https://doi.org/10.1002/tesq.129

Kormos, J., Csizér, K., & Iwaniec, J. (2014). A mixed-method study of language-learning motivation and intercultural contact of international students. *Journal of Multilingual and Multicultural Development, 35*(2), 151–166. http://doi.org/10.1080/01434632.2013.847940

Lajtai, Á. (2020). Hungarian EFL learners extramural contact with English. In R. Geld & S. L. Krevelj (Eds.), *UZRT 2018: Empirical studies in applied linguistics* (pp. 128–149). FF Press. https://doi.org/10.17234/UZRT.2018.3

Lee, J. S., & Hsieh, J. C. (2019). Affective variables and willingness to communicate of EFL learners in in-class, out-of-class, and digital contexts. *System, 82*, 63–73. https://doi.org/10.1016/j.system.2019.03.002

Lee, J. S., & Lee, K. (2019). Affective factors, virtual intercultural experiences, and L2 willingness to communicate in in-class, out-of-class, and digital settings. *Language Teaching Research, 24*(6), 813–833. https://doi.org/10.1177/1362168819831408

Leona, N. L., van Koert, M. J. H., van der Molen, M. W., Rispens, J. E., Tijms, J., & Snellings, P. (2021). Explaining individual differences in young English language learners' vocabulary knowledge: The role of extramural English exposure and motivation. *System, 96*, 102402. https://doi.org/10.1016/j.system.2020.102402

Littlewood, W. (1996). "Autonomy": An anatomy and a framework. *System, 24*(4), 427–435. https://doi.org/10.1016/S0346-251X(96)00039-5

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

Mercer, S., MacIntyre, P., Gregersen, T., & Talbot, K. (2018). Positive language education: Combining positive education and language education. *Theory and Practice of Second Language Acquisition, 4*(2), 11–31. Retrieved March 8, 2023, from https://www.journals.us.edu.pl/index.php/TAPSLA/article/view/7011

Öveges, E., & Csizér, K. (2018). Vizsgálat a köznevelésben folyó idegennyelv-oktatás kereteiről és hatékonyságáról: Kutatási jelentés [Research of foreign language teaching in the Hungarian public education system]. Hungarian Educational Authority.

Pawlak, M., & Csizér, K. (2022). The impact of self-regulatory strategy use on self-efficacy beliefs and motivated learning behavior in study abroad contexts: The case of university students in Italy, Poland and Turkey. *System, 105*, 102735. https://doi.org/10.1016/j.system.2022.102735

Pawlak, M., Csizér, K., & Soto, A. (2020). Interrelationships of motivation, self-efficacy and self-regulatory strategy use: An investigation into study abroad experiences. *System, 93*, 102300. https://doi.org/10.1016/j.system.2020.102300

Pawlak, M., Kruk, M., Zawodniak, J., & Pasikowski, S. (2022). Examining the underlying structure of after-class boredom experienced by English majors. *System, 106*, 102769. https://doi.org/10.1016/j.system.2022.102769

Pérez-Vidal, C., & Sanz, C. (Eds.). (2023). *Methods in study abroad research: Past, present, and future*. John Benjamins Publishing Company. https://doi.org/10.1075/rmal.4

Ryan, S. (2019). Motivation as an individual difference. In M. Lamb, K. Csizér, A. Henry, & S. Ryan (Eds.), *Palgrave Macmillan handbook of motivation for language learning* (pp. 163–182). Palgrave. https://doi.org/10.1007/978-3-030-28380-3_8

Sampasivam, S., & Clément, R. (2014). The dynamics of second language confidence: Contact and interaction. In S. Mercer & M. Williams (Eds.), *Multiple perspectives on the self in SLA* (pp. 23–40). Multilingual Matters. https://doi.org/10.21832/9781783091362-004

Sayer, P., & Ban, R. (2014). Young EFL students' engagement with English outside the classroom. *ELT Journal, 68*(3), 321–329. https://doi.org/10.1093/elt/ccu013

Smid, D. (2018). German study abroad students' language learner identity and investment in English. *Meisei International Studies, 10,* 73–96. https://cir.nii.ac.jp/crid/1050001201664183552

Stamenkovska, T., Kálmán, C., & Győri, J. G. (2022). The motivational disposition of international students learning foreign languages in Hungary: Testing the L2 motivational self system in the Hungarian context. *Journal of Adult Learning, Knowledge, and Innovation, 5*(2), 101–110. https://doi.org/10.1556/2059.2022.00060

Szentpáli Ujlaki, E. (2007). Study-abroad motivation – Preliminary results of a pilot project. In J. Horváth & M. Nikolov (Eds.), *UPRT 2007: Empirical studies in English applied linguistics* (pp. 3–18). Lingua Franca Csoport.

Teimouri, Y., Plonsky, L., & Tabandeh, F. (2022). L2 grit: Passion and perseverance for second language learning. *Language Teaching Research, 26*(5), 893–918. https://doi.org/10.1177/1362168820921895

Tseng, W.-T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics, 27*(1), 78–102. https://doi.org/10.1093/applin/ami046

Whitehead, S. J. (2021). I thought we weren't in Spain! Discursive negotiations of realness in a foreign language classroom. *System, 100,* 102532. https://doi.org/10.1016/j.system.2021.102532

Zaykovskaya, I., Rawal, H., & De Costa, P. (2017). Learner beliefs for successful study abroad experience: A case study. *System, 71,* 113–121. https://doi.org/10.1016/j.system.2017.09.020

2

Bilingualism and Emotions: An Exploratory Study on Language Significance, Swear and Taboo Words, and the Phrase 'I Love you' in the First and Second Language

Dina Al-Madanat
dinamdanat@gmail.com

Ágnes Albert
albert.agnes@btk.elte.hu

Abstract

In recent years, there has been a growing interest among researchers in the relationship between bilingualism and emotions. Although many researchers have examined the influence of emotions on bilingual performance (Anooshian & Hertel, 1994; Koven, 2004) and second language learning and use (Dewaele, 2011; Swain, 2011), much remains to be explored regarding the way bilinguals perceive the emotional significance of each of their languages. The present study aims to explore the perception of emotions in bilinguals' different languages, the emotional force of swearwords and taboo words (S-T words), and the emotional weight of the phrase "I love you". The study also investigates participants' language preferences for expressing their feelings and using swear and taboo words. Data were collected using an adapted web questionnaire on bilingualism and emotions (BEQ; Dewaele & Pavlenko, 2001-2003). The sample consisted of 143 bilingual students residing in Hungary. Results indicated that participants perceived their L1 to be more emotionally charged than their L2. Additionally, the perceived emotional force of S-T words and the phrase "I love you" was found to be higher in participants' L1. Participants also preferred using their L1 for expressing feelings and S-T words. Statistical analyses suggested that the emotional significance of the L1 is associated with the L2 learning context, bilinguals' self-perceived L2 proficiency, and gender.

The findings imply that foreign language materials should include more emotionally resonant materials and that learners should be encouraged to use their L2 outside the classroom to enhance their ability to express their feelings and communicate effectively in the future.

Keywords: bilingualism, emotions, perception of emotion, swearwords, taboo words

Bilingualism and Emotions: An Exploratory Study on Language Significance, Swear and Taboo Words, and the Phrase 'I Love you' in the First and Second Language

Approximately, almost half of the world's population is bilingual (Grosjean, 2021). The topic of bilingualism has been heavily debated in the literature, and researchers' perceptions of who qualifies as bilingual vary considerably due to the complex nature of the phenomenon and the various factors associated with it. According to Lambert et al. (1959), individuals can be considered bilingual only if they can speak the two languages perfectly and with equal proficiency. Liebkind (1995) presented a simpler definition, claiming that a person can be considered bilingual if they feel that they are bilingual and are regarded as such by others.

Grosjean (2008) offered a widely accepted definition of bilingualism, stating that it is "the regular use of two or more languages (or dialects), and bilinguals are those people who use two or more languages (or dialects) in their everyday lives" (p. 10). He also proposed that individuals can be considered bilingual if they can use the two languages without necessarily being fully fluent in both of them. Participants in this study are international students residing in Hungary who use their L1 and L2 in their daily life and have varying levels of L2 proficiency; therefore, the current study considers them as bilinguals based on Grosjean's (2008) above-mentioned definition.

In the past, research on bilingualism considered language and emotions as two separate fields; however, in recent years, researchers and scholars have begun to examine the languages bilinguals use in relation to emotions. According to Dewaele (2005), the study of emotions in second language acquisition can improve second language learners' sociocultural competency. Shao et al. (2019) agreed that emotions are important, asserting that they play an important role in second language learning and achievement. Li et al. (2018) claimed that emotions are crucial to human interaction: the inability to effectively express emotions could cause misunderstandings, affect relationships, and lead to discord between speakers.

Given the crucial role that emotions play in communication and second language learning, the impact of bilingualism on speaker's perception

and expression of emotions merits further investigation. The present study attempts to explore how Arabic-English bilinguals who study and live in Hungary perceive and express their emotions. Specifically, it examines whether their perception of the phrase "I love you", as well as S-T words, is consistent across the languages they speak. The study also investigates the influence of various independent variables related to the languages participants speak (i.e., participants' age of acquisition, level of proficiency, frequency of language use, context of acquisition, and gender) on how they perceive and express emotions in their two languages.

Literature Review

Bilingualism and Emotions

Although several key questions have been raised regarding the relationship between bilingualism and emotions, little attention has been paid to emotions in second language acquisition with the exception of anxiety (Li et al., 2018). Various studies comparing the emotional connections people have with their first language (L1) and second language (L2) show that bilinguals often express their emotions differently depending on the language they are using (Altarriba, 2008; Caldwell-Harris, 2014; Dewaele, 2004, 2008, 2017; Pavlenko, 2008).

Previous research suggests that bilinguals often consider their L1 as the language of emotions, preferring to use it to more effectively express their emotions (Anooshian & Hertel, 1994; Guttfreund, 1990). Marcos (1976) posited that the L2 is typically associated with intellectual functioning and is less emotionally charged compared to the L1, which is more closely related to the emotional expression. In his study on the emotional weight of the phrase "I love you" across the languages spoken by individual multilinguals, Dewaele (2008) found that people perceive the phrase as more powerful in their first language.

Studies on bilinguals' emotions show that bilinguals often use their L2 as a protective mechanism to distance themselves from negative memories or intense emotions (Aycicegi & Harris, 2004). Anooshian and Hertel (1994) agreed that emotions are stronger in the L1. In their study on Spanish-Eng-

lish bilingual adults, they found that participants recalled emotional words more effectively than neutral words when using their L1, a finding not observed in the L2. Schrauf (2000) also agreed with other researchers that emotions are stronger in the L1, suggesting that memories tend to be significantly more emotional when discussed using the L1.

The studies cited above suggest that the L1 is the language of emotional expression, while the L2 functions as a language of emotional detachment; as a result, numerous researchers in the field have been motivated to investigate the different factors that could contribute to differing emotion experiences in the L1 and L2 (Dewaele, 2004, 2008, 2017).

The existing literature provides various insights into the factors contributing to the emotional significance of the L1 and L2 among bilinguals. King (2011) posited that stronger emotional experiences in the L1 are often linked to childhood memories, which makes individuals feel more connected to that language. Another proposed aspect influencing higher emotionality in a given language is proficiency level. According to Harris et al. (2006), a high level of proficiency in the L2 could expose learners to more emotional contexts via close contact with native speakers of the L2. Marian and Kaushanskaya (2008) argued that a bilingual's emotional language is usually their language of preference rather than being determined by proficiency. In their study on Russian-English narratives, participants expressed more emotions in their L2 than in their L1. Ferré et al. (2010) corroborated this finding, asserting that the language of preference is the one which is the most emotional, while factors such as the age of second language acquisition, the context of the acquisition, language dominance, and similarities between the L1 and L2 were not found to affect the memory of emotional words in the L2. In contrast, an alternative explanation for the emotional closeness of the L1 stems from the fact that it is usually acquired in emotionally rich settings, while the L2 is commonly acquired in formal classroom environments (Ivaz et al., 2019).

In light of these previous studies, it can be concluded that in general researchers agree that the mother tongue holds a closer connection and has greater emotional weight for bilinguals than their L2; however, this may not always be the case. On the other hand, there is less agreement

among researchers regarding the specific factors that affect the intensity of emotions. As such, the current study seeks to investigate the relationship between bilinguals' emotions and languages while taking into account various influencing factors.

Swear and Taboo Words in Bilingualism

The use of S-T words is often considered as a linguistic device employed by speakers to gain in-group membership while establishing social norms and boundaries in language use (Dewaele, 2004). In his study, Dewaele (2004) investigated the choice of language for swearing in connection with factors including the context of acquisition, age of acquisition, proficiency, and language dominance. Data were collected from 1,039 multilinguals using an online questionnaire. The results showed that participants preferred to express S-T words in their L1 rather than their L2, although this language choice was also dependent on the intention of the speaker. Moreover, results indicated that there was a positive relationship between the frequency with which the given language was used and the choice for swearing, while neither gender nor educational level was found to have a significant role in the language chosen for swearing.

Dewaele (2017) extended his research to investigate the preferred language for swearing among multilinguals who reported the same level of proficiency in their L1 and L2. The data showed that even among bilinguals who consider themselves equally proficient in both languages, the L1 remained the preferred language for swearing as it allowed participants to express their anger in a more powerful way. Colbeck and Bowers (2012) found similar results, concluding that the mother tongue is more emotional and that taboo words have a greater emotional effect in the L1. Similarly, reprimands were also found to be more intense in the L1 than in the L2. In her study on Turkish-English late bilinguals, Pavlenko (2012) observed that reprimands had a greater impact when expressed in the L1 compared to the L2. She noted that learning languages early in life causes participants to perceive taboo words and reprimands to be more powerful in those languages. Conversely, learning other languages later in life would produce

a sense of emotional detachment when processing such words in the L2 (Pavlenko, 2012).

The existing literature suggests that in general the L1 is perceived to be more emotional than the L2. The emotional weight of S-T words is stronger in bilinguals' L1, and this emotional dimension must be considered in relation to different variables that may significantly influence bilinguals' emotional experiences. Although extensive research has been conducted on bilinguals' emotional perception and expression, it remains unclear as to which factors have the greatest impact on the language use and perception of bilinguals. It is also noteworthy to mention that, to the best of our knowledge, there have been no studies conducted which have investigated the emotional aspects of language perception and expression among international students in Hungary. To address this gap, it is crucial to investigate how bilinguals perceive and express emotions and swear words while taking into account a range of influencing factors.

Methods

The present study aims to investigate emotions associated with the L1 and L2 using a quantitative approach, while also attempting to explore the role of various independent variables. As such, the research seeks to answer the following questions:

1) How do international students in Hungary emotionally perceive the different languages they speak?
2) How do international students in Hungary perceive the emotional weight of the phrase "I love you" in the languages they speak?
3) How do international students in Hungary feel about S-T words in the languages they speak?
4) How are different background variables, including the frequency of L1 and L2 use, age of acquisition, context of L1 and L2 acquisition, level of proficiency, and gender, related to the emotions international students residing in Hungary experience in connection with the languages they speak?

Participants

The questionnaire was completed by 143 participants, consisting of 41 males and 102 females who comprised 29% and 71% of the sample, respectively. The participants' ages ranged from 19 to 64 ($M = 29$; $SD = 7 .5$). 35 nationalities were represented in the sample, among these Syrians ($N = 21$), Jordanians ($N = 18$), Russians ($N = 12$), and other nationalities with fewer than 10 representatives each. The participants spoke a total of 25 different first languages. Arabic speakers formed the largest group ($N = 63$), followed by Russian ($N = 14$), Spanish ($N = 9$), and 21 other language groups with fewer than 8 participants each. Regarding their L2, the participants spoke a total of 19 languages. English speakers comprised the largest group ($N = 101$), followed by French ($N = 16$) and Russian speakers ($N = 4$), as well as other language groups with fewer than 3 participants each. Participants self-reported their L2 proficiency as follows: intermediate ($N = 28$), upper intermediate ($N = 73$), and advanced level ($N = 42$). The participants can be considered as a representative sample of the international student population in Hungary as they live and study in different cities and universities and represent various degree levels and majors.

Instrument

The Bilingualism and Emotions Questionnaire (BEQ, Dewaele & Pavlenko, 2001–2003) was adapted for use in this study. While the general structure of the original instrument was retained, some items were added and others that were not related to the current study were removed.

The adapted BEQ questionnaire included scales measuring the emotional weight of the L1 and L2, the emotional weight of the phrase "I love you", the emotional weight of S-T words, their language preference for expressing their feelings, as well as their language preferences for expressing S-T words.

The adapted BEQ questionnaire consisted of a total of 41 items. This included five items related to the emotional weight of the participants' L1 and four items related to the emotional weight of their L2. The questionnaire also featured six items related to the emotional weight of the phrase

"I love you", out of which three were related to the L1 and three to the L2. Additionally, there were 12 items related to the emotional force of S-T words, out of which six were related to their L1 and six to their L2. Seven items collected data on the participants' language choice for expressing their feelings (four related to their L1 and three to their L2), and another seven items related to their language choice for using S-T words (four related to their L1 and three to their L2). Examples of the BEQ items can be found in Table 1. Participants were required to provide responses to all 41 items using a 5-point Likert scale, where 1 stood for "strongly disagree "and 5 for "strongly agree". The final section of the questionnaire consisted of a range of items designed to collect demographic data. The adapted BEQ questionnaire can be found in the Appendix.

Table 1

Sample Items From the Adapted BEQ

Name of the scale	Sample item
The emotional significance of the L1	My L1 is warm.
The emotional significance of the L2	My L2 is cold.
The emotional weight of the phrase "I love you" in the L1	"I love you" is more emotional in my L1.
The emotional weight of the phrase "I love you" in the L2	"I love you" is more emotional in my L2.
The emotional force of S-T words in the L1	Swear words feel stronger in my L1.
The emotional force of S-T words in the L2	Swear words feel stronger in my L2.
Language choice for expressing their feelings in the L1	The language I prefer to express my feelings in is my L1.
Language choice for expressing their feelings in the L2	The language I prefer to express my feelings in is my L2.
Language choice for using S-T words in the L1	The language I usually swear in is my L1.
Language choice for using S-T words in the L2	The language I usually swear in is my L2.

Data Collection

Participants were recruited through purposive and convenience sampling. The researchers selected the best-suited participants for the purpose of answering the research questions, with new participants recruited by other participants. The questionnaire was converted into a Google Form for data collection purposes and was distributed among the 143 participants by sharing the link to the questionnaire via email and social media platforms (i.e., Facebook and Instagram). Data collection occurred exclusively among international students studying in Hungary and took place over a period of 15 days from 15 April 2021 to 30 April 2021. The questionnaire was administered in English, as all the respondents were international students taking part in English-medium instruction. Therefore, their level of English was deemed sufficient for responding to the questions in English without difficulty.

Data Analysis

The responses provided by the participants were automatically transferred into a spreadsheet; minor modifications were made when formatting the dataset. Subsequently, the data were imported into a computer program for statistical analysis (i.e., version 20 of the Statistical Package for the Social Sciences, SPSS) was used for this purpose. As a first step, Cronbach alpha values were calculated to assess the reliability of the scales employed to measure the constructs. Following this step, descriptive and inferential statistical procedures were employed to characterize the sample and to compute the levels of emotionality among the participants' languages. The statistical tests used included paired samples *t*-tests, independent samples *t*-tests, one-way analysis of variance (ANOVA), and Pearson correlation. In the above-mentioned statistical procedures, the level of significance was set at $p < .05$.

Results

Reliability and Descriptive Statistics

All 10 scales demonstrated an acceptable level of reliability. Table 2 includes the Cronbach alpha values for the BEQ scales as well as the corresponding means and standard deviations. Given that most of the Cronbach alpha values exceeded .70, and that those falling below that figure exceed .60 (the minimum acceptable level as per Dörnyei, 2007), it can be concluded that the scales measuring the constructs under investigation were reliable.

Table 2

Descriptive Statistics for the Adapted BEQ

Scales (numbers of items)	Cronbach alpha	*M*	*SD*
Language choice for using S-T words in the L2 (3)	.69	4.53	.59
The emotional weight of the phrase "I love you" in the L1 (3)	.67	4.35	.97
Language choice to express feelings in the L1 (4)	.76	4.27	.62
The emotional significance of the L1 (5)	.92	4.24	.61
The emotional force of S-T words in the L1 (6)	.89	4.22	.63
The emotional weight of the phrase "I love you" in the L2 (3)	.65	3.70	.80
The emotional significance of the L2 (4)	.78	3.70	.58
The emotional force of S-T words in the L2 (6)	.90	3.60	.59
Language choice for expressing feelings in the L2 (3)	.70	3.60	.59
Language choice for using S-T words in the L1(4)	.78	3.20	.62

Emotions in the L1 and L2

A Paired samples *t*-test was conducted on the *emotional significance* scales in order to compare the emotional weight that participants attributed to each language. The results indicate a significant difference between the L1 and L2 constructs, $t(78) = -1.97$, $p < .001$. Specifically, the mean value for L1 emotionality ($M = 4.24$, $SD = 0.61$) was significantly higher than that for L2 emotionality ($M = 3.70$, $SD = 0.58$). This finding suggests that the participants consider their L1 to be more emotional than their L2 of the participants.

Another paired samples *t*-test was conducted with the two scales measuring the emotional weight of the phrase "I love you" in order to compare the emotional resonance of the phrase in the L1 and L2. A significant difference was observed in the scores for the emotional weight of "I love you" in the L1 (*M* = 4.35, *SD* = 0.97) and L2 (*M* = 3.70, *SD* = 0.80); *t* (78) = 6.43, *p* <.001. The results suggest that the phrase "I love you" is perceived to have more emotional weight in the L1 than in the L2.

To compare the emotional force of S-T words in the participants' L1 and L2, a paired samples *t*-test was carried out with the two scales measuring the *emotional force of S-T words* in their two languages. Based on the results, there is a significant difference in the scores for the emotional force of S-T words in the L1 (*M* = 4.22, *SD* = 0.63) and L2 (*M* = 3.60, *SD* = 0.59); *t*(76) = 6.45, *p* <.001. In light of these findings, it can be concluded that international students in Hungary perceive the emotional force of S-T words to be stronger in their L1 than in their L2.

With regard to preferences related to language use, a paired samples *t*-test was conducted to determine whether there is a difference in the scores for the preference to express emotions in the L1 or L2. The comparison of the mean values of the scales measuring *language choice to express feelings* in the L1 (*M* = 4.27, *SD* = .62) and in the L2 (*M* = 3.60, *SD* = .59) showed a statistically significant difference *t*(78) = 6.45, *p*<.001. The results indicate that international students in Hungary prefer to express their emotions in their L1 rather than their L2.

To further investigate language use preferences, a paired samples *t*-test was carried out to determine the language participants prefer for using S-T words. The results reveal a significant difference in bilinguals' scores for *language choice to use S-T words* in their two languages; bilinguals appear to prefer using their L2 (*M* = 4.53, *SD* = .59) for this purpose rather than their L1 (*M* = 3.20, *SD* = .62); *t*(78) = 6.45, *p* < .001. These findings suggest that international students in Hungary are more likely to use their L2 for S-T words rather than their L1.

In view of the results, it can be concluded that international students in Hungary perceive their L1 to be more emotional; specifically, the phrase

"I love you" and S-T words are perceived to be stronger in the L1. The findings also indicate that these students prefer to express their emotions in their L1, while the L2 is the preferred choice for using S-T words.

Background Variables Affecting Language Choice

A one-way ANOVA test was conducted to compare the emotional significance of the L2 in relation to participants' *frequency of L2 language use* (categorized as being used "every week", "every day", and "most hours of the day"). The aim of this test was to investigate whether the frequency of language use is associated with the emotional significance of the L2. The ANOVA test revealed no significant differences between the three student groups based on the above categorizations [$F(2,76) = 1.8$, $p = .83$]. These results indicate that the amount of time participants spend using the L2 does not impact the emotional significance of the L2; in other words, even if participants use their L2 frequently, the L1 is still considered to be more emotional.

Another one-way ANOVA test was conducted to compare the emotional significance of the L2 based on the *age groups at which participants learned the language* (i.e., before 2, between 3-13, and after 13). These age groups are the same age groups used by Dewaele (2010). The ANOVA test revealed no significant differences between these three student groups [$F(2,76) = .028$, $p = .97$]. These results suggest that the age at which participants learned their L2 does not affect the emotional significance they attach to it. It can be concluded that whether the L2 is acquired at an early age or later in life has no relationship to its emotional significance.

A third one-way ANOVA test was conducted to compare the emotional significance of the L2 among students who learned their L2 in *different contexts* (i.e., at home, school, or in a mixed context). The test showed significant differences between the three student groups [$F(2,76) = 4.05$, $p < .001$]. To identify which groups were significantly different from each other, post hoc tests were performed. These tests showed that the mean score for the home context ($M = 3.30$, $SD = .72$) and mixed context ($M = 3.60$, $SD = .72$) differed significantly from the school context ($M = 3.10$, $SD = 0.84$). The

findings suggest that participants who learned their L2 in a school setting reported a lower level of emotional significance than those who learned it in a naturalistic context (home) or mixed context (home and school). As such, the context of L2 acquisition appears to affect its emotional significance of the L2. For example, if the L2 is acquired at home or in a mixed setting rather than at school, it is likely to be perceived as having a greater emotional impact.

In order to examine the role of *gender* as a background variable, an independent-sample *t*-test was conducted to compare the mean scores for the emotional significance of the L2 between male and female participants to determine whether gender is linked to the emotional significance of the L2. The results revealed a significant difference in the scores between males (*M* = 3.03, SD = .84) and females (*M* = 3.35, *SD* = .80) *t* (2.1) =14, *p*< .001. These findings suggest that gender is linked to the emotional significance of the L2, with female participants reporting a higher level of emotional significance for their L2 than male participants.

Finally, to assess the strength of the relationship between participants' *perceived level of proficiency* and the emotional significance of their L2, a Pearson product-moment correlation was calculated. The correlation coefficient indicated a significant, albeit moderate, positive correlation between participants' L2 emotional significance and their proficiency level, (r = .365, p < .001). The results show that proficiency level is associated with the emotional significance of the L2. In other words, it can be said that when the level of L2 proficiency increases, its emotional significance increases as well. Participants who are highly proficient in their L2 are more likely to consider it to be more emotional.

The results indicate that certain factors were found to be associated with the emotional significance of the L2, while other factors were not. Specifically, the context in which the language was acquired, the level of proficiency, and gender were shown to be linked with the emotional significance of the L2. In contrast, the frequency of use and the age of language acquisition do not appear to affect the emotional significance of the L2.

Discussion

The first research question aimed to shed light on how bilinguals feel about their L1 and L2. The findings suggest that participants perceive their L1 as more emotional compared with their L2. The results are consistent with earlier studies, which propose that the L1 is the language that is emotionally closer to bilinguals in comparison to the L2 (Altarriba, 2003; Dewaele & Pavlenko, 2002; Dewaele, 2010; Harris et al., 2003; Pavlenko, 2006).

The second research question investigated whether participants considered the phrase "I love you" to be more emotional in their L1 or L2. Results showed that the phrase carries more emotional weight in the L1. These findings align with a previous study by Dewaele (2008), who investigated the emotional weight of the phrase "I love you" in different languages spoken by multilinguals. In that study, participants had to choose between three options: to indicate whether the phrase was strongest in their L1, L2, or L3. Most speakers reported that the phrase held the most emotional weight in their L1.

The third research question examined how participants perceive S-T words. The results show that participants consider S-T words to be more powerful in their L1. Additionally, the study investigated which language participants prefer for using S-T words. The results suggest that participants prefer to swear and use taboo words in their L2. It can be hypothesized that participants use S-T words in their L2 because they carry less emotional significance in that language compared to their L1. These findings are consistent with the results of Dewaele's studies (2004, 2005), in which participants confirmed that they prefer to swear in their L2 and that they consider S-T words more powerful in their L1 than in their L2; therefore, they prefer to express them using the L2.

The final research question addressed the association between various factors and emotional perceptions. Participants' frequency of language use, age of acquisition, the context of L2 acquisition, gender, and proficiency level were investigated to determine their relationship with bilinguals' emotions in their two languages. The results suggest that some variables are associated with bilinguals' emotions, while other variables are not.

The results revealed that neither the frequency of language use nor the age of acquisition is associated with emotional perceptions related to the L1 and L2 of the participants. This suggests that the L1 is still perceived as more emotional than the L2, regardless of whether the L2 is learned at an early age or used more frequently. Results of the current study are consistent with Dewaele (2011), suggesting that participants who use both languages daily and acquired their L2 at an early age still perceive their L1 as more emotional compared to their L2. The existing literature presents mixed results regarding the influence of these factors: some studies suggest that the age of acquisition and frequency of language use in certain cases might have a significant effect on the perception of emotions in the L2 (Dewaele, 2004, 2007, 2008).

The final three variables examined in the current study were the perceived level of proficiency, the context of L2 acquisition, and gender. Results indicate that these variables have a significant association with emotional perceptions Regarding the context of L2 acquisition, existing research suggests that one factor which contribute to the perception of the L2 as less emotionally charged may be the school context. The L2 is usually learned in a formal context in which emotional topics are less prevalent, whereas the L1 is typically learned in the emotionally rich environment of the home. (Dewaele, 2004, 2005, 2006, 2007; Harris, 2004; Pavlenko, 2006).

Another variable found to influence emotional perceptions was gender. Previous research has presented two different views on the relationship between gender and emotions. One line of research argues that gender is not linked to emotions (Lutz, 1996; Tannen, 1990), while the other view claims that gender is linked to the experience of different emotions, with females being generally more emotional than males (Dewaele & Véronique, 2001; Holmes, 2013). In their study, Dewaele and Véronique (2001) found that female participants were more emotional than male participants when using their L2 (i.e., French). These results are in line with the current study, confirming that emotions are influenced by gender.

Finally, the level of proficiency was found to be significantly associated with emotions in the L2: participants who reported higher levels of proficiency also reported greater emotional significance for their L2 than

those who were less proficient. These findings align with previous research in which the level of proficiency was found to influence emotions in L2 (Dewaele, 2004, 2005, 2006,2007; Pavlenko, 2012); however, they contradict Graham et al. (2001), who found that the level of proficiency does not affect emotions.

Conclusion

The current study investigated how bilinguals perceive and express their emotions in the languages they speak, while also exploring various factors that could influence these emotions. Quantitative analysis of the data from 143 international students in Hungary indicates that there are differences in the perception and expression of emotions in their L1 and L2, with the L1 generally perceived as more emotionally loaded. Results showed that neither the frequency of language use nor the age of acquisition are associated with emotional preferences for the L2. Although participants reported learning their L2 at an early age and using it daily, they still considered their L1 to be more emotionally significant when compared to the L2. However, the study did find that the context of acquisition, gender, and the proficiency level influence emotions. Participants who are highly proficient in their L2 or who acquired it in a home setting tend to perceive it as more emotional compared to those with lower proficiency levels or those who acquired their L2 at school. Gender was also found to impact emotions, with females tending to perceive their L2 as more emotional.

In summary, there is a consensus among researchers that the L1 carries greater emotional intensity compared to L2. Researchers also agree that "I love you" and S-T words are considered more emotional in the L1 than in the L2. In addition, the L1 is the preferred language for expressing emotions, while the L2 tends to be used to express S-T words, likely due to its lower emotional weight. However, the literature presents inconsistent results regarding the specific factors that have an impact on emotions and cause the L2 to be perceived as less emotional than the L1. While our results lend support to the role of context, gender, and self-perceived proficiency level, they do not offer proof for the influence of the age of acquisition or

frequency of use. Nevertheless, we hope that our research contributes to a broader understanding of how international students in Hungary perceive and express their emotions, and not only complements results from previous studies but also highlights the importance of examining emotions together with other aspects which still require further investigation.

As with all research, the current study has some limitations. One potential issue is that self-reported data can introduce social desirability bias, which can threaten the validity of the instrument (Li et al., 2018). Another shortcoming is that the study's quantitative approach fails to provide in-depth insights into the underlying reasons behind some of our findings. To address this, future research can complement previous studies by adopting a mixed-methods design; by combining both quantitative and qualitative approaches, researchers may be able to gain deeper insights into this area (Dörnyei, 2007).

References

Altarriba, J. (2003). Does cariño equal "liking"? A theoretical approach to conceptual nonequivalence between languages. *International Journal of Bilingualism*, *7*(3), 305–22. https://doi.org/10.1177/13670069030070030501

Altarriba, J. (2008). Expressions of emotion as mediated by context. *Bilingualism: Language and Cognition, 11*(2), 165–167. https://doi.org/10.1017/S1366728908003295

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory. *Cognition and Emotion, 8*(6), 503–514. https://doi.org/10.1080/02699939408408956

Aycicegi, A., & Harris, C. (2004). Brief report bilinguals' recall and recognition of emotion words. *Cognition and Emotion*, *18*(7), 977–987. https://doi.org/10.1080/02699930341000301

Caldwell-Harris, C. L. (2014). Emotionality differences between a native and foreign language: theoretical implications. *Frontiers in Psychology*, *5*, 1055. https://doi.org/10.3389/fpsyg.2014.01055

Colbeck, K. L., & Bowers, J. S. (2012). Blinded by taboo words in L1 but not L2. *Emotion*, *12*(2), 217–222. https://doi.org/10.1037/a0026387

Dewaele, J.-M. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of Multilingual and Multicultural Development*, *25*(2–3), 204–222. https://doi.org/10.1080/01434630408666529

Dewaele, J.-M. (2005). Investigating the psychological and emotional dimensions in instructed language learning: Obstacles and possibilities. *The Modern Language Journal*, *89*(3), 367–380. https://doi.org/10.1111/j.1540-4781.2005.00311.x

Dewaele, J.-M. (2007). The effect of multilingualism, sociobiographical, and situational factors on communicative anxiety and foreign language anxiety of mature language learners. *International Journal of Bilingualism*, *11*(4), 391–409. https://doi.org/10.1177/13670069070110040301

Dewaele, J.-M. (2008). The emotional weight of I love you in multilinguals' languages. *Journal of Pragmatics*, *40*(10), 1753–1780. https://doi.org/10.1016/j.pragma.2008.03.002

Dewaele, J.-M. (2011). Reflections on the emotional and psychological aspects of foreign language learning and use. *Anglistik: International Journal of English Studies, 22*(1), 23–42.

Dewaele, J.-M. (2017). Self-reported frequency of swearing in English: do situational, psychological and sociobiographical variables have similar effects on first and foreign language users? *Journal of Multilingual and Multicultural Development, 38*(4), 330–345. https://doi.org/10.1080/01434632.2016.1201092

Dewaele, J.-M., & Pavlenko, A. (2001–2003). *Web questionnaire on bilingualism and emotion*. University of London.

Dewaele, J.-M., & Pavlenko, A. (2002). Emotion vocabulary in interlanguage. *Language Learning*, *52*(2), 263–322. https://doi.org/10.1111/0023-8333.00185

Dewaele, J.-M., & Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: A cross-sectional study. *Bilingualism: Language and Cognition*, *4*(3), 275–297. https://doi.org/10.1017/s136672890100044x

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.

Ferré, P., García, T., Fraga, I., Sánchez-Casas, R., & Molero, M. (2010). Memory for emotional words in bilinguals: Do words have the same emotional intensity in the first and in the second language? *Cognition & Emotion*, *24*(5), 760–785. https://doi.org/10.1080/02699930902985779

Graham, C. R., Hamblin, A. W., & Feldstein, S. (2001). Recognition of emotion in English voices by speakers of Japanese, Spanish and English. *International Review of Applied Linguistics in Language Teaching, 39*(1), 19–37. https://doi.org/10.1515/iral.39.1.19

Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press. https://doi.org/10.1093/oso/9780199281282.001.0001

Grosjean, F. (2021). *Life as a bilingual: Knowing and using two or more languages.* Cambridge University Press. https://doi.org/10.1017/9781108975490

Harris, C. L. (2004). Bilingual speakers in the lab: Psychophysiological measures of emotional reactivity. *Journal of Multilingual and Multicultural Development*, 25(2-3), 223–247. https://doi.org/10.1080/01434630408666530

Harris, C. L., Ayçíçeğí, A., & Gleason, J. B. (2003). Taboo words and reprimands elicit greater autonomic reactivity in a first language than in a second language. *Applied Psycholinguistics*, 24(4), 561–579. https://doi.org/10.1017/s0142716403000286

Harris, C. L., Gleason, J. B., & Ayçiçeği, A. (2006). When is a first language more emotional? Psychophysiological evidence from bilingual speakers. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression, and representation* (pp. 257–283). Multilingual Matters.

Ivaz, L., Griffin, K. L., & Duñabeitia, J. A. (2019). Self-bias and the emotionality of foreign languages. *Quarterly Journal of Experimental Psychology*, 72(1), 76–89. https://doi.org/10.1177/1747021818781

King, G. (2011). Ensuring the data-rich future of the social sciences. *Science*, 331(6018), 719–721. https://doi.org/10.1126/science.1197872

Koven, M. (2004). Getting 'emotional' in two languages: Bilinguals' verbal performance of affect in narratives of personal experience. *Text & Talk*, 24(4), 471–515. https://doi.org/10.1515/text.2004.24.4.471

Li, C., Jiang, G., & Dewaele, J.-M. (2018). Understanding Chinese high school students' foreign language enjoyment: Validation of the Chinese version of the foreign language enjoyment scale. *System*, *76*, 183–196. https://doi.org/10.1016/j.system.2018.06.004

Liebkind, K. (1995). Bilingual identity. *European Education*, *27*(3), 80–87. https://doi.org/10.2753/EUE1056-4934270380

Lutz, C. A. (1996). Engendered emotion: Gender, power, and the rhetoric of emotional control in American discourse. In R. Harré & W. G. Parrott (Eds.), *The emotions: social, cultural and biological dimensions* (pp. 151–170). Sage. https://doi.org/10.4135/9781446221952.n11

Marian, V., & Kaushanskaya, M. (2008). Words, feelings, and bilingualism: Cross-linguistic differences in emotionality of autobiographical memories. *Mental Lexicon, 3*(1), 72–91. https://doi.org/10.1075/ml.3.1.06mar

Pavlenko, A. (2006). Bilingual selves. In A. Pavleko (Ed.), *Bilingual minds: Emotional experience, expression and representation* (pp. 1–33). Multilingual Matters. https://doi.org/10.21832/9781853598746-003

Pavlenko, A. (2008). Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and Cognition, 11*(2), 147–164. https://doi.org/10.1017/S1366728908003283

Pavlenko, A. (2012). Affective processing in bilingual speakers: Disembodied cognition? *International Journal of Psychology, 47*(6), 405–428. https://doi.org/10.1080/00207594.2012.743665

Schrauf, R. W. (2000). Bilingual autobiographical memory: Experimental studies and clinical cases. *Culture & Psychology, 6*(4), 387–417. https://doi.org/10.1177/1354067X0064001

Shao, K., Pekrun, R., & Nicholson, L. J. (2019). Emotions in classroom language learning: What can we learn from achievement emotion research? *System, 86*, 102–121. https://doi.org/10.1016/j.system.2019.102121

Tannen, D. (1990). Gender differences in topical coherence: Creating involvement in best friends' talk. *Discourse Processes, 13*(1), 73–90. https://doi.org/10.1080/01638539009544747

Appendix

Adapted from: **Dewaele, J.-M. and Pavlenko, A. (2001-2003) Web questionnaire *Bilingualism and Emotions*. University of London.**

Emotions Questionnaire

Hello, my name is Dina AL Madanat. I am a PhD student at Eötvös Loránd university. I would like to ask you to help me by answering the following questions concerning emotions in the languages you speak. I am interested in your personal opinions. Please answer sincerely, as this will help ensure the success of my research. I realize how valuable your time is, which is why I have designed this questionnaire to take no more than 10 minutes to complete. Your responses to questionnaire will remain completely anonymous.

I appreciate your input!

Dina

\---
Linguistic information

1ˢᵗ Language (L1)   2ⁿᵈ Language (L2)
Age at which you started learning the language
Context of Acquisition (outside of school), (at school) or mixed setting (at school and outside of school)
\---
On the scale from 1 (least proficient) to 5 (fully fluent), how do you rate yourself in terms of speaking, listening, reading, and writing in all of the languages in question?
L1
L2
\---

How frequently do you use each of the languages? Never=0, every year=1, every month=2, every week=3, every day=4, most hours of the day=5)
Frequency
L1
L2

The following is a series of statements about the languages you speak with which some people may agree or disagree. Please indicate your opinion after each statement by putting an "X" in the box that best reflects the extent to which you agree or disagree with the statement. There are no wrong or right answers.

		Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
A	**Emotional weight of I love you**					
1	I think (I love you) has more emotional weight in my first language than in my second language					
2	I find the phrase (I love you) is more emotional in my L1 than in my L2					
3	The phrase (I love you) has a deeper effect in my L1 than in my L2					
4	(I love you) sounds more romantic in my L2 than in my L1					
5	I love you is more meaningful in my L2 than in my L1					
6	I feel that the phrase (I love you) is more emotional in my L2 than my L1					
B	**Emotional significance of the L1 and L2**					
1	I think my L1 is warm					
2	I think my L1 is passionate					
3	I think my L2 is cold					
4	I think my L2 is passionate					
5	My L1 feels closer to my heart than my L2					
6	My L2 feels closer to my heart than my L1					
7	My L1 is emotional					
8	My L2 is emotional					
9	My L1 is romantic					
C	**Language preference to express emotions**					
1	I prefer to use my L1 rather than my L2 in intimate conversations					
2	I find it easier to talk about emotional topics in my L1 than my L2					
3	I find it easier to express my feelings in my L1 than my L2					

		Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
4	The language I prefer expressing my emotions in is my L2					
5	The language I prefer expressing my emotions in is my L1					
6	I prefer to use terms of affection in my L2 than in my L1					
7	I prefer using emotional terms in my L2 than my L1					
B	**Emotional weight of taboo and swear words in the L1 - L2**					
1	Swear words in my L1 have more emotional weight for me than in my L1					
2	Taboo words in my L1 have more emotional weight than in my second language					
3	Swear words in my L1 feel more serious than in my L2					
4	Taboo words in my L1 feel more serious than in my L2					
5	Swear words in my L1 have a stronger effect than in my L2					
6	Taboo words in my L1 have a stronger effect than in my L2					
7	Swear words in my L2 feel more serious than in my L1					
8	Taboo words in my L2 feels more serious than in my L1					
9	Swear words have a deeper effect in my L2 than in my L1					
10	Taboo words have a deeper effect in my L2 than in my L1					
11	I feel that swear words are more shameful in my L2 than in my L1					
12	Swear words feel stronger in my L2 than in my L1					

		Never	Rarely	Sometimes	Frequently	All the time	(I don't swear)
C	**Expression of swear words**						
1	The language I usually swear in is my L1						
2	The language I usually swear in is my L2						
3	The language I usually taboo words in is my L1						
4	The language I usually use taboo words in is my L2						
5	I prefer using swear words in my L1						
6	I prefer using taboo words in my L2						
7	I find it easy to use swear words in my L1						

If you have any questions or comments about the questionnaire, please email me: dinamdanatl@email.com

Thank you very much for your help ☺

3

Enjoyment and Anxiety in Two Foreign Languages: An Idiodynamic Approach

Kholoud Al-Madanat
kholoud@student.elte.hu

Abstract

Humans experience a diverse spectrum of emotions on a daily basis. Since emotional responses can be brief, traditional psychometric approaches based on retrospective data and averaging out results across groups may not effectively capture the dynamic nature of emotions (Dewaele et al., 2018). For this reason, this study deviates from traditional methods of collecting and analyzing data by employing an idiodynamic approach (MacIntyre, 2012) and exploring emotions from a Complex Dynamic Systems Theory perspective on a moment-to-moment timescale. The rapidly changing relationship between two emotions, anxiety and enjoyment, is investigated in regard to two foreign languages: English and French. The participants consisted of two international students studying in Hungary who speak French and English in addition to their mother tongue, Arabic. The participants were required to perform oral tasks in the two languages and then rate their per-second fluctuations in the two emotions for each language while watching video recordings of their tasks. After completing the ratings, the participants were interviewed in order to gain an in-depth understanding of the reasons behind the fluctuations in their ratings. The findings show that the dynamic relationship between the two emotions was evident in both languages and for different tasks. The fluctuations of emotions in both English and French were attributed to a number of internal and external factors.

Keywords: emotion, enjoyment, anxiety, idiodynamic approach, Complex Dynamic Systems Theory

Enjoyment and Anxiety in Two Foreign Languages:
An Idiodynamic Approach

Recent trends in SLA research have led to increased attention being dedicated to examining emotions. Emotions play a prominent role in our lives, with individuals experiencing a diverse spectrum of emotions throughout a given day, including anxiety or enjoyment. Emotions can act as a motivating or demotivating force which can incite, hinder, or stall individual action. The powerful impact of emotions on people's lives has made it a subject of interest for research in various disciplines, including education and applied linguistics. Swain (2013) argued that "emotions are the elephants in the room – poorly studied, poorly understood, seen as inferior to rational thought" (p. 195). With the exception of anxiety, little emphasis has been directed to emotions in second language acquisition (SLA) studies (Dewaele & Li, 2018). Prior (2019) contends that scholarly interest in the emotional components of language learning and teaching is witnessing significant growth, and that the time has come to acknowledge this influential aspect of SLA.

The term *affect* has been used to indicate aspects related to emotions in SLA. Affect can be defined as "broadly aspects of emotion, feeling, mood or attitude which can condition behavior and influence language learning" (Arnold & Brown, 1999, p. 1). This broad term encompasses a number of different concepts, including feelings, personality factors, mood, attitudes, and judgment. Most of the studies on affective factors have been quantitative in nature, investigating the cause-effect relationship between affective variables and learning outcomes (Ganschow & Sparks, 2001; Horwitz, 2001). Other researchers have examined the role of emotions through the lens of bilingualism and explored how emotions are expressed in different languages (Pavlenko, 2006, 2007).

Since emotional responses can be brief, traditional approaches based on retrospective data and averaging out results across groups may not be the most effective way to study emotions, as such methods may fail to capture their unique dynamic nature (Boudreau et al., 2018). For this reason, this study deviates from traditional methods data collection and analysis

by employing the idiodynamic approach to study emotions in the framework of Complex Dynamic Systems Theory (CDST).

The study examines the relationship between one positive emotion (i.e., enjoyment) and one negative emotion (i.e., anxiety) for two multilingual individuals, focusing on the fluctuations of these emotions for each participant when using their two foreign languages: English and French. Enjoyment and anxiety were chosen for the purposes of the current study due to their unique relationship. According to Dewaele and MacIntyre (2014), rather than viewing these two emotions as opposite ends of an emotional spectrum, where one implies the absence of the other, they should be considered as separate emotions, each with distinct trajectories that may align or diverge over time.

While the idiodynamic approach has been used in a number of studies to investigate the fluctuations of two emotions when using a single foreign language, to the best of the author's knowledge, no studies to date have employed the idiodynamic approach to examine the fluctuations of these emotions among participants using two foreign languages. This study aims to decipher the unique and dynamic nature of emotions on a moment-to-moment basis, linking these emotional fluctuations to the second language (L2) processes (see MacIntyre & Serroul, 2014).

Evolution of Emotion Research in SLA and Language Learning

The last decade has witnessed a surge in studies on emotions in the fields of SLA and language learning. Dewaele and Li (2020) categorized this research into three distinct phases. In the first phase, the Emotion Avoidance Phase, emotions were recognized but regarded as an irrational aspect of language learning. Cognitive processes were favored during this period, being viewed as more scientific (Prior, 2019). The next phase, the Anxiety Prevailing Phase, was marked by an increased awareness that emotion and cognition were related, and that emotion plays a critical role in language learning. However, a predominantly negative focus centered on anxiety prevailed (MacIntyre, 2017). The third stage, the Positive and Negative Emotions Phase, has witnessed a shift from focusing solely on learners' anxiety to exploring both the positive and negative emotions in language

learning and teaching (MacIntyre, 2016). The significance of this broader emotional range is evidenced by Dewaele and MacIntyre (2016), who compare positive and negative emotions to "the right and left foot of a learner" (p. 215).

Complex Dynamic Systems Theory and the Idiodynamic Method

The concept of change is central to CDST. The beginnings of the framework can be traced back to the 1960s, when it was used in various fields such as physics, biology, and demography. The application of CDST to SLA can be attributed to Larsen-Freeman (1997), who described second language acquisition as a "dynamic, complex and non-linear process" (p. 142). Since then, many SLA researchers have followed suit and integrated CDST into their research (de Bot et al., 2005; de Bot & Makoni, 2005; MacIntyre & Legatto, 2011; Wind, 2020). According to de Bot et al. (2007), language development exhibits key characteristics found in complex dynamic systems. These include sensitivity to initial conditions (Verspoor, 2015), interconnectedness between systems and sub-systems (Wind, 2020), the emergence of attractor states during development (Waninge et al., 2014; Wind & Harding, 2020), and variation both within individuals and across individuals (Boudreau, et al., 2018; Larsen-Freeman, 2006; Wind, 2014; Wind, 2021). Lowie (2013) further summarized the characteristics of language development which make it well-suited to be studied from a CDST perspective, highlighting "the existence of interconnected subsystems; the tendency to self-organization; and the occurrence of nonlinear, chaotic patterns of development" (p. 1).

Investigating individual differences in second language learning within the framework of CDST was first advocated by Dörnyei (2009, 2010) and Dekeyser (2012). Instead of investigating isolated relationships between variables, the approach has shifted the focus to variability, interaction, and change. Among the individual differences that have been investigated from a CDST perspective are L2 motivation (Waninge et al., 2014) self-regulation (Han & Hiver, 2018; Wind & Harding, 2020) and willingness to communicate (MacIntyre & Legatto, 2011). Considering the dynamic nature of emotions (Boudreau, et al., 2018; Dewaele, et al., 2020), applying CDST to SLA studies on emotions appears to be a well-justified approach.

According to Boudreau et al. (2018), emotions have four characteristics that render them an ideal subject for research within a CDST framework. First, they are not stable, fluctuating over time as a result of their interactions with other systems. Emotions can also be described as being emergent, as they result from the interactions between a number of elements such as physical changes, behavioral tendencies, and cognition. They are also an open system which can be influenced by both internal and external factors. Yet, despite the many interconnected components of the mechanisms underpinning emotion, perception, language, and communication, there is a sense of coherence in our emotions, even when they are in conflict. These components self-organize, converging in meaningful ways and persisting over certain periods of time. Another distinguishing feature of emotions is their adaptive nature, as they can guide individuals to take specific actions that help them adapt to various situations.

In order to investigate individuals' moment-to-moment experiences of enjoyment and anxiety while using two foreign languages, English and French, in the framework of CDST, the current study employs the idiodynamic method. Through this method, researchers can gain deeper insights into the dynamic changes of variables on a per-second timescale. Moments of change can be identified and traced through participants' self-ratings (MacIntyre & Legatto, 2011). As outlined by Gregersen and MacIntyre (2017) and MacIntyre and Legatto (2011), the method comprises four stages:

1. A video of a communication task is recorded which is shown to participants as soon as they finish the task.
2. The participants use special software to complete a moment-by-moment self-rating of the investigated variable.
3. The participants view a graph produced by the software and an interview is held in order to explore the reasons behind the fluctuations in emotions depicted in the graph.
4. The interview is transcribed for analysis.

MacIntyre and Legatto (2011) were the first to apply this method in their study exploring the dynamics of willingness to communicate. The findings of their study revealed that fluctuations during conversations covering

diverse topics were evident not only across individuals but also within each individual. The approach has since gained traction, with many scholars in the field utilizing this method in their research (e.g., Gregersen & MacIntyre, 2017; Gregersen et al., 2014, 2017; MacIntyre, 2012; MacIntyre & Legatto, 2011).

The Broaden and Build Theory

When using a foreign language, speakers experience both positive and negative emotions. Prior to the advent of positive psychology in SLA, negative emotions such as anxiety received the most attention from researchers, who associated the absence of negative emotions with the emergence of positive ones. Positive psychology challenged this perspective, with recent findings in this field revealing that positive emotions (e.g., enjoyment) and negative ones (e.g., anxiety) arise from a range of experiences and serve several functions (Fredrickson, 2013; MacIntyre & Gregersen, 2012).

Positive psychology was first introduced to the field of applied linguistics through a paper co-authored by MacIntyre and Gregersen (2012). In this paper, they made a reference to Fredrickson's (2001, 2003, 2004) Broaden-and-Build Theory which focuses on the impact of positive emotions on an individual's thought–action repertoire. Its central proposition is that positive emotions broaden an individual's thought-action resources and help to build the social, physical, psychological, and intellectual foundation needed for current and future achievement (Fredrickson, 2001, 2013). Fredrickson (2006) argued that positive emotions "(a) broaden people's attention and thinking, (b) undo lingering negative emotional arousal, (c) fuel psychological resilience, (d) build consequential personal resources, and (e) trigger upward spirals toward greater well-being in the future" (p. 97).

According to the Broaden-and-Build Theory, negative emotions narrow individuals' thought-action repertoires, which are linked to specific action tendencies (Lazarus, 1991). For example, Reeve (2014) notes that the negative emotion of contempt triggers social hierarchies. Fredrickson (1998) stated that positive emotions broaden specific action tendencies rather than narrowing them down. For example, it is difficult to pinpoint a specific action tendency linked to the positive emotion of enjoyment.

Applying the principles of this theory to English language learning has significant implications for the field of English language teaching (ELT). Teachers should aim to create an environment that not only reduces negative emotions but also enhances positive ones. Although positive emotions have been shown to direct one's thinking and actions in a way that can help combat and regulate negative ones (Fredrickson, 2004), the benefits of fostering positive emotions should not be marginalized and restricted solely to this role. The presence of positive emotions can improve the learning process by raising students' awareness of the linguistic input they receive, helping them better understand and absorb language forms and become more open to using diverse problem-solving techniques (Boudreau et al., 2018; Piechurska-Kuciel, 2017), In the long run. This can enhance learners' foreign or second language competence (Elahi Shirvan & Therian, 2018).

Enjoyment and Anxiety Relationship

Enjoyment is a particularly salient positive emotion that has attracted scholarly attention in the field of SLA (Dewaele & MacIntyre, 2014; Gregersen et al., 2017). It is an emotion that is not only experienced through the satisfaction of one's needs, but also arises when an individual looks beyond their needs and achieves something new, unexpected, and surprising (Csíkszentmihályi, 2008); this characterizes enjoyment by a "sense of novelty and accomplishment" (p. 46). As such, the emotion plays a vital role in long-term well-being and personal development (Seligman & Csíkszentmihályi, 2000). In addition, enjoyment has also been linked to higher self-assessed foreign language ability and improved scores on reading comprehension tests (Brantmeier, 2005).

The concept of enjoyment has been examined from various perspectives (Dewaele & MacIntyre, 2016; Li., 2018), including its origins and impact (Dewaele & Alfawzan, 2018; Dewaele et al., 2018) and its relationship with anxiety (Dewaele & MacIntyre, 2014, 2016). In their investigation of the different types of positive and negative emotions that Hungarian students experience when using their L2 (i.e., pride, contentment, comfort, relaxation, enjoyment, and anxiety), Piniel and Albert (2018) found that enjoyment and anxiety were the most recurring emotions. These findings

were corroborated by other studies (Elahi Shirvan & Taherian, 2018; Li, 2018). Research studying the relationship between enjoyment and anxiety has also found that enjoyment can reduce levels of anxiety, and that enjoyment and anxiety can co-occur (Dewaele & Alfawzan, 2018, Dewaele & MacIntyre, 2014, 2016; Dewaele et al., 2016).

The Current Study

The current study aims at describing the dynamic relationship between enjoyment and anxiety during the use of two foreign languages: English and French. This relationship is examined on a per-second timescale to gain a more comprehensive picture of how these emotions fluctuate during second language communication tasks. The participants carried out two tasks in a controlled setting: a story telling task and a structured interview.

The research questions for the present study are as follows:

1. What is the relationship between enjoyment and anxiety, and does this relationship change over time within each individual for each language?
2. To what do participants attribute the observed fluctuations in emotion?

Methods

Participants

The participants in this study were two international students, aged 26 and 28, studying at a large Hungarian university. One was a master's student while the other one was a PhD student. Both spoke Arabic as a first language and English and French as foreign languages. The selection process aimed to ensure that both individuals were proficient in both English and French. To meet this requirement, the researcher chose two Tunisian students, as Tunisians are known for their proficiency in French, a language which is commonly spoken alongside Arabic and which occupies a prominent position in Tunisia's education system. Concerning English, the two participants were completing their post-graduate studies in English, which

guaranteed an advanced level of English proficiency. Additionally, both participants self-assessed their language proficiency levels in both languages to be at an advanced level. To maintain anonymity, both participants were assigned pseudonyms: Nadia and Noor; therefore, the names appearing in this research are fictitious.

Materials

The participants used the Anion Variable Tester V2 software (Macintyre & Legatto, 2011) to rate their moment-to-moment levels of anxiety and enjoyment while viewing their recorded tasks. The scale of the ratings ranged from -5 to +5, with -5 indicating the very low enjoyment or anxiety and +5 indicating the highest value for each emotion. Once the participants completed the ratings, the program produced graphs illustrating the participants' emotional fluctuations during the communication tasks. Before starting the experiment, a full description of the program and the rating process was provided to the participants. They were then given time to practice using the program until they were confident in their ability to navigate the rating process effectively.

Procedure

The data collection took place on two separate occasions, with a four-week interval between them. Since the experiment was conducted online, the participants were asked to send the researcher consent and video release forms via email before their first sessions. On the day preceding the experiment, the researcher contacted the participants via email and asked them to choose a picture they find enjoyable to discuss in English for their upcoming session. For their first session, the participants were contacted through video call on Microsoft Teams. The session began with a short formal introduction from the researcher, who addressed ethical issues such as anonymity and the option to withdraw from the experiment at any time. Following these preliminary steps, the participants proceeded with the required tasks. A similar procedure was followed four weeks later, with the notable difference that the participants were informed that they would be using French for the second session.

The tasks the participants had to perform were as follows:

1. Photo narrative task: the participants were asked to talk about the photo they had chosen for three to five minutes. The narration was done in English for the first session and in French for the second.

2. Oral interview task: the participants were required to answer five oral interview style questions. The questions were adapted from MacIntyre and Legatto's (2011) study. The questions were read in English for both the English and French sessions. The questions were:
   a) Describe what you are wearing.
   b) Discuss the education system in Tunisia in some detail.
   c) Count to 100 by tens.
   d) Give directions from your dorm to the nearest supermarket.
   e) Discuss the role of the parliament in the Tunisian system of government.

The two sessions were recorded on Microsoft Teams. The participants completed the photo narrative task followed by the oral interview questions. After finishing the two tasks, the video recordings were converted using an online application and uploaded to the Anion Variable Tester V2 software. Next, the participants were asked to rate their levels of enjoyment and anxiety. The software then produced a graph which visually represented the emotional fluctuations experienced during each communication task. The participants and the researcher discussed the graphs, focusing on the spikes and dips in their ratings. At this stage, the researcher asked questions drawn from (Boudreau et al., 2018). The questions included:

1. You rated your (anxiety/enjoyment) as particularly high at this particular interval, can you explain why?
2. Your (anxiety/enjoyment) ratings were particularly low here, why do you think that was?
3. Your ratings remained stable and neutral over this time interval. Can you explain why?
4. Do you have anything else to add about your overall experience with the task?

The interviews were conducted in the participants' first language and were later transcribed. Only the excerpts quoted in the study were translated into English. In order to explore the emotional dynamics for each individual in each language, the data were analyzed within the participants rather than across participants. First, in order to gain a general understanding of the strength of the relationship between the two emotions, the overall correlation between enjoyment and anxiety was calculated for each communication task completed by each participant in both languages. Next, in order to gain a more detailed understanding of the dynamic relationships between the two emotions experienced during each task, interviews with the participants were carried out which provided insights on why they rated their emotions as high, low, or neutral during the tasks. Thematic analysis was used to analyze the data. The initial analysis of the data involved reading and re-reading the transcripts, after which categories and themes were established (Lemke, 1998). Question 1 sought responses related to the factors influencing low anxiety or enjoyment ratings. Question 2 focused on explanations for stable and neutral ratings, including feelings of familiarity or emotional detachment. With Question 3, the researcher sought further insights into the participants' overall task experience, considering factors like satisfaction, engagement, or suggestions for improvement. By employing the thematic analysis, the researcher gained a comprehensive understanding of the participants' perspectives, enabling the extraction of meaningful insights and conclusions regarding their emotional experiences.

Results

In order to explore the relationship between anxiety and enjoyment, the data was analysed on a per-individual basis. First, the correlations for the photo narrative task and the oral interview task were calculated for each participant and in each language separately (see Table 1).

Table 1

Correlations Between Per-Second Ratings of Enjoyment and Anxiety for Participant 1 (Nadia) and Participant 2 (Noor) for Each Task Type in English and French.

Participant	Task	Correlation
Nadia	English photo narrative	-.895
	French photo narrative	-.502
	English oral interview	-.664
	French oral interview	-.206
Noor	English photo narrative	-.428
	French photo narrative	-.439
	English oral interview	-.726
	French oral interview	-.397

Nadia's correlation results in the English photo narrative task show a very strong negative relationship between the two variables under investigation ($r = -.895$, $p < 0.05$). This implies that as one variable increases, the other decreases, and vice versa. This suggests a consistent dynamic relationship between enjoyment and anxiety, as can also be seen in the corresponding graph (see Figure 1 below). Since enjoyment and anxiety are often perceived as independent or even opposing emotions, it makes sense to assume that as enjoyment increases, anxiety would decrease. However, this negative relationship is stronger than what was anticipated by the author. This might be an indication that Nadia experienced a trade-off between these emotions during the task. For instance, as she felt more engaged and enthusiastic (i.e., experienced higher enjoyment), she might have felt less apprehension or nervousness (i.e., experienced lower anxiety). On the other hand, when faced with challenges or uncertainty, her enjoyment levels might have decreased while her anxiety levels increased.

Figure 1

The Idiodynamic Ratings During the English Photo Narrative Task for Nadia
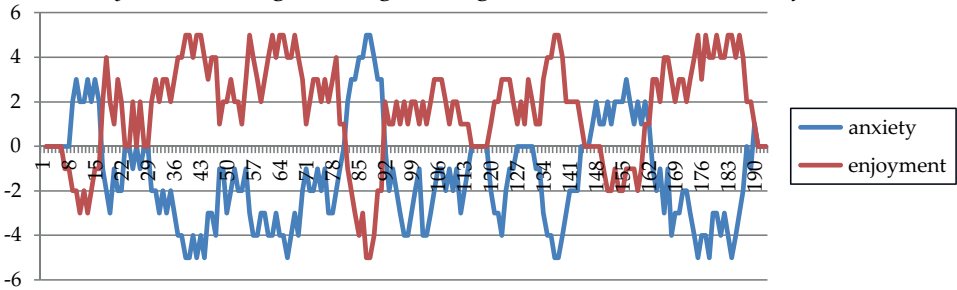
It can be noticed from the figure above that the participant was neutral for a few seconds at the beginning of the task, and then experienced a spike her anxiety level followed by another spike in enjoyment. During the interview, the participant provided an explanation for this trajectory:

> In the beginning, I could feel neither anxiety nor enjoyment. I was just describing the time and the place of the photo, then my anxiety grew as I did not know what exactly to say, I was just trying to say anything to fill time. Once I collected my thoughts, my enjoyment started to grow.

Although enjoyment predominated over anxiety throughout the later part of the task, there were two short segments during which anxiety overcame the participant again, once for 12 seconds (80–92 sec) and another for 15 seconds (148–163 sec). The participant noted the following in regard to these moments:

> My main source of anxiety was finding a new issue to talk about. For example, during seconds (80–92 s), I was experiencing moments of uncertainty about what I should say next. I have described the card game we played and some of the beautiful memories we shared. But then it seemed, I had to find more to talk about. The same thing happened near the end; in seconds (148–163).

For the photo narrative task in French, the results show a moderate negative correlation (r = -.502, $p < 0.05$). Anxiety was observed to be a dominant

emotion at the beginning of the task (see Figure 2 below). However, this did not persist throughout the photo narrative task, and only intertwined with enjoyment for a brief 5 seconds later in the task (134–139 sec). During the interview, the participant explained this observation:

> I was worried in the beginning, because I have learned the card game in the picture with my international friends. I have always played it in English, never in Arabic or French. So describing it in French was a bit weird in the beginning. However, as moments went by, I felt it was under control and I that could do it. But those feelings of anxiety came back for a very short period of time when I totally got confused about how to describe one of the English terms we used in French.

It should be noted that although the overall correlation for this task was moderate, in the final 32 seconds a very strong negative correlation ($r = -.929$, $p < 0.05$) was detected.

Figure 2

The Idiodynamic Ratings During the French Photo Narrative Task for Nadia
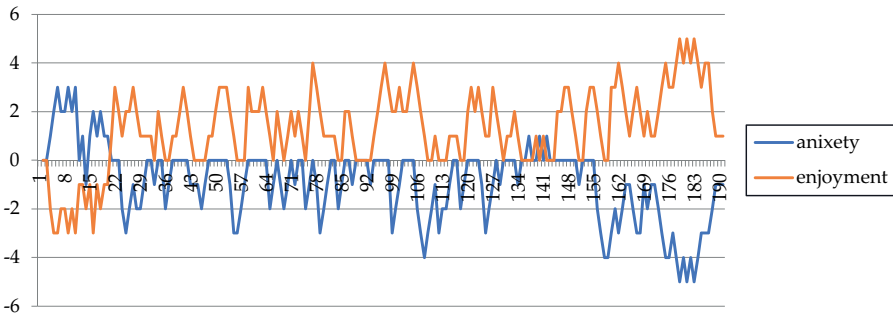
As can be seen in Table 1, the data from the same participant for the English oral narrative task yielded a moderate negative correlation ($r = -.664$, $p < 0.05$). Upon examining the corresponding graph (see Figure 3 below), it can be noticed that Nadia experienced periods of neutrality at different times during the task, specifically at seconds (1–11), (41–72), (81–101). During the interview, she revealed the following: "I had neither enjoyment

nor anxiety when I described the jeans and the black shirt I was wearing, the same while I was counting, and I also experienced similar feelings when I started describing the directions."

Spikes in anxiety and enjoyment can also be seen in Figure 3. From seconds (11–41), Nadia reported experiencing a high level of anxiety: "I don't feel comfortable at all talking about the educational system in my country. I only had little to say and I started thinking ahead and of how I would be able to fill in time and whether I should give more details." On another occasion, seconds (71–75), the participant experienced a very brief spike in enjoyment: "While I was describing the directions, the image of a cherry tree that has very dear memories to me flashed into my mind and brought me much joy."

Towards the end of the task, Nadia reported experiencing low levels of both enjoyment and anxiety. She elaborated on this during the interview: "I felt little enjoyment and little anxiety. It was not fun, but I was not worried as I have given a presentation before about Tunisia in my university, during which I have talked about the parliament. But it was not fun to talk about and it brought me no joy."

Figure 3

The Idiodynamic Ratings During the English Oral Interview Task for Participant 1
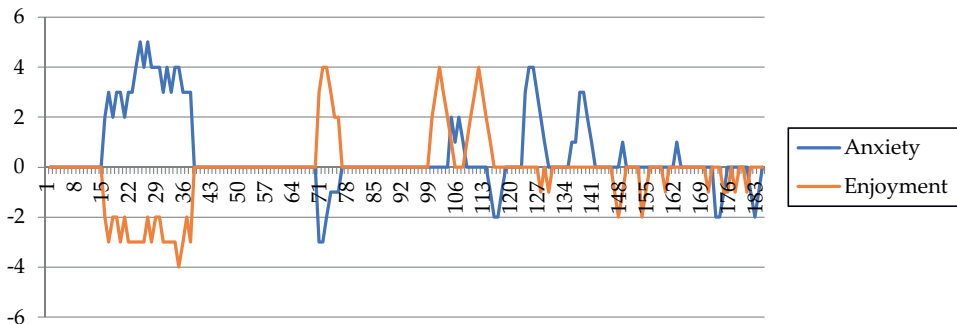
Contrary to the high and moderate negative correlation results observed in the previous tasks, (see Table 1), the French oral interview task yielded a relatively weaker negative correlation (r = -.206, p < 0.05).

Turning to the second participant, Noor, Table 1 shows a low negative correlation for the photo narrative task in English (r = -.428, p < 0.05),

with a similar correlation present for the French photo narrative task (*r* = -.439, *p* < 0.05). However, a strong negative correlation was found between enjoyment and anxiety (*r* = -.726, *p* < 0.05) in the English oral interview. The graph in Figure 4 below reveals that although anxiety was the dominant emotion, there were instances when enjoyment took over, specifically during seconds (4–25), (57–67), and (112–121). During the interview, Noor explained these fluctuations between enjoyment and anxiety:

> I enjoyed talking about my clothes. Especially that the T-shirt has my favorite slogan printed on it. Once, I started talking about the educational system in my country, I lost all interest and even started getting anxious. I think that the educational system in my country needs many reforms, so maybe this was going through my head as I have encountered many problems at school and university. This anxiety somehow stopped as I started counting and my enjoyment grew. I found counting by tens more like a mental exercise. When I first started giving directions, I was enjoying it; I remembered the long, beautiful street you have to take. Then my enjoyment dropped, and I started feeling anxious, as I'm very bad at giving directions, but once I could see the streets more clearly in my mind, I started enjoying it again. The last question has made me feel anxious again because I don't have enough English vocabulary related to this political topic.

Figure 4

The Idiodynamic Ratings During the English Oral Interview Task for Noor
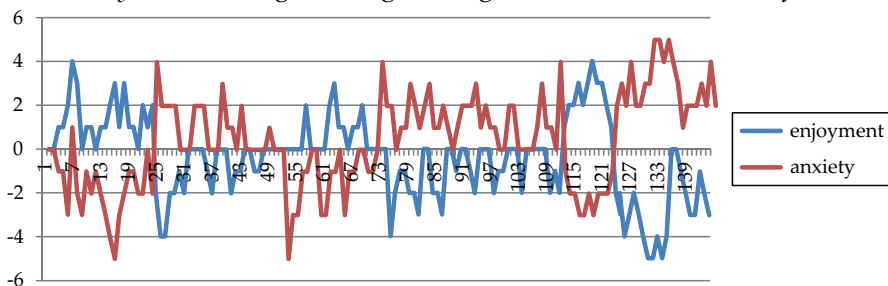
As can be seen in Table 1, the overall correlation for the French photo narrative task is relatively weak (*r* = -.397, *p* < 0.05). Noor remained emotionally

neutral during certain periods of the task; this included 11 seconds at the beginning of the task (1–11 sec) and about 22 seconds in the middle (52–74 sec). When interviewed, the participant provided an explanation for this neutrality, stating the following:

> When I first started describing my clothes, I was focused on translating the English words on my t-shirt from English to French. It was not fun, but it did not cause worry. For part of the discussion of the educational system in my country, I felt neutral, and the same applies to counting by tens; neither worry nor enjoyment.

The last seconds of the task (121–128 sec) were characterized by the dominance of anxiety and a decrease in enjoyment. As Noor explained, "At first, I said all I had about the parliament and its role in my country, and then I felt like I needed to add more, but had nothing to add, so I felt a bit anxious for a brief time, and there was no joy in that of course."

Figure 5

The Idiodynamic Ratings During the French Oral Interview Task for Noor
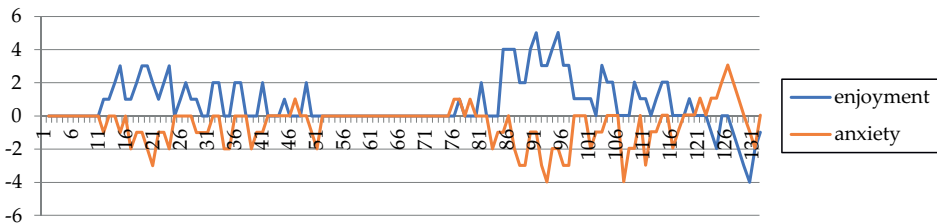
Discussion

The first research question of this study aimed at understanding the nature of the relationship between anxiety and enjoyment and whether it changes over time within participants and for each language. Our findings show varying levels of negative correlations between enjoyment and anxiety across different task types. This is consistent with the results of previous studies (see Boudreau et al., 2018; Dewaele & MacIntyre, 2014, 2016). How-

ever, it is important to avoid a simplistic interpretation of the relationship between the two emotions, as there were instances when they were intertwined or showed levels of correlation which differed from those observed for the task as a whole. Boudreau et al. (2018) found that these two emotions could also positively correlate, leaving us with the conclusion that the relationship between them is neither linear nor easily predictable. Rather, it is a complex, dynamic relationship that can change from one moment to the next.

The second research question sought to identify the causes of the fluctuations in enjoyment and anxiety for each participant in both languages. This qualitative part of the study enriched the quantitative data by incorporating the insights of the participants into the results and providing participants with an opportunity to articulate their emotional experiences during specific segments of the communication tasks. The results showed that the fluctuations were attributed to various internal and external factors, corroborating the findings of (Elahi Shirvan & Therian, 2017, 2020). The nature of the topic being discussed, the language being used, and the online setting of the task were all among the reasons that participants cited to explain the variability they experienced in regard to their emotions.

Conclusion

The present study expands on existing research investigating the per-second, complex, and dynamic relationship between enjoyment and anxiety in L2 communication. The present study is set apart through its investigation of a novel dimension, that is, the relationship between the two emotions across two foreign languages: French and English. The dynamic relationship between the two emotions was evident in both languages and for the different tasks, with the overall negative correlations between enjoyment and anxiety showing different patterns within each communication task for each participant. The idiodynamic method adopted by this study proved effective for exploring the complex relationship between the two emotions, as averaging data across individuals or taking the overall correlations of each communication task as a whole would have obscured the

unique relationship between these two emotions as well as the variation in their fluctuations. The emotional fluctuations in both languages were attributed to various internal and external factors that were reported by the participants during the interviews. By examining the fluctuations of emotions and considering participants' attributions, this research contributes to deepening our understanding of the complexity and variability of emotions in the context of foreign language communication.

This study has several limitations worth noting: first, while it aimed at exploring emotions experienced during the completion of two L2 communication tasks, the setting was more artificial than natural. The participants were not engaged in conversations, and were merely speaking in the form of monologues or answering a set of predetermined questions. This might have impacted the anxiety levels that they reported. Another limitation of the study is the online setting. Some participants may experience discomfort during online tasks due to concerns about internet connectivity and technical issues that might arise, or due to a dislike of the online atmosphere in general.

References

Arnold, J., & Brown, H. D. (1999). A map of the terrain. In J. Arnold (Ed), *Affect in language learning* (pp. 1–10). Cambridge University Press.

Boudreau, C., MacIntyre, P. D., & Dewaele, J.-M. (2018). Enjoyment and anxiety in second language communication: An idiodynamic approach. *Studies in Second Language Learning and Teaching*, *8*(1), 149–170. https://doi.org/10.14746/ssllt.2018.8.1.7

Brantmeier, C. (2005). Anxiety about L2 reading or L2 reading tasks? A study with advanced language learners. *Reading*, *5*(2), 67–85.

Csíkszentmihályi, M. (2008). *Flow: The psychology of optimal experience*. Harper Perennial.

De Bot, K., Lowie, W., & Verspoor, M. (2005). *Second language acquisition: An advanced resource book*. Routledge.

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

De Bot, K., & Makoni, S. (2005). *Old and new perspectives on language and aging*. Multilingual Matters. https://doi.org/10.21832/9781853598425-011

Dekeyser, R. (2012). Interactions between individual differences, treatments, and structures in SLA. *Language Learning*, *62*(2), 189–200. https://doi.org/10.1111/j.1467-9922.2012.00712.x

Dewaele, J.-M. & Alfawzan, M. (2018). Does the effect of enjoyment outweigh that of anxiety in foreign language performance? *Studies in Second Language Learning and Teaching*, *8*(1), 21–45. https://doi.org/10.14746/ssllt.2018.8.1.2

Dewaele, J.-M. & Li, C. (2018). Special issue 'Emotions in SLA' [Editorial]. *Studies in Second Language Learning and Teaching*, *8*(1), 15–19. https://doi.org/10.14746/ssllt.2018.8.1.1

Dewaele, J.-M. & Li, C. (2020). Emotions in second language acquisition: A critical review and research agenda. *Foreign Language World*, *196*(1), 34–49.

Dewaele, J.-M. & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, *4*(2), 237–274. https://doi.org/10.14746/ssllt.2014.4.2.5

Dewaele, J.-M. & MacIntyre, P. D. (2016). Foreign language enjoyment and foreign language classroom anxiety: The right and left feet of FL learning? In P. D. MacIntyre, T. Gregersen, & S. Mercer (Eds.), *Positive psychology in SLA* (pp. 215–236). Multilingual Matters. https://doi.org/10.21832/9781783095360-010

Dewaele, J.-M. MacIntyre, P. D., Boudreau, C., & Dewaele, L. (2016). Do girls have all the fun? Anxiety and enjoyment in the foreign language classroom. *Theory and practice of second language acquisition*, 2(1), 41–63. Retrieved from https://journals.us.edu.pl/index.php/TAPSLA/article/view/3941

Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford University Press.

Dörnyei, Z. (2010). The relationship between language aptitude and language learning motivation: Individual differences from a dynamic systems perspective. In E. Macaro (Ed.), *Continuum companion to second language acquisition* (pp. 247–267). Continuum.

Elahi Shirvan, M., & Taherian, T. (2017). Exploring the fluctuations of foreign language enjoyment in conversation: An idiodynamic perspective. *Journal of Intercultural Communication Research*, *47*(1), 21–37. https://doi.org/10.1080/17475759.2017.1400458

Elahi Shirvan, M., & Taherian, T. (2018). Longitudinal examination of university students' foreign language enjoyment and foreign language classroom anxiety in the course of general English: Latent growth curve modelling. *International Journal of Bilingual Education and Bilingualism, 24*(1), 31–49. https://doi.org/10.1080/13670050.2018.1441804

Elahi Shirvan, M., & Taherian, T. (2020). Rapid changes in foreign language learning anxiety caused by a multiplicity of topics: An idiodynamic approach. *Journal of Language and Education, 6*(1), 83–102. https://doi.org/10.17323/jle.2020.9684

Elahi Shirvan, M., Taherian, T., & Yazdanmehr, E. (2020). The dynamics of foreign language enjoyment: An ecological momentary assessment. *Frontiers in Psychology, 11*, 1391–1405. https://doi.org/10.3389/fpsyg.2020.01391

Ellis, N. C., & Larsen-Freeman, D. (2006). Language emergence: Implications for applied linguistics—Introduction to the special issue. *Applied Linguistics, 27*(4), 558–589. https://doi.org/10.1093/applin/aml028

Fredrickson, B. L. (1998). What good are positive emotions. *Review of General Psychology, 2*(3), 300–319. https://doi.org/10.1037/1089-2680.2.3.30

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist, 56*(3), 218–266. https://doi.org/10.1037/0003-066X.56.3.218

Fredrickson, B. L. (2003). The value of positive emotions: The emerging science of positive psychology is coming to understand why it's good to feel good. *American Scientist, 91*(4), 330–335. https://doi.org/10.1511/2003.26.330

Fredrickson, B. L. (2004). The broaden–and–build theory of positive emotions. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 359*(1449), 1367–1377. https://doi.org/10.1098/rstb.2004.1512

Fredrickson, B. L. (2006). The broaden–and–build theory of positive emotions. In M. Csíkszentmihályi & I. Selega (Eds.), *A life worth living: Contributions to positive psychology* (pp. 85–103). Oxford University Press. https://doi.org/10.1093/oso/9780195176797.003.0006

Fredrickson, B. L. (2013). Positive emotions broaden and build. In P. Devine & A. Plant (Eds.), *Advances in experimental social psychology* (pp. 1–53). Academic Press. https://doi.org/10.1016/B978-0-12-407236-7.00001-2

Ganschow, L., & Sparks, R. (2001). Learning difficulties and foreign language learning: A review of research and instruction. *Language Teaching, 34*(2), 79–98. https://doi.org/10.1017/S0261444800015895

Gregersen, T. S., & MacIntyre, P. (2017). Idiodynamics: An innovative method to build emotional intelligence through systematic self-assessment/reflection/critique. In T. S. Gregersen & P. D. MacIntyre (Eds.), *Innovative practices in language teacher education* (pp. 33–53). Springer. https://doi.org/10.1007/978-3-319-51789-6_3

Gregersen, T., Macintyre, P., & Meza, M. (2014). The motion of emotion: Idiodynamic case studies of learners' foreign language anxiety. *The Modern Language Journal, 98*(2), 574–588. https://doi.org/10.1111/modl.12084

Gregersen, T., MacIntyre, P. D., & Olson, T. (2017). Do you see what I feel? An idiodynamic assessment of expert and peer's reading of nonverbal language anxiety cues. In C, Gkonou, M, Daubney, & Dewaele, J.-M. (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 110–134). Multilingual Matters. https://doi.org/10.21832/9781783097722-008

Horwitz, K. (2001). Language anxiety and achievement. *Annual Review of Applied Linguistics, 21*, 112–126. https://doi.org/10.1017/S0267190501000071

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics, 18*(2), 141–165. https://doi.org/10.1093/applin/18.2.141

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied Linguistics, 27*(4), 590–616. https://doi.org/10.1093/applin/aml029

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press. https://doi.org/10.1093/oso/9780195069945.001.0001

Lemke, J. L. (1998). Analysing verbal data: Principles, methods, and problems. In K. Tobin & B. Fraser (Eds.), *International handbook of science education* (pp. 1175–1189). Springer.

Li, C. (2018). *A positive psychology perspective on Chinese students' emotional intelligence, classroom emotions, and EFL learning achievement* [Unpublished doctoral dissertation]. Xiamen University.

Lowie, W. (2013). Dynamic systems theory approaches to second language acquisition. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 1–9). Blackwell. https://doi.org/10.1002/9781405198431.wbeal0346

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

MacIntyre, P. (2016). So far so good: An overview of positive psychology and its contributions to SLA. In D. Gabryś-Barker & D. Gałajda (Eds.), *Positive psychology perspectives on foreign language learning and teaching* (pp. 3–20). Springer International Publishing AG. https://doi.org/10.1007/978-3-319-32954-3_1

MacIntyre, P. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou., M. Daubney, & Dewaele, J.-M. (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 11–30). Multilingual Matters. https://doi.org/10.21832/9781783097722-003

MacIntyre, P., & Gregersen, T. (2012). Emotions that facilitate language learning: The positive-broadening power of the imagination. *Studies in Second Language Learning and Teaching*, *2*(2), 193–213. https://doi.org/10.14746/ssllt.2012.2.2.4

MacIntyre, P., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied Linguistics*, *32*(2), 149–171. https://doi.org/10.1093/applin/amq037

MacIntyre, P., & Serroul, A. (2014). Motivation on a per-second timescale: Examining approach-avoidance motivation during L2 task performance. In Z. Dörnyei, P. MacIntyre, & A. Henry (Eds.) *Motivational dy-*

namics in language learning (pp. 109–138). Multilingual Matters. https://doi.org/10.21832/9781783092574-013

Pavlenko, A. (2006). Bilingual Selves. In A. Pavlenko (Ed.), *Bilingual minds: Emotional experience, expression, and representation* (pp. 1–33). Multilingual Matters. https://doi.org/10.21832/9781853598746

Pavlenko, A. (2007). *Emotions and multilingualism*. Cambridge University Press. https://doi.org/10.1017/CBO9780511584305

Piechurska-Kuciel, E. (2017). L2 or L3? Foreign language enjoyment and proficiency. In G. D. Gabryś-Barker, D. Gałajda, A. Wojtaszek & P. Zakrajewski (Eds*.), Multiculturalism, multilingualism and the self* (pp. 97–111). Springer. https://doi.org/10.1007/978-3-319-56892-8_7

Piniel, K., & Albert, A. (2018). Advanced learners' foreign language-related emotions across the four skills. *Studies in Second Language Learning and Teaching*, *8*(1), 127–147. https://doi.org/10.14746/ssllt.2018.8.1.6

Prior, M. (2019). Elephants in the room: An "affective turn," or just feeling our way. *The Modern Language Journal*, *103*(2), 516–527. https://doi.org/10.1111/modl.12573

Reeve, J. (2014). *Understanding motivation and emotion*. John Wiley & Sons.

Seligman, M., & Csíkszentmihályi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55*(1), 5–14. https://psycnet.apa.org/doi/10.1037/0003-066X.55.1.5

Shirvan, M. E., & Taherian, T. (2018). Longitudinal examination of university students' foreign language enjoyment and foreign language classroom anxiety in the course of general English: latent growth curve modeling. *International Journal of Bilingual Education and Bilingualism*, *24*(1), 31–49. https://doi.org/10.1080/13670050.2018.1441804

Swain, M. (2013). The inseparability of cognition and emotion in second language learning. *Language Teaching*, *46*(2), 195–207. https://doi.org/10.1017/S0261444811000486

Waninge, F., de Bot, K., & Dörnyei, Z. (2014). Motivational dynamics in language learning: Change, stability, and context. *The Modern Language Journal*, *98*(3), 704–723. https://doi.org/10.1111/modl.12118

Wind, A. M. (2014). Second language writing development from a Dynamic Systems Theory perspective. In B. P. Olmos-Lopez, J. Huang, & J. Almeida (Eds.), *Proceedings of the 8th Lancaster University Conference in Linguistics & Language Teaching* (pp. 90–123). University of Lancaster. https://www.lancaster.ac.uk/fass/events/laelpgconference/papers/v08/Attila-M-Wind.pdf

Wind, A. M. (2020). A dynamic usage-based approach to the longitudinal development of verb argument constructions in second language writing: Two case studies. In C. Kálmán (Ed.), *DEAL 2020: A Snapshot of Diversity in Applied Linguistics* (pp. 39–70). *Eötvös* University Press. https://www.eltereader.hu/media/2020/12/Kálmán-Csaba_WEB.pdf#page=58

Wind, A. M. (2021). Nonlinearity and inter- and intra-individual variability in the extent of engagement in self-reflection and its role in second language writing: A multiple-case study. *System*, *103*, Article 102672. https://doi.org/10.1016/j.system.2021.102672

Wind, A. M., & Harding, L. W. (2020). Attractor states in the development of linguistic complexity in second language writing and the role of self-regulation: A longitudinal case study. In W. Lowie, M. Michel, A. Rousse-Malpat, M. Keijzer, & R. Steinkrauss (Eds.), *Usage based dynamics in second language development* (pp. 130–154). Multilingual Matters. https://doi.org/10.21832/9781788925259-009

4

The Interrelationship Between Foreign Language Learning Motivation, Extroversion-Introversion, and Willingness to Communicate: A Quantitative Investigation

Wijdene Ayed
wijdene@student.elte.hu

Abstract

In the field of English language education, willingness to communicate (WTC), personality traits, and motivation are salient features that influence learners' foreign language learning process. As these variables are hypothesized to be complex constructs, this study aims to quantitatively explore their interrelationship. Specifically, it investigates WTC as the first construct, second/foreign language learning motivation as the second, and extroversion-introversion as the third. The study points to the importance of investigating WTC, motivation, and extroversion-introversion in concert as integral parts of the foreign language learning process. In terms of research methods, the study employs an online questionnaire administered through Google Forms which was developed and piloted based on previous studies and adapted for this particular research. Data were collected from a sample of N = 155 participants aged between 17 and 40 years from different nationalities and were analyzed with the Statistical Package for Social Sciences (SPSS) software 20.0 version. The results indicate that extroverted learners exhibit higher levels of motivation to learn English as a foreign language and a greater WTC in the target language compared to introverted students, who are less willing to communicate in English. These findings hold significant implications for learners of foreign languages, especially those who are concerned about their personality (i.e., introversion) being a demotivator or barrier to their English learning process and to their WTC.

Keywords: English language learning motivation, willingness to communicate, extroversion-introversion, quantitative research, correlation

The Interrelationship Between Foreign Language Learning Motivation, Extroversion-Introversion, and Willingness to Communicate: A Quantitative Investigation

Individual differences among learners shape each person's learning experience, a notion that aligns with a broader understanding that humans are inherently different from one another. As Baghaei and Dourakhshan (2012) stated, "individual differences (IDs) among people play an important role beyond general theories which are advanced by social scientists" (p. 55). In the context of applied linguistics, Dörnyei (2003) defined individual differences as "characteristics or traits in respect of which individuals may be shown to differ from each other" (p. 1). These individual differences encompass numerous variables that can influence the process of second language learning, including affective, psychological, cognitive, and contextual factors.

Considering the expansive research on individual differences and the rapid pace of change in the world, the field of education has been witnessing enormous transformations. Contemporary approaches to foreign language teaching are increasingly focused on simulating real-life situations, tasks, and activities and emphasize meaningful and authentic communication "as an ultimate goal of language learning" (Gholami, 2015, p. 1). This trend in language education has shifted toward the study, development, and implementation of a communicative approach. MacIntyre and Charos (1996) supported the idea that language pedagogy should "reflect the belief that one must use the language to develop proficiency, that is, one must talk to learn" (p. 3). Similarly, Alqahtani (2015) asserted that the "main purpose of learning a foreign language is to utilize it for meaningful and effective communication" (p. 1195). This trend aligns with the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2020) guidelines on foreign languages' communicative competence and its importance in the European context and beyond (Simić, 2014, p. 1). Given this transition from teacher-centered to learner-centered classes and the challenges therein and taking into consideration the role of communicative competence as an indispensable component of the foreign language learning process, this research aims to study aspects of

WTC and their relations to other individual difference variables, namely motivation and personality traits.

Studies carried out on second language (L2) motivation since the 1960s have demonstrated that motivation is a key factor and strongly influences the success of second and foreign-individual language learning (Dörnyei, 1998, 2005). Dörnyei and Guilloteaux (2008) asserted that if motivation is low or if there is no motivation, even competent learners with access to effective curricula and teachers will struggle to achieve their long-term objectives. Kanat-Mutluoglu (2016) echoed this sentiment, stating that "of all the individual differences related to L2 learning, motivation is regarded as one of the most important elements" (p. 27). Given its significance, L2 motivation is investigated in this study as a possible antecedent variable to WTC in the L2 learning process.

In the context of L2 learning processes, personality traits can play a significant role in understanding learner behaviors and strategies. As highlighted by Özbay et al. (2017), personality type is one of the important factors influencing foreign language learning. The authors list extroversion and introversion as two dimensions of personality that influence language learning. Extroverts are usually more comfortable communicating in the target language, whereas introverts may excel in the academic study of the language (Altunel, 2015).

Considering the critical roles of extroversion-introversion, motivation, and WTC in the process of learning new languages, this study aims to explore the interrelations among these three constructs. Previous studies offer there are insufficient and inconsistent results regarding this interrelationship and their collective impact on the L2 learning process. Such a study can prove beneficial in providing a better understanding of the complex interactions between these variables and their influence on the language learning process (Kanat-Mutluoglu, 2016),

Accordingly, this study seeks to connect three lines of research on individual differences, focusing on the interplay between L2 motivation, WTC, and extroversion-introversion traits. The research's main objective concentrates on the study of English as a second (L2) and foreign language (EFL), and as such these two terms are used interchangeably. It should be noted that the scope of the study does not extend to first language learning (L1).

Review of the Literature

Studies on Second/Foreign Language Learning Motivation

This study adopts Dörnyei's (1998) definition of motivation as "a process whereby a certain amount of instigation force arises, initiates action, and persists as long as no other force comes into play to weaken it and thereby terminate the action, or until the planned outcome has been reached" (p. 118). Numerous studies have emphasized the complexity of language learning motivation as an intricate and multifaceted construct (Anjomshoa & Firooz, 2015; Dörnyei, 1998; Schumann, 1986). In terms of measuring L2 motivation, there is no simple scale that can be used to assess the motivational construct; therefore, its measurement entails the use of multiple approaches and scales. This diversity of approaches for researching motivation in second and foreign language learning does not hinder but rather reinforces the understanding of this complex construct, which can be viewed from different angles (Dörnyei, 1998). The following paragraphs will review studies which shed light on the various perspectives researched in regard to foreign language learning motivation.

The social context model, proposed by Clément (1980), and the social cognitive theory, as elucidated by MacIntyre et al. (2010), share a common foundation in their utilization of scientific and cognitive approaches to observe, measure, test, and comprehend the social dimensions of individuals' experiences and their interactions. In this context, Giles and Byrne (1982) underscored the significance of ethnolinguistic vitality and its impact on foreign language acquisition, while McCroskey and Baer (1985) delved into the construct of willingness to communicate (WTC) and how it is influenced by various factors, including motivation. Additionally, Clément (1986) introduced the notion of self-confidence in language learning, which comprises two key components: anxiety and perceived competence, as further detailed by Darling and Chanyoo (2018).

Schumann's (1986) acculturation model predicts that the degree to which learners acculturate to the target language group correlates with their acquisition of the target language. In the context of motivation, Markus and Nurius (1986) introduced the concept of possible selves, describing

them as "ideal selves that we would very much like to become, selves that we could become, and selves we are afraid of becoming" (p. 954). Building upon this concept, Higgins (1987) provided a more intricate framework for this three-part self-model:

> (a) the actual self-representing the beliefs that you or others hold about who you really are at a given point in time; (b) the ideal self-representing the beliefs about who you would like to be and how you would wish others to view you; (c) the ought to self-representing the beliefs that you attribute to others about who you should be. (p. 320)

Gardner (1985, 2001) introduced and developed the socio-educational model which is illustrative of the social aspects of second language acquisition which sheds light on two important motivational characteristics affecting language learning. These characteristics include integrative motivation, which is referred to as "a sincere and personal interest in the people and culture represented by the other language group" and instrumental motivation, which is "the practical value and advantages of learning a new language" (Gardner & MacIntyre, 1991, p. 58). Integrative motivation is defined as the learner's inclination to communicate, interact, and identify with the second or foreign language community and is made up of three components: integrativeness, attitudes toward the learning situation, and motivation or interest in the foreign language (Peng, 2007).

The L2 motivational self-system, proposed by Dörnyei (2005), is divided into three constituents: (1) the Ideal L2 Self, which is drawn from Gardner's concept of integrativeness and explained by Dörnyei (2005) as "a powerful motivator to learn the L2 because of the desire to reduce the discrepancy between our actual and ideal selves" (p. 105), (2) the Ought-to L2 Self, which is based on Higgins (1987) conceptualization and refers to "to the attributes that one believes one ought to possess" (p. 105), and lastly (3) the L2 Learning Experience, based on Noels' (2001) intrinsic category "which concerns situation-specific motives related to the immediate learning environment and experience" (p. 106).

Self-determination theory (Deci & Ryan, 2004; Noels, 2001; Noel & Clément, 1996; Noels et al., 1999, 2001; Noels, et al., 2000) is centered on the role of individual choice and the challenges of personal growth. These free choices and decisions offer insight into individuals and their adaptability to their surrounding environment (MacIntyre et al., 2011). MacIntyre et al. claimed that "the most interesting actions a person undertakes are those that are freely chosen; those actions reveal details about the person and how they choose to fit into their environment" (p. 85). A language learner's experience is thus enriched when competence, relatedness, and autonomy are all combined, reflecting the self-determined and intrinsically motivated behavior of the learner. Self-determination theory will be the theoretical and empirical basis for motivation in the present study.

Studies on Willingness to Communicate in a Second/Foreign Language
The concept of WTC finds its roots in the construct of unwillingness to communicate introduced by Burgoon (1976). Burgoon described this unwillingness "as a predisposition to chronically avoid oral communication" (as cited in MacIntyre et al., 2001, p. 370). In their 1977 study, Mortensen and his team explored the inherent traits of communication. Their research yielded evidence suggesting that the presumed idea that individuals' communication tendencies remain consistent across various settings may not be accurate, as indicated by Zohrabi and Yousefi in 2016.

Following the revitalization of WTC in the L2 context and renewed interest in studying the construct, MacIntyre et al. (1998) described WTC as "readiness to enter into discourse at a particular time with a specific person or persons, using a L2" (p. 547). This definition will be applied in the present research. In the second and foreign language classroom, Dewaele and Dewaele (2018) claimed that WTC is the second most important construct contributing to successful learning besides motivation based on the responses of L2 teachers.

Alqahtani (2015) emphasizes the significance of willingness as "an important factor that is instrumental in carrying out an activity successfully" (p. 1165). This perspective aligns with the notion that interaction serves as the cornerstone of language learning. In classroom settings, where teach-

ers and learners engage in conversations, they create a microcosm of real and authentic social environments. This, in turn, enhances the language learning process, as noted by Şener (2014). Moreover, and MacIntyre and Charos explain, "the primary reason for language learning often is to use it to communicate" (p. 4).

The communicative approach widely used in language teaching stresses the primacy of interaction and communication and their crucial role in individual development (Gholami, 2015). As such, one goal of teaching a second or a foreign language should be to enhance learners' enthusiasm and eagerness to seek out communication opportunities and to take advantage of them to interact and empower their communicative competence (Hişmanoğlua & Özüdoğrub, 2017). Consequently, WTC is an essential part of foreign language learning education.

To identify the different variables affecting WTC as a volitional process, MacIntyre et al. (1998) developed a model of WTC integrating various variables. This conceptual pyramid model is considered to be a heuristic and comprehensive model, including individual differences aspects and background aspects such as linguistic, psychological, personal, motivational, attitudinal, educational, social, contextual, intercultural, communicative, and conversational variables.

Dörnyei and Skehan (2003) acknowledged that research on Willingness to Communicate (WTC) is closely connected to the broader field of motivation research, as they describe WTC as being situated "at the intersection of motivation and communicative competence research" (Dörnyei, 2003, p. 13). This perspective aligns with the idea that the WTC approach to the topic of motivation, as expressed by MacIntyre et al. (2001), explores motivation from various angles. Therefore, all cited works emphasize the intrinsic link between WTC and motivation, with Dörnyei and Skehan highlighting the bridge between WTC, motivation, and communicative competence. These perspectives frame WTC as an additional composite of individual differences variables relative to communication tendencies (Gharibi & Seyyedrezaeia, 2016).

Dörnyei (2005) emphasized the complexity of the WTC construct, which is built upon organically organized and integrated factors. This

view is furthered by Peng and Woodrow (2010), who described WTC as being "influenced by the joint effect of variables both internal and external to individual learners" (p. 835). To further accentuate this complexity, Mystkowska-Wiertelak and Pawlak (2014) asserted that WTC is "believed to stem from a combination of proximal and distal variables comprising psychological, linguistic, educational and communicative dimensions of language" (p. 245), which is similarly supported by Dewaele and Dewaele (2018).

It is important to note that these factors have typically been studied independently. Clément and Gardner (2001) claimed that the study of WTC as an independent construct implicitly involves motivation. The study conducted by Dörnyei and Skehan in 2003 provided evidence that Willingness to Communicate (WTC) in a second or foreign language is intricately connected to motivation. They argued that motivation, while often seen as directly linked to achievement in language learning, also plays a crucial role in facilitating effective communication. This perspective is further supported by Mystkowska-Wiertelak and Pawlak (2014), who emphasized that although motivation is commonly associated with language learning success, its significance in enabling effective communication cannot be overlooked. As highlighted above, motivation plays an important role in determining the rate of willingness to communicate within an individual (Altiner, 2018).

Prior research conducted by MacIntyre and Charos (1996) paved the way for understanding the relationship between the two constructs. MacIntyre et al. (2001) claimed that WTC is a variable that "approach(es) the topic of motivation from different directions" (p. 462). Learners who are more interested in using a second or foreign language in a classroom setting, for instance, showed higher motivation toward learning that language (Hashimoto, 2002; MacIntyre et al., 2001). This study investigated how these two variables are interrelated among international learners of English as a second/foreign language. For the present study I made the decision to measure WTC using two scales. WTC inside the classroom items were combined from the studies of MacIntyre et al. (2001) and Peng and Wood-

row (2010), whereas items for WTC outside the classroom were taken from McCorskey's (1992) study.

Studies on Extroversion-Introversion as a Personality Trait Scale
Individuals vary in personality, and these differences manifest in distinct language learning behaviors and processes. Some studies have focused on the role of personality traits and their potential influence on WTC (Yashima, 2002). The result of one such study (Yashima, 2002) showed that motivation influenced learners' self-confidence (one of the personality trait-related scales measured) in using the target foreign language, which subsequently led to greater WTC in this language.

Lalonde and Gardner (1984) posited that personality traits have an indirect effect on L2 achievement which can be observed through attitudes and motivation. According to Darling and Chanyoo (2018), all of Goldberg's (1992) Big Five personality traits are conducive in motivating both L2 learning and communication. These global traits are assessed based on five scales: extraversion (or surgency), agreeableness, conscientiousness, emotional stability, and intellect (Goldberg, 1992). The Introversion-extroversion trait has been explored in several investigations (MacIntyre & Charos, 1996).

The terms extrovert and introvert were first introduced in the 1980s and rapidly became recognized and used within personality research. This scale has been explored from both social and biological aspects. The focus of the present study will be on the social perspective. Wakamoto (2009) explained that extroverts tend to draw energy from interacting with the outside world and other people and are usually active and outgoing. Introverts, on the other hand, draw energy by concentrating on their inner world and thoughts, and are usually more reflective. While extroverts are more sociable and willing to initiate conversations, introverts tend to process their thoughts before speaking and avoid making linguistic mistakes. In foreign language learning, personality variables tend to be omitted from research, and there are a limited number of studies that have identified the effects of extroversion or introversion on foreign language learning.

Regarding academic achievement, it is believed that introversion is more desirable (MacIntyre & Charos, 1996).

However, the relation between language learning, communicative contact, and personality traits is generally conceptualized as an indirect one (MacIntyre et al., 1998). While one group of people may enjoy communicating with others without reservations, another group would speak out only if necessary. The former group may be more willing to take the opportunity to learn and communicate in the target language (Alqahtani, 2015), as it has been found that "that having more opportunities for interaction in L2 affects the frequency of L2 use directly and also indirectly through perceived competence and WTC" (Hashimoto, 2002, p. 40). MacIntyre et al. (1998) supported this claim, highlighting the fact that context and personality traits are WTC-influencing variables.

Extroversion-introversion as a trait not received as much attention in foreign language studies as other individual difference variables (Altunel, 2015; Dörnyei, 2006; Wakamoto, 2009). The aim of the present study is to use the extroversion-introversion trait drawn from Goldberg's (1992) Big Five personality traits (there labelled as "surgency") and to examine its influence on WTC and the motivation to learn English. The choice of this scale lies in the fact that it includes items expressing individuals' extroversion, energy, talkativeness, activeness, and assertiveness, aspects which are likely to be associated with learning new languages. The review of the literature demonstrated the importance of WTC, motivation, and extroversion-introversion in foreign language learning. Hence, the study aims to investigate their correlation.

Research Questions
Based on the review of the relevant literature outlined above, I set out to answer the following research questions in my study:
1. What characterizes the motivation, WTC, and extroversion-introversion of the English learners participating in the present study?
2. What is the relationship between language learners' extroversion-introversion, WTC, and motivation to learn English?

Methods

To fulfill the aim of the study and answer the above research questions, a quantitative research methodology has been proposed to investigate the interrelations among the three constructs under study. A questionnaire was developed based on previous studies and its design is detailed in the following section. A brief description of the measures is provided below.

Participants and Setting

The study presented here involved a total of 155 English language learners who participated by completing the main instrument. The data collected revealed that participants were from 26 different countries. Tunisian nationals constituted 52.9% of the total with 82 responses to the questionnaire. The remaining participants were from Egypt (11), Jordan (10), Syria (8), Hungary (5), Vietnam (5), France (4), Russia (3), Libya (3), Algeria (2), Morocco (2), Pakistan (2), and Ukraine (2). The following countries were represented by one participant each (i.e., .06%): Albania, Brazil, India, Italy, Kazakhstan, Kurdistan, Moldova, Mongolia, Montenegro, Nigeria, Norway, Rwanda, and Tanzania. Out of the 155 respondents, 84 were female (54.2%), 70 were male (45.2%), while one did not reveal their gender (0.6%). The participants' ages varied between 17 and 40 years old ($M = 24.76$, $SD = 3.820$), with the largest age group being 23-year-olds, accounting for 27.7% of the participants. One participant did not reveal their age.

The Instrument

The WTC construct is divided into two scales and comprises 32 items (See appendix). These two scales are designed to measure WTC inside the classroom (14 items) and outside the classroom (18 items), with items selected from Peng and Woodrow (2010), MacIntyre et al. (2001), and McCroskey (1992). MacIntyre et al. (2001) focused on four skill areas inside the classroom: reading, writing, speaking, and listening.

Next, foreign language learning motivation items (33 in total) were compiled from two studies (i.e., Csizér & Kormos, 2008; Noels et al., 2000).

These items are grouped into 8 scales: amotivation, external regulation, introjected regulation, integrated regulation, intrinsic motivation including accomplishment, knowledge and stimulation, and the intended effort (See appendix).

Lastly, a limited set of 7 items related to the personality trait scale were adopted from a study by Goldberg (1992). This scale is labeled in Goldberg (1992) as introversion-extroversion (surgency), and it includes positive and negative items related to being energetic, active, talkative, bold, and assertive (See Appendix). Examples of these items include 'I am an extroverted person' and 'I am assertive (communicating with confidence)'.

It is noteworthy that the items from all these studies were revisited and modified to best suit the purpose of this research. A total of 79 items were used in the main questionnaire (See appendix). It also included two separate sections of biographical questions gathering information on the view of the status and importance of the English language, length of its learning process and use, English proficiency, nationality, age, gender, and occupation (7 items).

Procedure for Data Collection and Measures

The questionnaires, containing measures of the aforementioned constructs were developed and administered in Google form format for both the pilot and main studies. These questionnaires were both distributed online, and participants were required to submit their responses based on a 1 to 5-point Likert scale ranging from 1 for "strongly disagree" to 5 for "strongly agree". Version 20.0 of the Statistical Package for Social Sciences (SPSS) software was used to analyze the data collected. To answer the research questions, reliability and descriptive statistics are presented, followed by multiple calculations of correlations between the different scales of the study.

Results and Discussion

Reliability of the Scales and Descriptive Statistics

Using SPSS, reliability analysis was performed to measure and ensure the internal consistency and reliability of the instrument (Dörnyei, 2007). The

reliability results obtained using Cronbach's Alpha confirmed that all of the 11 scales measuring the three major constructs were reliable. The motivation construct was reliable with Cronbach's Alpha ranging from α =.777 to α = .890 for all the subconstructs. For example, the intrinsic motivation scale related to accomplishment was found to be internally consistent α = .777 (M = 3.43, SD .992). The amotivation scale was also found to be reliable, with α = .890 (M = 1.45, SD = .730). The WTC inside the classroom construct was also shown to be reliable α = .958 (M = 3.89, SD = .939), as well as the outside the classroom scale α = .976 (M = 3.91, SD = .981). The extroversion-introversion scale was reliable with α = .779 (M = 3.57, SD = .655). Overall, the three scales under investigation (i.e., motivation, WTC, and extroversion-introversion) were proven to be reliable and demonstrated the internal consistency of the items.

Table 1

Descriptive Data of Clusters in the Main Study

	Cronbach's α	Mean	SD
Motivation			
Amotivation	0.89	1.45	0.73
External regulation	0.829	3.62	1.01
Introjected regulation	0.864	2.97	1.17
Integrated regulation	0.879	3.7	0.993
Intrinsic motivation-accomplishment	0.777	3.43	0.992
Intrinsic motivation- knowledge	0.79	3.82	0.897
Intrinsic motivation- Stimulation	0.849	3.83	0.884
Intended effort	0.868	3.52	0.86
WTC			
Inside the classroom	0.958	3.89	0.939
Outside the classroom	0.976	3.91	0.981
Extroversion-introversion	0.779	3.57	0.655

Note. Cronbach's alpha significance level α > .70

Correlational Analyses

It is hardly surprising that the strongest correlation was found between the two WTF scales (inside and outside the classroom) (r= .804, $p < 0.01$). The participants' willingness to communicate in the classroom setting significantly relates to their willingness to communicate outside of it.

Another relatively strong correlation can be seen between the integrated regulation scale and the inside-the-classroom WTC (r = .349; $p < .05$), as well as the outside-the-classroom WTC scales (r = 302; $p < .001$). Integrative motivation refers to the learners' intrinsic orientation or desire to communicate in a foreign language with a community of users. Learners who are integratively motivated about learning the language are more willing to take risks and use the language inside and/or outside the classroom settings. When they are integratively motivated enough, learners can succeed in learning languages and trying to integrate themselves into different communities using their acquired language. These results confirm Peng's research (2007) regarding immersion programs, which revealed a significant correlation between WTC and integrative motivation to learn a second language.

A similar notable correlation was found between the extroversion-introversion and WTC scales (r = .322; $p < .001$). The higher the WTC is, the more frequent communication is in the target language (Gharibia & Seyyedrezaei, 2016). Being an extrovert, who is talkative and assertive, may be highly beneficial for acquiring English through communicating with teachers, peers, and classmates. A study conducted by Baghaei and Dourakhshan (2012) claimed that "the choice to speak or to remain silent seems to be a factor in the success of a second language learner" (p. 55) and this choice or decision can be influenced by an individual's personality characteristics. Furthermore, it's highly encouraging to note that displaying high energy, a spirit of adventure, and a readiness to embrace risks correlates with one's level of WTC as in social warmth and friendliness in both formal interactions with unfamiliar individuals and casual encounters with friends and family (r = .336; $p < .001$).

Table 2

Correlation of Extroversion-Introversion, Motivation, and WTC Scales

Scales	Extroversion-Introversion	Motivation External regulation	Motivation Intro-jected regulation	Motivation integrated regulation	Motivation Inte-grated regulation	WTC-Inside the classroom
Extroversion-introversion	-					
Motivation external regulation	0.026	-				
Motivation introjected regulation	-.182*	.387**	-			
Motivation integrated regulation	0.089	.192*	.190*	-		
WTC- inside the classroom	.322**	0.153	0	.349*	-	
WTC- Outside the classroom	.336*	0.066	-0.062	.302**	.804**	

Scales	Extroversion Introversion	Motivation Accomplishment	Motivation Knowledge	Motivation Stimulation	WTC-Inside the classroom
Extroversion introversion	-				
Motivation- accomplishment	0.008	-			
Motivation – knowledge	0.051	.612**	-		
Motivation – stimulation	-0.03	.494**	.711**	-	
WTC- inside the classroom	.322**	.215**	.397**	.448**	-
WTC- outside the classroom	.336**	0.142	.363**	.353**	-

Scales	Extroversion Introversion	Amotivation	Intended effort	WTC-Inside the classroom	WTC- Outside the classroom
Extroversion introversion	-				
Amotivation	-0.035	-			
Intended effort	0.095	-0.117	-		
WTC- inside the classroom	.322**	-.244**	.364**	-	
WTC- outside the classroom	.336**	-.192*	.239**	-	-

Note. Significance *p<.05, two tailed. **p<.001, two-tailed.

The correlation between external regulation and introjected regulation within the motivational construct was statistically significant (r = .387; p < .001). External regulation describes the motivation to learn a language for the sake of rewards or fear of punishment. Introjected regulation is also a form of motivation in which a learner feels pressure to learn the language due to fear of failure. This explains the significant correlation between both external and introjected motivations. For example, learning the English language is indeed highly beneficial: it is a portal to new opportunities and a better future, and for many of those who are not able to learn it, the inability to interact with other English speakers can be embarrassing and serve as a barrier to improving their circumstances.

It was also found that the external regulation and introjected regulation subscales have a weaker correlation with integrated regulation. The benefit of learning the English language provides are only modestly influenced by its role as a fundamental part of the lives of its users (r = .192; p = .017). Furthermore, the negative feelings aroused by not being able to learn and master the language do not necessarily imply that it should be regarded as integral and necessary in all aspects of life (r = .190; p = .018). In other words, just because someone may feel frustrated or upset about their language learning difficulties, it doesn't necessarily mean that they should perceive language mastery as a fundamental requirement in all aspects of their life. This statement highlights the importance of considering context and individual circumstances when evaluating the significance of language proficiency.

Some studies have demonstrated that the correlation between motivation and foreign language WTC is only indirect (Ma et al., 2019). In the present study, a weaker correlation was found between external regulation and WTC in (r = .066; p = .412) and outside the classroom scales (r = .153; p = .058). Knowing the English language and seeking better opportunities do not necessarily influence the WTC in English with colleagues at work, as the correlation results showed they are not correlated, or in English language classes.

The was no significant correlation obtained among the introjected regulation of motivation and WTC inside the classroom scales (r = .000;

p = .996). Introjected regulation describes a controlling form of motivation in which behavior is regulated by internal forms of pressure, such as the participant's feelings of shame, embarrassment, unease, or guilt. The data suggest that these feelings did not influence the participants' willingness to use English during class.

The extroversion-introversion scale showed one statistically significant result: a negative correlation with introjected regulation (r = -.182; p = .024). This suggests that the more extroverted a learner is, the higher their introjected regulation would be. As such, learners may view reaching a higher level of proficiency in the language as a reward for themselves. This can be related to ego enhancement and pride, which are aspects of introjected regulation that extroverts typically attempt to satisfy. Considering this, personality traits have only an indirect effect on individuals' motivation level to learn a second or foreign language (MacIntyre & Charos, 1996).

Many studies have shown that "the quality of experience and performance can produce different results when one is behaving for intrinsic versus extrinsic motives." (Zohrabi & Yousefi, 2016, p. 17). The present study supports this claim, as there were weak correlations between the above-mentioned scales.

The Knowledge and stimulation scales were shown to highly correlate with one another (r = .711; p < .001). For the participants, acquiring knowledge about English-speaking communities and discovering new things appears to be particularly enchanting, as they consider that this process of learning is fun and enriching.

A strong correlation was also found between the knowledge accomplishment scales of intrinsic motivation (r = .612; p < .001). This suggests that learning about English-speaking communities around the world and discovering new things via the use of English as a foreign language appears to be a pleasurable experience for participants. Learners may be intrinsically motivated when the activities are interesting for them (Tan et al., 2016). This motivation may further stimulate them to learn the language, providing positive feelings regarding the ability to try new things.

The accomplishment scale also shows a high correlation with the stimulation scale (r = .494; p < .001). As such, can be concluded that overcom-

ing the challenges and difficulties that are present in the English learning process can stimulate learners and enhance the learning experience due to the positive feelings associated with using the language. Tan et al. (2016) also found that "The need for the feeling of accomplishment is linked to ego-boosting where students will have a sense of accomplishment by being able to write, read, listen and speak the target language" (p. 285).

Peng (2012) found that motivation is linked to learners' WTC. This study advocates the existence of a link between intrinsic motivation (i.e., stimulation) and WTC. The more enjoyable and positive the process of learning English is, the more language learners are willing to communicate in their classes ($r = .448$; $p < .001$). Whether they are willing to introduce themselves, give a short speech about their hometown or their favorite game, or simulate real-life situations using English, these enjoyable activities boost their proficiency in the target language. It is not surprising that in real-life settings, learners are also willing to communicate using English with strangers or friends ($r = .353$, $p < .001$).

Furthermore, gaining deeper knowledge about English communities induces learners to share this information in the same language in front of their teachers and peers ($r = .397$; $p < .001$). They also appear willing to share this knowledge and make use of it either in a work meeting or a friendly gathering (r = .363; p < .001). It also should be noted that this relationship is reciprocal: learners might enjoy learning from those they interact with because "the more one communicates, the more practice one has in talking and the more one learns (Brown, 1987; Rubin & Thompson, 1994)" (as cited in Yashima, 2002, p. 55). This learning does not only concern shared knowledge but also practicing the foreign language.

A weak correlation was observed between WTC and accomplishment as part of intrinsic motivation. The findings suggest that grasping or accomplishing difficult tasks while learning English does not influence the level of in-class WTC ($r = .215$; $p = .007$). It appears even less influential in the case of out-of-class WTC ($r = .142$; $p = .078$).

Negative correlations were also identified between the constructs. Amotivation towards learning English was shown to be associated with a decrease in in-class ($r = -.244$; $p = .002$) and out-of-class WTC ($r = -.192$; $p = .017$).

Therefore, the higher the participants' amotivation, the less they are willing to communicate in the target language. In addition, this amotivation did not correlate with participants' extroversion-introversion.

Furthermore, the findings showed that intended effort did not correlate with extraversion. As such, for the participants in the study there appears to be no relationship between being assertive, active, or energetic and making a concentrated effort to learn and communicate in English. However, there is a significant correlation between the intended effort and WTC scales. Inside the classroom, learners are more willing to expand their efforts and try their best to interact with their peers and instructors (r = .364; p < .001). When they are outside the classroom, they appear to be less willing to make an effort to communicate using the language (r = .239; p = .003).

Conclusion

In summary, the present research intended to answer two research questions. Overall, it is important to conclude that English language learners' motivation, WTC, and extroversion-introversion are three different and complex constructs that are intricately interrelated. The correlation findings revealed that a relationship exists between the WTC construct and participants' extraversion. The WTC scales also correlated with the intrinsic motivation subscales of knowledge and stimulation and correlated moderately with the extrinsic motivation subscales of external regulation and introjected regulation. Although there was no apparent correlation between extroversion and the motivational scales, it requires further investigation.

The research showed that there is a correlation between WTC, motivation, and extroversion-introversion among learners of English as a foreign language. Based on the theoretical framework and data collection through the online questionnaire, the diversity of the constructs alludes to the complexity inherent in the process of learning a new language. The study also provided an understanding of the nature of foreign language learning motivation, extroversion-introversion, and WTC among the respondents.

The analysis of correlations revealed noteworthy connections among the various elements examined in this study. A robust correlation was ob-

served between Willingness to Communicate (WTC) in both classroom and non-classroom settings and intrinsic motivation factors (specifically, stimulation and knowledge) as well as the extroversion-introversion tendencies of individuals. Furthermore, a moderate correlation was identified between the WTC measures and the subcategories of extrinsic motivation, including external regulation and introjected regulation.

Extroversion was not found to be related to the English language learning motivation of the participants, although correlations within the scales from the same constructs showed strong relationships. WTC inside the classroom strongly correlated with WTC outside the class, and extrinsic motivation also strongly correlated with intrinsic motivation. English language learners in this study were found to be both intrinsically and extrinsically motivated, thus revealing a moderate level of self-determination. These findings shed light on the significant effect that motivation has on WTC, showing the two factors to be tightly woven constructs related to second and foreign language acquisition (Peng, 2007). The study demonstrated that WTC and motivation to learn English are complementary and inseparable in the learning and teaching process. Mastering a foreign language such as English is indeed a complex and multi-faceted process that requires a combination of skills and competencies from the learner.

Limitations of the present study must also be addressed. While this study depicted how the motivation to learn English is related to WTC and extroversion-introversion, it did not explore how this correlation of the three constructs could contribute to the participants' success or failure in language learning. Methodologically, the sample size (N=155) is large enough to allow for generalizations over a population. However, participants were recruited from different national contexts in varying numbers, limiting the extent to which the findings can be generalized. Limiting the sample to one study context could be a further development for further exploration of the research topic. For example, as the majority of participants in this study were from the Republic of Tunisia, an investigation on the influence of motivation, WTC, and extroversion-introversion on English language learning in the Tunisian context could follow.

Alternative methods could be employed to assess the interrelationship between motivation and WTC as well as to test the influence of age, duration of studies, or proficiency levels of the participants. Selecting a specific group of participants and carrying out long-term observations, consecutive and regular follow-ups, as well as self-reports might provide a more insightful, comprehensive, and accurate monitoring of the language learning process. The study could have revealed more information if, for instance, qualitative research methods had been utilized.

The questionnaire survey that informed the present study offered a simplified view of the topic under investigation. Future investigations can be directed to the various components of WTC and can explore other factors influencing including social, cultural, or ethnic dimensions. Additional aspects of extroversion-introversion and motivational can also be incorporated to provide a more holistic understanding of the concepts. Exploring these different perspectives would offer a thorough exploration of the variables and their interrelationships. Furthermore, other research might examine other foreign languages besides English.

Given the recent tendencies in language pedagogy and teaching methodologies that emphasize the communication approach, teachers and language instructors might consider adopting and implementing interactive methods that would bolster communicative skills and make language learning more successful and rewarding, given the fact that learners are willing to communicate when they are motivated and interested in classroom content. Fostering willingness to use the target language both inside and outside of the classroom context as well as switching from teacher-centered classes to learner-centered can enhance learners' level of proficiency in English as a foreign language and provide learners with increased opportunities to communicate and interact during class, which will boost their communicative competencies. Second and foreign language educators can thus work to ensure that the learning experience is successful by enhancing students' intrinsic and extrinsic motivation, which will consequently influence their WTC and their linguistic competence as "a higher level of self-determination means a higher level of WTC" (Altiner, 2018, p. 184).

Departments of education and linguistic development within academic institutions, including within ministries of education and higher education, as well as national agendas and organizations, should work hand in hand to introduce policies and implement strategies to raise awareness of the vivacity and importance of communicative approaches as well as fostering and scaffolding motivation in educational settings.

References

Alqahtani, M. (2015). Saudi students' willingness to communicate and success in learning English as a foreign language. *Journal of Language Teaching and Research*, 6(6), 1195–1205. https://doi.org/10.17507/jltr.0606.06

Altiner, C. (2018). Turkish EFL learners' willingness to communicate in L2 and motivation. *Journal of Education and Training Studies*, 6(11a), 181–185. https://doi.org/10.11114/jets.v6i11a.3815

Altunel, V. (2015). *The impact of extroversion and introversion on language learning in an input-based EFL setting* (pp. 1–93) [Unpublished master's dissertation]. University of Kansas. http://hdl.handle.net/1808/21595

Anjomshoa, L., & Firooz, S. (2015). The importance of motivation in second language acquisition. *International Journal on Studies in English Language and Literature, 3*(2), 126–137. https://www.arcjournals.org/pdfs/ijsell/v3-i2/12.pdf

Baghaei, P., & Dourakhshan, A. (2012). The relationship between willingness to communicate and success in learning English as a foreign language. *The Modern Journal of Applied Linguistics*, 4(2), 82–92. http://site.iugaza.edu.ps/nmasri/files/2010/09/Mobile-Phone-Technology.pdf

Burgoon, J. K. (1976). The Unwillingness-To-Communicate Scale: Development and validation. *Communication Monographs, 43*(1), 60–69. https://doi.org/10.1080/03637757609375916

Clément, R. (1980). Ethnicity, contact, and communicative competence in a second language. In H. Giles, W. P. Robinson, & P. M. Smith (Eds.), *Language: Social psychology perspectives* (pp. 147–154). Pergamon. https://doi.org/10.1016/B978-0-08-024696-3.50027-2

Clément, R., & Gardner, R. C. (2001). Second language mastery. In H. Giles & W. P. Robinson (Eds.), *The new handbook of language and social psychology* (pp. 489–504). Wiley.

Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment. Companion volume*. https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4

Csizér, K., & Kormos, J. (2008). Age-related differences in the motivation of learning English as a foreign language: Attitudes, selves, and motivated learning behavior. *Language Learning, 58*(2), 327–355. https://doi.org/10.1111/j.1467-9922.2008.00443.x

Darling, W. E., & Chanyoo, N. (2018). Relationships of L2 motivational self-system components and willingness to communicate in English among Thai undergraduate Students. *Journal of Language and Communication, 23*(33), 1–22 https://so04.tci-thaijo.org/index.php/NJLC/article/view/135383/101151

Deci, E. L., & Ryan, R. M. (2004). Overview of self-determination theory: An organismic dialectical perspective. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 3–33). The University of Rochester Press. https://www.elaborer.org/cours/A16/lectures/Ryan2004.pdf

Dewaele, J.-M., & Dewaele, L. (2018). Learner-internal and learner-external predictors of willingness to communicate in the FL classroom. *Journal of the European Second Language Association, 2*(1), 24–37 http://doi.org/10.22599/jesla.37

Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching, 31*(3), 117–135. https://doi.org/10.1017/S026144480001315X

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Blackwell. https://doi.org/10.1002/9780470756492.ch18

Dörnyei, Z. (2003). Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. *Language Learning, 53*(S1), 3–32. https://doi.org/10.1111/1467-9922.53222

Dörnyei, Z. (2005). *The psychology of the language learner revisited*. Routledge.

Dörnyei, Z. (2006). Individual differences in second language acquisition. *AILA Review, 19*(1), 42–68. https://doi.org/10.1075/aila.19.05dor

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.

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

Gardner, R. C. (2001). Integrative motivation and second language acquisition. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 1–20). University of Hawai'i.

Gardner, R. C. (2007). Motivation and second language acquisition. *Porta Linguarum, 8*, 9–20. https://doi.org/10.30827/Digibug.31616

Gardner, R., & MacIntyre, P. D. (1991). An instrumental motivation in language study: Who says it isn't effective? *Studies in Second Language Acquisition, 13*(1), 57–72. https://doi.org/10.1017/S0272263100009724

Gharibi, S., & Seyyedrezaei, H. S. (2016). The importance of willingness to communicate and self-esteem in language learning and teaching. *Journal of International Education in Business, 4*(1), 194–197. http://www.psp-ltd.com/JIEB_56_4_2016.pdf

Giles, H., & Byrne, J. L. (1982). An intergroup approach to second language acquisition. *Journal of Multilingual and Multicultural Development, 3*(1), 17–40. https://doi.org/10.1080/01434632.1982.9994069

Gholami, L. (2015). Willingness to communicate and its relationship with emotional intelligence and gender differences. *International Letters of Social and Humanistic Sciences, 52*, 87–94. https://doi.org/10.18052/www.scipress.com/ILSHS.52.87

Guilloteaux, M. J., & Dörnyei, Z. (2008). Motivating language learners: A classroom-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly, 42*(1), 55–77. https://doi.org/10.1002/j.1545-7249.2008.tb00207.x

Goldberg, R. L. (1992). The development of markers for the big-five-factor structure. *Psychological Assessment, 4*(1), 26–42. https://psycnet.apa.org/doi/10.1037/1040-3590.4.1.26

Hashimoto, Y. (2002). Motivation and willingness to communicate as predictors of reported L2 use: The Japanese ESL context. *Second Language Studies, 20*(2), 29–70. http://hdl.handle.net/10125/40649

Higgins, E. T. (1987). Self-discrepancy: A theory relating to self and affect. *Psychological Review, 94*(3), 319–340. https://doi.org/10.1037/0033-295X.94.3.319

Hişmanoğlu, M., & Özüdoğru, F. (2017). An investigation of university students' willingness to communicate in English in relation to some learner variables. *Karabük Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 7*(2), 449–461. https://dergipark.org.tr/tr/download/article-file/374276

Kanat-Mutluoğlu, A. (2016). The influence of ideal L2 self, academic self-concept and intercultural communicative competence on willingness to communicate in a foreign language. *Eurasian Journal of Applied Linguistics, 2*(2), 27–46. https://doi.org/10.32601/ejal.461012

Lalonde, R. N., & Gardner, R. C. (1984). Investigating a causal model of second language acquisition: Where does personality fit? *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement, 16*(3), 224–237. https://psycnet.apa.org/doi/10.1037/h0080844

Ma, X., Wannaruk, A., & Lei, Z. (2019). Exploring the relationship between learning motivation and L2 WTC in an EFL classroom among Thai EFL learners. *English Language Teaching, 12*(7), 33–45. https://doi.org/10.5539/elt.v12n7p33

MacIntyre, P. D., & Charos, C. (1996). Personality, attitudes, and affect as predictors of second language communication. *Journal of Language and Social Psychology, 15*(1), 3–26. https://doi.org/10.1177/0261927X960151001

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

MacIntyre, P. D., MacMaster, K., & Baker, S. C. (2001). The convergence of multiple models of motivation for second language learning: Gardner, Pintrich, Kuhl, and McCroskey. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 461–492). University of Hawai'i.

MacIntyre, P. D., Baker, C. S., Clément, R., & Conrod, S. (2001). Willingness to communicate, social support, and language-learning orientations of immersion students. *Studies in Second Language Acquisition, 23*(3), 369–388. https://doi.org/10.1017/S0272263101003035

MacIntyre, P. D., Burns, C., & Jessome, A. (2011). Ambivalence about communicating in a second language: A qualitative study of French immersion students' willingness to communicate. *The Modern Language Journal, 95*(1), 81–96. https://doi.org/10.1111/j.1540-4781.2010.01141.x

MacIntyre, P. D., Noels, K., & Moore, B. (2010). Perspectives on motivation in second language acquisition: Lessons from the Ryoanji Garden. In M. T. Prior, Y. Watanabe, & S. K. Lee (Eds.) *Selected proceedings of the 2008 Second Language Research Forum: Exploring SLA perspectives, positions, and practices* (pp. 1–9). Cascadilla Press.

Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist, 41*(9), 954–969. https://doi:10.1037/0003-066X.41.9.954

McCroskey, C. J., & Baer, J. E. (1985, November 7–10). *Willingness to Communicate: The construct and its measurement* [Conference Presentation]. Annual Meeting of the Speech Communication Association, Denver, CO, United States. https://files.eric.ed.gov/fulltext/ED265604.pdf

McCroskey, C. J. (1992). Reliability and validity of the willingness to communicate scale. *Communication Quarterly, 40*(1), 16–25. https://doi.org/10.1080/01463379209369817

Mortensen, C. D., Arntson, P. H., & Lustig, M. (1977). The measurement of verbal predispositions: scale development and application. *Human Communication Research, 3*(2), 146–158. https://doi.org/10.1111/j.1468-2958.1977.tb00513.x

Mystkowska-Wiertelak, A., & Pawlak, M. (2014). Fluctuations in learners' willingness to communicate during communicative task performance: Conditions and tendencies. *Research in Language, 12*(3), 245–260. https://doi.org/10.2478/rela-2014-0019

Noels, K. A. (2001). New orientations in language learning motivation: Towards a model of intrinsic, extrinsic and integrative orientations and motivation. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and second language acquisition* (pp. 43–68). University of Hawai'i.

Noels, K. A., & Clément, R. (1996). Communicating across cultures: Social determinants and acculturative consequences. *Canadian Journal of Behavioural Science, 28*(3), 214–228. https://doi.org/10.1037/0008-400X.28.3.214

Noels, K. A., Clément, R., & Pelletier, G. L. (1999). Perceptions of teachers' communicative style and students' intrinsic and extrinsic motivation. *The Modern Language Journal, 8*(1), 23–34. https://doi.org/10.1111/0026-7902.00003

Noels, K. A., Clément, R., & Pelletier, G. L. (2001). Intrinsic, extrinsic, and integrative orientations of French Canadian learners of English. *Canadian Modern Language Review, 57*(3), 424–440. https://doi.org/10.3138/cmlr.57.3.424

Noels, K. A., Pelletier, L. G., Clément, R., & Vallerand, R. J. (2000). Why are you learning a second language? Motivational orientations and self-determination theory. *Language Learning*, *50*(1), 57–85. https://doi.org/10.1111/0023-8333.00111

Özbay, A. Ş., Aydemir, T., & Atasoy, Y. (2017). Extroversion—Introversion in the context of foreign language learning. *The International Journal of Educational Researchers, 8*(3), 1–21.

Peng, C. (2012). Self-regulated learning behavior of college students of science and their academic achievement. *Physics Procedia*, *33*, 1446–1450. https://doi.org/10.1016/j.phpro.2012.05.236

Peng, J. (2007). Willingness to communicate in an L2 and integrative motivation among college students in an intensive English language program in China. *University of Sydney Papers in TESOL, 2,* 33–59.

Peng, J. E., & Woodrow, L. (2010). Willingness to communicate in English: A model in the Chinese EFL classroom context. *Language Learning, 60*(4), 834–876. https://doi.org/10.1111/j.1467-9922.2010.00576.x

Schumann, J. H. (1986). Research on the acculturation model for second language acquisition. *Journal of Multilingual and Multicultural Development, 7*(5), 379–392. https://doi.org/10.1080/01434632.1986.9994254

Şener, S. (2014). Turkish ELT students' willingness to communicate in English. *ELT Research Journal, 3*(2), 91–109. https://dergipark.org.tr/en/download/article-file/63640

Simić, I. (2014). *Willingness to communicate: A comparison between Austrian and Serbian students.* [Unpublished master's thesis]. University of Graz.

Tan, S. M., Ong, S. C., Sim, C. P., Wei, W., & Hoi, Y. (2016). Willingness to communicate in English and motivation level among Chinese secondary students in northern Malaysia. *International Journal of Humanities and Management Sciences, 4*(3), 283–287.

Wakamoto, N. (2009). *Extroversion/introversion in foreign language learning: Interactions with learner strategy use.* Peter Lang.

Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal, 86*(1), 54–66. https://doi.org/10.1111/1540-4781.00136

Zohrabi, M., & Yousefi, M. (2016). The relationship between reflective teaching, willingness to communicate (WTC), and intrinsic motivation of Iranian advanced learners. *International Journal on Studies in English Language and Literature, 4*(2), 12–28. https://doi.org/10.20431/2347-3134.0402003

Appendix

English Language Learning Questionnaire
My name is _____ I am currently a _____ specialized in _____ at _____, _____.
I would like to ask for your help with my research.

The following questionnaire aims to examine how extroversion-introversion, motivation, and willingness to communicate influence the English language learning experience.

For later purposes, the questionnaire answers will be stored, and the data collected from them will be analyzed. The investigation is anonymous and voluntary. You have the right to withdraw at any time.

I am interested in your personal opinion, so please make sure your answers are sincere. This is fundamental for the quality of my research.

The research is divided into four major sections:

1	Extroversion-introversion
2	English language learning motivation
3	Willingness to communicate in the English language.
4	Bio questions

It takes between 10 and 15 minutes to fill out all required information.

Extroversion-Introversion	
1	I am an extroverted person.
2	I am an energetic person.
3	I am a talkative person.
4	I am willing to take risks.
5	I am an active person.
6	I am assertive (communicating with confidence).
7	I am an adventurous person.

English language learning motivation	
1	I cannot see why I study the English language; I couldn't care less about it.
2	I can't think of any good reason why I study the English language.

3   Honestly, I don't know; I truly have the impression of wasting my time studying the English language.

4   I don't know; I can't come to understand why I am studying the English language.

5   I am learning the English language to get a more prestigious job later on.

6   I am learning the English language to gain the benefits that talking in English will provide me.

7   I am learning the English language to have a better salary later on.

8   I am learning to gain the benefits that entrance into the English language community will provide me.

9   I am learning English because I would feel embarrassed or ashamed if I didn't know this language.

10 I am learning because I would feel ashamed if I couldn't speak to my friends from an English-speaking community English.

11 I am learning the English language because I would feel bad if I didn't know the language.

12 I am learning the target language because I would feel guilty if I didn't know the language.

13 I am learning the English language because knowing the language is a part of who I am and what I do.

14 I am learning the English language because it is an integral part of my life.

15 I am learning the English language because it is an important part of how I define myself.

16 I am learning the English language because it is a fundamental part of who I am.

17 I am learning the English language for the enjoyment I experience when I grasp a difficult construct in it.

18 I am learning the English language for the satisfaction I feel when I am in the process of accomplishing difficult exercises in it.

19 I am learning the English language for the pleasure I experience when surpassing myself in my English language studies.

20 I enjoy the feeling of acquiring knowledge about the English language community and their way of life.

21 I enjoy the feeling of acquiring knowledge about the English language community and their way of life.

22 I am learning the English language for the satisfaction I get in finding out new things.

23 I am learning for the pleasure I experience in knowing more about the English language.

24 I am learning the English language because I love doing it; it's fun.

25 I am learning for the "high" I feel when hearing the English language spoken.

26 I am learning because of the positive feeling that I experience when using the English language.

27 I am learning because it's a great feeling to be able to use the English language.

28 If a course in the English language was offered at university or somewhere else in the future, I would like to take it.

29 I would like to study the English language even if I were not required.

30 I would like to spend lots of time studying the English language.

31 I would like to concentrate on studying the English language more than any other topic.

32 I am prepared to expend a lot of effort in learning the English language.

33 I am working hard at learning the English language.

34 I think that I am doing my best to learn the English language.

Willingness to communicate in the English language

1 I am willing to do a role-play standing in front of the class in the English language.

2 I am willing to give a short self-introduction without notes to the class in the English language.

3 I am willing to give a short speech about my hometown in the English language to the class.

4 I am willing to translate a spoken utterance from my native language into the English language in my group.

5 I am willing to ask the teacher in the English language to repeat what he/she just said in English because I didn't understand.

6 I am willing to do a role-play in the English language at my desk, with my peer.

7 I am willing to ask my peer sitting next to me in the English language about the meaning of an English word.

8 I am willing to ask my group mates in the English language the meaning of a word I do not know.

9 I am willing to ask my peer sitting next to me how to say in the English language a phrase to express the thoughts in my mind.

10 I am willing to speak in a group about my summer vacation in the English language.

11 I am willing to speak with my teacher about my homework assignment in the English language.

12 I am willing to ask for instructions/clarification in the English language if I am confused about a task I must complete.

13 I am willing to describe the rules of my favorite game in the English language.

14 I am willing to play a game in the English language.

15   I am willing to talk with a service station attendant in the English language.

16   I am willing to talk with a physician in the English language.

17   I am willing to present a talk to a group of strangers in the English language.

18   I am willing to talk with a salesperson in a store in the English language.

19   I am willing to talk in a large meeting of friends in the English language.

20   I am willing to talk with a police officer in the English language.

21   I am willing to talk with a small group of strangers in the English language.

22   I am willing to talk with a friend while standing in line in the English language.

23   I am willing to talk with a waiter/waitress in a restaurant in the English language.

24   I am willing to talk in a large meeting of acquaintances in the English language.

25   I am willing to talk with a stranger while standing in line in the English language.

26   I am willing to talk with a secretary in the English language.

27   I am willing to present a talk to a group of friends in the English language.

28   I am willing to talk with a small group of acquaintances in the English language.

29   I am willing to talk in a large meeting of strangers in the English language.

30   I am willing to talk with a spouse (or girl/boyfriend) in the English language.

31   I am willing to talk with a small group of friends in the English language.

32   I am willing to present a talk to a group of acquaintances in the English language.

Bio questions

1   What is your nationality?

2   How old are you?

3   What is your gender? Male Female

4   What do you do in your life? Student employee other

5   Is English your: 1st foreign language? 2nd foreign language? 3rd foreign language?

6   What is your level in the English language? Beginner Intermediate Advanced Other

7   How long have you been learning the English language? Please add the number of years (or months) only.

You made it. This is the end of the questionnaire. Thank you for taking part in my research. Your participation is highly valued.

5

An EFL Learner's Development of Metacognitive Listening Strategy Use and Listening Performance: A Complex Dynamic Systems Theory Approach

Thet Oo Khaing
thetoo@student.elte.hu

Abstract

Using a Complex Dynamic Systems Theory (CDST) approach allows us to trace dynamic changes in learners' language development as well as dynamic interactions of variables present within a complex learning system (de Bot et al., 2007). Using a mixed-methods CDST approach, this small-scale study aims to explore the developmental trajectories of an EFL learner's metacognitive listening strategy use and listening performance as well as investigate the dynamic interaction between the two variables. A 21-year-old EFL learner from Myanmar received metacognitive strategy training for seven days, and his listening strategy use and listening performance were tracked every two days over a two-week period. Regarding the data collection, First Certificate English (FCE) listening practice tests were used to assess the learner's listening performance, and semi-structured interviews were conducted to measure the listening strategy use. Self-reflections were also collected to gain a more detailed picture of the processes under investigation. Data was analyzed using CDST methods and techniques. The results indicated non-linear developmental patterns in listening strategy use and listening performance as well as dynamic correlations between metacognitive listening strategy use and listening performance. The findings provide EFL teachers with insights into the interaction between metacognitive listening strategies and listening performance in a dynamic system. Implications for pedagogical practices are discussed in relation to the development of EFL learners' metacognitive listening strategies and listening performance.

Keywords: Complex Dynamic Systems Theory, metacognitive listening strategy, listening performance, developmental trajectories, EFL learner

An EFL Learner's Development of Metacognitive Listening Strategy Use and Listening Performance: A Complex Dynamic Systems Theory Approach

In the English as a foreign language (EFL) context, the development of listening skills has received less systematic attention from teachers and instructional materials when compared to the development of reading, writing, or speaking skills (Vandergrift & Goh, 2012). Listening is often considered to be the most difficult skill to develop among EFL learners (e.g., Abdalhamid, 2012; Alzamil, 2021; Darti & Asmawati, 2017), and listening activities used in EFL classrooms tend to test learners' listening ability rather than teaching listening. As a result, most EFL learners lack listening skills and are not aware of the processes involved in listening (e.g., Gilakjani & Ahmadi, 2011; Goh, 2000; Wah, 2019). Students' listening problems were reported to be associated with cognitive procedures such as real-time processing (Goh, 2000). In this regard, several studies have examined how listening strategies can be used to develop learners' listening performance, presenting listening strategies as an effective way to improve listening processes and comprehension (e.g., Cross, 2011; Goh & Taib, 2006; Ozeki, 2006; Vandergrift & Tafaghodtari, 2010). It should be noted that these studies adopted pretest-posttest research designs, assuming that learners' use of listening strategies and development was linear and systematic. However, Chamot (1995) pointed out that students' use of strategies might not always be as systematic and predetermined as generally believed.

From the CDST perspective, language use and development phenomena are viewed as emergent systems that are complex, interconnected, dynamic, self-organizing, context-dependent, open, adaptive, and non-linear (Larsen-Freeman, 1997). With a CDST approach, developmental dynamics can be examined through variability in language development. Considering that traditional listening strategy research has tended to overlook dynamic patterns of L2 learners' listening strategy use and development, conducting a case study using a CDST approach which examines an EFL learner's metacognitive listening strategy use and listening performance

may provide a better understanding of the dynamics of the developmental processes involved in an EFL learner's metacognitive listening strategies and listening performance. Looking at the EFL context in Myanmar, there is a lack of studies in investigating metacognitive listening strategy use and listening development of Myanmar EFL learners, and CDST remains to be utilized in any empirical studies. The present study aims to fill this research gap, and its findings can be of use to English Language Teaching in Myanmar and in similar contexts.

Thus, the primary aim of the present study is to investigate the dynamics of metacognitive listening strategy use in an EFL learner in Myanmar as well as his listening performance. The sub-aims of the study are 1) to explore the developmental patterns in the EFL learner's use of metacognitive listening strategies following metacognitive strategy training; 2) to observe changes in the EFL learner's listening performance over a two-week period; and 3) to examine the dynamic relationship between the EFL learner's use of metacognitive listening strategies and listening performance over the study period.

Literature Review

Metacognitive listening refers to the learners' awareness regarding their understanding of listening demands, cognitive goals, and their self-perceptions, and consists of five strategies: 1) problem-solving, 2) planning and evaluation, 3) directed attention, 4) personal knowledge, and 5) mental translation (Vandergrift et al., 2006). These strategies overlap with the main components in Anderson's (2002) model of metacognitive listening strategy training, namely the components of preparing and planning, deciding when to use particular strategies, monitoring strategy use, learning how to orchestrate various strategies, and evaluating strategy use. The use of metacognitive strategies can activate an individual's thinking and lead to improved performance in learning in general (Anderson, 2002).

The effects of using metacognitive strategies on listening performance has been investigated in numerous empirical studies, the results of which tending to show that metacognitive listening strategies were effective in

raising learners' awareness of the listening process, developing learners' listening performance, and increasing the self-regulated use of comprehension strategies. For instance, Goh and Taib (2006) conducted a small-scale study with a group of 10 Chinese primary school ESL learners over eight listening lessons and found increases in motivation, confidence, and strategy knowledge among participants with weaker listening skills. Vandergrift and Tafaghodtari's (2010) study found that the less-skilled listeners among 106 tertiary-level high-beginner and lower-intermediate learners of French as an L2 who participated in the strategy-based study made more gains than their more-skilled peers over a semester. Similarly, the results of the study conducted by Cross (2011) revealed that three out of the four participants with weaker listening skills made substantial gains across five lessons, whereas only one of the four more-skilled listeners showed improvement.

On the other hand, some empirical studies have provided contradictory findings regarding the effectiveness of metacognitive strategy training on listening performance. For example, a study carried out by Ozeki (2006) found no significant differences between the post-test scores of the experimental and control groups of EFL learners after a listening strategies intervention. Similarly, Rahimi and Katal's (2013) study observed that there were no significant differences between the listening abilities of the experimental and control groups after a 16-week period of metacognitive instruction. Dong (2016) pointed out that in recent literature, there has been an ongoing controversy over the effectiveness and practicality of listening strategy training. It should be noted that all of the aforementioned studies were conducted using quantitative interventional designs. Despite extensive research into L2 listening, little attention has been paid to examining L2 listening from a CDST perspective.

CDST is a framework for examining the change in complex systems over time and to explore the non-linear and iterative development of natural phenomena (Lowie, 2013). A dynamic system consists of various inter-related variables that interact and change constantly over time, which are never entirely stable but rather show a great deal of variability (Schmid et al. 2011). This variability in development is seen as "the intrinsic property

of the developmental process" which is assumed to be at the core of CDST (Van Dijk et al., 2011, p. 55). Analysing variability can help us in understanding how and when different subsystems change and develop over time, as well as the dynamic relations between the relevant subsystems (Verspoor et al., 2008). Variability also allows us to see phase transitions, which represent new forms or properties that appear during developmental processes (van Dijk & van Geert, 2007). From a CDST perspective, an individual's language is a fully integrated system with no end state in language development (Lowie, 2013). Chang and Zhang (2021) also conceptualize language learning as a non-linear and complex dynamic process due to the intricate, complex, and unpredictable characteristics which constitute the process.

Although existing L2 listening studies employing a CDST perspective have largely focused on theoretical aspects, such as how listening strategies contribute to learners' listening performance (Chang & Zhang, 2021), there have been a few empirical studies on L2 listening which have been conducted using a CDST approach. For instance, Qiu and Huang (2012) explored the effects of dynamic image schema (DIS) on the systematic improvement of ESL learners' listening ability from a CDST perspective, finding that DIS facilitated the learners' listening comprehension. In the longitudinal study conducted by Chang and Zhang (2021), a CDST framework was adopted to explore the complex patterns and the degree of variability in Chinese EFL learners' listening development, with variability identified both between and within individuals.

A CDST perspective on metacognition considers "the dynamic aspects of how learners perceive themselves, learning tasks, learning processes, and how they value others' views of them and how to complete the learning tasks in specific learning and teaching environments" (Zhang & Zhang, 2013, p. 116). Metacognition refers to the knowledge and control that an individual has over their cognitive processes (Baker, 2002). In other words, it can be regarded as a self-regulatory process as it is related to the awareness of one's comprehension process, use of strategies, and control of such strategies. In Wind and Harding's (2020) study, self-regulation was shown to be a dynamic construct. Regarding the CDST research re-

lated to listening strategies, Dong (2016) investigated the development of an EFL learner's use of metacognitive, cognitive and social/affective listening strategies as well as their listening performance over 40 weeks using a DST approach. The findings revealed non-linear developmental patterns in the learner's listening strategy use and listening performance, and a downward trend in the relationship between listening strategy use and listening performance.

While Dong's study indicated dynamics patterns in the use of three listening strategies and listening development over an extended period of time, the present study focuses on metacognitive listening strategies with an aim to determine whether a shorter period of strategy training can have an impact on an EFL learner's listening strategy use and development over a limited period of time. To achieve that aim, the current research adopted similar procedures used by Dong (2016). There are three main reasons for focusing specifically on metacognitive listening strategies: 1) metacognitive strategies alone can be used in strategy training, as seen in a number of previous studies (e.g., Cross, 2011; Goh & Taib, 2006); 2) no empirical studies examining metacognitive listening strategies have been conducted using a CDST approach; and 3) Myanmar EFL learners appear largely aware of any of the aspects of the listening process. The study therefore aimed to explore the dynamic developmental processes involved in a Myanmar EFL learner's metacognitive listening strategy use and listening performance as well as the interaction of those two variables in the dynamic system over a short study period.

Research Questions

To achieve the above-mentioned aim, the present study intends to address the following research questions:

1. What are the developmental patterns that can be observed in an EFL learner's use of metacognitive listening strategies following a short strategy training?

2 . How does the EFL learner's listening performance change over a two-week period?

3 . What is the dynamic relationship between the EFL learner's use of metacognitive listening strategies and listening performance over the study period?

Method

A mixed methods time-series research design was used in this study to investigate the dynamic changes in an EFL learner's metacognitive listening strategy use and listening performance and to examine the interrelationship between the two variables. The research design also allowed for evidence to be gathered regarding factors or past events which might explain the dynamic patterns observed among the variables.

The Participant

The study was conducted with a 21-year-old male undergraduate student in Myanmar majoring in law through a distance education program. The participant was selected using convenience and purposive sampling (Dörnyei, 2007). As an EFL learner who finds listening the most difficult among the four skills, and who has never been instructed on listening strategies in a systematic or explicit way, he met the criteria for the study. The participant has been learning English for about 15 years, starting in kindergarten. In addition, he joined English language classes (from elementary to pre-intermediate) at the British Council in Yangon, Myanmar, where his English developed from an elementary to a pre-intermediate level. At the time of the study, he was preparing to study at the B.A. marketing program at Bangkok University in Thailand, and he has been learning English with a private English teacher (the researcher) for 30 days. He has shown great interest in learning English and made extensive effort to improve his English skills, including listening, and was actively engaged in the metacognitive listening strategy training utilized in this study.

Research Instruments

In line with its mixed-method design, the present study adopted three main instruments for data collection: 1) listening practice tests, 2) semi-struc-

tured interviews, and 3) self-reflections. Detailed descriptions of each instrument are presented in the subsections below.

Listening Tests

The listening tests used to assess the participant's listening performance were selected from the First Certificate in English (FCE), one of the Cambridge English exams. The FCE test is widely recognized as a reliable English language test and is accepted by numerous businesses and educational institutions worldwide (Cambridge English, n.d.). The difficulty level of the listening test is B2 (Upper Intermediate) and the listening test takes about 40 minutes, containing four parts with 30 items in total (see Appendix B). In Part 1, there is a series of eight short extracts from monologues or exchanges between interacting speakers with one multiple-choice question for each situation. In Part 2 there are 10 fill-in-the-blank sentences related to a 3-minute monologue or text involving interacting speakers. In Part 3, five speakers talk about different topics, and multiple matching tasks are included in relation to the different speakers and topics. In Part 4 there is a longer recording accompanied by seven multiple choice questions, each worth 1 point. For the current study, FCE listening practice tests from Engxam.info were used, and it can be assumed that all the practice tests were homogeneous in terms of their level of difficulty.

The Interviews

Semi-structured interviews were employed in the present study to allow the participant to elaborate on the topic based on the guiding questions, leading to greater depth and breadth of the data provided by the participant (Dörnyei, 2007). An interview guide was prepared based on the Metacognitive Awareness Listening Questionnaire (MALQ; Vandergrift et al., 2006) (see Appendix C). It consists of 15 questions based on five main constructs: planning and evaluation, directed attention, person knowledge, mental translation, and problem solving. The interview was conducted in the participant's L1 to provide the participant with the opportunity to express their ideas freely without any language barrier.

Self-Reflections

During the study period, the participant was also asked to write self-reflections about his listening difficulties, listening experiences, feelings, and anything related to his listening strategy use and listening development. These reflections offered a clearer understanding of the participant's listening processes. From a CDST perspective, Wind (2021) found that engagement in self-reflections changes dynamically over time and that self-reflection can be traced using time-series analysis.

Data Collection Procedures

Prior to the data collection the participant filled in the consent form provided by the researcher. The metacognitive listening strategy training was carried out for one hour every day over a one-week period (7 hours in total). The researcher examined the developments in the learner's metacognitive listening strategy use and listening performance over two weeks, including the training period.

Using Anderson's (2002) metacognitive strategy training model, the teacher-researcher began by providing an explicit explanation of the metacognitive listening strategies in the L1 of the participant, Myanmar. The participant was then given strategy training during the listening lessons over the course of seven days. The teacher-researcher employed a 7-day intensive metacognitive listening strategy program aimed at improving the participant's English listening skills and metacognitive strategies, which were adapted from Anderson's (2002) model of metacognitive strategy training. This model was chosen as the strategies it provides align with the items in the listening strategies interview guide used in the study. Detailed descriptions of the strategies are presented in Appendix A.

The listening materials used for the in-class and after-class exercises were chosen from the coursebook entitled "New English File" (Intermediate Student's book) published by Oxford University Press (2006). Listening exercises for the training were selected from this book to ensure that different types of listening tasks were covered, such as multiple choice, picture matching, and question response. The participant was instructed on how to prepare and plan before listening activities, what strategies to use for

a given task, how to monitor strategy use, how to make use of the available strategies, and how to evaluate his strategy use based on his experiences. In order to reinforce the listening strategies and ensure additional practice opportunities, after-class exercises were also assigned to the participant after each lesson during the strategy training.

In the two-week observation period, the participant's strategy use and listening performance were measured six times in order to trace the changes in those two variables. The participant's listening performance was assessed with the FCE listening tests prior to the semi-structured interviews. The participant was then asked to write self-reflections based on his listening experiences. The first listening test and interview took place one day prior to the training with the aim of examining the learner's existing listening performance and listening strategy use before the training commenced. From that point forward, the listening test with the same tasks and the interviews were carried out every three days. The listening scores were entered into a Microsoft excel sheet after each instance of data collection, while the interviews were recorded with the consent of the participant. Self-reflections were collected after the observation period, and all of the research sessions took place online.

Data Analysis Procedures

Data was analyzed using CDST methods to investigate the development of the learner's use of metacognitive listening strategies and listening performance, as well as to explore the dynamic relationship between the two variables.

First, data from the listening tests were computer coded using SPSS software (version 23), while data about the participant's metacognitive strategy use obtained from the interviews data were coded manually. The codes were analyzed by the researcher and verified by a co-coder affiliated with the same university as the researcher. The values of all the available data were recalculated to a 0-1 scale in order to facilitate the visualization and comparison of the original values and interactions between different constructs (Verspoor et al., 2011). After recalculating the values, line graphs were generated to visualize the data representing the participant's use of

metacognitive strategies and overall listening performance. Since the use of smoothing techniques can help researchers visualize general developmental patterns present in data, polynomial trendlines were added to each graph and were set to the 3rd degree, as Verspoor et al. (2011) state that "to see the general trend a polynomial to the 2nd degree is useful" (p. 166). As the next phase of analysis, between-session variability was measured for both listening strategy use and performance in order to identify the developmental state of the learner's listening skills and strategy use as well as his adaptability to the changing environment around him. Between-session variability refers to the difference between an observation and the preceding observation for a given variable (Bassano & Van Geert, 2007) and can be calculated based on the absolute differences between consecutive measurement points over time (Dong, 2016). Finally, the dynamic relationship between strategy use and listening performance was examined by calculating the moving correlations between the two variables and creating a moving window of correlations on an excel sheet.

Results and Discussion

The Developmental Patterns of Metacognitive Listening Strategy Use
The raw data and smoothed trajectories of the individual metacognitive listening strategies are illustrated in Figures 1 (a) – (e) below. As shown in the figures, the learner's developmental trajectories regarding his metacognitive listening strategies were characterized by noticeable variation. With the help of polynomial trendlines (set to the 3rd degree), general dynamic patterns were found in the smoothed curves of the metacognitive strategy use trajectories. The analysis revealed that the participant's use of metacognitive listening strategies prior to the strategy instruction (T1) was very low. However, in the initial stage following the strategy training (T2), there was a sudden increase in the strategy use related to planning and evaluation as seen in in Figure 1(a), mental translation as seen in Figure 1(d), and problem solving, shown in Figure 1(e). In contrast, the use of the directed attention strategy, as seen in Figure 1(b), experienced a decline, while the person knowledge strategy, shown in Figure 1(c), remained constant. Sub-

sequent stages of the study period (T3 and T4) revealed fluctuations in strategies such as planning and evaluation as shown in Figure 1(a), mental translation as seen in Figure 1(d), and problem solving as seen in Figure 1(e). Interestingly, the learner's directed attention and person knowledge strategies, as seen in Figure 1(b) and Figure 1(c), respectively, showed an upward trend during the same stages. In the last two stages of the study (T5 and T6), which took place after the training, the use of planning and evaluation strategies again showed an increase and reached a high level, as seen in Figure 1(a). Mental translation strategy use, as shown in Figure 1(d), remained stable; although it stayed at a rather low level, a slight increase can be identified in comparison to the levels at the beginning of the study. On the other hand, the learner's use of problem-solving strategies, as seen in Figure 1(e), continued to decline, yet remained at a medium level. Both directed attention and person knowledge, shown in Figure 1(b) and Figure 1(c), respectively, showed concave trajectories, demonstrating a decrease in the rate of change towards the end of the study period. Furthermore, the learner's personal knowledge related to listening skills was reported to be rather low compared to his directed attention. A similar finding appeared in Dong (2016): planning, evaluation, and problem identification strategies showed an increase during the initial stages of the training.

Figure 1

The Dynamic Developmental Trajectories for Each Metacognitive Listening Strategy

(c) Person Knowledge

(d) Mental Translation

(e) Problem solving
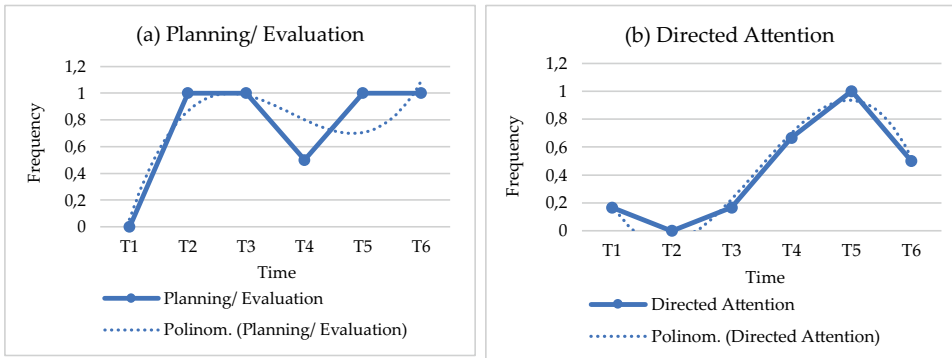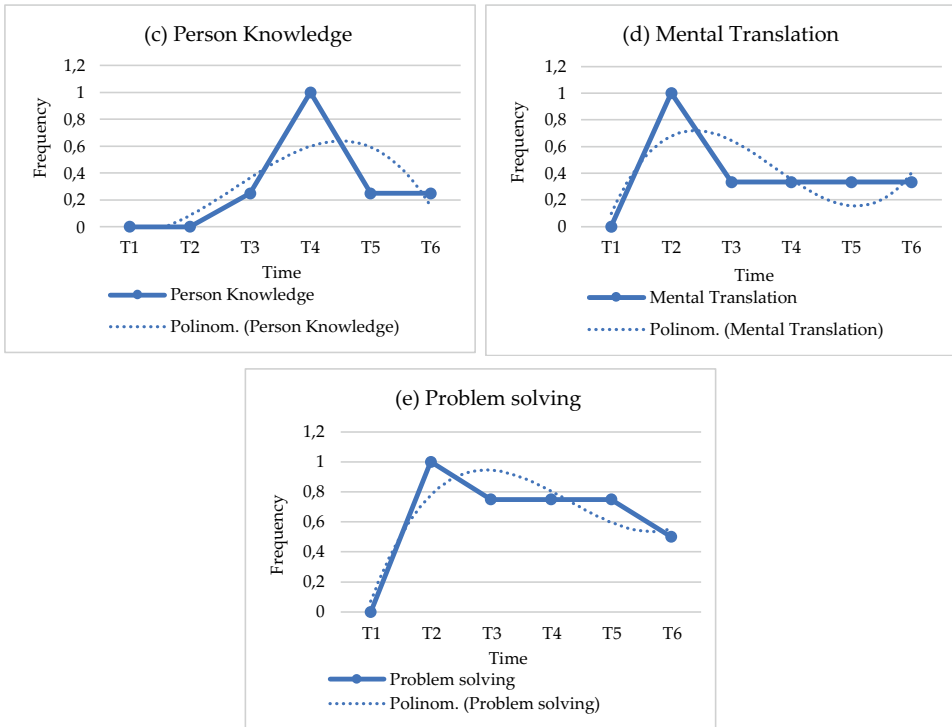
Figure 2 below illustrates the overall change in metacognitive listening strategy use, which was supplemented with a polynomial trendline (3rd degree) to showcase the general trend of the data. The examination of the overall strategy use indicated that the learner's use of metacognitive listening strategies increased notably after the strategy training and reached its peak at T4 and T5. Despite showing a small fluctuation midway through the observation period (at T3) and a slight decline at the end of the study (T6), the learner's overall strategy use at Times 3 and 5 was found to be much higher in comparison with their use (T1) prior to the instruction. This may reflect the effectiveness of the strategy instruction in increasing the learner's ability to apply the strategies learned during listening tasks and raising his metacognitive listening strategy awareness. Overall, the results indicate non-linear developmental patterns of listening strategy use with periods of progress and regression. The findings of the current study are compatible with those of previous studies (e.g., Dong, 2016; Siegler, 2006) in which the students' strategy development trajectories displayed

complexity, with considerable fluctuations and diverse patterns over time. This demonstrates that L2 learners develop in a non-linear way when it comes to listening strategy use, which is affected by a number of internal and external factors.

Figure 2

The Overall Use of Metacognitive Listening Strategies with Polynomial Trendline
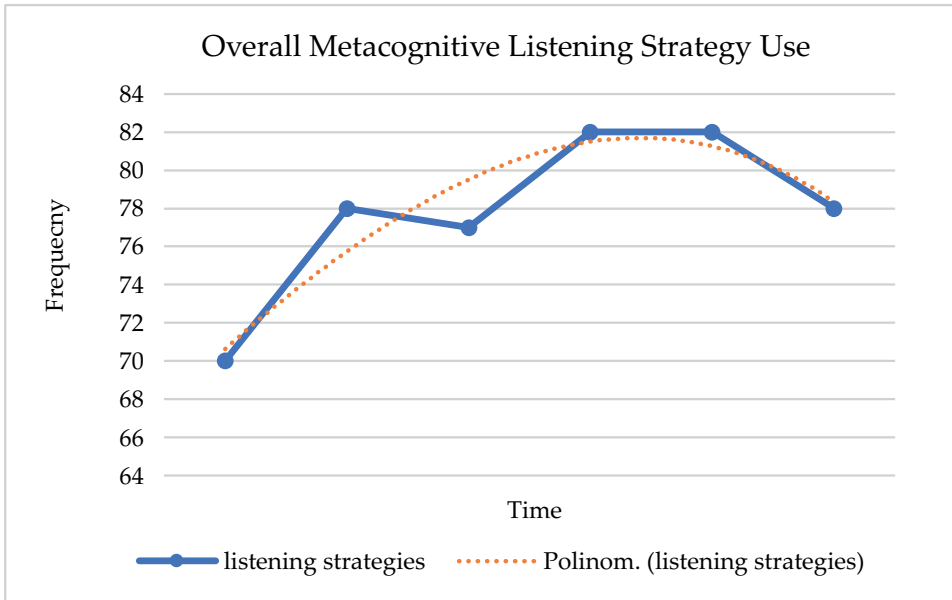
The trajectories of the participant's metacognitive listening strategies correspond to the self-reflections regarding his strategy use during the training. In one of the comments that he made during the first few days of the training, he stated that "after I was instructed with strategies, I came to understand a little more how to listen and how to find answers when doing a listening task." He also mentioned his difficulties with listening: "I don't understand when they use a lot of linking sounds. When I tried to pay attention to the first part, I didn't catch up with the later part of the listening." Similarly, he wrote that "I think I am weak at auditory monitoring and listening comprehension. I know that I sometimes missed the key words when listening".

These show the participant's heightened awareness of his own listening problems. His solutions for dealing with such problems were also mentioned: "I pay attention to the linking words. Then I think about what I heard again and decide whether my answer was appropriate or not. Sometimes, it works and sometimes it doesn't." Towards the end of the study period, the participant observed that "after learning the listening strategies, I understand about listening in general. I need to pay attention to self-monitoring and to comprehending the concepts in order to develop my listening skills." These self-reflections provide clear evidence that the learner became more aware of the listening process and employed additional listening strategies when carrying out listening activities following the seven days of strategy training.

The Developmental Patterns of Listening Performance

The raw data and the smoothed trajectory of the learner's listening performance is illustrated in Figure 3. As can be observed, the learner's developmental trajectory for listening performance was characterized by dynamic patterns. The smoothed curve for listening performance produced using a polynomial trendline (3rd degree) shows non-linear development in the learner's listening performance. The analysis demonstrates that the participant's listening ability prior to the strategy instruction (T1) was relatively low. In the initial stage after the strategy training (T2), there was a remarkable increase in the learner's listening performance, implying a positive impact of strategy instruction. However, the later stages of the training show that the learner's listening performance experienced a sudden decline (T3) which remained stable in the following stage (T4) and was lower than his prior performance measured before the strategy training. In the last two stages of the study (T5 and T6), which took place after the training, steady growth can be observed in the participant's listening performance. As a whole, the results indicate a non-linear developmental path for listening performance with periods of progress, stability, and regression. The findings of the current study are compatible with those of previous studies (Chang & Zhang, 2021; Dong, 2016), which also noted dynamic variability patterns in their participants' listening development.

Figure 3

The Listening Performance With A Polynomial Trendline (3rd degree)
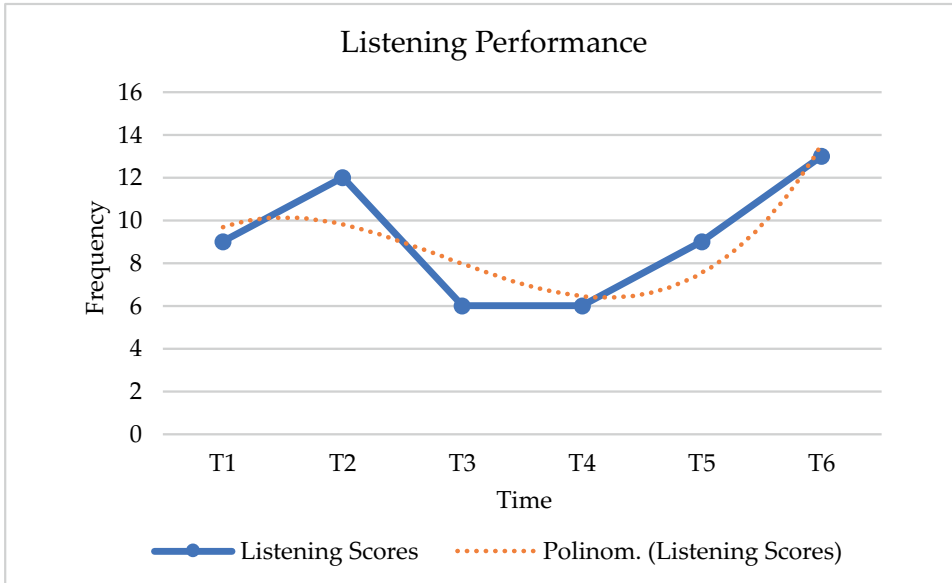
The decline in listening performance can be explained with the help of the participant's self-reflections. The analysis suggests that the learner struggled with several aspects involved in processing the listening tasks: reading the questions, focusing on the audio text, predicting the audio information, paying attention to the missing or asked information in the given task, and linking what he heard with his background knowledge to formulate answers. Furthermore, the participant had to select his strategies and monitor his strategy use during the listening process. Dong (2016) claimed that "resource competition" (p. 159; see Larsen-Freeman, 2009) took place during his participant's interaction with listening strategies in an attempt to understand the decline that he also observed in listening performance.

The Relationship Between Metacognitive Listening Strategies and Listening Performance

This section examines the dynamic relationship between the metacognitive listening strategy use and listening performance of the participant. Figure 4 presents the smoothed trajectories of both variables obtained

through exponential smoothing are illustrated in. In the first phase, from the beginning of the study period to T2 and in the last phase from T4-5, both listening strategy use and listening performance show a co-occurring increase. However, in the middle phase the listening strategies continued to steadily increase, while the listening performance decreased to a lower level. It is interesting to note that the middle phase took place during the strategy training, which may suggest that although the learner's application of metacognitive strategies increased after receiving guidance on strategy use, he may have encountered difficulties in coping with the use of strategies and processing listening tasks. This finding is partially aligned with Dong (2016), in which the participant's listening strategies and listening performance underwent three co-occurring phases in which both variables showed a gradual increase in the first phase followed by noticeable fluctuations in the second phase and ended with an upward trend in listening performance and a low-level plateau for listening strategies over a 40-week study period. Similarly, in the study conducted by Ozeki (2006), the listening strategy training resulted in overall low listening performance.

Figure 4

Smoothed Trajectories of Listening Strategy Use and Performance
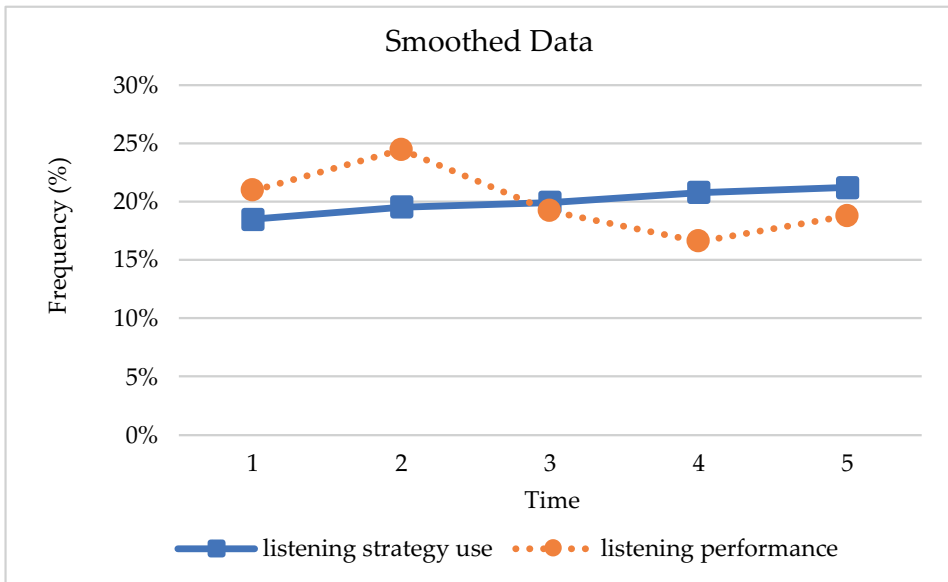
In order to gain insights regarding the temporal association between listening strategy use and performance, a moving correlation coefficient (with a window of 2 measurements) was computed, the result of which can be seen in Figure 5. The moving correlation analysis showed a U-curve in the relationship between listening strategies and listening performance throughout the study period. In the first moving window, the correlation coefficient r was just below zero, then fell gradually and reached its lowest point at the third window. In the last moving window, the trend went up again but remained below zero.

Figure 5

The Moving Correlation (window of 2 measurement) Between Strategy Use and Performance
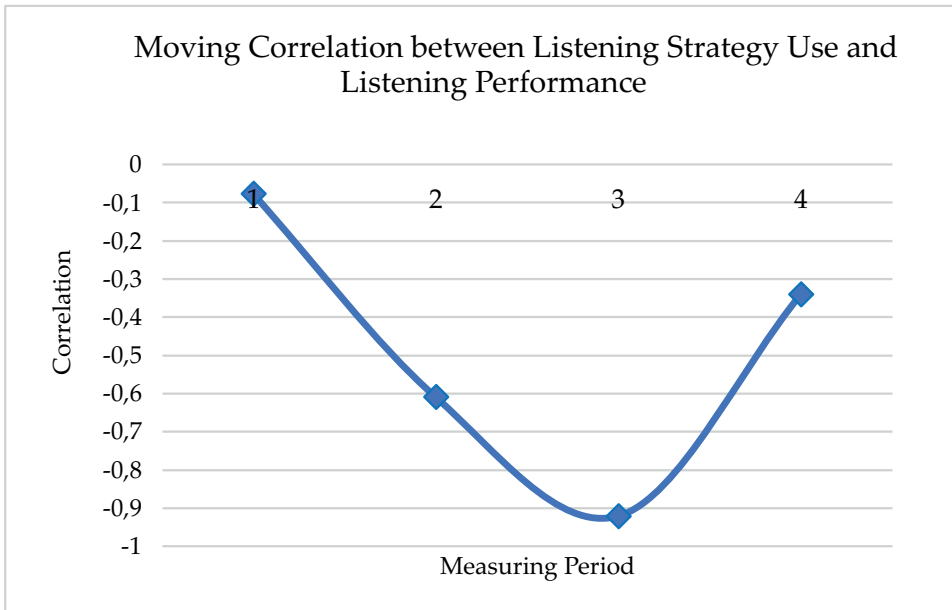
As presented in the graph above, the correlations across the study period remained below zero, which is corroborated by the statistical correlation analysis showing a (non-significant) negative correlation coefficient ($r = -.605$, $p = 0.279$). In other words, a moderate overall negative association can be observed between the use of listening strategies and listening perfor-

mance throughout the study period. One possible reason for this finding is that the emphasis placed on employing metacognitive strategies may have distracted the learner from effectively performing the listening tasks.

In summary, the findings of the present study reveal the developmental trajectories of the participant's metacognitive listening strategies characterized by noticeable diversity as well as his listening development, showing dynamic patterns and a moderate overall negative association between the use of listening strategies and listening performance over the study period. Using the CDST approach allows us to observe dynamic changes in a learner's use of listening strategies and performance as well as their dynamic relationship over time. However, the study also underscores that metacognitive listening strategy instruction developed the learners' awareness and use of listening strategies in performing listening tasks and activities such as planning for a listening task, directed attention, and evaluating strategy use and performance. Vandergrift and Goh (2012) stated that listening strategies could help learners develop their own self-regulated learning habits.

Conclusion

This small-scale study traced the dynamic developmental processes of an EFL learner's metacognitive listening strategy use and listening performance over a two-week period. With the help of CDST techniques, this study also examined how the two variables interact within a dynamic system. In general, non-linear developmental patterns were found for both the metacognitive listening strategies and listening performance of the participant across the study period, which might be due to the fact that the learner did not effectively apply the listening strategies introduced during the training. Results related to between-session variability for both variables revealed several peaks over the study period which reflected the learner's experiences mentioned in his self-reflections. The study also found that the correlation between the participant's metacognitive listening strategies use and listening performance was characterized by dynamic developmental patterns, especially during the period of strategy training. It was surprising

to find that the moving window correlation revealed a downward trend (except in the case of the last window) and a negative correlation between strategy use and listening performance during the study.

It is hoped that this study will contribute both theoretically and methodologically to research into metacognitive listening strategies from a CDST perspective. The findings of this study can enrich our understanding of the dynamics present in the development of metacognitive listening strategies and listening performance as well as in their interactions. Moreover, the results may also provide EFL teachers with a better understanding of their learners' dynamic developmental processes related to listening strategy use and performance. This can lead to improved teaching practice and strategy training, and can inform the selection of materials to cater to learners' developmental learning characteristics and requirements. However, there are limitations of the study which must be mentioned. First, the study focused on only one EFL learner, and as such it should be noted that the way learners approach listening might differ from individual to individual (Larsen-Freeman, 2009). Second, the study was conducted over a very short period due to the time constraints of the research. Third, the strategy training was carried out for only seven days although Chamot et al. (1993) claimed that strategy training implemented over an extended period of time enhances the effectiveness of students' strategy learning. Lastly, due to the time constraints and nature of a single case study, the research instruments were not piloted for this present study. Considering these limitations, generalizing the findings of this study to the development of metacognitive listening strategies and listening performance of other learners should be done with caution. To remedy this, follow-up studies could employ a greater number of participants from different contexts and could be conducted over a longer period of time using a DST approach in order to gain further insight as to whether the developmental patterns in metacognitive listening strategy use and listening performance may vary. In addition, learners' dynamic trajectories as well as variation in their listening strategy use and listening performance could be explored from a holistic view using a DST approach, such as that suggested by de Bot et al. (2007).

References

Abdalhamid, F. (2012). *Listening comprehension strategies of Arabic-speaking ESL learners*. [Unpublished master's dissertation]. Colorado State University. http://hdl.handle.net/10217/65198

Alzamil, J. (2021). Listening skills: Important but difficult to learn. *Arab World English Journal, 12*(3), 366–374. https://dx.doi.org/10.24093/awej/vol12no3.25

Anderson, N. J. (2002). The role of metacognition in second language teaching and learning. *ERIC Digest*, 1–7.

Baker, L. (2002). Metacognition in comprehension instruction. In C. C. Block & S. R. Parris (Eds.), *Comprehension instruction: Researched-based best practices* (pp. 65–79). The Guilford Press.

Bassano, D., & Van Geert, P. (2007). Modeling continuity and discontinuity in utterance length: A quantitative approach to changes, transitions and intraindividual variability in early grammatical development. *Developmental Science, 10*(5), 588–612. https://doi.org/10.1111/j.1467-7687.2007.00629.x

Cambridge English. (n.d.). *English language assessment*. https://www.cambridge-english.org/exams-and-tests/first/

Chamot, A. U. (1995). Learning strategies and listening comprehension. In D. J. Mendelsohn & J. Rubin (Eds.), *A guide for the teaching of second language listening* (pp.13–30). Dominie Press.

Chamot, A. U., Barnhardt, S., El-Dinary, P. B., Carbonaro, G., & Robbins, J. (1993). Methods for teaching learning strategies in the foreign language classroom and assessment of language skills for instruction: Final report. *ERIC Clearinghouse on Languages and Linguistics*.

Chang, P., & Zhang, L. J. (2021). A CDST perspective on variability in foreign language learners' listening development. *Frontiers in Psychology, 12*, 1–17. https://doi.org/10.3389/fpsyg.2021.601962

Cross, J. (2011). Metacognitive instruction for helping less-skilled listeners. *ELT Journal*, *65*(4), 408–416. https://doi.org/10.1093/elt/ccq073

Darti, & Asmawati, A. (2017). Analyzing students' difficulties toward listening comprehension. *Eternal*, *3*(2), 211–228. https://doi.org/10.24252/Eternal.V32.2017.A9

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

Dong, J. (2016). A dynamic systems theory approach to development of listening strategy use and listening performance. *System*, *63*, 149–165. https://doi.org/10.1016/j.system.2016.10.004

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.

Gilakjani, A. P., & Ahmadi, M. R. (2011). A study of factors affecting EFL learners' English listening comprehension and the strategies for improvement. *Journal of Language Teaching and Research*, *2*(5), 977–988. https://doi.org/10.4304/jltr.2.5.977-988

Goh, C. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, *28*(1), 55–75. https://doi.org/10.1016/S0346-251X(99)00060-3

Goh, C., & Taib, Y. (2006). Metacognitive instruction in listening for young learners. *ELT Journal*, *60*(3), 222–232. https://doi.org/10.1093/elt/ccl002

Larsen-Freeman, D. (1997). Chaos/Complexity science and second language acquisition. *Applied Linguistics*, *18*(2), 141–165. https://doi.org/10.1093/applin/18.2.141

Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, *30*(4), 579–589. https://doi.org/10.1093/applin/amp043

Lowie, W. (2013). Dynamic systems theory approaches to second language acquisition. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Lin-*

guistics (pp. 1–9). Blackwell. https://doi.org/10.1002/9781405198431. wbeal0346

Ozeki, N. (2006). *Listening strategy instruction for female EFL college students in Japan*. Macmillan.

Qiu, A., & Huang, J. (2012). The effects of dynamic image schema on ESL students' systematic improvement of listening comprehension: A dynamic system theory perspective. *International Journal of Learning and Development*, *2*(1), 241–254. https://doi.org/10.5296/ijld.v2i1.1320

Rahimi, M., & Katal, M. (2013). The impact of metacognitive instruction on EFL learners' listening comprehension and oral language proficiency. *The Journal of Teaching Language Skills*, *5*(2), 69–90. https://doi. org/10.22099/jtls.2013.1555

Schmid, M. S., Verspoor, M. H., & MacWhinney, B. (2011). Coding and extracting data. In M. H. Verspoor, K. de Bot, & W. Lowie (Eds.). *A dynamic approach to second language development* (pp. 39–54). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.29.03sch

Siegler, R. S. (2006). Microgenetic analyses of learning. In D. Kuhn, R. S. Siegler, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology* (pp. 464–510). Wiley. https://doi.org/10.1002/9780470147658.chpsy0211

van Dijk, M., & van Geert, P. (2007). Wobbles, humps and sudden jumps: A case study of continuity, discontinuity and variability in early language development. *Infant and Child Development*, *16*(1), 7–33. https:// doi.org/10.1002/icd.506

van Dijk, M., Verspoor, M., & Lowie, W. (2011). Variability and DST. In M. Verspoor, K. de Bot, & W. Lowie (Eds.), *A dynamic approach to second language development* (pp. 55–84). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.29.04van

Vandergrift, L., & Goh, C. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge. https://doi.org/10.4324/9780203843376

Vandergrift, L., Goh, C., Mareschal, C., & Tafaghodtari, M. H. (2006). The Metacognitive Awareness Listening Questionnaire (MALQ): Devel-

opment and validation. *Language Learning*, *56*(3), 431–462. https://doi.org/10.1111/j.1467-9922.2006.00373.x

Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning*, *60*(2), 470–497. https://doi.org/10.1111/j.1467-9922.2009.00559.x

van Geert, P., & van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development*, *25*(4), 340–374. https://doi.org/10.1016/S0163-6383(02)00140-6

Verspoor, M., de Bot, K., & Lowie, W. (2011). *A dynamic approach to second language development*. John Benjamins Publishing Company. https://doi.org/10.1075/lllt.29

Verspoor, M., Lowie, W., & van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *The Modern Language Journal*, *92*(2), 214–231. https://doi.org/10.1111/j.1540-4781.2008.00715.x

Wah, N. N. (2019). Teaching listening skills to English as a foreign language students through effective strategies. *International Journal of Trend in Scientific Research and Development*, *3*(6), 883–887.

Wenden, A. L. (2002). Learner development in language learning. *Applied Linguistics*, *23*(1), 32–55. https://doi.org/10.1093/applin/23.1.32

Wind, A. M. (2021). Nonlinearity and inter- and intra-individual variability in the extent of engagement in self-reflection and its role in second language writing: A multiple-case study. *System*, *103*, Article 102672. https://doi.org/10.1016/j.system.2021.102672

Wind, A. M., & Harding, L. W. (2020). Attractor states in the development of linguistic complexity in second language writing and the role of self-regulation: A longitudinal case study. In W. Lowie, M. Michel, A. Rousse-Malpat, M. Keijzer, & R. Steinkrauss (Eds.), *Usage based dynamics in second language development* (pp. 130–154). Multilingual Matters. https://doi.org/10.21832/9781788925259-009

Zhang, L. J., & Zhang, D. (2013). Thinking metacognitively about metacognition in second and foreign language learning, teaching, and research: toward a dynamic metacognitive systems perspective. *Contemporary Foreign Language Studies, 12*, 111–121.

Appendix A

Metacognitive Listening Strategy Training

Metacognitive strategies	Description	Training period
Preparing & Planning	The learner prepares in relation to the given listening task(s) and plans how to listen, thinking about which strategies he should use in order to accomplish the given task(s).	Day 1-7
Deciding when to use particular strategies	The learner thinks and makes conscious decisions about the learning process and chooses the best and most appropriate strategy in a given situation.	
Monitoring strategy use	The learner asks himself periodically whether or not he is still using those strategies as intended.	
Learning how to orchestrate various strategies	The learner coordinates, organizes, and makes associations among the various strategies available.	
Evaluating strategy use	The learner evaluates and reflects on his own use of strategies, his listening comprehension and all the previous stages	

Appendix B

Listening Model Test One

Part 1 (it is played twice in the audio)

You will hear people talking in eight different situations. For questions **1-8**, choose the best answer (**A**, **B** or **C**).

1 You hear a customer talking to a shop assistant about a coat she bought. What does she want?		
A a different kind of item.	**B** the same item but in a different size.	**C** her money back.
2 You hear a weather forecast on the radio. Tomorrow, the weather in the east of the country will be		
A stormy in the morning.	**B** sunny in the afternoon.	**C** foggy in the evening.
3 You hear an office worker talking about cycling to work. What does she enjoy most about it?		
A getting some exercise each morning.	**B** avoiding the traffic into town.	**C** thinking about the day ahead.
4 You hear a radio announcer talking about a competition for writers of short stories. The man says that one of the rules is that		
A you have to be over sixteen to enter.	**B** you can submit more than one entry.	**C** your entry must be emailed.
5 You hear a conversation about reading. The man enjoys reading books which		
A have characters that remind him of people he knows.	**B** describe situations that he finds highly amusing.	**C** are set in places that he is unlikely ever to visit.
6 You hear two people talking about watching films on the Internet. What do they agree about?		
A the advantages of buying films online.	**B** the usefulness of reading film reviews.	**C** the pleasure of watching films at home.
7 You hear a woman at an airport talking on the phone. Why did she miss her flight?		
A She was held up by traffic.	**B** There was a long queue at check-in.	**C** She went to the wrong terminal.
8 You hear a man talking about his new job. What attracted him to this job?		
A the type of work.	**B** the opportunities for promotion.	**C** the salary offered.

Part 2 (it is played once in the audio)

You will hear an expert snowboarder called Brad Mitchell talking about the sport of extreme snowboarding. For questions **9-18**, complete the sentences with a word or short phrase.

Extreme snowboarding

Brad says there are no **9** _____ to warn extreme snowboarders of dangers. Brad advises snowboarders always to follow the **10** _____ when descending. Brad always wears a **11** _____ when he goes into the mountains. According to Brad, you need a lot of **12** _____ to set off down the mountain. Brad particularly enjoys doing several **13** _____ when he is going down a slope. Brad says at first he found it difficult to do a good **14** _____ on steep slopes. Brad says you must never **15** _____ if you feel you're about to fall. Brad advises against putting your weight on your **16** _____ in a fall. Brad always carries a **17** _____ in case he is in difficulty following a fall. In the future, Brad would most like to try **18** _____ snowboarding.

Top of Form

Bottom of Form

Part 3 (it is played once in the audio)

You will hear five short extracts in which people talk about habits they find difficult to control. For questions **19-23**, choose from the list (**A-H**) the habit each person has. Use the letters only once. There are three extra letters which you do not need to use.

A eating unhealthy food	Speaker 1 **19** _____
B doing too much exercise	Speaker 2 **20** _____
C buying unnecessary items	Speaker 3 **21** _____
D watching too much television	Speaker 4 **22** _____
E spending too much time online	Speaker 5 **23** _____
F oversleeping	
G working too hard	
H arriving late for everything	

Part 4 (it is played once in the audio)

You will hear Leonie Steiner talking to an interviewer about her work as a music teacher in a school. For questions **24-30**, choose the best answer (**A**, **B** or **C**).

24. Leonie first started learning the piano		
A with a relative.	**B** at primary school.	**C** with a private teacher.
25. Leonie started giving music lessons		
A for the pleasure of seeing others learn.	**B** because she needed some extra money.	**C** to see if she was suited to teaching.
26. Leonie most likes to teach students who		
A have great natural talent at an early age.	**B** need good teaching to develop their talent.	**C** have previously been taught badly.
27. Leonie thinks that schools should		
A employ far more music teachers.	**B** buy good musical instruments.	**C** ensure that all their pupils pass music exams.
28. Leonie thinks the problem with singing in schools is that		
A many students are too embarrassed to sing.	**B** few students want to learn how to sing.	**C** singing is not often taught in them nowadays.
29. Leonie believes her success as a music teacher is a result of		
A choosing a particular age group of children to teach.	**B** the training she received as a student teacher.	**C** a natural ability to communicate with young people.
30. What decision did Leonie find difficult to make?		
A to turn down the offer of a job abroad.	**B** to refuse promotion in the school.	**C** to continue teaching when she felt tired.

Appendix C

Metacognitive Awareness Listening Interview guide

1. Before you start to listen, do have a plan in your head for how you are going to listen?
2. After listening, do you think back to how you listened, and about what you might do differently next time?
3. Do you have a goal in mind as you listen?
4. When you have trouble understanding, what strategies do usually you do to help you understand?
5. When your mind wanders, do you try to get your concentration right away? How?
6. When you have difficulty understanding what you hear, do you give up and stop listening?
7. Do you feel that listening comprehension in English is a challenge for you?
8. Do you find that listening is more difficult than reading, speaking or writing in English?
9. Do you feel nervous when you listen to English?
10. Do you translate in your head as you listen?
11. Do you translate word by word as you listen?
12. Do you translate key words as you listen?
14. Do you use the words you understand to guess the meanings of the unknown words?
15. Do you compare what you understand with what you know about the topic?
16. Do you use the general idea of the text to help you guess the meaning of the unknown words?

6

Problems of Translatability in Linguistic and Cultural Jokes: The Case of Arabic-English Translation

Duha Alsayed Ahmad
alsayedahmad@student.elte.hu

Abstract

This exploratory qualitative study aimed to carry out an in-depth investigation of the translatability of Arabic jokes into English by identifying the challenges that translators encounter when such texts. In addition, it emphasized the strategies translators used to overcome these challenges In order to achieve these aims, two translators were asked to translate eight Arabic jokes into English and were then invited to retrospectively fill in an open-ended questionnaire on the difficulties they encountered during translation and the strategies they used to overcome them. The analysis of the translations and the comments provided by the translators indicate that translators encounter linguistic problems, such as the translation of homonyms and false cognates, as well as cultural problems when translating Arabic jokes into English, which result in a loss of the humorous effect. Moreover, the use of literal translation was also found to occur frequently when translating Arabic jokes into English. Based on the research results, it can be concluded that there are certain aspects of Arabic cultural and linguistic jokes that are untranslatable into English.

Keywords: translatability, translation problems, translation strategies, cultural jokes, linguistic jokes.

Problems of Translatability in Linguistic and Cultural Jokes: The Case of Arabic-English Translation

Over the past century, linguists have been focused on theories and strategies to facilitate the translation of texts across different genres languages (e.g., Gaber, 2005; Newmark, 1988; Pedersen, 2005). With jokes being an integral part of daily life, there is a growing recognition of the need to investigate strategies that can be used in translating jokes and the problems that translators might encounter when translating them from one language to another (Zabalbeascoa, 1996).

To joke is "to deliberately intend something, but only where that 'something' is an attempt to make others laugh" (Attenborough, 1979, p. 143). When translating jokes from one language to another, translators attempt to produce an equivalent translation in the target language that is comprehensible and as humorous as the original, which can prove to be a challenging task (Saleh, 2018). As such, translating jokes is a complex process which involves the transfer of linguistic and cultural aspects of the original joke to the target language.

The problem of whether jokes are translatable or not has been addressed by several researchers and investigated across various languages (e.g., Attardo, 2002; Newmark, 1988; Popa, 2005; Raphaelson-West, 1989; Servaitė, 2005). However, little research has been done on the translatability of Arabic jokes into English. Therefore, this qualitative exploratory study aimed at examining the translatability of Arabic jokes by examining the linguistic and cultural barriers translators face when translating jokes from Arabic into English. More specifically, the study aimed at emphasizing the strategies used to translate linguistic and cultural Arabic jokes into English. In this paper, I first provide a brief, but comprehensive, overview of the constructs related to translation and jokes, followed by a summary of the empirical research which has been carried out in connection with the translatability of jokes. I then report on an exploratory qualitative study in which I attempted an in-depth investigation into the problems and strategies involved in translate Arabic jokes into English.

Review of the Literature

Translation

Scholars have defined translation in various ways over the years. Catford (1965) viewed translation as "the replacement of textual material in one language (SL) by equivalent textual material in another language (TL)" (p. 20). Similarly, Nida and Taber's (1982) definition also highlighted the goal of achieving the closest possible equivalence with the source text (p. 12). However, these definitions primarily focused on the process of achieving an equivalent target text (TT) and failed to take the intended meaning of the source text (ST) into consideration. In contrast, Newmark (1988) underscored the importance of preserving the author's intended meaning when translating a text in a process he referred to as dynamic equivalence. Recent perspectives on translation emphasize not just linguistic transfer, but also the conveyance of the cultural connotations of the ST as well. For example, Dweik and Suleiman (2013) and Wang and Sunihan (2014) argued that translation should encompass both language and cultural aspects, ensuring that both are reflected in the target language.

Translation and Culture

As highlighted above, both language and culture play an important role in the translation process. Nida (1964) defined culture as "the total beliefs and practices of a society" (p. 157), whereas Poirier (1968) concluded that culture is the sum of the experiences and social behaviors that have been acquired or conveyed across generations. Drawing from these definitions, culture can be conceptualized to include social behaviors, practices, beliefs, values, traditions, clothing, and various other aspects of a given society. Given this broad understanding of the term, the translation process involves not only a rendering of the language of the ST, but also its culture, taking into account factors such as cultural traditions, behaviors, beliefs, and values of the TT and ST cultures.

The considerations above underscore the intricate relationships between language, culture, and translation. Language is transmitted through culture, and vice versa, establishing a strong relationship between the two

(Janfaza et al., 2012). Echoing this sentiment, Newmark (1988) posited that the language of a nation or society is influenced by its culture, and that the relationship between the two is characterized by complexity. Given this intricate relationship, translators must be aware of both the linguistic and cultural references of the text. In this view, translation serves as a tool for 'intercultural interchange' (Almubark, 2017, p. 1). This suggests that in order for translators to successfully transfer meaning from the ST to TT language, they have to be bicultural/multicultural in addition to being bilingual/multilingual (Bassnett & Lefevere, 1990; Chiaro, 2004; Leppihalme, 1997).

Humor and Jokes in Translation

Humor is defined by The Online Oxford English Dictionary (2019) as "the quality of being amusing or comic, especially as expressed in literature or speech", while Popa (2005) views it as "the ability to appreciate the situations when wordplay is funny or amusing" (p. 54). Jokes, as a form of humor and an example of the above-mentioned wordplay, are defined by Wilson (1979) as "any stimulation that evokes amusement and that is experienced as being funny" (p. 2). Jokes often employ double meanings, metaphors, ambiguity, symbolic meaning, or absurdity to evoke humor, and usually target individuals, social groups, norms, or beliefs (Mahdjoubi & Djafour, 2015). Essentially, a joke can be considered a written or spoken utterance with the goal of amusing the audience.

Jokes have been categorized in several ways. For example, Raphaelson-West (1989) and Schmitz (2002) classified jokes into three categories: universal jokes, linguistic jokes, and cultural jokes. Universal jokes, or reality-based jokes, are those which rely on a universal sense of humor and remain funny even when they are translated into another language (Schmitz, 2002). Linguistic jokes, on the other hand, are language-based and employ linguistic means to produce a humorous effect. Linguistic jokes can also be broken down into several types. According to Servaitė (2005) and Mahdjoubi and Djafour (2015), linguistic jokes can be divided into syntactic, phonological, morphological, and lexico-semantic jokes. In the first type, that is, syntactic jokes, the context is what makes the joke humorous. In phonological jokes, the humor is derived from ambiguity. This ambiguity is

due to a phonetic nuance in the word or expression used, such phonetically similar words or expressions which carry completely different meanings; such jokes are often deemed untranslatable (Laurian, 1992). Similarly, in morphological jokes, the ambiguity is caused by the morphological structure of the joke, that is, the expression used may possess a different meaning from the intended one. Finally, lexico-semantic jokes are consisting of combined features, relying on linguistic phenomena such as synonymy, antonymy, polysemy, and homonymy to create a humorous effect. According to Zabalbeascoa (1996), translating such jokes is particularly challenging. Whereas linguistic jokes are language-based, cultural jokes are related to the beliefs, values, or norms of a specific nation or culture (Tisgam, 2009). These may include jokes which have religious or ethnic references.

Translation Problems: The (Un)Translatability of Jokes

Newmark defined the term "translation problem" as "a stretch of text of any length which is not readily amenable to literal or word for word translation" (1993, p. 2). Ghazala (2008) similarly considers any difficulty that hinders the translation of a text as a translation problem. Translating jokes is a particularly challenging task for translators due to the difficulty in rendering the cultural and linguistic elements of the joke from the SL to the TL without losing the sense of humor conveyed by the original joke (Ghussain & Al-Latif, 2003; Mahdjoubi & Djafour, 2015). A major concern in joke translation is the issue of untranslatability. Translatability, as Pym and Turk (1998) defined it, is "the capacity for some kind of meaning to be transferred from one language to another without undergoing radical change" (p. 273) When a translator is unable to transfer the language and culture of the ST into that of the TT, the ST is said to be untranslatable. Bahameed (2008) suggests that untranslatability emerges from a lack of intercultural equivalence (p. 5); in other words, untranslatability occurs due to the vast linguistic or cultural differences between the languages of the ST and TT.

The distinction between linguistic and cultural translation problems is applied to humor by Catford (1965), who posited that there are jokes which cannot be translated, citing two main barriers: linguistic and cultural untranslatability. These two barriers correspond to Abdellah's (2002)

classification of translation problems (i.e., linguistic and cultural problems). Linguistic untranslatability arises when "functionally relevant features" of the ST language do not have equivalent linguistics features in the target language, which causes ambiguity or misunderstanding of the ST (p. 98). On the other hand, cultural untranslatability arises when "a situational feature, functionally relevant for the SL text, is completely absent from the culture of which the TL is a part" (p. 99). According to the hypotheses of many linguists (e.g., Abdalkareem et al., 2021; Ghazala, 1995; Nord, 1991), culture-specific items or culture-bound references are the most difficult to translate. Consequently, it can be hypothesized that jokes which have culture-bound items will be particularly challenging and may even be untranslatable (Bekkai, 2010; Low, 2011; Tisgam, 2009).

Strategies for Translating Linguistic and Cultural Jokes
Translating jokes requires a different approach from the translation of other genres, with translators having the liberty to modify, omit, or even completely replace elements from the original joke to preserve the humorous effect in the TT language (Alkadi, 2010). Various strategies have been presented for translating different types of jokes. In regard to linguistic jokes, Panek (2009) identified six strategies for translating humor, which include literal translation, condensation, name deletion, substitution, transposition, and paraphrasing, while Newmark (1988) introduced a three-step strategy to translate jokes that depend on the double meaning of a word. Firstly, the translator seeks a TL word that has the same two meanings of the ST word. If this method has failed, then they should either give up one of the two intended meanings or use two or more lexical items to express the two intended meanings. Finally, if neither method proves fruitful, the translator should try to identify a synonymous word which approximates the dual-meanings of the ST word.

Although linguists and translation theorists may have different translation approaches and techniques, they generally agree that it is crucial to take cultural connotations into account when translating texts. To ensure that the target audience can relate to the original text, the transfer of cultural connotations is equally as important as language transfer (Newmark, 1988;

Nida, 1964; Venuti, 2001). For instance, Nida's (1964) and Newmark's (1988) strategies of dynamic and cultural equivalence emphasized the significance of conveying the cultural context of the ST in order to enable the target audience to comprehend the text in a manner that is comparable to the original audience and to be able to produce an equivalent effect on the TT audience.

Different strategies have been identified for translating culture-bound jokes, which contain cultural expressions, concepts, or idioms. Translating such jokes requires transferring the cultural references of the ST to the TT culture in order to preserve the humorous effect. As mentioned above, culture-bound jokes may prove particularly difficult to translate. Tisgam (2009) suggests that translators attempt to retain the cultural context of the ST and convey the humorous aspect of the joke to the target language. If the joke loses its humorous effect in the translation, the translator may opt for creating a TL cultural context in which the joke can be understood rather than transferring the cultural context of the ST. Popa's (2004) strategy for translating challenging culture-bound jokes involves the identifying culture-bound expressions and looking for dynamic equivalents in the target culture, that is, equivalent forms that are very close to the SL forms and which evoke the same effect on the TT readers (Nida, 1964). If there is no dynamic equivalent, the translator needs to decide whether to translate the culture-bound expression literally or to explain it by adding extra information to the joke or by using footnotes.

However, while using footnotes might prove helpful in eliminating ambiguity and explaining culture-specific terms, the excessive use of footnotes might detract from the humor of the joke. The intended humorous interpretation of the joke may not be immediately apprehended by the target audience due to the need to process additional information related to the explanation of the joke. As such, translators are advised to avoid using footnotes whenever possible. Furthermore, omission is also viewed negatively as a solution for translation problems, as it can be considered as a sign of lack of effort or expertise on the part of the translator (Leppihalme, 1994).

Pedersen (2005) also offers comparable strategies, highlighting techniques for rendering what Pedersen refers to as "extralinguistic culture-bound references" (p. 2):

1. Using culturally or functionally equivalent expressions.
2. Glossing, or what Pedersen calls specification, which involves using footnotes or adding words to explain an ambiguous cultural concept.
3. Borrowing words from the source text. This strategy is called "retention" by Pedersen.
4. Substitution of the SL cultural item with a different TL item or paraphrasing it using words or expressions from the TL.
5. Direct translation, or literal translation, is translating the ST culture-bound expressions without altering them or adding extra information.
6. Generalization implies replacing a specific SL culture-bound item with a more general one from the TT language.
7. Omission (i.e., deleting the ST culture-bound item).

Linguistic and Cultural Considerations in Translation Between Arabic and English

Translation between Arabic and English, originating from two distinct language families and representing vastly different cultures, present distinct challenges to translators (Bahameed, 2008; Sekhri, 2016). On the one hand, Arabic is a Semitic language that is highly inflection and offers flexibility in its sentence structure (Baffaqeeh, 2009), allowing for both nominal and verbal sentences (i.e., "subject + verb + object" and "verb + subject + object", respectively). English, on the other hand, is a West Germanic Anglo-Saxon language, which is markedly less inflectional and possesses a more rigid syntactical structure, predominantly adhering to a "subject + verb + object" order (Ghazala, 1995). These differences, which extend to features such as pronouns, gender, clauses, and number, all present a challenge when translating jokes between the two languages (Al-Sohbani & Muthanna, 2013).

As put forward by Mahdjoubi and Djafour (2015) translators encounter three main problems when translating cultural jokes between Arabic and English. First, literal translations of cultural jokes might result in the loss of the humorous effect. Secondly, jokes may contain taboo subjects (e.g., sex or religion) which might prove to be offensive to the TT audience. Many translators tend to omit such taboo elements, which might also weaken the humorous effect of the joke. Lastly, it can be difficult to transfer stereotypes held by the ST culture to the TT.

Aim of the Study

Given the issues highlighted in the literature review, there are a number of challenges associated with the translation of jokes from Arabic into English. However, previous studies have not provided sufficient evidence for the untranslatability of Arabic jokes which are markedly language- or culture-specific. Hence, the current study attempts to identify the problems encountered by the translators who participated in this study when translating cultural and linguistic jokes from Arabic into English and determine the degree of translatability of these jokes. Moreover, it aims to highlight the strategies used to translate these jokes from Arabic to English. The research questions to be answered in this study are as follows:

1. To what degree are cultural and linguistic jokes translatable from Arabic into English?
2. What problems did the translators participating in the study encounter when translating these jokes from Arabic into English?
3. What were the strategies the translators used when rendering cultural/linguistic jokes from Arabic into English?

Research Design and Methods

The Translators

In order to translate the Arabic jokes into English, two translators were approached and asked to participate in the study. The translators were selected based on their availability and willingness to participate in the research. Moreover, the researcher ensured that each had a minimum of two years' experience in translating texts from Arabic into English. The first translator was a 25-year-old female student who obtained her master's degree from a reputable university in the UK and who completed her bachelor's degree at a large university in Hungary, majoring in English and American Studies. The translator's native language was Arabic, and she was fluent in English with more than 15 years of English language learning experience. Moreover, she had three years of experience in translating different types of texts from Arabic into English and vice versa.

The second translator was a 25-year-old female translator who finished her BA in Translation Studies and her MA in English and American Studies from a large university in Hungary. Similarly, the translator's native language was Arabic, and she was fluent in English with more than 15 years of English language learning experience. In addition, she had been working as a translator for four years at the time of the study. Given their qualifications, these two translators were chosen to participate in the translation of the eight Arabic jokes that were included in this research.

Data Collection and Analysis

In order to answer the abovementioned research questions, an exploratory qualitative study was conducted. The rationale behind using this research method was its suitability for investigating relatively unexplored phenomena, as indicated by Dörnyei (2007): "If very little is known about a phenomenon, the detailed study of a few cases is particularly appropriate because it does not rely on previous literature or prior empirical findings" (p. 30). The data for this qualitative study came from a mini-corpus composed of eight Arabic jokes along with their English translations. Due to length constraints, the present paper discusses only eight jokes. These jokes were collected from several websites that are popular among Arabs based on their relatively high number of subscribers. The jokes are contemporary were still circulating among Arabic language speakers at the time when the study was conducted (see Appendix A for references). These jokes were selected based on their type (e.g., linguistic jokes and cultural jokes) and the various elements they possess which might cause translation difficulties such as homonymy, false cognates, and culture-specific phrases. The analyzed jokes were autonomous texts and were not part of any larger works.

After translating the eight Arabic jokes, the translators were asked to detail the difficulties they encountered when translating these jokes in a retrospective open-ended questionnaire. In addition, in the retrospective questionnaire, translators were asked to point out which translation strategies they used in translating each joke.

For data analysis, Nida's (1969) and Hatim and Munday's (2019) text analysis method was used in order to extract meaning from the jokes collect-

ed and to draw conclusions relevant conclusions. This method of analysis in translation studies involves the analysis of the linguistic, grammatical, and communicative features and functions of the source text which will be transferred and restructured to form a translation in the TL. First, the translations for each joke provided by the co-translators were compared to identify any differences between them. Interestingly, both translators provided almost the same translations for the eight jokes using the same translation techniques. Next, the author referred to the retrospective open-ended questionnaire to gain information about the translation problems and difficulties that both translators encountered while translating each joke. Each joke was categorized according to the previously discussed joke types, and the translations of these eight jokes were analyzed in terms of their linguistic and cultural problems. For identifying translation problems and joke types, Servaitė's (2005) classification of linguistic jokes and Schmitz's (2002) classification of jokes and Abdellah's (2002) and Nord's (1991) classifications of translation problems were used as guiding frameworks when analyzing the eight jokes.

Subsequently, the translation strategies used in translating each joke as indicated by the translators in the retrospective open-ended questionnaire were analyzed according to the existing literature on translation strategies. To this end, I referred back to Panek's (2009) techniques for translating humor, Newmark's (1988) strategies for rendering linguistic jokes, and Popa's (2004) strategies for translating culture-specific jokes. In addition, Gaber's (2005) and Pedersen's (2005) strategies were used for analyzing cultural references within cultural jokes. The to refer to these particular translation strategies and approaches was due to their compatibility with the strategies employed by the translators.

Results

The Analysis of Arabic Jokes

In this section, the problems encountered by the translators when translating the jokes are identified and analyzed in depth, and the strategies they used in translating these jokes are highlighted. The section concludes with a summary of the most significant outcomes of the analysis of these Arabic jokes.

The eight Arabic jokes are classified into two main categories: linguistic jokes and cultural jokes. For each category, the jokes are listed and analyzed without following any particular order. Regarding the translation of jokes, the original joke in Arabic is provided first, followed by its literal translation (LT) in italics, and then the proposed translation (PT)provided by the translators. Footnotes added by the translators are indicated with an asterisk (*). The Arabic jokes references can be found in Appendix A. Some problematic Arabic words are provided in italics and transliterated (i.e., rendering from Arabic script into the Latin alphabet; see Appendix B).

The Translation of Arabic Linguistic Jokes

Joke 1

أستاذ طلب من طالب لبناني أن يضع رقم 14 بجملة مفيدة. قال الطالب: مبارح أنا وماما كنا 'فورتين' من الضحك.

LT: An English teacher asked a Lebanese student to use the number 'fourteen' in a meaningful sentence. The student answered: Yesterday Mum and I were laughing to tears.

PT: An English teacher asked a Lebanese student to use the number 'fourteen' in a meaningful sentence. The student answered: Yesterday Mum and I were fourteen**.

* In Lebanese dialect fourteen means laughing to tears.

According to Servaitė's (2005) classification of linguistic jokes, this is a so-called lexico-semantic joke which depends on the use of false cognates. Although such jokes are becoming increasingly popular, they have not yet been mentioned in previous research (e.g., Al-Sohbani & Muthanna, 2013; Mahdjoubi & Djafour, 2015; Tisgam, 2009). This joke is particularly difficult to translate as the two words fourteen (the English number) and fourteen (laughing to tears), which sound and look similar, are drawn from English and Arabic, respectively. As we can understand from the joke, the student manipulated the similarity of the English number *fourteen* with the

---

* In Lebanese dialect fourteen means laughing to tears.

Lebanese word فورتين *fourteen,* meaning "laughing to tears". To translate the joke, both translators used Newmark's (1988) strategy for rendering linguistic jokes that depend on double meanings. They opted for one of the meanings of the word *fourteen* and sacrificed the other meaning, that is, the number 14. Furthermore, as seen in the PT, the translators employed Panek's (2009) strategy of adding extra information represented by a footnote explaining what *fourteen* means in Arabic. However, even if the linguistic features of such a joke have been transferred successfully, the TT joke is not considered as humorous as the ST joke.

Joke 2

محشش قال لصاحبه: عندي لغز: شو الفرق بين الأسبوع والصحراء؟ ما جاوب صاحبه.
قال المحشش: الأسبوع فيه أحد, بس الصحراء ما فيها أحد!

LT: A stoned man told his friend: I have a mystery: Do you know what the difference between the week and the desert is?'. His friend did not answer. The stoned person said: there is Sunday in the week, but there is no one in the desert!

PT: A junkie told his friend: I have a mystery: Do you know what the difference between the week and the desert is? His friend did not answer. The junkie said: there is one Sunday in the week, but there is no one in the desert!

This joke also falls under the lexico-semantic category (Servaitė, 2005). As reported by the translators, the ambiguity in this joke is generated by the Arabic word *aḥad*, which has two meanings: Sunday and one (as in no one). Thus, the stoned man's mystery revolves around this homonymy. According to Zabalbeascoa (1996), in most cases jokes which are based on homonymy lose their humorous effect when they are translated. Here, the translators decided to sacrifice one of the two meanings of the word *aḥad* following Newmark's (1988) translation strategies for jokes that feature a word with double meanings as they could not find one word which carries the two meanings of the Arabic word *aḥad*. As a result, the joke is not easy to understand and is likely not funny for the TT audience.

Joke 3

دخل بخيل على حمام عمومي ولم يخرج, لأنه كتب على الباب من الخلف 'ادفع'.

LT: A stingy person entered a public toilet and did not come out, because it is written on the door from the inside «push »!

PT: A stingy person entered a public toilet and did not come out, because it is written on the door from the inside «push »!

This is another example of a lexico-semantic joke which exploits the ambiguity of homonyms (Servaitė, 2005). In Arabic, the words *pay* and *push* can both be expressed using the word *edfa'* (i.e., ادفع). The stingy man in the joke, not wanting to pay for using the toilet, mistakes the intended meaning of the word *edfa'* (i.e., *push*) with its other meaning (i.e., *pay*). The linguistic untranslatability of this joke is the result of the ambiguity arising from the use of homonyms (Al-Sohbani & Muthanna, 2013). Both translators employed Panek's (2009) literal translation strategy in rendering this joke from Arabic into English. Moreover, they chose Newmark's (1988) translation strategy to deal with the double meaning of the word *edfa'*,sacrificing one of the two meanings of the Arabic word *edfa'*, as there is no word in English that means both push and pay. As a result, the TT joke loses its humorous effect.

Joke 4

<div dir="rtl">مرة واحد حب، طحنوه.</div>

LT: *Once a guy fell in love, they grind him.*
PT: Once a guy fell in love, they grind him.

This example also falls under the category of lexico-semantic jokes (Servaitė, 2005), and similarly presents translation difficulties as it depends on homonymy (Zabalbeascoa, 1996). The Arabic word حب , which sounds like *ḥab*, can mean either *fell in love* or *grain*. When the audience hears or reads 'once a guy' followed by *ḥab*, they would normally expect *ḥab* to mean *fell in love*. However, when the word *ḥab* is followed by the word طحنوه / *taḥanūh*, which literally means *they grind him*, the audience is forced to reinterpret the meaning as *grain*. Thus, the joke can be understood as *once a guy was a grain, they grind him*. The translators chose Newmark's (1988) strategy of rendering linguistic jokes that rely on words with double-meaning, sacrificing one of the two meanings of the word *ḥab*. Consequently, the joke loses

its humorous effect and becomes incomprehensible. This is also interesting from a pragmatic point of view, as the interpretation of the word *ḥab* depends on the context in which it is used and on the intended meaning of the speaker or writer.

The Translation of Arabic Cultural Jokes

Joke 5

واحد واثنان وثلاثة دخلوا على مطعم وأكلوا مقلوبة, أصبحوا ثلاثة, اثنان, واحد!

LT: One, Two, Three entered a restaurant and they ate Maqlūba, they became Three, Two, One!

PT: One, Two, Three entered a restaurant and they ordered Maqlūba, they became Three, Two, One!

Understanding the joke above requires the receiver to be familiar with the Palestinian dish *Maqlūba*; therefore, according to Schmitz (2002) joke taxonomy, it is an example of a cultural joke. Maqlūba literally means *upside down*, deriving its name from the action of flipping the pot upside down after the rice inside has finished cooking. In the joke, the names of the three figures (i.e., one, two, and three) become reversed after eating Maqlūba, referencing the upside-down serving style of the dish. The translators used Gaber's (2005) and Pedersen's (2005) borrowing strategy to render the cultural item in the ST translation, that is, they borrowed the word Maqlūba from Arabic and used it in the TT joke without modification. However, without previous knowledge about Maqlūba and the way it is served, people from other cultures who are unfamiliar with this dish will not be able to make sense of the joke. This is in line with Tisgam's (2009) and Low's (2011) claim that translating jokes with culture-bound items and references is a challenging task for translators.

Joke 6

أراد أستاذ تربية اسلامية في المدرسة أن يتعرف على أسماء طلبته, فشرع بسؤالهم واحدا تلو الاخر: ما أسماؤكم؟ فأجابت الطالبة الأولى: أنا مريم, فقال لها: مريم, اقرئي سورة مريم, قال التالي: أنا يوسف, فقال له الأستاذ: اقرأ سورة يوسف. وأجاب الطالب الثالث: أنا محمد, فقال له

الأستاذ: اقرأ سورة محمد. نظر الأستاذ بعدها الى طالب
خائف وسأله عن اسمه, فأجاب الطالب: أنا فاتحة!

LT: A new Islamic Education teacher in the school wanted to know all his students' names. He asked them one-by-one: What is your name? The first student replied: My name is Mariam. The teacher asked her to read Mariam chapter from the Quran. The next student said: My name is Yusuf. The teacher asked him to read Yusuf chapter from the Quran. The next said: My name is Muhammed. The teacher asked him to read Muhammed chapter from the Quran. The teacher saw an anxious student sitting behind, he asked him: What's your name? The student replied: My name is Fātiḥa!

PT: There was a new Islamic Education teacher in the school, and he wanted to know all his students' names. He asked and they began to respond. The first said: My name is Mariam. The teacher asked her to read Mariam chapter from the Quran. The next student said: My name is Yusuf. The teacher asked him to read Yusuf chapter from the Quran. The next said: My name is Muhammed. The teacher asked him to read Muhammed chapter from the Quran. The teacher saw an anxious student sitting behind, he asked him: What's your name? The student replied: My name is Fātiḥa!

To understand this joke, the reader/listener must have knowledge about the Quran, rooting the joke in the Islamic context. Therefore, the humor fits neatly into the category of cultural jokes defined by Schmitz (2002). In the joke, the names of the first three students who introduced themselves to the teacher were similar to the names of three long chapters of the Quran. The teacher asked each student to read the chapter that has the chapter bearing their name. This was stressful for the last student, most likely because he also carries the name of a long Quran chapter which he might have difficulties reading. Therefore, he decided to replace his name with the name of a very short chapter of the Quran, entitled *Al-Fātiḥa*. The humorous aspect of the joke lies in the fact that the name *Fātiḥa* is used only as a female name among Arabs, and thus it sounds odd and funny that the name of the last (male) student is *Fātiḥa*. The joke reflects a cultural phenomenon that is common in Muslim countries, that is, borrowing proper names from the Quran. If the TT audience is not familiar with this

norm or the knowledge that *Fātiḥa* is a woman's name, the humor will be lost. According to the translators, the translation of this joke is problematic as it contains culture-bound references (Low, 2011; Tisgam, 2009). As such, they provided a translation that was very similar to the literal translation of the joke as suggested by Panek (2009) and Popa (2004), since it is difficult to add extra information or a footnote to the TT joke to explain all the cultural references that can be found in the ST joke. However, the use of literal translation has resulted the humorous effect of the joke being lost.

Joke 7

واحد مسمي أولاده: أسد, نمر, فهد, وذيب. سألوه ليش, قال: علشان لما أناديهم اقلهم مرة وحدة تعالوا يا حيوانات!

LT: A man named his kids: lion, tiger, leopard and wolf. They asked him why, he said: So when I call them, I can tell them at once: Come here, animals!

PT: A man named his kids: lion, tiger, leopard and wolf. They asked him why, he said: So when I call them, I can tell them at once: Come here, animals!

Based on Schmitz's (2002) typology, the example above would be classified as a cultural joke, as it represents a norm observed in Arab societies. In the Arab world, it is not unusual to encounter individuals named after animals, such as eagle, wolf, leopard, tiger, or even lion. Such names are normally given to males and are chosen based on special qualities associated with the animals, such as bravery and intelligence. The translators reported that the difficulty in translating this joke lied in the decision to retain the proper names from the ST joke or to translate them literally. In this cultural joke, the translators literally translated the children's names (which are the names of animals) from Arabic into English in order to retain the humorous effect of the joke (Panek, 2009; Popa, 2004). However, this approach fails to transfer the cultural significance of these names into the TT translation. This problem was addressed by Mahdjoubi and Djafour (2015), who suggested that the translator should indicate that these names are not absurd but actually do exist in Arabic cultures, otherwise the TT audience might not understand that this joke is grounded in reality. The translators

explained that they did not want to overburden the TT joke, which consists of only two lines, by adding additional information or a footnote that might be longer than the joke itself in order to clarify its cultural references.

Joke 8

اراد معلم جديد في المدرسة أن يعرف أسماء طلابه فبدأ يسألهم وهم يجيبون. سأل الأول: ما اسمك؟ أجاب الطالب: أنا محمد. وسأل الثانية: ما اسمك؟ أجابت: أنا نورة. سأل الثالث: ما اسمك؟ أجاب: أنا آسم. نظر اليه المعلم وقال مستغربا: هذا ليس اسما, أنت اسمك قاسم. فقال الطالب: هذا ما قلته لك. وقد كان الطالب الثالث من مدينة فاس المغربية. ظل المعلم يضرب الولد حتى نطق اسمه صحيحا, بعدها سأل الولد الذي كان خائفا عن اسمه فأجاب: اسمي قحمد.

LT: A new teacher in the school wanted to know the names of his students, so he started asking them about their names and they started answering. He asked the first student: What is your name? The first student replied: My name is Mohammad. He asked the second student: What is your name? The second student replied: My name is Noura. He asked the third student: What is your name? The third student replied: My name is Aasim. The teacher looked at him suspiciously and said: That is not your name! Your name is Qasim not Aasim. The student said: This is what I said! But the student was from the Moroccan city, Fez. The teacher became mad and he started to hit the boy until he pronounced his name correctly. The teacher looked at the next student who was frightened, and he asked about his name. The student said: My name is Qahmad!

PT: A new teacher in the school wanted to know the names of his students, so he started asking the first student: what is your name? the first student replied: My name is Mohammad. The second student replied: My name is Noura. The third student, who was from the Moroccan city- Fez where they pronounce the letter 'q' as 'a' replied: My name is Aasim. The teacher said: That is not your name! your name is Qasim not Aasim. The student said: this is what I said. The teacher became mad, and he started to hit the boy until he pronounced his name correctly. The teacher looked at the next student who was frightened, and he asked about his name. The student said: My name is Qahmad!

To understand this joke and to find it funny, the receiver must be familiar with two of its elements: common proper names used by Arabs and the

different dialects of Arabic. As such, it is a cultural joke based on Schmitz's (2002) taxonomy and can be considered culturally untranslatable as it is loaded with cultural references (Low, 2011; Tisgam, 2009).

Similar to English, the Arabic language has many different dialects that vary from one region to another. In this joke, the student was from the Moroccan city of Fez, where speakers replace the letter 'q' with 'a' while speaking. Therefore, the student mispronounced his name from the teacher's perspective, so the teacher hit the student to encourage him to learn the correct pronunciation. This frightened the next student, who was apparently did not understand what was wrong with his friend's name and why the teacher insisted on him replacing 'a' with 'q'. In order not to be hit by the teacher, he replaced the 'a' in his name with 'q' . The humor derives from student's correct name being *Ahmad,* with 'a', a common male name among Arabs and other Muslims, whereas *Qahmad* is not. The translator translated this joke using Gaber's (2005) and Pedersen's (2005) glossing strategy, adding extra information to explain the cultural references of the SL joke, (i.e., explaining that people in Fez pronounce the letter 'q' as 'a'). However, adding extra information to the joke might result in diminishing its humorous effect (Leppihalme, 1994).

Discussion

The findings gained from the in-depth analysis of the proposed translations of the eight Arabic jokes above highlight the significant challenges associated with translating Arabic jokes into English. These difficulties suggest that such jokes may be deemed untranslatable due to the potential loss of their intended humorous effect during the translation process. This aligns with the conclusions drawn by, Alharthi (2016), Mahdjoubi and Djafour (2015), and Saleh (2018). This conclusion is also partially supported by El-Yasin (1997), who concluded that Arabic cultural jokes pose minor problems when translated into English, whereas linguistic jokes are often untranslatable. In contrast, the finding of this research contradicts Newman's (1988) and Popa's (2005) hypotheses that all jokes are translatable.

In connection with the second research question, which aimed to examine the translation problems encountered by translators when translating jokes of this nature, both translators reported facing numerous linguistic and cultural translation issues. The most frequent problem encountered by the translators when translating Arabic linguistic jokes was the translation of homonyms, which reinforces Zabalbeascoa's (1996) claim that linguistic jokes which contain homonyms are particularly difficult to translate. In addition, based on the analysis of these jokes, it can also be argued that phonological jokes are untranslatable, a finding which resonates with the findings of Laurian (1992), who concluded that phonological jokes are untranslatable. The most interesting linguistic problem pointed out by the translators when translating the Arabic jokes was the difficulty of translating false cognates. In reference to the analysis of Joke 1, jokes which depend on false cognates to produce a humorous effect appear to be untranslatable. Surprisingly, this problem has not been addressed in previous studies researching the translation of jokes. This calls for further investigation into this particular issue in future research. As noted in the literature, all of these linguistic problems result from the significant difference between the linguistic systems of Arabic and English (e.g., Al-Sohbani &Muthanna, 2013; Ghussain & Al-Latif, 2003).

Regarding the translation of cultural jokes, the most frequent problems reported by the translators were the translation of proper nouns, followed by the translation of religious concepts from Arabic into English. Moreover, the obscenity of some Arabic cultural expressions and references from the TL culture made the translators' task more difficult. These findings correspond to the results from Tisgam's (2009) and Bekkai's, (2010) studies, which pointed out the difficulty of translating cultural jokes from Arabic into English. These problems can largely be attributed to be the vast differences between the cultures of English and Arabic speakers. Moreover, Arabic culture is strongly tied to the Islamic religion and affected by Islam in every aspect of life. This can be also observed on the linguistic level, as those in the Arab world often use religious expressions in their everyday communication, including in jokes.

With respect to the third research question of the current study focusing on translation strategies, one of the most frequently used strategies by the two translators in translating the eight Arabic jokes was the use of literal translation. The reason behind this choice might stem from the translators' concerns about being unfaithful to the SL joke. However, the translators noted clearly that the use of literal translation in translating Arabic jokes into English usually results in the loss of the joke's humorous effect. This finding confirms what Mahdjoubi and Djafour (2015) concluded regarding the use of literal translation in rendering cultural and linguistic jokes, as they claimed that the literal translation of jokes results in the loss of their humorous effect.

Newmark's (1988) strategy for translating jokes which depend on wordplay (i.e., sacrificing one of the two meanings of an ambiguous word) was also used frequently in rendering Arabic linguistic jokes containing homonyms and false cognates. The reason behind the frequent use of this translation strategy might be that it is very difficult, if not impossible, to find an English word which serves the same functions or reflects the same meanings as the original Arabic words. Moreover, the translators used footnotes, (or glossing, as described by Gaber, 2005) in some jokes, such as the footnote included in Joke 1 and the extra information added to Joke 8 to clarify the ambiguity caused using cultural and linguistic expressions. However, as mentioned in the literature review, the use of footnotes may not have helped the TT audience appreciating the humor of the jokes. Other strategies which were used to render cultural references in Arabic jokes are provided by Pedersen (2005) and Gaber (2005), including the use of culturally equivalent expressions, and borrowing ST cultural items to create an equivalent joke.

Conclusion

The purpose of the present study was to thoroughly examine the translatability of Arabic linguistic and cultural jokes in order to highlight the problems that translators encounter when translating these jokes into English as

well as the strategies they use to overcome these problems. The uniqueness of this study lies in the fact that, as shown in the literature review, there is a very limited number of studies which have investigated the translation of Arabic jokes into English The current investigation contributes to the existing literature on translating jokes by identifying problems that translators may encounter when translating cultural and linguistic jokes and examining which translation strategies they use in dealing with these problems. In addition, most of these studies (e.g., El-Yasin, 1997; Mahdjoubi & Djafour, 2015) analyze Arabic jokes which are outdated and no longer used by Arabs.

Based on the results of the joke analysis, it was found that the translation of Arabic linguistic and cultural jokes into English poses significant challenges. Translators of Arabic jokes might encounter linguistic problems such as the use of homonyms and false cognates. Another type of translation problems which translators of Arabic jokes into English might encounter is the rendering of cultural references from the ST joke into the TT joke. These cultural references include the rendering of proper names, culture-bound expressions, and idioms. Moreover, the sensitive nature of some concepts or items from the TL (e.g., religious concepts) might result in the untranslatability of the Arabic joke into English.

Furthermore, several translation strategies were employed by the translators in this study to translate Arabic jokes into English. The use of literal translation and Newmark's (1988) strategy for rendering linguistic jokes, in which the translator sacrifices one of the double meanings of a word, appeared to be used frequently in translating language-bound jokes from Arabic into English. Other translation strategies used by the translators included glossing, using a cultural equivalent expression, and borrowing.

Based on the obtained results, it can be concluded that many Arabic jokes are untranslatable into English due to the difficulty in rendering their linguistic and cultural references. As evidenced by the translations of the sample jokes in the present study, the translation of Arabic jokes into English often results in the loss of the comedic element, as the TT fails to fulfil the function of the ST.

Despite its contributions, the present research also has its limitations. First, since the analysis was based on only eight Arabic jokes were translated into English, the sample cannot be considered to be representative. Secondly, since the jokes were only translated by two translators, the translations provided might not be fully representative, and there might be other translations for the same jokes proposed by other translators using different translation strategies. Thus, additional research should be conducted in this area to further investigate the translatability of Arabic jokes into English using a larger number of jokes and involving more than two experienced translators. Moreover, further research should be devoted to the examination of false cognates in Arabic jokes and the challenges they cause for translators.

References

Abdalkareem, N., Mahdi, M., & Omer, H. (2021). *Exploring some problems that face translators in translating subtitling culturally expressions from English to Arabic*. Sabha University. http://41.208.72.220:8080/handle/1/2273

Abdellah, A. S. (2002). What every novice translator should know. *Translation Journal*, *6*(3). https://translationjournal.net/journal/21novice.htm

Alharthi, A. A. A. (2016). *Challenges and strategies of subtitling humour: A case study of the American sitcom Seinfeld, with particular reference to English and Arabic*. (28467715) [Doctoral dissertation, University of Salford]. Proquest. https://www.proquest.com/openview/e47745ca83939fb3a6b7b7175fd9e71d/1?pq-origsite=gscholar&cbl=2026366&diss=y

Alkadi, T. (2010). *Issues in the subtitling and dubbing of English-language films into Arabic: Problems and solutions* (326) [Doctoral dissertation]. Durham University. http://etheses.dur.ac.uk/326/

Almubark, A. A. (2017). The influencing factors of cultural knowledge in translating cultural specific concepts from Arabic into the English at Jazan University in Saudi Arabia. *International Journal of Applied Linguistics and English Literature*, *6*(2), 106–113. http://dx.doi.org/10.7575/aiac.ijalel.v.6n.2p.106

Al-Sohbani, Y., & Muthanna, A. (2013). Challenges of Arabic-English translation: The need for re-systematic curriculum and methodology reforms in Yemen. *Academic Research International*, *4*(4), 442–450. http://www.savap.org.pk/journals/ARInt./Vol.4(4)/2013(4.4-47).pdf

Attardo, S. (2002). Translation and humour: An approach based on the General Theory of Verbal Humour (GTVH). *The Translator*, *8*(2), 173–194. https://doi.org/10.1080/13556509.2002.10799131

Attenborough, F. (2014). Jokes, pranks, blondes and banter: recontextualising sexism in the British print press. *Journal of Gender Studies*, *23*(2), 137–154. https://doi.org/10.1080/09589236.2013.774269

Baffaqeeh, N. (2009). *Strategies for translating idioms and culturally-bound expressions within the human development genre* (822875) [Unpublished master's thesis]. University of Birmingham. https://www.birmingham.ac.uk/documents/college-artslaw/cels/essays/translationstudiesdiss/noorbalfaqeeh822875diss.pdf

Bahameed, A. (2008). Hindrances in Arabic-English intercultural translation. *Translation Journal*, *12*(1), 1–16. https://translationjournal.net/journal/43culture.htm

Bekkai, M. (2010). *The influence of culture on Arabic/English/Arabic translation of idioms and proverbs* [Unpublished master's thesis]. Mentouri University. http://site.iugaza.edu.ps/wamer/files/2017/02/The-Influence-of-Culture-on-ArabicEnglishArabic-Translation-of-Idioms-and-Proverbs.pdf

Catford, J. C. (1965). *A linguistic theory of translation: An essay in applied linguistics.* Oxford University Press.

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.

Dweik, B. S., & Suleiman, M. (2013). Problems encountered in translating cultural expressions from Arabic into English. *International Journal of English Linguistics*, *3*(5), 47–60. http://dx.doi.org/10.5539/ijel.v3n5p47

El-Yasin, M. (1997). The translatability of Arabic jokes into English. *Meta: Journal des traducteurs/Meta: Translators' Journal*, *42*(4), 670–676. https://doi.org/10.7202/002706ar

Gaber, J. (2005). *A textbook of translation: Concept, method, practice*. University Book House.

Ghazala, H. (1995). *Translation as problems and solutions.* ELGA Publication.

Ghussain, A., & Al-Latif, R. A. (2003). *Areas of cultural and linguistic difficulty in English-Arabic translation* (1416) [Doctoral dissertation]). Durham University. http://etheses.dur.ac.uk/1416/

Hatim, B., & Munday, J. (2019). *Translation: An advanced resource book for students*. Routledge. https://doi.org/10.4324/9780429266348

Laurian, A. M. (1992). Possible/impossible translation of jokes. *Humor: International Journal of Humor Research*, *5*(1-2), 111–127. https://doi.org/10.1515/humr.1992.5.1-2.111

Leppihalme, R. (1994). Translating allusions: When minimum change is not enough. *Target. International Journal of Translation Studies*, *6*(2), 177–193. https://doi.org/10.1075/target.6.2.04lep

Leppihalme, R. (1997). *Culture bumps: An empirical approach to the translation of allusions*. Multilingual Matters. https://doi.org/10.21832/9781800417908

Low, P. A. (2011). Translating jokes and puns. *Perspectives: Studies in Translatology*, *19*(1), 59–70. https://doi.org/10.1080/0907676X.2010.493219

Mahdjoubi, N., & Djafour, S. (2015). *Translation of jokes: The possibility of linguistic and cultural rendering* [Unpublished master's thesis]. Kasdi Merbah University.

Newmark, P. (1988). *A textbook of translation*. Prentice Hall.

Nida, E. A. (1964). *Toward a science of translation*. E. J. Brill. https://doi.org/10.1163/9789004495746

Nida, E. A. (1969). Science of translation. *Language*, *45*(3), 483–498. https://doi.org/10.2307/411434

Nida, E. A., & Taber, C. R. (1982). *The theory and practice of translation*. E. J. Brill. https://doi.org/10.1163/9789004669147

Nord, C. (1991). *Text analysis in translation: Theory, methodology, and didactic application of a model for translation-oriented text analysis*. Rodopi.

Panek, M. (2009). *Subtitling humor – The analysis of selected translation techniques in subtitling elements containing humor*. GRIN Verlag.

Pedersen, J. (2005). *How is culture rendered in subtitles?* MuTra 2005 – Challenges of Multidimensional Translation: Conference Proceedings. Stockholm, Sweden. https://www.euroconferences.info/proceedings/2005_Proceedings/2005_Pedersen_Jan.pdf

Poirier, J. (1968). *Ethnologie générale*. Gallimard.

Popa, D. E. (2004). Language and culture in joke translation. *Ovidius University Annals of Philology*, *15*, 153–159. https://www.ceeol.com/search/article-detail?id=99161

Popa, D. E. (2005). Jokes and translation. *Perspectives: Studies in Translatology*, *13*(1), 48–57. https://doi.org/10.1080/09076760508668963

Raphaelson-West, D. (1989). On the feasibility and strategies of translating humour. *Meta: Journal des traducteurs/Meta: Translators' Journal*, *34*(1), 128–141. https://doi.org/10.7202/003913ar

Pym, A., & Turk, H. (1998). Translatability. In M. Baker (Ed.), *Routledge encyclopedia of translation studies* (pp. 273–277). Routledge.

Saleh, S. (2018). *The difficulties that MA students face in translating jokes from English to Arabic in Middle East University and Petra University* [Unpublished master's thesis]. Middle East University. https://meu.edu.jo/libraryTheses/5adc5a679b349_1.pdf

Schmitz, J. R. (2002), Humor as a pedagogical tool in foreign language and translation courses. *Humor, 15*(1), 89–113. https://doi.org/10.1515/humr.2002.007

Sekhri, O. (2016). Problem of terminology of the Arabic language: The Arabic language and the translation of emotive expressions into English. *Timsal N Tamazight*, *8*, 232–251. https://doi.org/10.52360/1717-000-003-010

Servaitė, L. (2005). The anatomy of a joke. *Tiltai: humanitariniai ir socialiniai mokslai*, *4*, 81–85. https://vb.ku.lt/object/elaba:6090220/

Tisgam, K. H. (2009). Translating cultural humour: Theory and practice. *Wasit Journal for Humanities, 5*(9), 79–121. https://www.iasj.net/iasj?aId=19938&func=fulltext

Venuti, L. (2001). Strategies of translation. In M. Baker (Ed.), *Routledge encyclopedia of translation studies* (pp. 240–244). Routledge. https://rb.gy/m8nez

Wang, J., & Sunihan, (2014). An analysis of untranslatability between English and Chinese from intercultural perspective. *English Language Teaching, 7*(4), 119–125. https://doi.org/10.5539/elt.v7n4p119

Wilson, C. P. (1979). *Jokes: Form, content, use, and function*. Academic Press.

Zabalbeascoa, P. (1996). Translating jokes for dubbed television situation comedies. *The Translator, 2*(2), 235–257. https://doi.org/10.1080/1355650 9.1996.10798976

APPENDIX A

Sources of the Jokes Included in the Corpus

Joke 1: https://arabia.babycenter.com

Joke 2: http://www.yafeta.com

Joke 3: http://hrafesh.blogspot.com/2013/02/blog-post_22.html

Joke 4: https://vb.3dlat.com/showthread.php?t=270866

Joke 5: https://arabia.babycenter.com

Joke 6: https://vb.3dlat.com/showthread.php?t=211912

Joke 7: http://www.iqr3.com/vb/showthread.php?t=78077

Joke 8: http://www.iqr3.com/vb/showthread.php?t=78077

APPENDIX B

Scheme of Transliteration

ص	ش	س	ز	ر	ذ	د	خ	ح	ج	ث	ت	ب	ء
ṣ	sh	s	z	r	dh	d	kh	ḥ	j	th	t	b	'

س	و	ه	ن	م	ل	ك	ق	ف	غ	ع	ظ	ط	ض
y	w	h	n	m	l	k	q	f	gh	'	ẓ	ṭ	ḍ

Vowels

ـَ	ـُ	ـِ	اَ	ي۔ِ	وُ
a	u	i	ā	ī	ū

7

A Pragmatic Analysis of ELF Communication: A Group Discussion in an International University Context From Etic and Emic Perspectives

Árpád Farkas
farkas.arpad@btk.elte.hu

Abstract

This study was carried out with the primary aim of investigating a communicative event from multiple perspectives. Specifically, the paper reports on the findings of empirical research into the pragmatics of English as a lingua franca communication in an international university context. The participants were five postgraduate students from five different countries. The study is unique in that it presents both the researcher's etic analysis of and the participants' emic views on a speech event. The researcher's pragmatic analysis focused on interaction as well as on language use. The features of interaction identified in the data included agreement, apology, discussion moderation, off-topic interaction, questions, and silence. Additional features of the participants' language use such as discourse reflexivity, echo, hedging, and nonstandard usage were identified and analysed. The emic perspectives of the students were explored through stimulated-recall interviews, which provided insights into the participants' different perceptions of various aspects of the speech event. Overall, the results of the research indicate that a combination of etic and emic perspectives may result in an enhanced understanding of pragmatic phenomena.

Keywords: pragmatics, English as a lingua franca, emic data, stimulated recall

A Pragmatic Analysis of ELF Communication:
A Group Discussion in an International University Context
From Etic and Emic Perspectives

The English language has been instrumental in the internationalisation of education, which is especially the case for higher education, with many international students relying on English to study outside of their home countries (Graddol, 2006). The use of English in such international contexts and elsewhere has come to be known as English as a lingua franca (ELF), which can be understood "as a way of referring to communication in English between speakers with different first languages" (Seidlhofer, 2005, p. 339). Among applied linguists, the widespread use of ELF has generated considerable research interest, which started in the 1990s (e.g., Firth, 1996; Jenkins, 1998) and has continued to the present day (e.g., Illés & Bayyurt, 2024; Walkinshaw, 2022).

Research has also been conducted on the ways in which ELF is used specifically in international contexts of higher education. Based on a corpus of 1,000,000 spoken words, ELFA, which denotes English as an Academic Lingua Franca, is a pioneering research project aimed at discovering and describing features of ELF used in academic settings (Mauranen & Ranta, 2008). Some of the results of the ELFA project include descriptive findings on the characteristics of academic ELF that distinguish it from the academic language use of native English speakers, ELF lexis and structure, and ELF discourse features such as repetition and explicitness strategies (Mauranen, 2012). Another investigation with a similar focus was undertaken by Smit (2010), who carried out research into ELF classroom discourse in higher education to explore some of the pragmatic features of interaction in that context. The main findings provided insights into repair and directness/indirectness in ELF communication as well as into code switching and the negotiation of meaning. An attempt was made to incorporate the participants' emic views, which Smit obtained by conducting semi-structured interviews, though the study was longitudinal in nature (i.e., with a focus on participants' general views). A somewhat more emic approach was adopted by Kalocsai (2011), whose research examined the use of ELF in communi-

ties of practice at a Hungarian university. In the course of data collection, Kalocsai employed a range of ethnographic methods including "'playback sessions' or 'retrospective interviews'" (p. 68) with the purpose of eliciting comment on the interaction in which her participants had engaged. The playback sessions afforded Kalocsai the opportunity to become acquainted with the emic perspectives of her participants and thus arrive at a more accurate interpretation of the data at large.

Despite the valuable insights yielded by emically oriented ELF research, the focus within the profession appears to have remained largely etic (i.e., centred on the researcher's analytical perspective). Ideally, etic methods of data analysis should be used in conjunction with emic data from research participants as the combination of the two perspectives offers the potential for gaining a more comprehensive understanding of qualitative language data. Similar views were expressed by Illés (2020), who highlighted the need "to capture participants' reality using ethnography" (p. 9) and pointed out that observation on its own might fall short of this aim. Thus, it may be necessary to broaden the scope of the research methods customarily employed in the field of applied linguistics if the emic realities of individual language users are to be studied. This methodological innovation in ELF pragmatics would seem to be justified also in that it is in line with a suggestion made by Pitzl (2022), who argued that "for pragmatic research, … we need to expand the methodologies we use for the description of lingua franca encounters" (p. 63). A research method that appears to be particularly well suited to the collection of emic data is stimulated recall, which facilitates introspection through exposure to either audio or video stimulus (Ryan & Gass, 2012). Although stimulated recall has been used in education research (e.g., to compare different teachers' thought processes during decision making; Westerman, 1991), it has thus far been underutilised in research on ELF pragmatics and discourse analysis. This, then, creates a niche for empirical investigations into the use of ELF supplemented with emic data to be carried out. Therefore, the present study of ELF classroom communication was conducted with a dual focus of analysis: Interactional data were analysed by the researcher, and the analysis was complemented with the participants' emic perspectives

explored through stimulated recall. Following an overview of the theoretical considerations taken into account in the pragmatic analysis employed in this study and a description of the research methods, the results of the analysis are presented below.

Theoretical Background

The pragmatics of ELF communication has been researched extensively in recent years (e.g., Cogo & House, 2018; House, 2022; Illés, 2020; Kecskes, 2019; Taguchi & Ishihara, 2018; Walkinshaw, 2022). However, the purpose of this section is to present a brief overview of some key notions in order to introduce the pragmatic theories that informed the data analysis. Because pragmatics "explores how listeners can make inferences about what is said in order to arrive at an interpretation of the speaker's intended meaning" (Yule, 1996, p. 3), pragmatic analysis necessarily focuses on how meaning emerges from a given context of interaction. One of the ways in which interactional meaning can be studied is through examining how speakers deploy words to achieve their communicative goals. The model perhaps most closely associated with the analysis of the expression of speakers' intentions is Speech Act Theory (Tsohatzidis, 1994).

What is known today as Speech Act Theory grew out of a series of lectures delivered by John Langshaw Austin, who endeavoured to explain how speakers can perform actions through the use of language. A speech act, broadly, is to be understood as an action that is performed by making an utterance. Austin (1962) referred to sentences through which actions are taken as performative utterances and provided examples such as naming a ship or entering into matrimony, both of which are performed by uttering particular words in specific contexts. In essence, the model of speech acts that Austin devised consisted of three elements: a locutionary act, an illocutionary act, and a perlocutionary act. Of these, the locutionary act, which is frequently shortened to locution, simply consists of what is said. For example, a speaker who wishes to know the time might produce the following locution: "Can you tell me what time it is please?" When uttering these words, the speaker will have made a request for information. The speaker's intention (i.e., request for information) is the illocutionary force

of the utterance, and if it is understood in the same way as it was intended by the speaker, it will have the desired perlocutionary effect on the hearer (i.e., the hearer will tell the speaker what time it is). The same request could be made by uttering more direct (e.g., "Tell me the time") or more indirect (e.g., "Do you have a watch?") locutions, but it is the perceived illocutionary force upon which the hearer's interpretation depends: If the hearer believes that the speaker wants to know the time, then that is how the locution will be interpreted irrespective of what form the utterance takes.

However, it is important to note that speech acts are not easily distinguishable from instances of language use that are not speech acts because such a distinction could be made only if there were a finite number of speech acts. According to Flowerdew (1990), the application of Speech Act Theory may be beset by potential challenges, the first of which is that the number of speech acts remains undetermined. Thus, it is not possible to draw up a finite list of speech acts as their number may be infinite. If intention (i.e., the illocutionary force) is at the core of speech acts, then speech acts may be equated with meaning. In support of this assumption, Searle (1971) argued that meaning does not exist without communicative intention:

> In order to regard [a message] as an instance of linguistic communication one must suppose that its production is what I am calling a speech act. It is a logical presupposition, for example, of current attempts to decipher the Mayan hieroglyphs that we at least hypothesize that the marks we see on the stones were produced by beings more or less like ourselves and produced with certain kinds of intentions. If we were certain the marks were a consequence of, say, water erosion, then the question of deciphering them or even calling them hieroglyphs could not arise. (Searle, 1971, p. 40)

In Searle's conception, then, meaning is derived from a sign based on the assumption that the sign was produced with communicative intention; therefore, every utterance is a speech act because no utterance is devoid of communicative intention. What follows from this is that a speech event could be analysed as a succession of speech acts. However, such an approach

would not be practicable as the high number of speech acts would render data reduction for analysis impossible, and even the separation of one speech act from another would be unachievable, hence the difficulty to which Flowerdew (1990) alluded. What is more, such an absolutist approach to the analysis of speech acts would, in all likelihood, fall short of being representative of the experience of language users, who do not conceive of every utterance they make or hear as a speech act. To reduce this apparent discrepancy between theory and practice, it may be necessary to call upon language users whose linguistic output is analysed to augment the analysis by providing insights into their emic perspectives on different aspects of speech events. Although emic data may not directly address the issues raised by Flowerdew, the input language users can provide may prevent inappropriate interpretations from being imposed on interactional data through revealing participants' true communicative intentions.

Politeness theory is also frequently utilised in research conducted on the pragmatics of linguistic interaction (e.g., Haugh & Culpeper, 2018; Liu, 2020; Sorlin, 2017). Principally, politeness theory centres around the notion of face, which Brown and Levinson (1987) described as "the public self-image that every member wants to claim for himself" (p. 61). Used in common parlance in phrases such as "lose face" or "save face", the concept of face in politeness theory is comprised of two constituents: negative face and positive face. Negative face can be understood as a person's "need to be independent, not imposed on by others" (Yule, 1996, p. 131), whereas positive face is related to the approval that a person may receive from their interlocutors (i.e., the perception of a person by others communicated through, e.g., compliments). For example, if a speaker were to ask a favour of a hearer, it would be an imposition on the hearer, thereby posing a threat to the hearer's negative face by depriving the hearer of the freedom to act as they please (provided that they oblige and do the favour). Similarly, if criticism were levelled at the hearer, it would be seen as a threat to the hearer's positive face given that an expression of disapproval can be taken, by extension, as an indication that the person being criticised is not as well liked as previously thought (i.e., their self-image is sullied). Utterances that pose a threat to an interlocutor's face are known as face-threatening acts

(Brown & Levinson, 1987). The basis of the theory is the principle that "face is something that is emotionally invested, and that can be lost, maintained, or enhanced, and must be constantly attended to in interaction" (Brown & Levinson, 1987, p. 61). What this presupposes, then, is that interactants are by default conscious of the need to maintain their face and refrain from threatening the face of others. However, participation in a communicative event entails performing a balancing act between interaction and the preservation of everyone's face. It is the process of satisfying these conflicting demands into which insight may be gained by exploring the emic perspectives of those engaged in communication.

In addition to speech acts and politeness, conversational turns can also be a focus of analysis when interaction is studied. Originally a sociological research endeavour (Sacks et al., 1974), the study of turn-taking has since been the mainstay of conversation analysis. Studies focused on the analysis of turn-taking tend to rely on naturally occurring data obtained from instances of genuine interaction (Wilkinson & Kitzinger, 2011). One of the aspects of turn-taking that can be studied is how one turn is succeeded by another. Couples of turns, when analysed in relation to one another, are known as adjacency pairs, which "are utterances produced by two successive speakers in a way that the second utterance is identified as related to the first one as an expected follow-up to that utterance" (Paltridge, 2021, p. 100). A question and an answer would be a typical example of an adjacency pair in that the speaker's utterance necessitates a response. However, various other types of adjacency pairs can occur in conversation (e.g., apologies, offers, requests, etc.). Levinson (1983) posited that utterances made by the second speaker (i.e., in response to the initial turn of the first speaker) can be categorised either as "preferred" or as "dispreferred" second parts. In this dichotomy, the preferred second parts would be expected and thus unmarked (e.g., agreement in response to a suggestion), whereas the dispreferred second parts would be marked as they deviate from the expected, default answer (e.g., the refusal of a request rather than the acceptance of it). Adjacency pairs have implications for the structure of conversational turns more broadly: "Whereas preferred responses tend to occur without delay, and to be short and to the point, dispreferred re-

sponses are likely to be delayed and elaborated" (Wilkinson & Kitzinger, 2011, p. 26). Thus, preference in turn-taking can be established in a general sense externally, without probing into the views of those who converse. Nevertheless, added layers of understanding of interactants' actual preferences may be gained if their emic views on the exchange of turns are incorporated into the analysis.

Finally but by no means less importantly, it should be acknowledged that the background knowledge which the participants of speech events possess is likely to be of paramount significance for communication as a whole. Such knowledge may also be referred to as schema, which "can be defined as a cognitive construct, a configuration of knowledge, which we project on to events so as to bring them into alignment with familiar patterns of experience and belief" (Widdowson, 2004, p. 43). It follows from Widdowson's definition that each individual will have a unique interpretation of "events" as no two interlocutors have the same set of beliefs and background knowledge. An additional dimension of complexity is created by the need for interactants to gauge the amount of background knowledge the other(s) may have. As Widdowson (2007) explained, "a first-person party (a speaker or writer, **P1**) produces a text, which keys the second-person party (listener or reader, **P2**) into a context assumed to be shared" (p. 22). As such, meaning will be understood as it was intended only if P1's assumption about the amount of shared knowledge is correct, and the context is indeed shared by P1 and P2. Because successful communication requires the schemata of both parties to align with one another to some extent, it is to be expected that misunderstandings occur frequently and unbeknownst to the participants, though it is something of an oversimplification to label situations when P1 and P2 draw on different schemata as outright misunderstandings. What is evident, nonetheless, is that intended meaning may differ radically from how it is understood, and the schema used by P2 seems to be the key determinant in the process of deciphering the message. It would be an overambitious goal for this (and arguably any) study to account for the use of schemata by empirical means, but insights into this complex phenomenon can be gleaned by juxtaposing the emic perspectives of both P1 and P2 parties.

Methods

Data Collection

The data analysed in this study were collected in an international university context from a seminar discussion in which postgraduate students studying for an advanced degree in the humanities conversed. The topic of the lesson—and that of the conversation—was related to teacher education, which the students were discussing based on research papers that they were required to have read in preparation for the discussion. Although the university at which the students were enrolled was located in Hungary, the context of the discussion was not tied to any one location as the seminar was held online due to safety measures necessitated by the COVID-19 pandemic. In addition, none of the participants was Hungarian, and not all of them were in Hungary at the time of the lesson. The choice of this context for data collection was twofold: On the one hand, the data constitute a purposive sample (Dörnyei, 2007) in that the discussion under analysis is a specific type of communication (i.e., interaction among international students through the medium of ELF) that aligns with the aims of this study as opposed to a random sample of classroom talk. On the other hand, it ought to be acknowledged as per the suggestion made by Cohen et al. (2007) that it was also a convenience sample, inasmuch as the researcher had access to the group of students through their teacher; therefore, generalisability was neither made possible by the data nor pursued by the researcher.

The lesson that yielded the data was conducted entirely online via Microsoft Teams, which aided the researcher: The online delivery simplified the data collection as the conversation was recorded using the built-in function of the communication platform that enables users to record meetings. Furthermore, all participants were clearly visible on video and speaking into their own microphones, which greatly enhanced the quality of the data compared to what could have been collected in a physical classroom setting. Prior to the data collection, the students were informed of what their participation in the study would entail, and written consent was obtained from each of them. Not only were the participants anonymised to

prevent them from being recognised, but they were also given the opportunity to request that utterances they deemed sensitive be excluded from the study, though no such request was made. The duration of the lesson was 90 minutes, but it was not recorded in its entirety. Instead, a shorter segment in which the teacher assigned the students the task of discussing research papers they had read was recorded. The students carried out the speaking task in breakout rooms (i.e., separate channels of discussion dedicated to groupwork), and the recorded conversation from one of the two breakout rooms constituted the stimulus analysed in the present paper. The conversation lasted for 14 minutes and 17 seconds and involved only five students. Notably, neither the teacher nor the researcher was present during the discussion. The participants' background information is summarised in Table 1 below.

Table 1

Background Information About the Participants

	Student 2	Student 4	Student 5	Student 7	Student 9
Age	43	27	29	28	24
Gender	Female	Male	Female	Female	Female
First language	Mongolian	English	Persian	Spanish	Arabic

Note. The identifiers assigned to the students do not ascend in a linear fashion because the speech event does not include every student from the seminar (i.e., those who are not listed in the table were conversing in another breakout room).

Every participant spoke English as it was the medium of communication. As Table 1 shows, one of the participants was a native speaker of English, which in some respects made the data even more unique as the participation of native English speakers in academic ELF communication is understudied: Mauranen and Ranta (2008) reported that native speakers of English had produced only 5% of the data in a corpus of academic ELF that included about 650 participants (p. 200). The other participants in this study were proficient users of English based on their self-declared CEFR levels of proficiency (Council of Europe, 2020), though there was one participant who estimated their command of English to be around the B2 level. Neverthe-

less, the differences between the students' levels of proficiency in English did not seem to hinder communication.

Once recorded, the stimulus was used to probe into the participants' emic perspectives through stimulated-recall interviews (SRI). Efforts were made to minimise the length of time that would elapse between the recording of the stimulus and the SRIs because prolonged intervals pose a threat to the dependability of the SRI data (Ryan & Gass, 2012). Each participant, therefore, was invited to view the stimulus and comment on it at their earliest convenience; however, the students were not available for SRIs immediately after the lesson. Although one participant was able to take part in an SRI on the same day as the stimulus was recorded, the students waited 6 days on average (M = 5.6, SD = 3.9) before they watched the stimulus. The SRIs were also conducted via Microsoft Teams and were video recorded similarly to the stimulus. Even though the participants were occasionally asked general questions by the researcher, they were encouraged to comment on all aspects of the stimulus that they found remarkable. This open-ended approach to the SRIs was intended to facilitate genuine recall and to allow the participants' emic perspectives to be explored through their own explanations. SRI data were collected from each student, and the combined length of the SRIs was approximately 88 minutes, with an average SRI duration of 18 minutes (M = 17.6, SD = 4.0).

Data Analysis

Because producing a transcript of a speech event requires some degree of interpretation, transcription itself is considered a facet of data analysis (Bailey, 2008). In the case of the conversation (i.e., the stimulus) analysed in the present study, little interpretation went into the creation of the transcript as it was created through the use of transcription software rather than by a human transcriber. It was decided that the layers of nuance that the researcher's transcription could have added to the transcript were not required because the transcript was used in conjunction with the video footage to analyse the discussion. It was primarily the original video that was analysed, with the transcript used as an adjunctive source of information to facilitate the analysis (e.g., it was utilised when the analysis necessitated

that the conversation be searchable). The length of the transcript of the stimulus was 2,085 words. The interpretative nature of transcription was more pronounced in the case of the SRIs, which the researcher transcribed manually and selectively: There was no need to transcribe every word of the SRIs as the interviews included irrelevant information (e.g., introductory remarks made by the researcher or occasional digressions from the main purpose of the interviews) as well as redundant information (e.g., repetition). Therefore, sections of the SRIs that did not seem pertinent to the research aims were not transcribed. The rest (i.e., everything deemed relevant) was transcribed, and the combined length of the transcripts for the five SRIs amounted to 4,827 words, with an average length of just below 1,000 words per SRI (M = 965, SD = 374). In addition, field notes were also compiled by the researcher; the purpose of the notes was to enumerate the conversational turns taken by the participants and to highlight features of the speech event that would subsequently be of interest to the analyst. Thus, the creation of the field notes can also be considered as a stage of the analysis. Similarly to the transcripts, the field notes were used alongside the video footage to assist in the analysis.

Although elements of conversation analysis (Hutchby & Wooffitt, 2008) were made use of in the analysis of the stimulus data, the researcher's analytical approach more closely resembled what Canagarajah et al. (2020) referred to as an expansive interactional analysis. The analysis itself consisted in multiple viewings of the recording and the identification of the salient points of the speech event. Salience was often established on the basis of the research literature on pragmatics: Features of interaction that appeared to bear out existing theories of language use were deemed significant. In this sense, then, the analysis was deductive as it was informed by the work of others, thereby allowing links to be created between the broader context of pragmatic research and the results of the present study. However, an exclusively deductive approach to data analysis would have undermined the exploratory aims of this study, which is why the researcher decided against limiting the focus of analysis to pragmatic phenomena directly connected with previous research. In order to allow new insights to emerge, the stimulus was analysed in an open-ended manner, whereby phenomena that

appeared to be important in the context of the discussion were examined. In sum, the analysis of the stimulus was inductive and deductive in equal measure. By contrast, the SRIs were analysed in a fully inductive fashion. Given that the purpose of the interviews was to explore the participants' perspectives, the researcher did not impose a theoretical framework of any kind on the SRI data. The information from the SRIs was used to understand the emic experience of the students who took part in the speech event and to contrast it with the researcher's etic analysis. It was hoped that this combination of perspectives would give rise to an enhanced understanding of the speech event that could not have been obtained by other means.

Results

The Researcher's Perspective
The features of the speech event identified by the researcher as salient can be grouped into two broad categories: interaction and language use. Although interaction is inextricably linked with language use (i.e., interaction is a type of language use), the terms used in this study denote pragmatic differences that were observed in the data. The distinction between the categories is based on the presence or the absence of perceived interpersonal communicative intent in the utterances categorised under each feature. For example, questions were considered to be features of interaction as a student who asked a question was likely to expect their interlocutors to become involved in the discussion by providing an answer; on the other hand, hedging was seen purely as a feature of language use because a student who hedged was less likely to do so out of a desire to engage their interlocutors in conversation. It is to be acknowledged that the categorisation is rooted in the analyst's perception of the participants' communicative intentions; therefore, the outsider's etic perspective is likely to differ from the insiders' emic experience. Nevertheless, the researcher's assumptions regarding the intentions behind utterances can be deemed to be just as valid as assumptions made by any other listener. Table 2 below provides a summary of the features of the speech event based on the analysis of the stimulus data.

Table 2

Salient Features of the Speech Event Based on the Researcher's Analysis

Features of interaction	Features of language use
agreement (7)	discourse reflexivity (3)
apology (3)	echo (2)
discussion moderation (5)	hedging (3)
off-topic interaction (1)	nonstandard usage (4)
question (9)	
silence (2)	

Note. The numbers in brackets indicate the rate of occurrence of each feature in the data.

Despite the highly language-based nature of the discussion (i.e., nonverbal communication played a lesser role than in a face-to-face conversation), the features of interaction that emerged from the data were more numerous than the features of the participants' use of language. As such, the features of interaction are discussed first and in greater detail. However, as it is typical with qualitative data, the numbers do not tell the whole story. Therefore, the results are presented through a thick description, "which inevitably eats up more space" (Howitt, 2016, p. 394) than a numerically focused presentation might do, but it is hoped that a better understanding of the features can be gained by examining them in the context in which they occurred.

Silence is a case in point: On the face of it, silence does not seem to have been a prominent feature of the discussion with only two instances having been identified. Nonetheless, it stood out as one of the most salient features of the speech event from the researcher's perspective due to its conspicuousness. Out of the two instances of silence, the first one was particularly noticeable. It occurred fairly early in the speech event after Student 5 had provided a definition of teacher identity—a key term from the set text—in response to the discussion question. Then, at 1 minute and 20 seconds, the conversation ground to a halt, and silence descended upon the breakout room: No one spoke for the next 18 seconds. Eventually, Student 4 broke the silence by asking, somewhat tentatively, the following question: "Erm, so in regard… regarding teacher identity … yeah, so, what, what did the

text say about identity? Why, why is it so important?" (Speech Event, 1:38). In the interim, only Students 4 and 5 appeared to be looking directly at their screens; the others, who showed no inclination to join in, were looking away. As Student 5 had just finished her turn, the onus was on the rest of the group to continue the conversation. Considered in isolation, the episode provides few contextual clues as to why Student 4 was the one who initiated the resumption of talk, but a few indicators emerge when the segment is viewed within the broader context of the discussion. Firstly, Student 4's active involvement in the conversation at this point was consistent with the overall patterns of participation observed in the speech event. Based on the students' conversational turns, Students 4 and 5 were the most active participants with 14 and 10 contributions to the discussion respectively. Secondly, Student 4 appeared gradually to assume responsibility for the management of the talk exchange, which is discussed in more detail below as it relates to another feature of interaction.

Although the division of roles amongst the discussants was not evident at the outset, it became apparent by the end of the speech event that Student 4 had taken on the role of discussion moderator, thereby becoming the de facto coordinator of the conversation. In total, five instances of discussion moderation were identified in the data, all of which were attributable to Student 4. Variance was observable in the realisation of those utterances whose illocutionary force was identified as discussion moderation. In terms of syntax, discussion moderation was realised through both interrogative and declarative sentences. Student 4 asked questions in three out of five of these speech acts, which lent an air of indirectness and informality to his discussion moderation: Instead of instructing his groupmates to comment on the reading, he was asking open-ended questions about the text to encourage the others to contribute. However, Student 4 abandoned his indirect approach at one point and opted to nominate the groupmate whom he wanted to involve in the discussion. Having listened to a point made by Student 5 about the fluidity of teacher identity, Student 4 said, "yeah, I think that's… that kind of covers it. Erm, [Student 9], how did, how did you like this article?" (Speech Event, 3:12). Considering that Student 9 had barely spoken up to this point, the question asked by Student 4 might have been a well-meant

attempt to allow Student 9 to express her views. Good intentions notwith-standing, Student 4's act of discussion moderation posed a threat to Student 9's negative face in that it deprived her of the freedom to decide whether to speak or not. As it soon transpired, Student 9 was not prepared to speak as she had been caught off guard by the question; thus, she was forced to apologise for being unable to respond immediately, and she promised to provide a response later ("I'll get back to you", Speech Event, 3:37). The question asked by Student 4 and the response provided by Student 9 can also be analysed as an adjacency pair, wherein Student 9's inability to answer the question constituted a dispreferred second part (Levinson, 1983). Student 9 appears to have been aware of the dispreferred nature of her response given that she promised—in addition to apologising for giving an unsatisfactory answer—to respond more appropriately later on.

The differences between form and function (i.e., speech acts with a broader illocutionary force than the locution, e.g., a question, would syntactically indicate) mean that discussion moderation is a feature of the data which overlaps with other features, namely questions and an instance of off-topic interaction. In other words, questions were sometimes asked with the apparent purpose of moderating the discussion, but interrogative sentences were also uttered by participants who did not seem to have such intentions. The questions that did not produce the perlocutionary effect of discussion moderation included technical enquiries (e.g., "Is it bothersome? Can you hear anything?", Speech Event, 0:04) and questions asked in the course of off-topic interaction.

The fact that off-topic interaction was identified only once in the data is illustrative of the inadequacy of numerical measures to describe interaction in a given context of communication. What the single occurrence of the feature conceals is that the change of topic—driven by Student 4—was a pivotal moment in the speech event. Up until 12 minutes and 12 seconds, the students had been discussing their views on the set text. Then, Student 4 changed the topic by suggesting that those who were physically in Hungary at the time could meet in person, and the discussion irretrievably veered off course. From this moment on, the participants were engaged in off-topic interaction, which lasted until the very end of the breakout room session.

Although an in-depth analysis of the segment in question is outside the scope of this paper due to the richness of detail it would require, it should be noted that the change of topic appeared to enliven the atmosphere, and what the participants shared with one another at this stage of the discussion was indicative of their awareness of the broader context (i.e., the classroom context vs. the outside world). Objects outside the classroom and experiences unrelated to the set text were suddenly made relevant to the discussion by means of the participants' sharing stories with each other and asking questions about them.

In addition to questions, agreement was also a prominent feature of the interaction in terms of frequency ($N = 7$). Some instances of agreement seem to have been related to discussion moderation in the sense that all but one of the utterances expressing agreement were made by Student 4. Agreement can be seen as a precursor to discussion moderation in some cases: Student 4 often responded to what his interlocutors had said by first expressing agreement and then suggesting that they talk about something else. An example of the overlap between agreement and discussion moderation is shown in the excerpt quoted above (Speech Event, 3:12). However, not all occurrences of agreement culminated in discussion moderation. What was universal was the basic structure of the utterances that were classed as expressions of agreement. In all instances (including the one produced by Student 5), agreement was signalled by the participants' use of the word "yeah", which was typically accompanied by a follow-up reaction of some description. In some cases, the agreement was made more explicit by the follow-up; for example, it was the case when Student 4 responded to Student 7's conversational turn by saying, "yeah. I, I, I, I would agree with you" (Speech Event, 5:19). In other instances, the agreement of the speaker was emphasised by a follow-up comment on their interactant's turn with which they agreed. For instance, in response to Student 7's remarks on the harm that may be caused if research were conducted by ill-informed or unscrupulous researchers, Student 4 replied, "yeah, that could definitely be problematic, I think" (Speech Event, 10:33). The number of times agreement was identified in the data is, once again, deceptive as it may convey the impression that the participants of the speech event agreed with one

another seven times in total; however, this was not the case. Agreement was sometimes indicated only in a nonverbal fashion (e.g., by nodding). These instances of agreement were not included in the analysis as the focus was on linguistic communication only.

Aspects of the participants' language use that emerged as salient included discourse reflexivity (Mauranen, 2010), echo, hedging, and nonstandard usage. Although none of the features was so frequent as to be declared typical of the conversation, nonstandard usage—with four occurrences—was the most striking characteristic of the participants' language use. It ought to be stressed that the participants predominantly used Standard English. The other observation that must be made is that at no point did nonstandard usage compromise comprehensibility or give rise to ambiguity. Almost all of the instances in which deviation from Standard English was identified concerned the use of the indefinite article. It was inserted before the noun "research" on two occasions (by two different participants), creating the impression that the noun was used in a countable sense. Another example of nonstandard usage was the omission of the indefinite article from a position where it would have been obligatory in Standard English (i.e., before a noun phrase that consisted of an adjective and a singular countable noun). A different omission discovered in the data was that of the verb in an interrogative sentence produced in the course of the off-topic interaction described above ("How many of us in Hungary?", Speech Event, 12:57). As mentioned previously, none of these instances of nonstandard usage appeared to hinder communication in any way. Hedging and discourse reflexivity were also features of the participants' language use that occurred in the data more than a couple of times: The former was mostly realised through the participants' use of the adverb "maybe", whereas the latter manifested itself in the verbalisation of the illocutionary force behind the students' utterances (e.g., "I wanted to ask", Speech Event, 13:04).

The Participants' Perspectives

A feature of the speech event around which the participants' emic remarks coalesced was the period of silence that was also identified as salient by the researcher on account of its length. As Student 5 had been speaking before

the participants lapsed into silence, she was technically the speaker who enabled the silence when she finished her turn. In her SRI, Student 5 retrospectively revealed that she had stopped talking in order for her groupmates to have the floor: "Because in the beginning I started to talk, I wanted other participants to start talking too so that all the conversation not completed by me and let others speak" (SRI 5, 6:22). However, the other participants were reluctant to start talking immediately. Student 4, who broke the silence in the end, was unsure of what to say. As he explained, "I myself was probably not following [Student 5] here, hence the long silence after she finished" (SRI 12, 30:00). For others, it seems, the discussion had reached a juncture where they found it difficult to contribute without remembering specific information from the articles that they had read in preparation for the seminar. Student 2, for instance, was "thinking about … the differences and similarities of the three articles" (SRI 6, 9:11) to formulate a response. In addition, Student 2 felt that she did not necessarily need to speak because she had already contributed to the discussion at that point. Student 7, for whom the prospect of having to talk about the set text loomed large, was unable to concentrate as she was attempting to recall what she had read: "I just remember not even paying attention to what [Student 5] was saying. I was just gathering my thoughts about the reading because I didn't really remember" (SRI 3, 4:59). Interestingly, Student 5 was also engaged in reviewing the articles despite the fact that she would not be expected to respond to her own point: "In this period of silence, I always wanted to study the articles that the professor sent for us … I was looking for the highlighted parts" (SRI 5, 5:48). The only participant who remained largely unperturbed by the protracted period of silence was Student 9, who said that she had noticed "that it was silent, and I did look at my screen, like to see what happened—maybe my internet or something—but then I just ignored it" (SRI 8, 19:51).

As noted earlier, the resolution to the situation came courtesy of Student 4, whose intervention (quoted above as Speech Event, 1:38) provided the impetus for the conversation to resume. Despite taking the initiative to ask a question that would facilitate the discussion, Student 4 was less confident than he may have appeared to be. His principal motivation for asking the question was the avoidance of silence:

I think I was just trying to find something to… somewhere to take the conversation. I don't want to sit in silence with these guys, so I wanted to, I guess, take it somewhere. And here, I also have to admit … I didn't read this text either, actually, so I'm trying to [inaudible] in the dark, like, OK identity appears to be important in this section; I'm gonna talk about that. (SRI 12, 31:17)

Student 4's admission that he had not read the text that was being discussed helps explain the tentativeness with which he formulated his question. The lack of sufficient familiarity with the reading may also account for the open-ended nature of Student 4's question: Instead of asking his groupmates about something specific in the text, he asked in a broader sense what the text had said about identity. From the researcher's perspective, the question was seen as one that was asked with the intention of moderating the discussion; however, that initial view of the illocutionary force may need to be reassessed in the light of what Student 4 reported in his SRI. Even though the utterance does appear to have been characterised by an element of discussion moderation, Student 4 may have genuinely wanted to learn about the information he was seeking. Another reason for his decision to ask a general question rather than one addressed to a particular participant was revealed by Student 4 when he discussed how his schematic knowledge of the broader classroom context had shaped his expectations for the speech event. Student 4 had participated in several seminar discussions with the same fellow students at the time of the conversation in question; therefore, his background knowledge of the participants was a resource on which he drew to arrive at an interpretation of how the speech event was unfolding. In particular, it was Student 9's reluctance to engage in conversation that took Student 4 by surprise based on his prior experience:

Usually [Student 9]'s more vocal, and she's usually more active in the conversation, and … I was expecting maybe [Student 9] to say something 'cause usually she has this very, like, hard-worker vibe to her, so I was like, "oh she'll say something". But no, apparently, she did not, so then, after the long silence, I jumped in there to do something. (SRI 12, 32:34)

Even though the fact that no one—including Student 9—was willing to break the silence ran counter to Student 4's expectations at this point of the discussion, Student 4 seems to have continued to rely on his background knowledge of the participants when addressing or instructing his peers. For instance, Student 4's previous experiences of interaction with his group-mates are likely to have acted as a catalyst for nomination 94 seconds later, when he asked Student 9 to express her opinion on the set text (quoted above as Speech Event, 3:12). This was because Student 4 himself was unable to critique the reading owing to having been insufficiently acquainted with the text, and he believed that Student 9 would have something of substance to contribute to the discussion. As Student 4 put it, "I was maybe turning to [Student 9] as this beacon of hope" (SRI 12, 35:39).

However, Student 4's hopes of gleaning insights from Student 9 were short-lived as the latter promptly refused to speak and apologised for doing so. Ostensibly content with the response, Student 4 smiled and nodded in acknowledgement of Student 9's excuse. Nonetheless, it emerged from the SRI data that neither of the interactants had revealed their true feelings in the course of this exchange. Student 4 must have been somewhat displeased with Student 9's reply as he said that he had "expected [Student 9] to have something to say because usually she's, as I said, more vocal, more active, but … my expectations were not met" (SRI 12, 34:55). In a similar vein, Student 9 voiced her displeasure at having been nominated by her groupmate: "I don't like the fact that somebody points at me to speak; I just don't like that. If I want to say something, I'll say it, but don't tell my name and call for me" (SRI 8, 21:34). Student 9 also explained that her failure to respond in an appropriate manner to Student 4's question had been caused by her busy schedule as she was having to work (i.e., doing her day job) from home whilst participating in the seminar discussion. Although Student 4 could not have possibly known what Student 9 was doing on the side, the question he asked caused Student 9 to lose face as she had to explain in front of the whole group why she was not complying with Student 4's request.

Student 5's participation in the speech event was an aspect of the conversation that gave rise to a convergence of opinions. Student 5, who

was the second most talkative participant based on the number of her conversational turns, conveyed the impression that she was not always speaking spontaneously. Once again, the participants' schematic knowledge of the broader classroom context came to the fore: When Student 9 verbalised her views on Student 5's contributions to the conversation, she drew on her background knowledge of how their group discussions were ordinarily conducted and compared it to what had happened on this particular occasion. Student 9 summarised her impressions in the following way:

> I feel like [Student 5]'s just reading out the things that she's seeing in front of her. It's just something that I usually feel; I don't know if it's like that or not, but it's what I remember at that moment what I thought of. (SRI 8, 17:16)

It seems plausible that the view which Student 9 expressed regarding Student 5's participation was also held by the other members of the group. There is evidence provided by Student 4 which indicates that this may have been the case. When describing his impressions of his peers, Student 4 mentioned in passing what he had thought about Student 5's contributions: "[Student 5] was just reading paragraphs from the text, and so I don't think she was taking us anywhere" (SRI 12, 35:32). The observations of the two participants can be supplemented with the researcher's perception (not detailed above), which coincided with those of Students 4 and 9. In spite of having been reached independently by three different observers, the conclusion that Student 5 was reading out segments of the set text instead of speaking spontaneously would remain somewhat speculative without input from the person about whom the observations were made. Fortunately, Student 5 lent credence to this supposition by describing an instance when she had used the reading to answer a question:

> When [Student 4] asked about the meaning of teachers' identity, I remember at the time I found some highlighted parts in the article… So, then I saw that the others are not saying anything and talking; I started to say the meaning of teachers' identity. (SRI 5, 8:59)

What is perhaps the most striking aspect of the differences between the students' perspectives on this particular episode is the way in which reading rather than speaking was seen by the different participants. Student 5 appears to have been pleased with what she was doing: Not only was she transparent in retrospect about having used the article to answer Student 4's question, but she also stressed that she had done it at a time when no one was speaking. At no point did Student 5 imply that she had been embarrassed when reading out passages during the discussion, which may indicate that she considered her response to be a valuable addition to what had been said and possibly even a springboard for further discussion. In other words, she believed she was taking the discussion forward. This, however, is not the impression that Student 5 seems to have left on her fellow interlocutors. Student 9 spoke of Student 5's habit of reading out excerpts from the set text in descriptive terms, but the very fact that she mentioned it may—on the principle of markedness—have been indicative of a degree of disapproval. By contrast, Student 4 was more openly critical of his peer's behaviour and said in no uncertain terms that Student 5's reliance on the text had not brought about progress in the discussion. These contrasting attitudes towards the act of reading out parts of the text appear to have been underlain by differences in the participants' schemata of a seminar discussion: Such a speech event would entail formulating and expressing an interpretation of a set text in the eyes of some, whereas for others the sharing of verbatim quotations from the text was not beyond the bounds of a discussion.

Conclusion

The objective of the present paper was to explore the possibility of researching the pragmatics of academic ELF interaction by incorporating participants' emic perspectives into the analysis. To this end, a study in which the researcher's pragmatic analysis of ELF interactional data was supplemented with the interactants' emic views on the speech event was carried out. The aspects of the discussion that took on salience from the researcher's etic perspective can be categorised into features of interaction

and features of language use. Out of the two main strands, the former, that is, the features of interaction were more numerous and more striking. In total, six main features of interaction were identified in the data which, in descending order of frequency, were the following: questions, agreement, discussion moderation, apology, silence, and off-topic interaction. However, it was found that the rate of occurrence for each feature was not indicative of its perceived salience in the context of the conversation. For example, silence was among the most notable aspects of the interaction from the researcher's perspective despite the fact that only a couple of instances of silence were observed in the data. Although there were obvious overlaps between interaction and language use, four features of the conversation were solely categorised as features of language use: nonstandard usage, discourse reflexivity, hedging, and echo. Instances of nonstandard usage, which was the most frequently identified language feature, pertained almost exclusively to the use of the indefinite article. The participants' language use was also characterised by adverbial hedging and by discourse reflexivity, in which they engaged through their explicit articulation of the illocutionary force of their utterances.

The interactants' emic perspectives, which were enquired into by means of conducting a video stimulated-recall interview with each student, were integrated into the study in order to illustrate the differences between the researcher's etic analysis of the speech event and the participants' emic experiences. The emic perspectives were a valuable addition in that they revealed discrepancies: Seldom did the participants' and the researcher's perspectives on the same speech event intersect. Moreover, the input from the students provided insight into each participant's unique perception of their involvement in the conversation. For example, an instance of silence analysed by the researcher resulted in a convergence of emic views, and it emerged that the episode had been experienced in markedly different ways by the different participants. While some students were silent because they were unsure of how to respond to what their groupmate had said, others were scanning the set text in search of something to say, and others still decided to ignore the situation altogether instead of going through the rigmarole of contributing to the conversation. On the whole, the participants'

emic remarks were not as detailed or methodical (and nor could they be expected to have been) as the researcher's categorisation of the discourse features, but the comments they made were in some ways more insightful than any analysis conducted by an external observer could have been.

Even though the results of the analysis appear to have fulfilled the initial aims of the study, the limitations of the findings ought to be mentioned. In spite of the deep understanding of specific aspects of the speech event which were gained by dint of studying the participants' inner perspectives, the arbitrariness of the data obtained via stimulated recall made the research outcomes difficult to anticipate. This is a trait which stimulated recall shares with other methods of qualitative data collection, but the participants of this study were provided with even less guidance on what they were expected to say than they would have been, for instance, in a semi-structured interview. Reality is rarely structured; consequently, stimulated recall also had to be fairly open-ended if the participants' experiences were to be explored. As a result, the researcher bore the burden of interpretation and synthesis, making the results unique to this iteration of the analysis. However, the analyst's potential subjectivity was not the only factor that influenced the results: The research participants are likely to have presented a version of their recollections which was particular to the moment of data collection. Although a reduction in the amount of time which elapses between the speech event and the stimulated-recall sessions may go some way towards mitigating the risk of misremembrance, it cannot be guaranteed that participants will manage to verbalise exactly what they were thinking at the time which they are requested to remember.

References

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

Bailey, J. (2008). First steps in qualitative data analysis: Transcribing. *Family Practice, 25*(2), 127–131. https://doi.org/10.1093/fampra/cmn003

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press. https://doi.org/10.1017/CBO9780511813085

Canagarajah, S., Kimura, D., Shahri, M. N. N., & Amory, M. D. (2020). Toward an expansive interactional analysis. In J. McKinley & H. Rose (Eds.), *The Routledge handbook of research methods in applied linguistics* (pp. 500–513). Routledge. https://doi.org/10.4324/9780367824471

Cogo, A., & House, J. (2018). The pragmatics of ELF. In J. Jenkins, W. Baker, & M. Dewey (Eds.), *The Routledge handbook of English as a lingua franca* (pp. 210–223). Routledge. https://doi.org/10.4324/9781315717173

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Routledge. https://doi.org/10.4324/9780203029053

Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment. Companion volume*. https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.

Firth, A. (1996). The discursive accomplishment of normality: On 'lingua franca' English and conversation analysis. *Journal of Pragmatics, 26*(2), 237–259. https://doi.org/10.1016/0378-2166(96)00014-8

Flowerdew, J. (1990). Problems of speech act theory from an applied perspective. *Language Learning, 40*(1), 79–105. https://doi.org/10.1111/j.1467-1770.1990.tb00955.x

Graddol, D. (2006). *English next: Why global English may mean the end of 'English as a foreign language'*. British Council. https://www.teachingenglish. org.uk/sites/teacheng/files/pub_english_next.pdf

Haugh, M., & Culpeper, J. (2018). Integrative pragmatics and (im)politeness theory. In C. Ilie & N. R. Norrick (Eds.), *Pragmatics and its interfaces* (pp. 213–239). John Benjamins Publishing Company. https://doi.org/10.1075/ pbns.294

House, J. (2022). The pragmatics of English as a lingua franca. *Applied Pragmatics*, *4*(2), 121–136. https://doi.org/10.1075/ap.00005.hou

Howitt, D. (2016). *Introduction to qualitative research methods in psychology* (3rd ed.). Pearson.

Hutchby, I., & Wooffitt, R. (2008). *Conversation analysis* (2nd ed.). Polity Press.

Illés, É. (2020). *Understanding context in language use and teaching: An ELF perspective*. Routledge. https://doi.org/10.4324/9780429274589

Illés, É., & Bayyurt, Y. (Eds.). (2024). *English as a lingua franca in the language classroom: Applying theory to ELT practice*. Routledge. https://doi. org/10.4324/9781003258698

Jenkins, J. (1998). Which pronunciation norms and models for English as an international language? *ELT Journal*, *52*(2), 119–126. https://doi. org/10.1093/elt/52.2.119

Kalocsai, K. (2011). *Communities of practice and English as a lingua franca: A study of Erasmus students in Szeged* [Doctoral dissertation, University of Szeged]. SZTE Repository of Dissertations. https://doktori.bibl.u-szeged.hu/id/eprint/775/1/Kalocsai_Dissertation_Final.pdf

Kecskes, I. (2019). *English as a lingua franca: The pragmatic perspective*. Cambridge University Press. https://doi.org/10.1017/9781316217832

Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press. https://doi. org/10.1017/CBO9780511813313

Liu, J. (2020). A pragmatic analysis of hedges from the perspective of politeness principle. *Theory and Practice in Language Studies*, *10*(12), 1614–1619. http://dx.doi.org/10.17507/tpls.1012.15

Mauranen, A. (2010). Discourse reflexivity – A discourse universal? The case of ELF. *Nordic Journal of English Studies*, *9*(2), 13–40. https://doi.org/10.35360/njes.216

Mauranen, A. (2012). *Exploring ELF: Academic English shaped by non-native speakers*. Cambridge University Press.

Mauranen, A., & Ranta, E. (2008). English as an academic lingua franca – The ELFA project. *Nordic Journal of English Studies*, *7*(3), 199–202. https://doi.org/10.35360/njes.108

Paltridge, B. (2021). *Discourse analysis: An introduction* (3rd ed.). Bloomsbury Academic. https://www.perlego.com/book/3028286/discourse-analysis-an-introduction-pdf

Pitzl, M-L. (2022). From *cross* to *inter* to *trans* – *cultural pragmatics on the move: The need for expanding methodologies in lingua franca research. In I. Walkinshaw (Ed.), *Pragmatics in English as a lingua franca* (pp. 55–80). De Gruyter Mouton. https://doi.org/10.1515/9781501512520-004

Ryan, J., & Gass, S. (2012). Stimulated recall. In R. Barnard & A. Burns (Eds.), *Researching language teacher cognition and practice: International case studies* (pp. 144–161). Multilingual Matters.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696–735. https://doi.org/10.2307/412243

Searle, J. R. (1971). What is a speech act? In J. R. Searle (Ed.), *The philosophy of language* (pp. 39–53). Oxford University Press.

Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal*, *59*(4), 339–341. https://doi.org/10.1093/elt/cci064

Smit, U. (2010). *English as a lingua franca in higher education: A longitudinal study of classroom discourse*. De Gruyter Mouton. https://doi.org/10.1515/9783110215519

Sorlin, S. (2017). The pragmatics of manipulation: Exploiting im/politeness theories. *Journal of Pragmatics*, *121*, 132–146. https://doi.org/10.1016/j.pragma.2017.10.002

Taguchi, N., & Ishihara, N. (2018). The pragmatics of English as a lingua franca: Research and pedagogy in the era of globalization. *Annual Review of Applied Linguistics*, *38*, 80–101. https://doi.org/10.1017/S0267190518000028

Tsohatzidis, S. L. (Ed.). (1994). *Foundations of speech act theory: Philosophical and linguistic perspectives*. Routledge. https://doi.org/10.4324/9780203206478

Walkinshaw, I. (Ed.). (2022). *Pragmatics in English as a lingua franca*. De Gruyter Mouton. https://doi.org/10.1515/9781501512520

Westerman, D. A. (1991). Expert and novice teacher decision making. *Journal of Teacher Education*, *42*(4), 292–305. https://doi.org/10.1177/002248719104200407

Widdowson, H. G. (2004). *Text, context, pretext: Critical issues in discourse analysis*. Blackwell Publishing. https://dx.doi.org/10.1002/9780470758427

Widdowson, H. G. (2007). *Discourse analysis*. Oxford University Press.

Wilkinson, S., & Kitzinger, C. (2011). Conversation analysis. In K. Hyland & B. Paltridge (Eds.), *The Continuum companion to discourse analysis* (pp. 22–37). Continuum.

Yule, G. (1996). *Pragmatics*. Oxford University Press.

8

A Comprehensive, Process-Based Model of L2 Listening Behaviors

Gyula Tankó
tanko.gyula@btk.elte.hu

Abstract

Listening is a key communication skill that is still in need of research in spite of the recent increased theoretical and empirical educational and testing research activity whose purpose has been to inform, at times jointly, teaching practice and assessment. This article aims to contribute to the body of literature available on the assessment of L2 listening ability. Focusing on transactional, one-way listening assessed with independent listening tasks, the present research reviews the most relevant literature available on teaching and assessing listening with the ultimate aim of building a comprehensive model of listening behavior types suitable for listening research, assessment development, test evaluation, and item writer training purposes. Based on the review, a cognitive processing model of listening is enhanced with the addition of a metacognitive component that can functionally accommodate the listening operations performed by listening behaviors subdivided into global and local strategies and skills. Given its inception, the proposed model arguably has a considerable degree of theoretical validity, but it needs to be empirically validated before it is applied in the domains of research, assessment development, test evaluation, and item writer training.

Keywords: model of listening behavior types, listening strategies, listening skills, assessing L2 listening, one-way listening

A Comprehensive, Process-Based Model of L2 Listening Behaviors

As evidenced by experiential data as well as reported estimates, listening is a core communication skill: Language users spend most of their communication time listening to others (e.g., Burley-Allen, 1995; Hedge, 2004; Wagner, 2013). However, not all of this time is spent engaged in effective listening, which is especially true for non-native speakers of a language. Consequently, listening has been researched, taught, and tested, but there is a long-standing consensus among language professionals that relative to the other basic communication skills, listening has not been taught (Buck, 2001; Elin, 1972; Goh, 2010; Landry, 1969) or assessed appropriately. Some of the most compelling reasons for the latter are that language testers disagree about how listening can be best assessed (Feyten, 1991) or find it challenging, which has resulted in the neglect and oversight of listening by language assessment researchers (Wagner, 2013).

This article aims to contribute to the body of literature on the assessment of L2 listening ability. Specifically, it focuses on transactional listening (i.e., listening employed to obtain information, Vandergrift & Goh, 2012) engaged by both task types differentiated by Underwood (1989): tasks performed while-listening (e.g., map labelling or spotting mistakes) and those performed post-listening (e.g., sequencing or establishing speaker attitude). When completing such one-way listening tasks, in which the listener has a non-collaborative role (Buck, 2001), listeners reproduce the message they hear in a new form; therefore, these transactional listening tasks require primarily deep processing (Nation & Newton, 2009). A final stipulation is that the article is concerned with the assessment of listening by means of independent listening tasks. The integrated assessment of listening engages processes and requires a construct definition that is beyond the scope of this article.

The article reviews the most relevant literature available on teaching and assessing L2 listening in order to build a comprehensive model of listening behavior types suitable for listening research, assessment development, test evaluation, and item writer training purposes. Consequently,

following the definition of listening in the next section, a review of the pertinent body of literature that conceptualizes listening as a process and of the taxonomies of listening sub-skills is presented. Based on the review, an expanded cognitive processing model of listening is proposed, which can accommodate the model of listening behavior types constructed to serve the above-stated purposes.

The Nature of Listening Comprehension

The skill most frequently used in everyday communication with which language users have been reported to spend at least 40-50% of their communication time is listening (Burley-Allen, 1995; Hedge, 2004; Wagner, 2013). Numerous definitions of listening have been formulated in the body of listening literature. A simple one states that it is the "process of understanding speech in a first or second language" (Richards & Schmidt, 2010, p. 344). The process involves such components as linguistic units provided in aural input ranging from phonemes that combine into words and that in turn form syntactic structures, listener expectations and goals, the topic and context of the input, or the background knowledge of the listener.

In a meta-analysis of the definitions proposed for listening, Glenn (1989) conducted a content analysis of key words in 50 definitions of the concept of listening collected from sources focusing on listening and speech communication published over a period of 64 years. She concluded that of the seven macroconcepts that emerged four were to be considered central to the definition of listening. Based on these findings, listening can be defined as a language activity that subsumes the processes of *perception*, *interpretation*, *attending* (i.e., listening with a purpose by making a conscious and concentrated effort), and *responding* (e.g., performing an evaluation, drawing a conclusion, and giving a verbal or non-verbal reply). The fifth most frequent macroconcept whose inclusion in a definition of listening ability was not warranted by the summary of commonalities is *remembering* (i.e., linguistic knowledge stored in long-term memory used for decoding and parsing, as well as for recalling and restating facts heard in a piece of

spoken discourse). Glenn noted that retrieving a large number of factual details that must be stored in long-term memory is very likely a separate cognitive skill from listening. This is especially important from an assessment perspective because based on this, a test of listening comprehension should not be a test of long-term memory—that is, a test of a large number of factual details that can only be remembered from aural input if they are stored in the listener's long-term memory.

Compared to those analyzed by Glenn (1989), more recent definitions of listening comprehension (e.g., Buck, 2001) describe the act of listening as a highly complex process that draws on both *linguistic knowledge* (i.e., knowledge about the sounds, patterns, and structures they form in a language; words and grammar and the meanings they express; and discourse structure) as well as on *non-linguistic knowledge* (i.e., topical knowledge, knowledge about context, and world knowledge). Furthermore, as Field (2009) pointed out, the listening process may be only partially controlled by the listener, and unlike reading—where the input text is present and can be re-read at will—it is not recursive. Instead, asking the speaker to clarify or repeat something and, if possible, listening to an audio material a second time during the listening test are the only options to which a listener may have recourse. As a consequence, "listening is a very individual activity in terms of the processes employed and the interpretations reached" (Field, 2009, p. 31).

Buck (2001, p. 114) proposed a formal definition of the listening construct. He described listening as the ability to:

- process extended samples of realistic spoken language, automatically and in real time,
- understand the linguistic information that is unequivocally included in the text, and
- make whatever inferences are unambiguously implicated by the content of the passage.

He admitted that this definition neither includes sociolinguistic or pragmatic knowledge—in other words, it lacks reference to the contextual aspects

of the listening process—nor strategic competence, which is in fact missing from all the above-reviewed definitions. However, he noted that this can be remedied with the inclusion in the test of such texts and tasks that have the characteristics of the target-language use situation intended to be assessed. Such an enriched construct entails various types of knowledge as well as a specific set of skills and strategies that assist the deployment of listening as a cognitive process.

Listening as a Cognitive Process

Anderson (2020) conceptualized language comprehension as a process consisting of three partly ordered and partly overlapping stages, namely the *perceptual*, *parsing*, and *utilization* stages. The first stage involves lower-level and the remaining two higher-level processes. In the first stage, perceptual processes encode (i.e., perceive) spoken input in the case of listening and written (i.e., visual) input in the case of reading. In the second stage, a mental representation of the meaning of a (spoken) utterance or a (written) sentence is built through the parsing process that integrates semantic and syntactic cues by employing semantic information as a guide for making syntactic decisions. In the third stage, the listener or speaker can utilize the newly built mental representation in different ways (e.g., an instruction can be followed by an action, or an assertion can be stored in memory or used to make an inference about the speaker or writer). Whereas psychologically plausible, and in its three-stage perception-parsing-utilization structure similar to other accounts of language comprehension (e.g., Field, 2013 discussed below or Goh & Vandergrift, 2021), this conception of the comprehension process is overly general to help distinguish, define, and operationalize listening behaviors. Therefore, related approaches—one focusing on discourse comprehension in general and two others specifically on listening ability—are discussed next.

Utterances and sentences are typically processed as the constituents of larger contexts, such as spoken or written texts that represent instances of discourse. A markedly influential model that explains discourse level reading and listening comprehension processes is the Construction-Inte-

gration Model of Text Comprehension (Kintsch, 1998; 2018; Kintsch & Van Dijk, 1978). According to this model, discourse comprehension is realized in two ordered phases through *knowledge construction* and *knowledge integration* that result in a *final text representation* (see Figure 1).

The exact wording of utterances or sentences, namely the form of the message derived by means of such lower-level decoding processes as word recognition or syntactic parsing, constitutes the *surface representation* of aural or visual input that is usually retained for a short time in memory. The surface representation constitutes the basis for *knowledge construction*. The actual words from the input are transformed into an idea-level constituent called a *proposition*, which is a minimal unit of thought defined as "an abstract, theoretical construct, which is used to identify the meaning, or what is expressed by a sentence under specific contextual restrictions (speaker, time, place), and which is related to truth values" (Van Dijk & Kintsch, 1983, p. 111). Several propositions combine into a *microstructure*, which is the local, sentence-level structure of a text. With the help of the operations of selection, generalization, and construction, known as *macrorules* (Van Dijk, 1980), globally relevant propositions can be derived from the ideas expressed with *micropropositions* in the *microstructure*. The resulting higher order units, called *macropropositions*, constitute the *macrostructure* of a text. Kintsch (2018) defined macropropositions as "summary statements at different levels of generality. They subsume what the different sections of a text are about" (p. 192). Consequently, macrostructure can denote the gist, global- or discourse-level structure of a text that is necessary for text comprehension.

Figure 1

A Representation of the Components of the Construction-Integration Model

Representation level	Discourse comprehension phase	Input	Output
			final text representation
			↑
Situational	knowledge integration	textbase & comprehender's relevant prior knowledge →	coherent & interpretable whole (elaborated propositional network)
↑			↑
Propositional	knowledge construction	microstructure & macrostructure →	textbase (propositional network)
↑			↑
Surface	decoding	words →	linguistic representation

visual/acoustic input ↑

In the *knowledge construction* phase, the propositions in the micro- and macrostructure are organized into a semantic propositional network of text meaning, that is, a propositional structure called textbase. At this point, the propositions and the relationships between them form a basic, undeveloped mental representation of discourse. Comprehenders remember the macropropositions from such a semantic or *propositional representation*—in other words, the message conveyed by a text—for a longer time than the surface representation of aural or visual input.

In the *knowledge integration* phase, the meanings conveyed by the text in the form of a network of propositions unconstrained by syntactic structure become integrated through the process of inference with information from the comprehender's prior knowledge. The resulting level of text representation is the *situational representation*, which is a coherent and interpretable whole, a refined representation of text content. It consists of a network of

propositions enriched with information retrieved from the knowledge bases of the comprehender. Among other factors, the reading or listening goals and the relevant prior knowledge a comprehender brings to the input determine the inferences they make and the situational representations they construct. The outcome of the knowledge construction and integration phases is a final text representation that is ready to be interpreted and evaluated, and that can be stored in the comprehender's long-term memory.

The Construction-Integration Model informed the conceptualizations of the listening process proposed by Field (2009, 2013), which is arguably the most comprehensive and detailed representation of the listening process alone that is available to date. The most recent version of Field's information-processing model (Field, 2013) differentiates five listening phases (see Figure 2): *input decoding*, *lexical search*, *parsing*, *meaning construction*, and *discourse construction*, corresponding to three sets of operations that process speech signals and words, and construct meaning.

In the first three phases that describe lower-level processes, the listener first turns the acoustic cues into phonemes and stressed or unstressed syllables. The sequences of sounds are then matched to the most probable word forms in the mental lexicon, and their meaning is accessed. Next, the words are organized into a syntactic structure; their meaning is further disambiguated; and an intonation pattern is identified, which helps determine the salience of message components as well as the speaker's communicative goals and attitude. In the first three phases, the processing is recursive: Each subsequent stage can inform the operations performed in an earlier phase.

Figure 2

Aggregated Representation of Field's (2013) Information-Processing Listening Model

Listening phase		Outcome		Knowledge source
discourse construction	→	**discourse representation**	←	discourse
↑				
meaning construction	→	meaning representation	←	pragmatic, external, topic
↑				
parsing	→	propositional information	←	syntactic
↑ ↓				
lexical search	→	word string	←	lexical
↑ ↓				
input decoding	→	phonological string	←	phonological
↑ ↓				
Acoustic input				
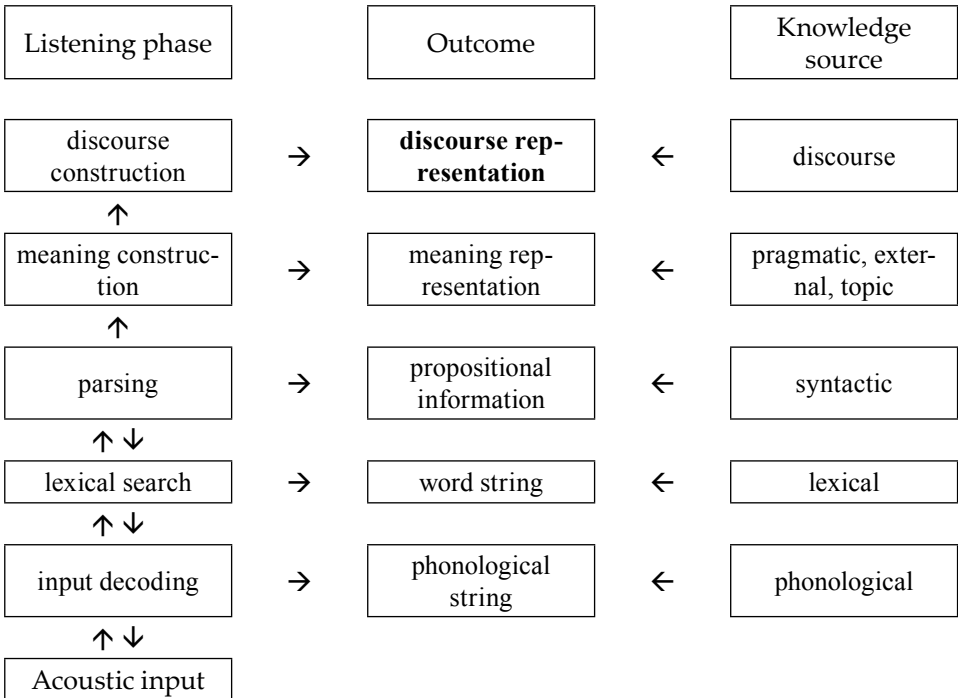
The remaining two phases account for higher-level processes. As discussed above in relation to the Construction-Integration Model, in the *meaning construction* phase a *meaning representation* is built when the propositions that represent the literal meaning of the parsed input are integrated with information retrieved from the listener's knowledge base. This meaning representation is enriched by taking into consideration the context, speaker intentions, reference, and inference. In the *discourse construction* phase, the listener relates the information obtained so far to the already available information related to the listening event and checks whether the two are consistent, decides how relevant the new information is to the emerging *discourse representation*, and then structures it. The processes involved are *information selection*, *integration*, and *self-monitoring*. Field's (2013) account of the use of the selection and integration processes is congruent with Rost's (1990) claim that the emerging discourse representation constructed by the listener is

reductive in the sense that the listener cannot retain in memory a verbatim representation of a text for more than several seconds. The representation is additive to the text in that the listener must bring in background information and assumptions in constructing a coherent representation. (p. 128)

The discourse level structure built by a proficient listener is hierarchical and consists of major and related subordinate points. In contrast, the one built by a less proficient listener is in essence linear. Its construction is also informed by the listener's discourse knowledge and is fundamentally affected by the effective use of the appropriate listening sub-skills.

Taxonomies of Listening Sub-Skills

A number of more or less detailed lists and taxonomies of operations related to listening ability and the process of listening comprehension have been proposed in the literature. Some of these (see Table 1), specifically those compiled for teaching (e.g., Richards, 1983; Rost, 1990; White, 1998) or for test development in general (e.g., Hughes, 1989; Munby, 1978; Weir, 1993) are theoretical, and as such, speculative. Others are based on research conducted in relation to a specific language test (e.g., Aryadoust, 2013; Bejar et al., 2000; Buck et al., 1997; Chapelle et al., 1997; Rukthong, 2014).

Table 1

Lists and Taxonomies of Operations Related to Listening Ability

Developed for		
Language teaching	Test development in general	Specific language tests
• taxonomy of micro-skills for conversational listening or for listening to lectures (Richard, 1983) • taxonomy of enabling skills and enacting skills in listening (Rost, 1990) • checklist of listening sub-skills (White, 1998)	• taxonomy of language skills (Munby, 1978) • list of macro- and micro-skills (Hughes, 1989) • checklist of listening comprehension operations (Weir, 1993)	• IELTS (Aryadoust, 2013) • TOEFL (Bejar et al., 2000, Chapelle et al., 1997) • TOEIC (Buck et al., 1997) • PTE Academic (Rukthong, 2014)

This section discusses those complementary and occasionally overlapping taxonomies compiled for both teaching and assessment purposes that are the most relevant for the purposes of this article. The first three of these were created for teaching. Rost (1990) differentiated three types of listening. *Selective listening* refers to deriving specific information from the input text and is typically measured with closed items. Students are expected to be able to extract specific information even when their language proficiency and topical knowledge fail to reach the complexity of the input. The explanation Rost (1990) provided for *global listening* foregrounds that there are different types of global listening: He stated that with global listening "students construct an overall sense, or gist, of a text" (p. 232) and then that they can also "identify topics and transition points between topics" (p. 232). These global listening types therefore encompass two types of listening for gist: understanding the topic requires understanding the overall sense of the input. In contrast, the identification of topics and topic shifts in the input requires the construction of a skeletal text level structure, which is the same activity as skimming in the case of reading (cf. Weir & Khalifa, 2008). The last type of listening is *intensive listening*, which earlier used to refer to a focusing-on-form activity—a grammatical consciousness raising activity that teachers use after students have gained a certain level of understanding of the input. However, this definition of intensive listening was modified more recently to the "ability to listen intensively *when required*—as in listening for specific details or to spot a particular word" (Rost, 2016, p. 171). As discussed below, Pan (2017) also conceptualized intensive listening as listening to linguistic forms (e.g., proper nouns) in order to reproduce them in writing.

Lund (1990) proposed a taxonomy of real-world listening behaviors that he intended to serve as a conceptual framework for the teaching of listening ability. It has three key components that define listening tasks: the *input text*, *listener function*, and *listener response*. Listener response is "what the listener does to demonstrate successful listening" (Lund, 1990, p. 109)—for example, selects one element from a number of alternatives, gives a written or spoken answer, or places pictures in order based on a story—and is closely related to listener function, which is directly relevant to

the development of a model of listening behavior types. Listener function determines "how the learner must approach the text and suggests what must be derived from it" (Lund, 1990, p. 109). It is dependent on the purpose of listening and determines the skills and strategies the listener uses.

Altogether six L2 listening behaviors are distinguished in the taxonomy, which can be used with any type of input text, specify the part of the input to be processed, and prompt the type or amount of information to be extracted. *Identification* refers to the recognition and discrimination of form (e.g., singular vs. plural form) rather than content and requires no actual message comprehension, while *Orientation* denotes the listener tuning in and identifying the context (e.g., participants, discourse topic, genre, and speaker function). Message comprehension becomes important with function three, *Main Idea Comprehension*. This is followed by *Detail Comprehension*, which denotes the understanding of specific information. It is to be specified here that listeners may listen only for details separately or in relation to the main ideas they want to elaborate. *Full Comprehension* entails understanding both the main ideas and details, namely the entire input text irrespective of its length. The last function is *Replication*, specifically the reproduction of input form, and it is, like Identification, less concerned with meaning.

Finally, like Lund (1990), Field (2009) prioritized listener goals and differentiated types of listening on the basis of these. Inspired by the work of Urquhart and Weir (1998) on reading, Field (2009) conceptualized listening types as varying based on the level of detail and attention required for the achievement of the listening goal. Accordingly, he distinguished two levels (i.e., global and local) and four focuses of attention, namely *shallow, medium, deep*, and *very deep* attentional focuses. Of the listening types he listed, those that are relevant from an assessment perspective are presented in Table 2 with the original wording of the listening type definitions and labels, as well as of the listener's goals formulated as questions and statements.

It is noteworthy that Field (2009) identified listening types at both the global and local level that require markedly different amounts of careful listening. The deeper the focus of attention is, the more the listener must stay tuned in and listen carefully. However, whereas he differentiated close

listening at both levels, search listening only occurs in the framework at the local level—without any explanation provided.

Taxonomies or models of listening skills and strategies have also been created for assessment purposes. Some of them, like the one proposed for the TOEFL 2000 listening test that specified competences and abilities, were general and therefore unsuitable for operationalization purposes. The competences planned to be assessed by the test were procedural, sociolinguistic, linguistic, and discourse competence (Chapelle et al., 1997). The listening abilities proposed to be assessed were as follows: comprehension of details and facts, vocabulary, main ideas and supporting ideas, content and relationship inferences, and communicative function of utterances (Bejar et al., 2000).

Table 2

Listening Types Determined by Listener Goals Relevant for Assessment (Field, 2009, p. 66)

Level of detail	Level of attention	Type of listening	Listener goals
Global	shallow	Skimming (listening generally) to establish discourse topic and main ideas.	*What is it about?*
	medium	Listening for plot; listening to commentary.	*What happened next?*
	deep	Close listening to establish the speaker's main points and to trace connections between them.	*What is important?*
Local	medium	Focused scanning to locate one area of information needed by the listener.	*When will the speaker mention X?*
		Search listening to locate and understand information relevant to predetermined needs.	*What is the answer to these questions?*
	deep	Close listening to record in depth the speaker's main points and supporting detail.	*I assume that everything is relevant.*
	very deep	Listening to vital instructions.	*I assume that everything is important.*

Comparatively more detailed was Weir's (1993) taxonomy of communicative listening sub-skills. In his checklist of operations intended to inform test design, Weir (1993) differentiated three types of meaning comprehension. *Direct Meaning Comprehension* refers to such listening activities as listening for and understanding gist, main ideas, important information, supporting details, or specific details. *Inferred Meaning Comprehension* pertains to activities like making inferences and deductions, recognizing communicative functions, and deducing word meaning from context. *Contributory Meaning Comprehension* denotes linguistic form related activities, for example, understanding phonological features, syntactic structure, lexis, discourse markers, or cohesion. Weir (1993) found it important to include inference in his taxonomy. In a more recent and thoroughly informative framework discussed next, Pan (2017) also included inference and noted that deducing implicit information increases the listener's processing load.

The comprehensive framework for the communicative testing of listening comprehension proposed by Pan (2017) conceptualized communicative listening ability as consisting of *listening macro-skills*, *linguistic competence*, and *strategic competence*. The linguistic sub-processes of listening comprehension, that is, grammatical, discoursal, functional, and sociolinguistic *micro-skills*, together with non-linguistic strategies (i.e., strategies used either to control the deployment of listening skills or to compensate for limitations in their use) form *macro-skills*. Whereas micro-skills are *bottom-up* processing operations applied to the linguistic input, non-linguistic strategies are *top-down* processing operations informed by non-linguistic knowledge (i.e., topical knowledge, and knowledge of communicative situations and events, which partially overlaps in this framework with contextual knowledge and its linguistic realization). Following closely the work of Van Dijk (1980), the framework differentiates three *discourse levels*: *top level*—e.g., discourse topic (i.e., the core proposition, which is to be understood as the global topic in Van Dijk's words); *global level*—namely, the macrostructure—and *local level*—that is, the microstructure of discourse. The framework further distinguishes *information types* (i.e., situational context and background knowledge, propositional meaning, communicative

function, and form) and *manners*—the explicit or implicit conveyance of information.

Based on these discourse levels, information types, and manners of conveyance, Pan (2017) defined five listening purpose types. The listener uses *tuning-in (top-level listening)* to understand the discourse topic and overall function of the input (i.e., the highest ranking macroproposition in the input and the overall/global speech act realized by it) together with the situational context. Whereas no reservations can be formulated about the psychological reality of this listening purpose, most probably few test items target it. The reasons for this may be, for example, that the topic of the input is usually given in the task instructions in order to help candidates to prepare for the effective processing of the input. Moreover, if the topic is not stated explicitly, candidates can infer it—along with the gist of the input text—based on the test items that they are encouraged to read prior to listening to an audio material so as to enhance their performance.

Extensive (top-global-level) listening makes it possible for the listener to understand the main ideas (i.e., macropropositions) and main functions (i.e., macrospeech acts) of the main parts of the input. When engaged in *full (top-global-local-level) listening*, the listener combines top and global level information with local level information, namely with micropropositions and microspeech acts. The listener uses *selective (local-level) listening* to extract specific, local-level details (i.e., micropropositions or microspeech acts). Finally, the listener focuses on linguistic form during *intensive listening* (e.g., when recording names); which is what Lund (1990) called *replication* and defined as follows:

> The listener's main objective is to reproduce the message in either the same or a different modality, as in oral repetition, dictation, or transcription. The listener's attention is focused more on the fidelity of the replication than on the depth of comprehension, more on surface processing than meaningful processing. Consequently, replication does not imply a higher proficiency than comprehension, but rather defines a different way of attending to the text. (p. 109)

Pan (2017) subsumed the five listening purpose types under four macro-skills. The first two of these are understanding explicit or implicit propositional meaning expressed and direct or indirect speech acts performed at the top, global, and local levels of discourse. The third macro-skill is recognizing the situational context and linguistic co-text of the discourse, and the last one refers to the recognition of linguistic and phonological forms.

Based on the above overview of the nature of listening comprehension, the process of listening, and taxonomies of listening sub-skills, in the next section an expanded cognitive processing model of listening is proposed. Two functionally closely related components are added to Field's (2013) information-processing listening model. The first of these is the goal setter, which is a metacognitive component that controls listening goals and attention level. Its inclusion in the model entails the second component: a model of listening behavior types.

An Expanded Cognitive Processing Model of Listening

The expansion of the processing model of listening proposed here was informed not only by the body of knowledge available on listening but also on reading. In spite of obvious differences, the two language abilities are often addressed jointly in the literature on researching, assessing, and teaching language comprehension. Some of these differences are as follows:

- the type, amount, and linguistic realization of the information encoded in written or spoken texts (Rost, 1990); and
- mode of presentation and interpretation of the input text, which is synoptic (i.e., the text is present as an object and is understood as a whole) for writing, whereas it is dynamic (i.e., the delivered text is process-like, happens in real time, and must be interpreted while heard) for listening (Halliday, 1989); closely related to
- the extent to which the comprehender can control the rate of incoming text, namely the possibility to re-read or ask for repetition (Munby, 1978).

Nevertheless, the numerous commonalities shared by the two comprehension processes warrant their joint discussion, and research on one ability

informs the conception of the other (e.g., Anderson, 2020; Buck, 2001; Field, 2009; Kintsch, 1998; Kintsch & van Dijk, 1978; Lund, 1990; Rost, 2016; Shohamy & Inbar, 1991). As Rost (1990) stated: "Although reading and listening involve different linguistic decoding skills (i.e., visual vs. aural), the cognitive strategies that underlie effective reading will have much in common with those that underlie effective listening" (p. 8). Consequently, the elaboration of Field's (2013) process model of listening proposed here follows the cognitive processing model of reading comprehension described in two sources (Khalifa & Weir, 2009; Weir & Khalifa, 2008).

As shown in Figure 3, the model proposed by Field (2013) can be expanded with the addition of the previously mentioned metacognitive *goal setter* component (Khalifa & Weir, 2009; Weir & Khalifa, 2008). The goal setter identifies the purpose of listening, and determines the attention level required and the type(s) of listening behavior to be used in order to comprehend spoken discourse. As entailed by the function of this metacognitive mechanism to control the deployment of listening behavior types during the listening comprehension process, Field's (2013) expanded model is now ready to be enhanced with a model of listening behavior types.

Figure 3

Field's (2013) Information-Processing Listening Model Expanded With a Goal Setter and Listening Behaviors

Listening behaviors	Listening phase	Outcome	Knowledge source
Goal setter Selecting appropriate type of listening behavior:	discourse construction	**discourse representation**	discourse
	↑		
Selective listening *Local* Scanning Local search listening *Global* Skimming for gist Global search listening	meaning construction	meaning representation	pragmatic, external, topic
	↑		
→	parsing	propositional information	syntactic
	↑ ↓		
Detailed listening *Local* Careful local listening	lexical search	word string	lexical
	↑ ↓		
Intensive listening *Global* Careful global listening	input decoding	phonological string	phono-logical
	↑ ↓		
	Acoustic input		

A Model of Listening Behavior Types

Based on how much of the input text the listener must work with, two scopes of operation can be differentiated as *Selective* and *Detailed* (see Figures 4 and 5). The level of comprehension can be *limited* (e.g., when the discourse topic is to be identified), *partial* (when only some of the content of one or several utterances is to be processed), *thorough* (when a substantial part of an utterance, several related utterances that develop one sub-topic, or the entire input text is to be understood), and *near complete* (when the listener must work with macro- and micropropositions, that is, main ideas and their supporting points).

The input text can be processed at both the local and global discourse levels. Following Weir and Khalifa (2008), local comprehension is taken to mean that the listener understands propositions at the microstructure level, within a single utterance. Utterance is to be understood in this taxonomy as "a complete thought, usually expressed in a connected grouping of words, which is separated from other utterances on the basis of content, intonation contour, and/or pausing" (Shewan, 1988, p. 124). It corresponds in writing to the sentence, and in the assessment context it is typically also available in writing in the form of a transcript/tapescript. Global comprehension refers to the processing of propositions from two or more adjacent or non-adjacent utterances. This can entail combining micropropositions from several utterances with or without the formulation of macropropositions, or the processing of macropropositions explicitly or implicitly present in the input text. Local comprehension is elicited by textually explicit test items, which derive "an answer from the same text sentence from which the question was generated" (Pearson, 1982, p. 21), that is, the information used to formulate a test item and the answer to the item are expressed in one utterance. Global comprehension is required in order to respond to textually implicit test items that derive "an answer from a text sentence different from the one from which the question was derived" (Pearson, 1982, p. 22), so the listener must combine information provided in separate utterances in order to produce an answer.

Figure 4

A Model of Selective Listening Behavior Types

Scope of operation	Discourse level	Level of comprehension	Type of listening behavior	Listening goal
SELECTIVE	Global	Limited	*Skimming for gist*	• Determine/Infer the discourse topic (i.e., the general/overall sense of the input text). • Identify and understand the main ideas/important details in the input text. • Construct a macropropositional structure (outline summary) based on the main ideas/important details of the input text.
		Partial	*Global search listening*	• Identify and understand main ideas/important details through semantic matching based on predetermined topic(s).
	Local	Limited	*Scanning*	• Identify specific information (e.g., a particular word, phrase, name, figure, or date).
		Partial	*Local search listening*	• Locate and understand predetermined information in an utterance through semantic matching.

Figure 5

A Model of Detailed Listening Behavior Types

Scope of operation	Dis-course level	Level of compre-hension	Type of listening behavior	Listening goal
DETAILED	Global	Thorough	*Careful global listening*	• Understand explicitly stated main ideas and supporting details across utterances. • Understand implicit main ideas expressed across utterances. • Decode (all/most of) the utterances in order to comprehend the text as a whole. • Carefully establish a macro-structure for the input text. • Infer the main point made by the speaker(s)/of the input. • Deduce the meaning of lexical items based on the context provided by several utterances. • Understand the speaker's mood or attitude to the topic.
	Local	Near complete	*Careful local listening*	• Understand an explicitly stated (main) idea & supporting details in one utterance. • Understand an implicit (main) idea expressed in one utterance. • Decode most of the information in one utterance in order to comprehend it as fully as possible. • Carefully establish the micro-structure of one utterance. • Understand syntactic structures in an utterance. • Understand lexical items in an utterance. • Deduce the meaning of lexical items from morphology or the co-text provided in an utterance.

Scope of operation	Discourse level	Level of comprehension	Type of listening behavior	Listening goal
DETAILED	Local	Limited	*Intensive listening*	• Listen for the linguistic form of the discourse so as to reproduce it in either the same or a different modality. • Listen to write down accurately a person's name or a place name that has been spelled, or an address (e.g., while taking a telephone message).

There are two *global selective* listening behavior types. A listener who is skimming for gist or is engaged in global search listening only needs to process parts of the input—which can be determined in both cases with the help of discourse knowledge. *Skimming for gist* is used when the goal is to determine the discourse topic (i.e., what the input text is about overall), or identify main ideas. *Global search listening* is used when the listener has to identify and understand one or more main ideas or important details (e.g., main supporting ideas) on a specific topic given in the test item. This listening behavior is deployed when the wording of the test item and of the relevant part of the input text are different. Consequently, the listener must perform semantic content matching with the use of paraphrasing, which allows two linguistic forms to be matched to one meaning. Whereas skimming for gist is unfocused in the sense that any part of the listening input may be relevant for the realization of the listening goal, search listening is focused because the test item that sets the listening goal determines the semantic content the listener is expected to locate and understand in the input text. Due to the level of awareness they require, skimming for gist and global search listening are *listening strategies*.

The two selective local listening behavior types, *scanning* and *local search listening*, are also strategies for the reason stated above. Both are engaged by a test item that can be answered by locating information in a single utterance in the input text. Both require the identification of information on a topic specified in the test item, so both are focused. However, whereas in the case of local search listening the wording of the test item and the

wording of the information in the targeted utterance do not match, and as a consequence the listener must establish the equivalence of two linguistic forms through their shared meaning, the listener only needs to perform a string match in the case of scanning. The wording of the test items that can be answered by scanning match the information targeted in the input text either exactly or to a large extent. This requires matching a string of graphemes or phonemes stored in echoic memory to a string of phonemes in the input text. By using scanning, the listener can locate predetermined information such as a particular word, phrase, name, or date. By means of semantic matching, local search listening allows the location and comprehension of predetermined information expressed in one utterance.

The local and global types of detailed listening behaviors require the careful and near complete processing as well as the thorough understanding of information provided explicitly or implicitly in one or more utterances. *Careful global listening* is elicited by test items that listeners can answer if they manage to identify and understand explicitly stated or inferred main ideas and their supporting details by combining information across utterances located either in a subsection of or in the entire input text. Therefore, careful global listening allows the listener to build the macrostructure of the input text. Conversely to careful global listening, *careful local listening* is elicited by those test items that can be answered with the help of the thorough comprehension of the lexis and structures in a single utterance that communicates either an explicit idea or an implicit one (i.e., not literal meaning). Careful local listening makes it possible for the listener to infer the meaning of a lexical item in an utterance based on its morphology or immediate co-text. If making an inference requires the combination of propositional content across utterances, the listener's careful global listening skill is engaged. *Intensive listening* is required for the surface processing of the input. The main goal of the listener engaged in intensive listening is not comprehension but an accurate replication, for example in written form, of names, numbers, short messages, or parts of the input with the original wording. For example, a test item that requires writing down words that are spelled measures the listener's phonological awareness, namely their ability to relate specific sounds to letters. Finally, given that the listener

needs to process without a predetermined focus one or more utterances carefully and thoroughly by relying on automatized processes in order to be able to answer test items that require the use of careful global and local listening, the detailed listening behaviors are *listening skills*.

Several of the listening behavior types featured in the proposed model can be deployed jointly. Skimming may require not only the location of lexis that indicates a topic shift but also the careful processing of the co-text of the topic shift cue. Similarly, the listener who identifies the relevant part(s) of the input text with the use of search listening or scanning must switch to careful listening to process the co-text in one utterance or across several. Whereas in the case of reading a change between reading behavior types means changing the speed of reading (cf. Carver, 1992), switching between listening behavior types means adjusting the level of attention and the duration of the time and stretches of the input text for which the listener stays tuned in.

Conclusion

This article enhanced an existing cognitive model of listening so that it could accommodate a comprehensive model of listening behavior types built on the basis of the relevant body of literature on listening. The proposed model of listening behavior types consists of four selective global and local listening strategies as well as of three global and local listening skills that constitute a model which is unique in its structural features and in terms of some listening behavior types that were missing from earlier models. It was designed for the purposes of listening research, assessment development, test evaluation, and item writer training concerned with transactional, one-way L2 listening assessed with independent listening tasks.

The fact that the model of listening behavior types was constructed based on information synthesized from the listening literature grants it a considerable degree of theoretical validity. Therefore, a model arguably marked by a high psychological reality is now ready to be empirically validated through the analysis of test items sampled from independent

listening tasks. Following that, once any revision that may be needed has been completed, the model will be usable for theoretical and such practical purposes as assessment development, test evaluation, and item writing.

References

Anderson, J. R. (2020). *Cognitive psychology and its implications* (9th ed.). Macmillan Learning.

Aryadoust, V. (2013). *Building a validity argument for a listening test of academic proficiency.* Cambridge Scholars.

Bejar, I., Douglas, D., Jamieson, J., Nissan, S., & Turner, J. (2000). *TOEFL 2000 listening framework: A working paper* (TOEFL Monograph Series, Report No. 19). Educational Testing Service.

Buck, (2001). *Assessing listening.* Cambridge University Press. https://doi.org/10.1017/CBO9780511732959

Buck, G., Tatsuoka, K., Kostin, I., & Phelps, M. (1997). The sub-skills of listening: Rule-space analysis of a multiple-choice test of second language listening comprehension. In A. Huhta, V. Kohonen, L. Kurki-Suonio, & S. Luoma (Eds.), *Current developments and alternatives in language assessment.* University of Jyväskylä. https://doi.org/10.1111/0023-8333.00016

Burley-Allen, M. (1995). *Listening the forgotten skill* (2nd ed.). John Wiley & Sons.

Carver, R. P. (1992). Reading rate: Theory, research, and practical implications. *Journal of Reading, 36*(2), 84–95. https://www.jstor.org/stable/40016440

Chapelle, C., Grabe, W., & Berns, M. (1997). *Communicative language proficiency: Definition and implications for TOEFL 2000* (TOEFL Monograph Series Report No. 10). Educational Testing Service.

Elin, R. J. (1972). Listening: Neglected and forgotten in the classroom. *Elementary English, 49*(2), 230–232. https://www.jstor.org/stable/41387072

Field, J. (2009). *Listening in the language classroom.* Cambridge University Press. https://doi.org/10.1017/CBO9780511575945

Field, J. (2013). Cognitive validity. In A. Geranpayeh & L. Taylor (Eds.), *Examining listening.* Cambridge University Press.

Glenn, E. C. (1989). A content analysis of fifty definitions of listening. *Journal of the International Listening Association*, *3*(1), 21–31. https://doi.org/10.1207/s1932586xijl0301_3

Goh, C. C. M. (2010). Listening as process: Learning activities for self-appraisal and self-regulation. In N. Harwood (Ed.), *English language teaching materials (pp.* 179–206). Cambridge University Press.

Goh, C. C., & Vandergrift, L. (2021). *Teaching and learning second language listening: Metacognition in action* (2nd ed.). Routledge. https://doi.org/10.4324/9780429287749

Halliday, M. A. K. (1989). *Spoken and written language*. Oxford University Press.

Hedge, T. (2004). *Teaching and learning in the language classroom*. Oxford University Press.

Hughes, A. (1989). *Testing for language teachers*. Cambridge University Press.

Khalifa, H., & Weir, C. J. (2009). *Examining reading: Research and practice in assessing second language reading* (Studies in Language Testing, Vol. 29). UCLES/Cambridge University Press.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge University Press. https://doi.org/10.4324/9781315110592-12

Kintsch, W. (2018). Revisiting the construction-integration model of text comprehension and its implications for instruction. In D. E. Alvermann, N. J. Unrau, M. Sailors, & R. B. Ruddell (Eds.), *Theoretical models and processes of literacy* (pp. 178–203). Routledge.

Kintsch, W., & Van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*(5), 363–394. https://doi.org/10.1037/0033-295X.85.5.363

Landry, D. L. (1969). The neglect of listening. *Elementary English*, *46*(5), 599–605. https://www.jstor.org/stable/41386549

Lund, R. L. (1990). A taxonomy for teaching second language listening. *Foreign Language Annals, 23*(2), 105–115. https://doi.org/10.1111/j.1944-9720.1990.tb00348.x

Munby, J. (1978). *Communicative syllabus design*. Cambridge University Press.

Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. Routledge. https://doi.org/10.4324/9780203891704

Pan, Z. (2017). *Assessing listening for Chinese English learners: Developing a communicative listening comprehension test suite for CET*. Routledge.

Pearson, P. D. (1982). *A context for instructional research on reading comprehension* (Tech. Rep. No. 230). University of Illinois, Center for the Study of Reading.

Richards, J. C. (1983). Listening comprehension: Approach, design, procedure. *TESOL Quarterly, 17*(2), 219–240. https://doi.org/10.2307/3586651

Richards, J. C., & Schmidt, R. (2010). *Longman dictionary of language teaching and applied linguistics* (4th ed.). Pearson Education.

Rost, M. (1990). *Listening in language learning*. Routledge.

Rost, M. (2016). *Teaching and researching listening* (3rd ed.). Routledge. https://doi.org/10.4324/9781315732862

Rukthong, A. (2014). Investigating the listening construct underlying listening-to-summarize tasks. In B. P. Olmos-López, J. Huang, & J. Almeida (Eds.), *Papers from the Lancaster University Postgraduate Conference in Linguistics & Language Teaching* (Vol. 8, pp. 67–89). Department of Linguistics and English Language, Lancaster University.

Shewan, C. M. (1988). The Shewan Spontaneous Language Analysis (SSLA) system for aphasic adults: Description, reliability, and validity. *Journal of Communication Disorders, 21*, 103–138. https://doi.org/10.1016/0021-9924(88)90001-9

Shohamy, E., & Inbar, O. (1991). Validation of listening comprehension tests: The effect of text and question type. *Language Testing, 8*(1), 23–40. https://doi.org/10.1177/026553229100800103

Underwood, M. (1989). *Teaching listening*. Longman.

Urquhart, A. H., & Weir, C. J. (1998). *Reading in a second language: Process, product and practice*. Longman.

Van Dijk, T. (1980). *Macrostructures: An interdisciplinary study of global structures in discourse, interaction, and cognition*. Lawrence Erlbaum.

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press.

Vandergrift, L., & Goh, C. C. (2012). *Teaching and learning second language listening: Metacognition in Action.* Routledge.

Wagner, E. (2013). *Assessing listening*. In A. J. Kunan (Ed.), *Companion to language assessment* (pp. 47–63). Wiley-Blackwell. https://doi.org/10.1002/9781118411360.wbcla094

Weir, C. (1993). *Understanding and developing language tests*. Prentice Hall.

Weir, C. J., & Khalifa, H. (2008). A cognitive processing approach towards defining reading comprehension. *Cambridge ESOL: Research Notes, 31*, 2–10.

White, G. (1998). *Listening*. Oxford University Press.

9

EFL Teachers' Perspectives on the Objectively Scored Parts of the Advanced Level English School-Leaving Examination

Zsuzsanna Andréka
andreka.zsuzsanna@btk.elte.hu

Abstract

This study aimed to explore the perspectives of Hungarian secondary school English teachers on the objectively scored components, namely the reading, use of English, and listening test papers, of the advanced level English school-leaving examination. Five EFL teachers from different regions in Hungary were interviewed. Data were collected through both online and face-to-face interviews and subsequently analyzed using thematic content analysis to identify recurring themes and patterns. Most of the criticism or dissatisfaction expressed by the interviewees were directed towards the pictures accompanying each task, certain task types (e.g., the True, False, Not mentioned type in the reading papers or the Error correction type in the use of English paper), and the scoring keys. However, the participants held positive views on other specific aspects of the examination, such as the content of source texts, some of the task types in the reading paper, and the word formation task in the use of English paper. Additionally, the format and design of the test papers were mostly viewed positively. These insights can inform future improvements in English language assessment, both within the context of the school-leaving examination and beyond, as well as curriculum development to better align with the interests and needs of Hungarian high school educators and students.

Keywords: Hungarian EFL school-leaving examination, EFL teachers, objectively scored tests, task types

EFL Teachers' Perspectives on the Objectively Scored Parts of the Advanced Level English School-Leaving Examination

In 2005, the current two-level school leaving examination system was introduced in Hungary. The original foreign language school-leaving examination had an intermediate level at A2/B1 and an advanced level test at B2 level as defined by the Common European Framework of Reference for Languages (Council of Europe, 2020). In 2017 the system was revised, and the intermediate level was changed from A2/B1 to B1 level only. However, the CEFR level of the advanced examinations did not change. The school leaving examination requirements are always based on the National Core Curriculum (NCC). A new NCC was introduced in 2020 (Hungarian Government, 2020) and many subjects saw changes to their school-leaving examinations; however, the foreign language examinations largely retained their existing format based on the exam requirements.

The school-leaving examination has five obligatory subjects: Hungarian, History, Mathematics, Foreign Language, and a compulsory elective subject. Thus, all secondary school leavers have to take a foreign language exam (Hungarian Government, 2021), and while they can choose from a variety of languages, English is one of the most popular choices, with 19,690 students taking the B2 level exam last year (Ministry of Education, 2022). In addition to being a popular choice, the English school leaving examination bears importance for those test takers who achieve a minimum of 60% on both the written and oral components of the advanced level test, since they can be issued a B2 level, nationally accredited language examination certificate. If they scored between 40 and 59%, they can still be awarded a B1 level certificate.

As a summative form of assessment, the school-leaving examination serves as the conclusion of secondary education in Hungary. The advanced level examination also serves as an entrance examination, as it is required for admission to English major tertiary education programmes in Hungary, such as the English or American Studies BA or the Undivided Teacher Training Master Programme. If a student applies for the English or American Studies BA, they must have a B2-level EFL secondary school-leav-

ing certificate. However, students applying for the Teacher of English as a Foreign Language majors only need to obtain a B2 level certificate if they choose a non-foreign language second major or a foreign language major for candidates who wish to teach minor languages (e.g., Slovenian or Japanese). If students choose to major in another major language (e.g., German, French, Dutch, Italian, or Spanish) besides English, having a B2 level school-leaving certificate from any of their two selected languages is sufficient.

As a high-stakes assessment tool with a number of stakeholder groups, the school-leaving examination needs to be thoroughly validated. While the discussion about test validation and test design is often directed at experts in the field (Vogt & Tsagari, 2022), it is important not to overlook the views of students and teachers in ongoing discussions on test design and validation.

Theoretical Background

Teachers' Role in the Hungarian Education System

As mentioned previously, teachers constitute one of the stakeholder groups involved in assessment. Not only do they create their own tests for classroom assessment, but they often prepare students for various tests using a range of resources, including sample tests or available past exams. In addition to these preparatory activities, teachers also act as raters or examiners in the evaluation process, be it a test administered in the classroom or the school-leaving examination.

In Hungary, students from 9th grade until 12th grade spend three to four hours a week learning their first foreign language and three hours learning their second foreign language (Hungarian Government, 2020). However, this is not entirely uniform; although each school is required to develop their pedagogical programme based on the NCC, they can modify these programmes to a certain extent. Any such local pedagogical programme needs to be approved by the head of the local institution.

According to Act CXC (Hungarian Government, 2011), teachers are obliged to impart knowledge objectively in a versatile and varied approach.

They are also responsible for planning and adapting the classroom activities to their group. Additionally, they must assess students' work in the form of grades or verbal feedback in accordance with the requirements of the pedagogical programme. Furthermore, teachers can be tasked with additional responsibilities such as preparing students for exams, organizing competitions, or holding facultative classes and study rooms, to name a few. Teachers are also required to take part in continuous trainings and administrative work. In the case of the school-leaving examinations, the head of the school assigns teachers to serve as examiners and raters (Hungarian Government, 1997).

Potentially Problematic Areas of the Objectively Scored Tests
Objectively scored test are those in which items are corrected using a scoring key, for example, multiple choice tests or true/false/not stated in the text items (Coombe, 2018). In the case of the school-leaving examination, the reading, use of English, and listening test papers each have their own scoring key with one or a set of acceptable solutions, and raters must evaluate students' answers based on these scoring keys. The following sections present previous research in relation to the objectively scored components of the school-leaving examination which have highlighted some potential issues with these test papers.

Test specifications
Einhorn (2009) argued that the test specifications drafted for the new school-leaving examinations in Hungary were not operationalized in the assessment tasks. She also noted that many language teachers did not understand the specifications and used the actual assessment tasks to infer the newly set performance standards. A questionnaire study (Vígh, 2012) on secondary school teachers of English, German, French and Italian from all types of secondary schools investigated the washback effect of the new school-leaving examination system on learning and teaching processes. The findings showed that teachers preferred the new examination system and modified their lesson plans and instructional content accordingly but did not implement notable methodological changes.

Operational Assessment Tasks

Tankó and Andréka (2021) investigated the use of English (UoE) sub-tests of the advanced school-leaving examination. Their findings showed that the majority of items were below B2 level. Moreover, the scoring keys were found to be incomplete and contained improbable or faulty answers. They also observed the selection of erroneous input (e.g., incoherent and/or incorrectly punctuated texts) and noted modifications made to the original text titles which could potentially hinder comprehension. Moreover, the UoE sub-tests designed for the autumn examination sessions were often of lower quality and featured more technical problems. The authors concluded that unless the remaining parts of the examination counterbalanced the weaknesses of the UoE sub-test, the examinations whose UoE sub-tests were analyzed misreported the true language ability of the test takers, which has negative implications for the tertiary educational institutions making admission decisions based on these scores.

Another study by Tankó and Andréka (2022) revealed that the reading test papers only measure a narrow range of the intended construct, contain poorly designed task types, and focus mainly on literal comprehension with excessive focus on local comprehension. There is also inconsistency in the amount of reading text in each paper, and the task context generated by the reading task instructions, images, and titles contribute to construct-irrelevant difficulty. One-quarter of the test task items contained C1 and C2 level lexis, and the input texts were sometimes at a substantially lower CEFR level than intended, leading to construct-irrelevant easiness or difficulty resulting from excessively challenging test task items. The low cut score set for the paper exacerbates these problems, which negatively affect secondary school leavers who are potential English major students.

A study on the development process of listening tasks (Király, 2012) revealed that the tasks were not tested beforehand, and the same team of nine people produced entire exams each year, with two of them responsible for the listening tasks. Authentic materials were used but were heavily edited. Another problem concerned the scoring key: Teachers complained that it was often incomplete and ambiguous, and that they did not receive information about the new task formats.

These previous studies have identified several weaknesses regarding the English school-leaving examination, including incomplete scoring keys, erroneous input texts, poorly designed assessment tasks, and construct-irrelevant difficulties. Given that teachers use the test specifications, and the test papers themselves in their classroom instruction to prepare students for the school-leaving examination, it is important from a test development perspective to gain direct insight from these educators regarding their user experience. Thus, the present study aims to answer the following research question:

> What kind of perspectives do high school English teachers hold on the objectively scored components of the advanced level English school-leaving examination?

Research Methods

To answer the research question above, an exploratory qualitative study was designed in the form of a semi-structured interview. In this section, the participants, the research instrument, the data collection, and the data analyses will be discussed.

Participants

In total, five high school English teachers from Budapest (n = 3), Miskolc (n = 1) and Zalaegerszeg (n = 1), were selected through purposive and convenience sampling (Dörnyei, 2007) and interviewed. At the time of data collection, they were teaching in five different high schools. The sample consisted of three female and two male participants, and their ages ranged between 38 and 62 (M = 50.8, SD = 9.6). The participants had an average of 28 years of teaching experience (SD = 7.3), and 24 years of experience teaching English in high schools (SD = 5.7) (Table 1). All of the participants have university degrees and hold not only regular classes but also facultative classes. Additionally, they have been carrying out examination duties as examiners and raters on the advanced level English school-leaving examinations.

Table 1

Information on the Five Participants

Pseudonym	Gender	Age	Years of teaching English	Years of teaching English in a high school
Paul	male	44	25	19
Hannah	female	38	17	17
Jane	female	62	36	29
Emily	female	55	30	26
Anthony	male	55	32	29

Research Instrument

In order to better understand high school English teachers' perspectives on the objectively scored components of the advanced level English school leaving examination, a semi-structured interview guide was selected. This format allows the interviewer to direct the interview with pre-set questions while providing the interviewee with the chance to share anything during the interview that they deem relevant to the topic (Howitt, 2016). The interview guide consisted of three parts: The first part consisted of questions about the interviewee's background; the second part focused on beliefs and experiences regarding the foreign language school leaving examination with an emphasis on the advanced level English school leaving examination; the last section addressed the objectively scored parts of the examination.

Data Collection Procedures

Data were collected through a combination of online and face-to-face interviews. For three participants, the interviews were conducted online due to geographical distance, while the remaining two participants were interviewed in person. During the interview the participants were handed or shown some test papers and scoring keys from the 2017-2022 period. Among the sample papers were some which contained potentially problematic tasks (Tankó & Andréka, 2021, 2022), but the participants were not informed about this to avoid influencing their opinions. The interviews

were conducted in the mother tongue of the participants (i.e., Hungarian) to ensure clear communication and understanding between the interviewer and interviewee. The interviews lasted an average of 60 minutes and were recorded using a microphone, the Dictaphone program on a laptop, and a smartphone. The recordings were made only after receiving consent from the participants and were transcribed verbatim after the interviews.

Data Analysis

Following transcription, the interviews were coded, and thematic content analysis (Clarke & Braun, 2014) was employed to identify recurring themes and underlying patterns in the data. Given that the study aimed to explore the attitudes of teachers towards the objectively scored test papers, the codes were classified as (a) positive, (b) neutral, and (c) critical attitudes, as well as (d) participants' suggestions.

Results and Discussion

This section presents the results and discussion of the interview analysis. The findings are organized according to the test paper constituents rather than the codes and themes for the sake of clarity. Thus, the results are presented in the following order: general information, format of the test papers, accompanying pictures, task instructions, task types, source texts, and scoring keys.

General Information

Overall, the preliminary guidelines at the beginning of each test did not attract substantial criticism, and were regarded as clear, simple, and comprehensible. The teachers shared the opinion that students generally fail to read these instructions independently, necessitating continuous reminders from the instructors during classroom practice. One interviewee stated that these instructions were too lengthy. A few respondents disagreed with the directive requiring that replies be limited to the appropriate box, claiming that it makes correcting the papers much harder because the question of accepting correct answers outside of these boxes (i.e., next to the box in cas-

es that the answer cannot fit due to earlier corrections) often arises during rater meetings and discussions with the examination board leader.

Format

The format of the school-leaving examination papers was generally regarded as acceptable, with participants deeming them to be clear, well organized, and easy to follow. The layout of the tasks and source texts was commended by the interviewees as it eliminated the need for test-takers to frequently switch pages. However, several concerns about the formatting were expressed. Many respondents regarded the excessive amount of paper used in the exam as a disadvantage. Furthermore, since the documents are tightly bound together, the addition of a 15-digit identifying number at the top of each page was seen as superfluous.

> The hapless student needs to write down this eight- or twelve-digit ID code twenty-five times. I think the students get tired of it by the time they get through writing it, and it's pretty meaningless, because they're [the test sheets] strung together, so there's little to no chance of it falling apart in a way that it's irrecoverable. It's a waste of time and you have to be very careful not to mess up a number. (Jane)

Another important stylistic concern was mentioned regarding the listening papers and the lack of boxes in which raters could mark their correction.

> Shame on them, even after 10-15 years they have not been able to add the extra one column that they put in all the others, as if we don't mention this to them every year. Seriously… it is always there, where you should write and what, scores and ticks. And here you write as you like it and wherever. It's baffling, I know this is a very small thing, but we really hate it. We really hate it… here too! [points to the test paper] What am I supposed to do now? I administered everything so nicely so far and now… you get it. I really hate this. (Hannah)

While they mentioned that the reason for this is probably related to editing and printing, and that adding the boxes could possibly disrupt the page

layout, the question remains as to why the test editors cannot format the tasks in a way that would align with the other two test papers and include the spaces for corrections.

Pictures

Previous research (Tankó & Andréka, 2021, 2022) has shown that the pictures on the test papers were problematic for a number of reasons, such as their quality, relevancy and usefulness. When asked about the visual aids on the exam papers, the respondents shared the opinion that they are a welcome addition that can help to break up the monotony of the test, lower anxiety, and activate schemata. In general, however, the teachers agreed that students frequently ignore the visuals since they are generally not relevant to the subject matter of the exam, and that they often advise the students to ignore the pictures during practice unless the picture is of good quality and can effectively introduce the topic.

As mentioned previously, the quality of the visual aids was the primary concern. Three responders bemoaned the photographs' poor resolution, small size, and black and white colour scheme. Another interviewee, while acknowledging that the pictures are usually there to help the students, noted that they can also mislead them.

> Well, I think it's good that they're trying to help the student visually to let them know what the text will be about. But it can be misleading… I've encountered problems like "but it was on the picture…". And so that's what we're talking about… that it comes with the exam routine that if it is reading comprehension then you cannot take any other source into consideration, only the reading task's source text. This is a very hard learning process. It's our job to teach students this exam English. (Paul)

Although suggestions were made for improving these visuals, such as enlarging them and adding colour, the teachers also acknowledged that doing so would be expensive and inefficient. Additionally, it was noted that the photographs could be removed entirely as they do not actually enhance the exam tasks.

> Unnecessary. And a waste, I mean waste of ink. I mean, I know that they want to spice up the text a bit, and that's how it normally works, but it's a final exam, so they're not going there to play or to have fun, they want to pass an exam. I'm not sure that's the way to entertain students. I don't think anyone's going to look at it. (Anthony)

Task Instructions

The instructions for the test tasks were generally well-received. The majority of the instructors expressed satisfaction with the instructions, finding them understandable and straightforward. They emphasized the value of strategically using italicization and underlining to attract attention to crucial features and aid student comprehension. A few responders did point out that the instructions were overly lengthy and that students often skipped over reading them completely. In the context of the listening exam paper, this issue was seen to be especially significant. One suggested solution for this issue was the introduction of clearer instructions.

> They're very long, very long and it's exactly because they're so damn long that the kid ignores them and you have to "beat it" into them that they should not, because there can be a moment when a new type of task comes up in the exam, which we didn't know about […], it should be shorter, I understand that again it is there so that there is something to refer to, but the student loses attention, especially the one who has practiced it, because if we are going to work on it, he knows that yes, ABC and that's all he will start with, because he will read only that, although you do not really need to know more than that, but sometimes you do. (Paul)

The teachers agreed that having the instructions in English was beneficial and that it is not the fault of the exam papers if the students chose not to read them. Additionally, they mentioned that pupils frequently fail to pay attention to and follow instructions despite the fact that teachers review them together with the students in class and continually repeat them.

Task Types

The test specifications for each part of the exam list a wide range of task types that can be included; however, there are only a handful of them that are used each year, with an occasional outlier. For example, in the past six years (2017-2023) there has been an error correction task at the end of each use of English test paper (Educational Department, 2023a). Similarly, there has been a true, false, not stated type task present in the reading papers over the same period. When asked about the task types, the interviewees focused mostly on these typically occurring task types in each part of the exam and generally agreed that these tasks function as intended. However, they did not mention the task types that are present in the test specifications but are not or seldom used in the tests. Lastly, the school leaving examination was criticized for not being very realistic or reflecting real-life language use; nevertheless, the instructors also admitted that such a test would be challenging to standardize.

Reading

When asked about the task types in the reading paper, the interviewees shared mixed views. They generally viewed the reading tasks as the least problematic and most effective. Although several of them stated that students frequently struggle with this activity, all of the participants commented positively on the summary exercise. The true, false, and not stated task type was also mentioned as challenging. While true or false activities often do not pose major problems for students, their difficulty can increase when the Not stated choice is added to the possible answers. This is because students often read too deeply into the content or have different viewpoints from those of the test designers. As a result, they find it difficult to distinguish between incorrect and unmentioned facts.

> What I think is a hateful thing, and they don't like it either, and I explain it very badly, is the true, false, not stated. That they shouldn't want to be smarter than the task. So that the not stated is really not stated because it is not in it, and that I should not think that it says this, and a statement

implies this because you can infer it. And they suffer greatly from this because they obviously read in a way that they want to infer underlying content. (Hannah)

Lastly, the sentence part matching task was mentioned as potentially problematic as many of its items may be solved by utilizing cohesive devices and grammatical understanding, suggesting that it may not only test reading but also use of English.

Use of English

This section of the test paper received the most criticism from the participants. The views of the participants reflected that this part is outdated, "ill willed" (Jane), should not exist as a separate section, and should be removed as a whole.

> I would take it out completely as it is from the exam. Why? Well, because I think that part of language knowledge is that I use it, so I'm not sure that a student should know how to be analytical language users. Maybe he doesn't even know his mother tongue like that. (Paul)
>
> This is the most problematic task. On the one hand, because let's say it's past its time. At the same time, I understand why it should be included, since theoretically if someone wants to continue his studies in a direction where they need grammar stuff like English studies, American studies, translation, interpretation, then you need to have this knowledge [...]. (Anthony)

Nonetheless, some positive aspects of the tasks were also mentioned. For example, the word formation task was mentioned as one that is easy to understand, and students can prepare for.

> But otherwise, I don't mind that it's included. I think they make the most of it because there are a lot of different tasks. Because it comes from the text. Because it actually highlights things really well that the student should be able to deduce from contextual cues. [...] And that I already know what kind of problems they have, so I can prepare them. (Hannah)

Nevertheless, the majority of their remarks were unfavorable. The teachers shared the opinion that expecting the students to know the language analytically at the B2 level might be too demanding, and that these tasks require knowledge that the test takers do not use consciously while using the language.

> Professionally speaking, it is again very unnatural that children have to think about things that they don't think about at all when using spontaneous language. They don't use these things, so it makes them feel unsure. It is a malicious, not language teaching friendly view, this intentional misleading and trying to trick them. (Jane)

There were other issues that were task-specific, such as the word formation tasks being overly demanding while only awarding one point, or the error correction task being awkwardly edited with line breaks that do not make sense to the reader. One participant suggested including a sentence transformation task, as it would be a more realistic task and something students frequently use to express themselves.

Listening

Even though attitudes towards the exercises were mostly neutral, the listening tasks did not garner much appreciation. Similar issues to those described in the reading papers arose in regard to the true, false, not mentioned type task, and the two words missing exercises also drew criticism for being unfair to pupils and testing precise hearing rather than listening comprehension. One participant stated that "this is a parrot task. If we only accept the words said, we only have to repeat things, even though there are a few thousand other possibilities for the same thing" (Anthony).

As for listening tasks in which students have to complete sentences with important missing information, the participants expressed frustration regarding the vagueness of the task. It was stated that oftentimes the instructions do not specify the amount of words that could be written as the answer, which can cause problems for the students as they might place importance on the same things that the task makers do. Some rarely oc-

curring task types, such as the matching pictures based on the audio task, were claimed to require more than just listening comprehension, involving integrated skills.

Source Texts

Regarding the source texts provided for each task in the different sections, the participants were mostly satisfied. They found the texts to be sufficiently long, and it was expressed that even if they are a bit longer than what is typically seen in language tests at the B2 level, students still have enough time to read them. The majority of teachers tended to agree that the texts were either at B2, occasionally B2+, or C1 level, with the listening texts displaying the most variety. According to their observations, the subjects covered in these texts are not very stimulating or interesting to the students, but from the viewpoint of the teachers, the selection of the source texts has been improving over the past few years. Nonetheless, some interviewees stated that the content might not always be appropriate for 16-18-year-old students.

Other remarks concerning the source texts were student- and test-paper-specific. Each participant agreed that the length of the texts was appropriate for the reading and use of English tasks, but many of them also noted that students' attention spans are generally shorter these days, that they often find reading boring, and that they frequently fail to concentrate on, realize, or understand the content of what they are reading. This is particularly true for the use of English tasks, during which students frequently concentrate on the gaps and skim the texts. In case of the listening, most participants mentioned the presence of music before and after each task. Opinions about it were divided, with some participants believing it to be beneficial for reducing the students' stress levels and making the tasks friendlier, while others viewed it as unnecessary, distracting, or even stress-inducing during the exam, when the students are already highstrung and hyper-focused on completing the task in front of them.

> They are always playing music. And it's so controversial. Because on the one hand it eases the tension, so there's a niceness to it. On the other hand, when

> the child is on pins and needles and the task should come already, he has his whole brain, eyes and ears focused on the task, and then instead of playing the task, they start playing music instead. Which has a certain charm and kindness to it, actually. So it's quite controversial, is it good or not? (Jane)

The level of noise and other distractions in the audio, as well as the authenticity of the recordings, were two additional contentious aspects of the source texts. According to several participants, the audios are excessively sterile, and they expressed preference for the inclusion of more natural noises.

> …there's no such thing as a sterile environment. You're talking in a restaurant, in a nursery, in a pub, wherever. So, it's not like someone is going to speak into the microphone nicely pronouncing things. So, really let's record a conversation at a football match with a microphone, so that the background noise just seeps through, so that the conversation can still be heard. Because that's what you'll encounter in real life. (Paul)

Others expressed appreciation for the variety of accents featured in the exercises and reported greater satisfaction with the recording quality. The fact that the recordings are segmented during their second playback was cited as another advantage of the recordings.

Scoring Keys
Generally, the interviewees found the scoring keys to be clear and easy to work with. However, they raised concerns in three specific areas: (1) the scoring of open-ended questions, (2) judging spelling mistakes, and (3) score calculation. All of the participants expressed some level of frustration regarding the correction of open-ended questions, as the scoring key often does not include all possible acceptable solutions. In such cases, a scoring key modification can be issued. Such alterations of the possible solutions do not only occur with tasks such as gap filling which can have two to three acceptable answers, but also in the case of true, false, not stated tasks as well, which raises concerns.

> It's hard to follow beyond a point, but that many times even in the true, false, not stated task a correction comes, in the reading and also in the listening. The sloppiness is astounding in my opinion. They can't afford that, that there is a task, and the child has to clearly decide, and then the exam board finds out that they can't decide either... (Hannah)

Since the teachers are tasked with correcting a large volume of tests in a short amount of time, such constant modifications can be frustrating. This increases their workload, as they need to revisit already-corrected papers multiple times. Although the scoring manual states that spelling errors "will only be taken into account if they obscure the clarity of the answer," one of the participants observed that this might be problematic.

> However, you have to communicate this clearly to the correcting teachers and we are often not able to decide. How many spelling mistakes can be there that would still leave the answer recognizable? If someone, let's say you, have already made 3 spelling mistakes. That's a big problem. Something goes wrong or by accident the letter o was left out, but if there is one more mistake then it is not recognizable and then the problem arises that for a teacher everything is recognizable. So, if the student writes only the vowels, it is still recognizable, because that is how we have been socialized as teachers. (Emily)

Thus, the decision about how many spelling mistakes can be accepted is ultimately determined by the scoring committee heads, whose strictness may vary: one might only accept two spelling mistakes, while another may permit three or more.

Lastly, many comments were directed at the score conversion table included in the scoring key and the discrepancy between test scores and exam scores. Students earn test scores in each task which are later on converted to exam scores that determine their final grade. As an example, in the 2023 May exam, the use of English section had 35 test points and 30 exam points (Educational Department, 2023a). Since there are more test points than exam points, some test points are converted to the same exam

points. For example, the test scores of those who earned 24 or 25 test points are converted to 21 exam points. This raises questions of equity, as achieving a lower test score (24 test points in the previous case) can result in the same exam score (21 exams points) as achieving a higher test score (e.g., 25 test points). Additionally, the interviewees noted that this conversion can be sometimes disorienting and more labor-intensive than a straightforward scoring system without the need of conversion.

Conclusion

This study aimed to explore the attitudes of Hungarian high school EFL teachers towards the objectively scored components of the advanced level English school-leaving examination. In order to achieve this, qualitative interviews were carried out with five high school EFL teachers. The teachers displayed positive attitudes towards the content of the source texts and certain task types, such as the summary task in the reading paper and the word formation task in the use of English paper. They were also largely satisfied or neutral towards the general information given at the beginning of each test paper, the task instructions, the length of the source texts, and the diversity of accents in the audio recordings as well as their quality. Additionally, the format and design of the test papers were mostly viewed as suitable and were even praised for their layout. Nevertheless, some found the test papers too lengthy and wasteful in their paper use, and the compulsory identification number on top of each page was also criticized as redundant.

Based on the interviews, the most critical aspects of the exams were the accompanying pictures, the task types, and the scoring key. The pictures accompanying the tasks were deemed useless by the participants, as their size, quality, color, and often their imagery were largely seen as unhelpful the students or even misleading. Although the interviewees understood the intended use of these pictures, based on their own perceptions and their experiences with students they viewed them as unnecessary. The task types that garnered the most comments were those featured in the use of English and listening paper. The reading paper was generally

found to be less problematic, although questions arose as to whether the matching tasks in the test measured reading or use of English skills. The listening task paper also received some criticism in regard to the sentence completion task type in which no word limit was given and the gap filling task, in which students need to write in two words that they heard. The former was thought to cause difficulties during rating, while the latter was questioned regarding its fairness and usefulness as a listening comprehension task. The least favored task paper was the use of English, which was viewed as outdated and requiring analytical skills that the teachers considered to exceed the B2 level. Most teachers suggested either removing it entirely or at least replacing it with different task types, such as sentence transformation. Lastly, three issues were mentioned concerning the scoring keys. The first was the difference between exam scores and test scores, as well as how they were calculated. This discrepancy was deemed troublesome, potentially inequitable, and perplexing, as teachers do not receive information or training on why the scores are calculated in this manner. The other two problems were the incomplete list of solutions and the frequent modifications to the keys. While the teachers acknowledged that it is likely impossible to include every possible and acceptable solution for the gap filling type tasks, they called for increased rigor in designing the scoring keys to reduce the number of modifications and the subsequent rescoring of test papers. Moreover, these frequent alterations often affect the task types where the solution should be definite, such as the true, false, not stated task, which point towards broader problems in assessment design given that these tests are considered objective.

The results of this study should be viewed in the light of its limitations, which concern both the number of participants and the interview guide. Due to time constraints and availability, only five out of seven participants could be included in this study. In addition, although the participants were presented with some of the test papers and scoring keys during the interviews, they mostly had to rely on their memory.

For future research, it would be beneficial to interview more teachers to gain an even more nuanced picture of their perspectives. Additionally, the selection of sample tests used during the interviews could be revised

to better represent the positive and negative aspects of these test papers. Finally, research on the identified issues from the students' perspective could also prove valuable in the design of the advanced level English school-leaving examination.

References

Clarke, V., & Braun, V. (2014). Thematic analysis. In Teo, T. (Eds), *Encyclopedia of critical psychology*. Springer. https://doi.org/10.1007/978-1-4614-5583-7_311

Coombe, C. (2018). *An A to Z of second language assessment: How language teachers understand assessment concepts*. British Council.

Council of Europe (2020). *Common European framework of reference for languages: Learning, teaching, assessment. Companion volume*. Council of Europe.

Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.

Educational Department (2023a). Központi írásbeli feladatsorok, javítási-értékelési útmutatók. [Centralized written test papers, scoring instructions]. https://www.oktatas.hu/kozneveles/erettsegi/feladatsorok

Educational Department (2023b). Statisztikák, vizsgaeredmények. [Publicly available statistical data and school leaving examination results]. https://www.ketszintu.hu/publicstat.php

Einhorn, Á. (2009). *Az idegen nyelvi érettségi vizsga reformja*. [The reform of the foreign language school-leaving examination]. Oktatáskutató és Fejlesztő Intézet.

Howitt, D. (2016). *Introduction to qualitative research methods in psychology* (3rd ed.). Pearson.

Hungarian Government (1997). A Kormány 100/1997. (VI. 13.) Korm. rendelete az érettségi vizsga vizsgaszabályzatának kiadásáról. [Government Decree 100/1997. (VI. 13.) on the issuance of the examination regulations for the school leaving examination.] https://net.jogtar.hu/jogszabaly?docid=99700100.kor

Hungarian Government (2002). OM rendelet az érettségi vizsga részletes követelményeir l [(Ministry of Education) on the Specific Requirements of Upper Secondary School Leaving Examinations] Ministerial Decree 40/2002. (V. 24.). http://njt.hu/cgi_bin/njt_doc.cgi?docid=65297.357363

Hungarian Government (2011). 2011 évi CXC. törvény a nemzeti köznevelés-ről [Act CXC of 2011 on national public education]. https://njt.hu/jogszabaly/2011-190-00-00

Hungarian Government (2020). *A Kormány 5/2020. (I. 31.) Korm. Rendelete a Nemzeti alaptanterv kiadásáról, bevezetéséről* és *alkalmazásáról szóló 110/2012. (VI. 4.) Korm. rendelet módosításáról* [Government decree 5/2020. (I.31.) on the modification of Government decree 1010/2012. (VI.4.) on the issuing, introduction, and application of the new National Curriculum]. https://magyarkozlony.hu/dokumentumok/3288b6548a740b9c8daf918a399a0bed1985db0f/megtekintes

Hungarian Government (2021). A Kormány 65/2021. (II. 15.) Korm. Rendelete egyes köznevelési és szakképzési tárgyú, valamint a jelnyelvoktatói névjegyzék vezetésével összefügg kormányrendeletek módosításáról. [Government Decree amending the government decrees on public education, vocational training, and management of the list of sign language teachers.] Magyar Közlöny, 2021 február 15.

Hungarian Government (2022). A Kormány 100/1997. (VI. 13.) *Korm. rendelete az érettségi vizsga vizsgaszabályzatának kiadásáról* [Government Decree 100/1997. (VI. 13.) on the issuance of the examination regulations for the school leaving examination]. https://njt.hu/jogszabaly/1997-100-20-22

Király, Z. (2012). Az angol érettségi „hallott szöveg értése" feladatának gyártási technológiája. *Modern Nyelvoktatás*, *18*(3), 55–66. [The production technology of the listening comprehension task of the English school-leaving examination].

Ministry of Education (2002). *40/2002. (V. 24.) OM rendelet az* érettségi *vizsga részletes követelményeiről* [Ministry of Education Decree on the detailed requirements of the secondary school leaving examination]. https://njt.hu/jogszabaly/2002-40-20-45

Tankó, G., & Andréka, Z. (2021). Probing the advanced level EFL school-leaving examination: The Use of English paper. In G. Tankó & K. Csizér, (Eds.), *DEAL 2021: Current explorations in English applied linguistics*

(pp. 65–105). Eotvos Lorand University. https://edit.elte.hu/xmlui/handle/10831/62555

Tankó, G., & Andréka, Z. (2022). Probing the reading paper of the advanced level EFL school-leaving examination. In G. Tankó & A. M. Wind (Eds.), *DEAL 2022: Challenges and opportunities in contemporary English applied linguistics*, (pp. 23–57). https://doi.org/10.21862/ELTE.DEAL.2022.2.23

Vígh, T. (2012). A kétszintű idegen nyelvi érettségi tanítási-tanulási folyamatra gyakorolt hatása: Egy tanári kérdőíves vizsgálat eredményei. [The impact of the two-level foreign language school-leaving examination on the teaching-learning process: results of a teacher questionnaire survey]. *Iskolakultúra, 22*, 62–79.

Vogt, K., & Tsagari, D. (2022). Language assessment literacy and test validation: Highlighting the role of language teachers. *Studies in Language Assessment, 11*(1), 1–15. https://doi.org/10.58379/UQVQ7101

Appendix

Interjú protokoll

Az interjú célja, hogy többet megtudjunk az emelt szintű angol érettségi vizsga objektíven értékelt részeiről (Olvasásértés, Nyelvhelyesség, Hallásértés) a középiskolai angoltanárok tapasztalatai és véleménye segítségével. Fel fogok tenni néhány kérdést, és szeretném, ha a feladatokkal kapcsolatos, pl. tanórai felhasználásuk vagy az érettségi vizsga felkészítés során gyűjtött tapasztalatai alapján válaszolna rájuk. Kérem, gondolatait ossza meg nyíltan és őszintén, ugyanis azok segíteni fognak nekem a kutatási kérdéseim megválaszolásában, és a feladatfejlesztési munka minőségét javíthatják.

A kérdések nyílt végűek, és nincsenek helyes vagy helytelen válaszok. Bármikor abbahagyhatja az interjút, és nem kell válaszolnia olyan kérdésekre, amelyek kényelmetlenül érintik. Szeretném biztosítani Önt arról, hogy ez az interjú bizalmas, és az Ön személyazonossága névtelen marad, ezért nyugodtan válaszoljon a kérdésekre.

Az interjú körülbelül 40 percig fog tartani. Nagyra értékelem önkéntes részvételét a kutatásomban, és szeretném megköszönni, hogy időt szakított rám. Az adatok elemzése és a jobb adatminőség érdekében az Ön előzetes beleegyezésével az interjút rögzítem.

Dátum:

Interjúztató:

Formátum: Személyes / Online

Interjú kezdte:

Interjú vége:

Interjúkérdések:

A. Bevezető kérdések

Megerősítené, hogy nem bánja, ha az interjút rögzítem? / Beleegyezik, hogy az interjút rögzítsem?

1. Hány éves?
2. Milyen (tanári) végzettsége van?
3. Mióta tanít angol nyelvet?
   - Mióta tanít angol nyelvet középiskolában?
4. Tanít más tantárgyat is az angolon kívül?
   - Mióta tanítja ezt/ezeket a tantárgy(akat)?
5. Tanít angol fakultációs órákat?
   - Heti hányszor?
   - Melyik évfolyamoknak?
   - Mióta?
6. Szokott érettségiztetni?
   - Milyen szint(ek)en?
   - Mióta?
   - Mennyire ismeri az emelt szintű angol érettségi tartalmát és formátumát?
   - Érettségiztetés előtt hogyan frissíti ismereteit az emelt szintű angol érettségiről?

B. Érettségihez kapcsolódó kérdések: A következő részben az emelt szintű érettségivel kapcsolatban fogok feltenni pár kérdést.

7. Véleménye szerint melyek az angol emelt szintű érettségi vizsga erősségei és gyengéi?
   - Ön szerint ezek hogyan befolyásolhatják a diákok nyelvtanulásának kimenetelét, eredményességét?
8. Ön szerint, milyen a teszttel kapcsolatos nehézségekbe ütköznek a diákok az emelt szintű érettségi felkészülés közben? Ilyen esetekben mit tesz, hogy segítse a diákokat?

9. A nyelvtanári karriere során tapasztalt-e bármilyen változást vagy trendet a diákok hozzáállásában az emelt szintű angol érettségi vizsgához, különösképpen az elmúlt 4-5 év során?
   Mesélne erről bővebben?

10. A nyelvtanári karriere során tapasztalt-e bármilyen változást vagy trendet a tanárok hozzáállásában az emelt szintű angol érettségi vizsgához, különösképpen az új két szintű érettségivel kapcsolatban?
    Mesélne erről bővebben?

C: Az egyes részvizsgákhoz kapcsolódó kérdések: A következő részben az emelt szintű angol érettségivel, pontosabban az olvasott szöveg értése, nyelvhelyesség, és hallásértés feladatlapokkal kapcsolatban fogok feltenni pár kérdést.

11. Mi a véleménye az Olvasott szöveg értése feladatlapról?
    a) Mit gondol a feladatlapon a vizsgázóknak adott tájékoztató információkról?
    b) Mit gondol a feladatlap formátumáról/kinézetéről?
    c) Mit gondol a feladat típusokról?
    d) Mit gondol a feladatok instrukcióiról?
       - Mennyire érthetőek az instrukciók a diákok számára?
       - Tapasztalatai szerint hogyan használják a diákok az instrukciókat?
    e) Mi a véleménye a feladathoz tartozó grafikákról?
       - Ön szerint mi a szerepe a grafikáknak ebben a feladat típusban?
       - Tapasztalatai szerint a diákok mennyire figyelnek oda a grafikákra? Hogyan használják azokat?
       - Hogyan lehetne az instrukciókon javítani, hogy a diákokat jobban segítsék?
       - Van bármilyen egyéb instrukciókkal kapcsolatos észrevétele?
    f) Mit gondol a feladat megoldásához elolvasandó szövegekről?
       - Mi a véleménye a szövegek tartalmáról?
       - Mi a véleménye a szövegek nehézségi szintjéről?
       - Mi a véleménye a szövegek terjedelméről?
    g) Milyen tapasztalatai vannak a javítási-értékelési útmutatóval kapcsolatban? Milyen az útmutatóval dolgozni?

- Mi segít?
- Mi okoz nehézségeket?
- Hogyan lehetne az útmutatót javítani, hogy az értékelő tanárok munkáját segítse?

12. Mi a véleménye az Nyelvhelyesség feladatlapról?
    a) Mit gondol a feladatlapon a vizsgázóknak adott tájékoztató információkról?
    b) Mit gondol a feladatlap formátumáról/kinézetéről?
    c) Mit gondol a feladat típusokról?
    d) Mit gondol a feladatok instrukcióiról?
       - Mennyire érthetőek az instrukciók a diákok számára?
       - Tapasztalatai szerint hogyan használják a diákok az instrukciókat?
    e) Mi a véleménye a feladathoz tartozó grafikákról?
       - Ön szerint mi a szerepe a grafikáknak ebben a feladat típusban?
       - Tapasztalatai szerint a diákok mennyire figyelnek oda a grafikákra? Hogyan használják azokat?
       - Hogyan lehetne az instrukciókon javítani, hogy a diákokat jobban segítsék?
       - Van bármilyen egyéb instrukciókkal kapcsolatos észrevétele?
    f) Mit gondol a feladat megoldásához elolvasandó szövegekről?
       - Mi a véleménye a szövegek tartalmáról?
       - Mi a véleménye a szövegek nehézségi szintjéről?
       - Mi a véleménye a szövegek terjedelméről?
    g) Milyen tapasztalatai vannak a javítási-értékelési útmutatóval kapcsolatban? Milyen az útmutatóval dolgozni?
       - Mi segít?
       - Mi okoz nehézségeket?
       - Hogyan lehetne az útmutatót javítani, hogy az értékelő tanárok munkáját segítse?

13. Mi a véleménye az Hallásértésértés feladatlapról?
    a) Mit gondol a feladatlapon a vizsgázóknak adott tájékoztató információkról?
    b) Mit gondol a feladatlap formátumáról/kinézetéről?
    c) Mit gondol a feladat típusokról?

d) Mit gondol a feladatok instrukcióiról?
- Mennyire érthetőek az instrukciók a diákok számára?
- Tapasztalatai szerint hogyan használják a diákok az instrukciókat?

e) Mi a véleménye a feladathoz tartozó grafikákról?
- Ön szerint mi a szerepe a grafikáknak ebben a feladat típusban?
- Tapasztalatai szerint a diákok mennyire figyelnek oda a grafikákra? Hogyan használják azokat?
- Hogyan lehetne az instrukciókon javítani, hogy a diákokat jobban segítsék?
- Van bármilyen egyéb instrukciókkal kapcsolatos észrevétele?

f) Mit gondol a feladathoz tartozó hallott szövegekről?
- Mi a véleménye a szövegek tartalmáról?
- Mi a véleménye a szövegek nehézségi szintjéről?
- Mi a véleménye a szövegek terjedelméről?
- Mit gondol a felvételek minőségéről?
- Van bármilyen egyéb szövegekkel kapcsolatos észrevétele?

g) Milyen tapasztalatai vannak a javítási-értékelési útmutatóval kapcsolatban? Milyen az útmutatóval dolgozni?
- Mi segít?
- Mi okoz nehézségeket?
- Hogyan lehetne az útmutatót javítani, hogy az értékelő tanárok munkáját segítse?

Van még valami, amit szeretne megemlíteni vagy kérdezni, mielőtt befejezzük az interjút? Köszönöm szépen, hogy részt vett a kutatásban. Ha érdeklik a kutatási eredményeink, elküldhetjük az Ön e-mail címére.

10

Variation in Syntactic Complexity Between Natural and Social Science Disciplines: A Corpus-Based Investigation

Attila M. Wind
wind.attila@btk.elte.hu

Abstract

Measuring second language (L2) writing development has been facilitated by rapid advancements in technology over the past few years. Numerous studies have investigated the development of syntactic complexity in L2 writing, using the Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC; Kyle, 2016) to calculate various indices. The TAASSC uses the Corpus of Contemporary American English (COCA) as a proxy to produce indices of syntactic sophistication. However, the COCA is a general corpus that includes research articles from a wide range of academic disciplines. Previous studies have shown that there is variation in academic text construction (e.g., Hyland, 1999, 2008; Hyland & Tse, 2007) and syntactic complexity across academic disciplines (e.g., Casal et al., 2021; Lu et al., 2021). Despite this, no previous studies have investigated if there is variation in syntactic complexity between natural and social science articles. To address this gap, two corpora, one for natural sciences and one for social sciences, were designed and a total of 14 indices were calculated with the L2 Syntactic Complexity Analyzer (Lu, 2010). Statistically significant differences were found between the natural and social science corpora in 10 indices, including the mean length of sentences and T-units, clauses per sentence, verb phrases and clauses per T-unit, subordination, Complex T-unit ratio, coordinate phrases per T-unit, and Complex nominal per T-unit indices.

Keywords: *syntactic complexity, corpus linguistics, natural and social science corpus, TAASSC*

Variation in Syntactic Complexity Between Natural and Social Science Disciplines: A Corpus-Based Investigation

In the past decade, there has been a rapid increase in the number of studies employing computational tools such as the Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TAASSC) to investigate second language (L2) writing development. For example, a cross-sectional study by Kyle and Crossley (2017) found that verb-argument construction (VAC) indices are more reliable indicators of writing proficiency than structure-based indices of syntactic complexity such as the mean length of T-units. In a later study, Kyle and Crossley (2018) found that fine-grained indices of phrasal complexity were better indicators of writing quality than traditional or fine-grained clausal indices. From a usage-based perspective, Kyle et al. (2021) explored developmental trajectories of L2 learners of English over two academic years, focusing on syntactic complexity and VAC sophistication. The indices in Kyle and Crossley (2017, 2018) and Kyle et al. (2021) were calculated with the TAASSC, with several VAC indices of syntactic sophistication in the TAASSC derived from the Corpus of Contemporary American English (COCA).

The COCA contains more the one billion words of texts from eight genres, including academic texts. However, as Kyle (2016) noted, a possible limitation of the sophistication indices is the utilization of COCA as a proxy to measure L2 writing development. As the COCA represents general English language use in the United States (Davies, 2009, 2010), it is likely that the COCA does not entirely represent the different types of language that L2 learners are exposed to, especially in countries other than the USA. Kyle (2016) also noted that a corpus that includes the types of language that language learners are more frequently exposed to might serve as a more reliable proxy for language experience and, consequently, may provide more robust and representative findings. Apart from Kyle's (2016) criticism, there is another possible limitation of the use of COCA as a proxy. The research articles (RAs) in the COCA were selected from more than 200 different peer-reviewed journals; although the RAs were chosen to cover the whole range of the Library of Congress classification system, different academic

disciplines have unique discourses (Hyland, 2004), and variation has been observed in syntactic complexity across disciplines (e.g., Casal et al., 2022).

In this study, corpora for two academic disciplines (i.e., natural and social sciences) were created in order to investigate the above issues more closely. The first corpus consists of academic RAs collected from natural science journals, while the second corpus contains RAs from social science journals. In order to understand the advantages of using specialized corpora for the researcher, it is important to understand the limitations of using the COCA as a proxy for analyzing L2 writing developmental.

Corpus of Contemporary American English

The Corpus of Contemporary American English (COCA) is one of the 17 core corpora provided by English-Corpora.org, a collection of 19 corpora from Mark Davies at Brigham Young University (BYU; Davies, 2023), formally known as the BYU Corpora. Among these, COCA is ranked sixth in terms of size. The largest corpus is the News on the Web corpus (NOW; 17,6 billion words), followed by the iWeb: the Intelligent Web-based Corpus (14 billion words), the Global Web-Based English corpus (GloWbE; 1,9 billion words), the Wikipedia Corpus (1,9 billion words), and the Coronavirus Corpus (1,5 billion words). However, unlike COCA, the iWeb, NOW, GloWbE, Wikipedia Corpus, and Coronavirus Corpus represent a specific type of language use and cover limited time periods. For example, the NOW contains texts from 2010, the iWeb includes text solely from 2017, the GloWbE contains text from 2012 to 2013, the Wikipedia Corpus includes texts from 2014, while the Coronavirus Corpus is inherently limited to texts between January 2020 to December 2022. When representativeness is considered, the largest corpus is the COCA, followed by the Corpus of Historical American English (COHA; 475 million words), the British National Corpus (BNC; 100 million words), and the Strathy Corpus (Canada; 50 million words).

According to Davies (2023), a representative corpus should meet three criteria. First, the corpus must contain texts from a wide range of genres. Second, the corpus must be large in terms of size (i.e., approximately one billion words). Third, the corpus should be recent; more precisely, texts should be drawn from sources composed between 1990 and 2023. It is im-

portant to note that Davies' (2023) criteria are limited to synchronic corpora. The COCA closely adheres to Davies' (2023) criteria, containing approximately 1 billion words, covering a period of nearly thirty years (1990–2019), and it is balanced (i.e., a wide range of genres is included in the corpus and their proportion). However, the COCA includes only one dialect, American English, as opposed to the NOW which contains 20 dialects.

Designed by Mark Davies and released online in 2008, the COCA is a constantly growing corpus with 24–25 million words added to it each year. In 2009, it was composed of more than 385 million words (Davies, 2009), while as of 2023 it is composed of more than one billion words from 485,202 texts (Davies, 2023). The corpus is evenly divided between eight different registers: (1) TV/movies, (2) spoken language, (3) fiction, (4) magazine, (5) newspaper, (6) academic language, (7) web/blog, and (8) web/general. In addition, the COCA contains texts from blogs and websites. Each year new words are added to the COCA while maintaining the same balance between genres. Despite its breadth, Egbert et al. (2022) criticized the COCA for having poor representativeness compared to the British National Corpus (BNC), defining representativeness as the extent to which a corpus can represent real-world language use. Davies (2023); however, disagreed with Egbert et al. (2022), claiming that the representativeness of the COCA is equal to or greater than that of the BNC. In 2023, the average word count for the eight genres in the COCA was 125,361,219 words, with the genre 'fiction' containing the fewest words (119,505,305 words), while the genre 'web/general' had the most (129,899,427 words).

Egbert et al. (2022) specifically criticized the representativeness of the academic sub-corpus of the COCA, claiming that it was not as academic as the BNC. Egbert et al. (2020) listed two main problems with the academic sub-corpus of the COCA. First, linking adverbials such as *however*, *thus*, *therefore* etc. are less frequent in the academic sub-corpus of the COCA than in the academic sub-corpus of the BNC, despite linking adverbials being highly characteristic of the academic genre. For example, the linking adverbial 'however' has a normed frequency of 1,220 in the BNC compared to 890 in the COCA, while the linking adverbial 'thus' has a normed frequency of 551 in the BNC compared to 477 in the COCA. Egbert et al. concluded

that, in this sense, the COCA could be considered defective. According to Davies (2023), Egbert et al. (2020) contested this comparison for two main reasons: (1) dialect and (2) time. The COHA shows that linking adverbials such as *however*, *thus*, and *therefore* were more frequent in American English thirty years ago, when the BNC was designed in 1991, than now. In addition, the COCA demonstrates a steady decrease in the frequency of the use of linking adverbials from 1990 to 2019, while the GloWbE shows that *linking adverbials* are more frequent in British than in American English. The second main criticism of Egbert et al. (2020) is related to the use of nominalizations, a characteristic feature of academic writing (e.g., Biber et al., 2011). The frequency of nominalization is 35,613 tokens in the academic sub-corpus of the BNC, compared to 33,636 tokens in the academic sub-corpus of the COCA. Egbert et al. (2020) concluded that, in this sense, the COCA can be seen as inferior to the BNC; however, Davies (2023) argued that the difference of 6% is negligible. In addition, the COHA demonstrates that the frequency of nominalizations has declined in American English since the 1990s.

The academic sub-corpus of the COCA is drawn from RAs from more than 200 different peer-reviewed journals covering a wide and balanced range of academic disciplines including the social sciences, education, history, law, humanities, medicine, religion/philosophy, science/technology, and business. The size of the academic sub-corpus of the COCA is 122,988,361 words consisting of 26,137 texts. The journals were selected to cover the full range of the Library of Congress Classification (LCC). The LCC, developed by the Library of Congress in the US, is a system of library classification that is widely used for shelving books in libraries. In the past decade, there has been an increase in the utilization of computational tools that use the COCA as a proxy for studying L2 writing development. One of these computational tools is the TAASSC.

Tool for the Automatic Analysis of Syntactic Sophistication and Complexity (TASSC)

The TASSC is a free computational tool designed for researchers, educational experts, and teachers to gauge a wide range of fine-grained indices of

syntactic complexity and sophistication. Although the TAASSC is mainly used to measure L2 writing development (Kyle & Crossley, 2017, 2018; Kyle et al., 2021; Park & Yoon, 2021; Thongyoi & Poonpon, 2020; Wind, 2020), it can also be used to trace L1 writing development (Alsahlanee & Jaganathan, 2023). The TAASSC offers a user-friendly graphical user interface and works on widely used operating systems such as Mac, Windows, and Linux. The TAASSC allows for batch processing, and data can be stored on the hard drive of the user, facilitating the storage and processing of sensitive data. In the TAASSC, plain text files are used as input, and it is recommended to store all the text files to be analyzed in a designated folder. The output is a comma-separated values (.csv) spreadsheet that can be opened by commonly used software such as Excel.

Four different groups of indices are available in the TAASSC and can be categorized according to the kind of parser used for their operationalization. The first group of indices is calculated by the L2 Syntactic Complexity Analyzer (L2SCA; Lu, 2010), which in turn uses the Stanford Parser (Klein & Manning, 2003) and Tregex (Levy & Andrew, 2006) to calculate the indices. The former creates a constituency representation of each sentence in a specific text, while the latter is employed to identify specific patterns in that representation. The remaining three groups of indices use the Stanford Neural Network Dependency (SNND) parser (version 3.5.1; Chen & Manning, 2014) and a Python XML parser to calculate structures and include fine-grained clausal complexity, fine grained phrasal complexity, and syntactic sophistication indices. The SNND parser can be considered highly accurate and fast, producing a dependency representation of each sentence that includes several functional categories (e.g., subject and direct object). The SNND outputs data in XML format as output, and an XML parser in Python is employed to read the file.

The TAASSC also includes 14 indices from the L2SCA as well as 31 fine-grained clausal complexity, 17 fine-grained phrasal complexity, and 15 syntactic sophistication indices. The last group of indices uses the COCA as a proxy and is rooted in usage-based theories of language development (Ellis, 2002; Goldberg, 1995; Langacker, 1987). In the TAASSC, the 15 syntactic sophistication indices include several variations, leading to a total of

38 different indices. These are calculated in reference to five sub-corpora (all written, academic, fiction, magazine, and newspaper) in the COCA. As such, the total number of indices is 190 (38 indices multiplied by the five different genres). Kyle (2016) argued that the COCA is more appropriate to use as a proxy than the BNC, citing its larger size and the inclusion of more recent texts. In addition, for research focused on American English, the COCA represents a more adequate proxy for language use in the American context than the BNC.

Studies Using the TAASSC

Since the launch of the TAASSC, it has been used in several studies to measure L2 writing development (e.g., Kyle & Crossley, 2017). One of the earliest studies using the tool was conducted by its creator himself, Kristopher Kyle, along with his supervisor Scott Crossley in 2017. Kyle and Crossley (2017) criticized previous studies on L2 writing development, claiming that over the past 50 years L2 writing syntactic sophistication had been measured from the perspective of absolute complexity (Bulté & Housen, 2012; Lu, 2011; Ortega, 2003; Wolfe-Quintero et al., 1998), defined by Bulté and Housen (2014) as the "objective inherent properties of linguistic units and/or systems thereof" (p. 43).

In contrast, Mazgutova and Kormos (2015) argued that L2 writing development should be measured from a relative complexity perspective. Relative complexity is defined by Bulté and Housen (2014) as "the cost and difficulty of processing" (p. 43). Mazgutova and Kormos (2015) claimed that the complexity of learners' output should be measured with reference to the mode, genre, and communicative demands of the specific task performed by the learner. Interestingly, although the authors argued for the measurement of relative complexity, in their 2015 study they used complexity indices informed by the absolute complexity perspective (such as the mean length of T-units).

One promising solution was recommended by Kyle and Crossley (2017), who introduced a corpus-based approach for gauging syntactic sophistication in L2 writing. They also claimed that the absolute complexity perspective is not in harmony with modern and prominent approaches

to L2 acquisition such as usage-based theories (e.g., Ellis, 2002). Kyle and Crossley (2017) adopted a frequency-driven approach to measure L2 writing development by calculating the frequency of verb argument constructions (VACs) and the strength of associations between VACs and the verbs used in VACs (i.e., verb-VAC combinations). They also compared the VACs indices to traditional indices of syntactic complexity such as the mean length of T-units. In their study, Kyle and Crossley sought to evaluate the efficacy of VACs and traditional indices to model one aspect of holistic scores of writing quality in Test of English as a Foreign Language (TOEFL) independent essays. It was found that usage-based indices explained greater variance (R^2 = .142) in holistic scores of writing quality than traditional indices (R^2 = .058). In a case study, Wind (2020) traced the writing development of two Hungarian English as a foreign language (EFL) learners', using the TAASSC to compute VACs indices. In a more recent study, Kyle et al. (2021) employed the TAASSC to compute usage-based indices to trace the longitudinal development of 20 Dutch EFL learners over two academic years.

Nevertheless, the use of the TAASSC to measure L2 writing development presents two main concerns. First, the TAASSC uses the COCA as a reference corpus; therefore, it reflects the language use in the United States. Consequently, using the TAASSC to assess Dutch (Kyle et al. 2021) and Hungarian learners (Wind, 2020) might paint a distorted picture of their writing development. Second, the academic sub-corpus of the COCA includes journals from many different disciplines. As different academic disciplines have their unique discourse (Hyland, 2004) and variation in syntactic complexity is present across academic disciplines (e.g., Casal et al., 2021), it can be presumed that there might be differences in syntactic complexity between two main academic disciplines: natural and social science. The reliability of using TAASSC to trace the development of syntactic sophistication might therefore be questionable.

Previous studies have looked at differences in linguistic complexity across academic writing and conversation (e.g., Biber, 1991). For example, Biber et al. (2011) found that most clausal subordination measures tend to be more frequent in conversation than in academic writing. Further, academic writing is characterized by complex noun phrase constituents rather than

clause constituents. Their research also showed that complex phrases are more common in academic writing, while clauses are more common in conversation. In addition, variation exists across academic disciplines. As Hyland and Tse (2007) found, disciplinary variation existed in regard to the range, frequency, collocations, and meaning of academic vocabulary items. This variation was explained by differences in the practices and discourses of various academic communities. Hyland (2008) also found variation in the frequencies and preferred uses of lexical bundles in research articles and degree papers across the fields of applied linguistics, biology, business studies, and electrical engineering. In addition, variation exists in the citation practices of research articles across disciplines (e.g., Hu & Wang, 2014; Hyland, 1999; Swales, 1990). For instance, variation has been observed between applied linguistics and general medicine research articles in terms of the frequencies of a specific type of dialogic engagement (Hu & Wang, 2014). Applied linguistics research articles were found to use more dialogically expansive citations, whereas those in general medicine used more dialogically contractive citations. Although academic writing is characterized by complex syntactic structures, few previous studies have systematically explored variations in syntactic complexity among different disciplines.

A rare example is Lu et al.'s (2021) study in which the introduction sections of research articles in two social and engineering disciplines were compared. The RAs were annotated for rhetorical move-steps and coded for syntactic complexity. The study found significant effects of discipline-specific, rhetorical move-step, and the association of the discipline and rhetorical move-step on the syntactic complexity of sentences in the introductions of RAs in the four disciplines. Effects were also observed regarding the interaction of the disciplines and rhetorical move-step on syntactic complexity of sentences. In another corpus-based study, Casal et al. (2021) investigated eight syntactic complexity indices across four research article sections (i.e., introduction, methods, results, and discussion) and three social science academic disciplines (i.e., applied linguistics, psychology, and economics). The researchers included 240 texts in their specialized corpus and analyzed the texts with a modified version of the L2SCA. The indices were compared across the different RA sections and three disciplines (i.e., applied linguistics, economics,

and psychology). The authors found a significant large effect of RA sections and discipline on all eight indices. Writers from the field of applied linguistics used more syntactically complex structures (as measured by overall T-unit complexity, overall sentence complexity, elaboration at the clause level, and complex noun phrases per clause) than writers from the other two disciplines. In a more recent study, Casal et al. (2022) found disciplinary variation in the frequency and contingency profiles of English VACs. The authors designed a specialized corpus of 400 articles across four disciplines (i.e., chemical engineering, electrical engineering, economics, and psychology) and analyzed 11 VACs. It was found that the VAC distributions are stable across the general corpus (BNC) and the authors' specialized corpus as far as normalized occurrence rates and frequency rank orders were concerned.

These findings (Casal et al., 2021, 2022; Lu et al., 2021) suggest disciplinary variation between the natural and social sciences. To explore this further, two corpora were designed in the present study to identify potential differences between the natural and social sciences in regard to syntactic complexity indices. The first corpus contains RAs from the natural sciences, while the second corpus contains social science RAs. This study aimed at answering the following two research questions:

> RQ1: Is there a difference between natural and social science corpora in terms of syntactic complexity?

> RQ2: Is there a difference between the abstracts, the introductions, and the literature reviews taken from the natural and social science corpus in terms of syntactic complexity?

Methodology

Corpus Design
Two corpora were created for the purposes of this study: (a) a natural science corpus and (b) a social science corpus. One of the main aims of this study was to compare three different sections of academic writing: (1) abstracts, (2) introductions, and (3) literature reviews. The rationale behind

choosing these three RA sections was motivated by pedagogical reasons: at the author's department, undergraduate students are required to write a non-empirical Bachelor of Arts (BA) thesis consisting of at least an abstract, an introduction, and a literature review section. As such, the first criterion in designing the two corpora was that the selected journals must have manuscripts containing these three different sections.

The first corpus contained RAs from journals on natural sciences, while the second corpus contained RAs from social science journals. The RAs were selected from the top journals based on the SCImago Journal Rank (SJR). At the time of the compilation of the two corpora, the 2021 version of the SJR was available; therefore, the 2021 ranking was used. The SJR indicator is a metric of how prestigious a scholarly journal is. SCImago ranks journals according to the number of citations a particular journal receives and the prestige of the journals from where the citations originate. A total of six journals were selected for this study. Three top journals were selected from the field of natural science, and three from the field of social science. However, in order to increase the representativeness of the corpus, only one journal was selected from a specific subfield (e.g., medicine, biochemistry, or physics) of natural science or social science. The first journal selected for the natural science corpus was the *Ca-A Cancer Journal for Clinicians* (Ca-ACJC) from the field of medicine (rank 1). The second journal was the *Nature Reviews Molecular Cell Biology* (NRMCB) from the field of biochemistry (rank 2), while the third journal was the *Reviews of Modern Physics* (rank 12). The journal *Cell and Nature Medicine* and *Nature Review Genetics* were omitted because they were from the same subfield, that is, biochemistry, as the NRMCB. The journal *MMWR Recommendations and Reports* and the *New England Journal of Medicine* were also left out as they represent the same subfield (medicine) as Ca-ACJC. The *Proceedings of the IEEE International Conference on Computer Vision* was omitted as it is not a journal but conference proceedings. As such, the third journal selected for the natural science corpus was the *Review of Modern Physics*. Regarding the social science corpus, the first journal was the *Quarterly Journal of Economics* (QJE). The *National Vital Statistics Reports* (rank 11) was omitted since its articles do not feature literature review sections, while the *Amer-*

ican Economic Review (rank 13) and the *Journal of Political Economy* (rank 16) were left out because they cover the same subfield (economy) as the QJE. Therefore, the *Administrative Science Quarterly* was selected as the second journal for the social science corpus. The third journal for the social science corpus was the *Modern Language Journal* (rank 563) for pedagogical reasons: since the author works at an applied linguistics department, important information could be collected to inform the academic writing courses taught by the department. Table 1 shows the first 20 journals from the SJR.

Table 1

SJR Journal Ranking in 2021

Rank	Journal	Area
1	Ca-A Cancer Journal for Clinicians	Medicine
2	Nature Reviews Molecular Cell Biology	Biochemistry, Genetics and Molecular Biology
3	Quarterly Journal of Economics	Economics, Econometrics and Finance
4	Cell	Biochemistry, Genetics and Molecular Biology
5	MMWR Recommendations and Reports	Environmental Science; Health Professions; Medicine; Social Sciences
6	New England Journal of Medicine	Medicine
7	Nature Medicine	Biochemistry, Genetics and Molecular Biology; Medicine
8	Nature Reviews Materials	Energy; Materials Science
9	Proceedings of the IEEE International-al Conference on Computer Vision	Computer Science
10	Nature Reviews Genetics	Biochemistry, Genetics and Molecular Biology; Medicine
11	National Vital Statistics Reports	Social Sciences
12	Reviews of Modern Physics	Physics and Astronomy
13	American Economic Review	Economics, Econometrics and Finance
14	Nature Biotechnology	Biochemistry, Genetics and Molecular Biology; Chemical Engineering; Engineering; Immunology and Microbiology

Rank	Journal	Area
15	Chemical Reviews	Chemistry
16	Journal of Political Economy	Economics, Econometrics and Finance
17	Nature	Multidisciplinary
18	Annual Review of Immunology	Immunology and Microbiology; Medicine
19	Administrative Science Quarterly	Arts and Humanities; Social Sciences
20	Nature Reviews Immunology	Immunology and Microbiology; Medicine

The final versions of the two corpora contained an almost equal number of words. However, the number of RAs and the means of the RAs from the two corpora differed significantly. Table 2 shows the number of RAs, the means, the standard deviations, and the number of words in the different sections (i.e., abstracts, introductions, and literature reviews) of the two corpora.

Table 2

The Natural and the Social Science Corpora

		Natural science corpus	Social science corpus
Abstracts	Number of abstracts	33	64
	Mean	187.7	187.88
	SD	53.96	33.64
	Total number of words	6,194	11,448
Introductions	Number of introductions	33	64
	Mean	917.94	804.52
	SD	932.26	502.76
	Total number of words	30,292	51,489
Literature reviews	Number of literature reviews	28	61
	Mean	6000.71	2265.20
	SD	3344.91	1070.72
	Total number of words	168,076	138,177
Total	Number of RAs	94	189
	Ca-A Cancer Journal for Clinicians	37	

	Natural science corpus	Social science corpus
Nature Reviews Molecular Cell Biology	36	
Reviews of Modern Physics	21	
Quarterly Journal of Economics		23
Administrative Science Quarterly		40
Modern Language Journal		126
Mean	2176.19	1064.1
SD	3149.85	1099.44
Total number of words	204,562	201,114

Data Analysis

After the selection of the three journals for the two corpora, the manuscripts were downloaded to the hard drive of the author. The RAs from the three different sections were copied and pasted from the downloaded PDF files into an MS Word file, with footnotes and the equations removed from the RAs. In several natural science journals, such as the *Review of Modern Physics*, footnotes are used rather than in-text citations. Since computational tools such as the L2SCA cannot process words that have numbers attached to them (e.g., *development6* or *particle5*), numbers were removed in such cases. In addition, RAs from the journal *Reviews of Modern Physics* contained several equations containing Greek letters and numbers that would have been difficult for the L2SCA to process. Therefore, these were removed from the RAs on the natural science. The MS Word DOC files were saved as txt files and submitted for linguistic complexity analysis to the L2SCA. The data were checked for normality, and an unequal variance t-test (Welch's test) was used to test whether the means of two corpora were equal. Welch's test is usually applied when there is a difference between the variations of two populations and when their sample sizes are unequal, as is the case in the present study. In addition, the nonparametric counterpart of the Welch test, the Mann-Whitney U test, was used to compare differences between the two independent groups (i.e., natural and social science corpora). The statistical tests were performed using JASP (JASP Team, 2023).

Results

RQ1: Is there a difference between natural and social science corpora in terms of syntactic complexity?

Table 3 shows the descriptive statistics for the natural and social science corpora, including abstracts, introductions, and literature reviews. RAs from the social science corpus were longer — as demonstrated by higher mean values in the MLS, the MLT, and the MLC indices — than the RAs from the natural corpus. RAs from the social corpus contained a higher number of clauses per sentence, and also contained a higher number of verb phrases per T-unit and clauses per T-unit. As far as subordination is concerned, the RAs from the social science corpus also included more dependent clauses per T-unit and dependent clauses per clause. The RAs from the social corpus also contained higher mean values regarding T-units per sentence and complex T-unit ratio indices. Although the RAs from the social science corpus contained a higher number of coordinate phrases per T-unit, the RAs from the natural science corpus contained a higher number of coordinate phrases per clause. Likewise, the RAs from the social science corpus contained a higher number of complex nominals per T-unit, whereas RAs from the natural corpus contained a higher number of complex nominals per clause. Table 3 also shows that the variables were not normally distributed, except in the case of the CT/T index. Therefore, the Mann-Whitney U test was used in the case of those variables that were not normally distributed, and the Welch test was used for the statistical analysis of the CT/T index.

Table 3

Descriptive Statistics for The Two Corpora (Combined)

Index	Corpus	Mean	SD	W	df	p
MLS	NAT	24.979	5.250	.946	93	<.001
	SOC	28.914	5.562	.975	188	0.002
MLT	NAT	24.191	4.183	.973	93	0.047
	SOC	27.295	5.236	.974	188	0.002
MLC	NAT	16.505	3.647	.905	93	<.001

Index	Corpus	Mean	*SD*	*W*	*df*	*p*
	SOC	15.823	3.082	.966	188	<.001
C/S	NAT	1.544	0.311	.986	93	0.397
	SOC	1.864	0.373	.981	188	0.012
VP/T	NAT	2.010	0.386	.955	93	0.003
	SOC	2.599	0.581	.983	188	0.025
C/T	NAT	1.492	0.226	.972	93	0.039
	SOC	1.755	0.328	.980	188	0.009
DC/C	NAT	0.329	0.098	.979	93	0.125
	SOC	0.412	0.103	.982	188	0.017
DC/T	NAT	0.509	0.218	.957	93	0.003
	SOC	0.753	0.309	.973	188	0.001
T/S	NAT	1.033	0.132	.950	93	0.001
	SOC	1.063	0.09	.877	188	<.001
CT/T	NAT	0.399	0.163	.978	93	0.111
	SOC	0.525	0.163	.991	188	0.293
CP/T	NAT	1.015	0.488	.929	93	<.001
	SOC	1.113	0.46	.973	188	0.001
CP/C	NAT	0.697	0.351	.923	93	<.001
	SOC	0.645	0.256	.987	188	0.088
CN/T	NAT	3.73	0.823	.982	93	0.241
	SOC	4.207	0.929	.982	188	0.014
CN/C	NAT	2.553	0.71	.907	93	<.001
	SOC	2.445	0.57	.985	188	0.045

Note. NAT = Natural science corpus, SOC = Social science corpus.

Table 4 shows the results of the Mann-Whitney U test for the 13 indices from the L2SCA. Nine indices out of the 13 were statistically significant, as well as the Welch test for the CT/T index (t = 6.13, p < .000).

Table 4

Mann Whitney U Test

Index	*U*	*z*	*P*
MLS	5292	5.537	< .000
MLT	5606.5	5.052	< .000
MLC	8061.5.	-1.266	.204

Index	U	z	P
C/S	4684.5	6.474	< .000
VP/T	3419.5	8.425	< .000
C/T	4521.5	6.725	< .000
DC/C	4804	6.29	< .000
DC/T	4543.5	6.692	< .000
T/S	7955.5	1.43	.153
CP/T	7458	2.197	.028
CP/C	8695.5	-0.288	.772
CN/T	6001.5	4.443	< .000
CN/C	8517	-0.564	.575

RQ2: Is there a difference between the abstracts, introductions, and literature reviews taken from the natural and social science corpora in terms of syntactic complexity?

Table 5 shows the descriptive statistics for the abstracts, introductions, and literature reviews for the 14 syntactic complexity indices.

Table 5

Descriptive Statistics (Abstracts, Introductions, and Literature Reviews)

Index	Corpus	Abstracts		Introductions		Literature reviews	
		Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
MLS	NAT	25.499	4.748	26.378	5.6	22.717	4.798
	SOC	25.942	4.693	31.801	5.847	29.005	4.426
MLT	NAT	24.531	4.264	25.365	4.104	22.406	3.693
	SOC	24.593	4.863	29.623	5.407	27.687	4.08
MLC	NAT	17.171	4.33	17.056	3.557	15.071	2.352
	SOC	14.631	3.086	16.926	3.328	15.916	2.308
C_S	NAT	1.547	0.386	1.573	0.28	1.507	0.247
	SOC	1.829	0.433	1.917	0.368	1.846	0.304
VP_T	NAT	2.029	0.472	2.035	0.362	1.96	0.301
	SOC	2.525	0.74	2.679	0.526	2.591	0.423
C_T	NAT	1.475	0.304	1.512	0.203	1.488	0.131
	SOC	1.73	0.414	1.779	0.305	1.756	0.241
DC_C	NAT	0.297	0.128	0.348	0.092	0.342	0.046
	SOC	0.379	0.125	0.428	0.09	0.431	0.082

Index	Corpus	Abstracts		Introductions		Literature reviews	
		Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
DC_T	NAT	0.472	0.283	0.542	0.212	0.514	0.112
	SOC	0.7	0.383	0.785	0.289	0.774	0.227
T_S	NAT	1.048	0.162	1.038	0.117	1.009	0.11
	SOC	1.063	0.108	1.077	0.094	1.048	0.059
CT_T	NAT	0.387	0.229	0.415	0.132	0.394	0.087
	SOC	0.504	0.209	0.541	0.146	0.53	0.121
CP_T	NAT	1.159	0.561	1.103	0.454	0.743	0.303
	SOC	0.877	0.46	1.278	0.434	1.187	0.384
CP_C	NAT	0.821	0.415	0.74	0.311	0.499	0.209
	SOC	0.528	0.267	0.727	0.243	0.682	0.214
CN_T	NAT	3.815	0.93	3.898	0.781	3.432	0.671
	SOC	3.845	0.969	4.578	0.94	4.198	0.712
CN_C	NAT	2.69	0.9	2.624	0.654	2.307	0.423
	SOC	2.304	0.645	2.625	0.613	2.405	0.36

Table 6 shows the results of the test of normality (i.e., the Shapiro-Wilk test). As far as the abstracts are concerned, Table 6 shows that six indices were normally distributed (i.e., MLS, MLT, DC/C, CT/T, CP/C, and CN/T), while eight indices were not (i.e., MLC, C/S, VP/T, C/T, DC/T, T/S, CP/T, and CN/C). In terms of the introductions, four indices were normally distributed, while 10 were not. As far as the literature reviews are concerned, seven indices were not normally distributed, while seven showed normal distributions. Therefore, the unequal variances t-test (Welch's test) was used in case of the normally distributed variables, while the Mann-Whitney U test was carried out for the non-normally distributed variables.

Table 6

Test of Normality (Abstracts, Introductions, and Literature Reviews)

Index	Corpus	Abstracts			Introductions			Literature Reviews		
		W	*df*	*ρ*	*W*	*df*	*ρ*	*W*	*df*	*ρ*
MLS	NAT	0.955	32	0.184	0.888	32	0.003	0.9	27	0.012
	SOC	0.986	62	0.709	0.958	63	0.028	0.971	60	0.163
MLT	NAT	0.988	32	0.971	0.917	32	0.015	0.935	27	0.081
	SOC	0.978	62	0.312	0.939	63	0.004	0.979	60	0.381

Index	Corpus	Abstracts			Introductions			Literature Reviews		
		W	*df*	*ρ*	*W*	*df*	*ρ*	*W*	*df*	*ρ*
MLC	NAT	0.863	32	< .001	0.932	32	0.041	0.959	27	0.335
	SOC	0.959	62	0.031	0.936	63	0.002	0.976	60	0.268
C_S	NAT	0.97	32	0.487	0.961	32	0.274	0.928	27	0.055
	SOC	0.948	62	0.009	0.976	63	0.256	0.986	60	0.736
VP_T	NAT	0.905	32	0.007	0.98	32	0.774	0.968	27	0.518
	SOC	0.947	62	0.008	0.949	63	0.011	0.983	60	0.58
C_T	NAT	0.948	32	0.118	0.976	32	0.668	0.983	27	0.921
	SOC	0.945	62	0.006	0.976	63	0.257	0.975	60	0.245
DC_C	NAT	0.981	32	0.815	0.967	32	0.404	0.963	27	0.42
	SOC	0.985	62	0.638	0.992	63	0.956	0.919	60	< .001
DC_T	NAT	0.937	32	0.056	0.935	32	0.048	0.929	27	0.059
	SOC	0.93	62	0.001	0.972	63	0.147	0.975	60	0.257
T_S	NAT	0.922	32	0.021	0.917	32	0.015	0.892	27	0.007
	SOC	0.729	62	< .001	0.945	63	0.007	0.97	60	0.137
CT_T	NAT	0.972	32	0.549	0.98	32	0.777	0.988	27	0.978
	SOC	0.97	62	0.119	0.992	63	0.957	0.923	60	< .001
CP_T	NAT	0.964	32	0.344	0.876	32	0.001	0.869	27	0.002
	SOC	0.896	62	< .001	0.972	63	0.15	0.96	60	0.045
CP_C	NAT	0.968	32	0.423	0.901	32	0.006	0.863	27	0.002
	SOC	0.961	62	0.041	0.974	63	0.192	0.966	60	0.084
CN_T	NAT	0.971	32	0.517	0.931	32	0.037	0.964	27	0.431
	SOC	0.984	62	0.574	0.946	63	0.007	0.987	60	0.787
CN_C	NAT	0.89	32	0.003	0.919	32	0.017	0.977	27	0.766
	SOC	0.988	62	0.804	0.948	63	0.009	0.989	60	0.866

Table 7 shows the results of the Welch tests. As far as abstracts are concerned, Table 7 shows that the difference between the natural and social science corpora was statistically significant in the dependent clause per clause (DC/C) and complex T-unit ratio (CT/T) indices. As far as the introductions are concerned, there were statistically significant differences in the clause per sentence (C/S), clause per T-unit (C/T), dependent clause per clause (DC/C), and complex T-unit ratio (CT/T) indices between the natural and social science corpora. As far as the literature reviews are concerned, the mean length of clause (MLC), clause per sentence (C/S), verb phrase per T-unit, clause per T-unit, dependent clause per T-unit, complex nominal per

T-unit, and complex nominal per clause indices showed statistically significant differences between the natural and social science corpora.

Table 7

Welch Tests

Index	Abstract			Introduction			Literature Review		
	t	*df*	*ρ*	*t*	*df*	*ρ*	*t*	*df*	*ρ*
MLS	-.438	64.879	.663						
MLT	-.046	73.447	.963						
MLC							-1.585	51.593	.119
C/S				-5.131	81.581	<.001	-5.573	63.707	<.001
VP/T							-8.042	71.783	<.001
C/T				-5.132	88.733	<.001	-6.777	84.272	<.001
DC/C	-2.935	64.025	.005	-4.067	63.732	<.001			
DC/T							-7.216	86.497	<.001
T/S									
CT/T	-2.357	59.323	.022	-4.270	70.760	<.001			
CP/T									
CP/C									
CN/T	-.123	67.881	.902				-4.907	55.392	<.001
CN/C							-1.064	45.578	.293

Table 8 shows the Mann-Whitney U tests for the abstracts, introductions, and literature reviews. In terms of the abstracts, the differences between the natural and social science corpus were statistically significant in seven syntactic complexity indices. As far as the introductions and the literature reviews are concerned, there were six syntactic complexity indices that showed statistically significant differences.

Table 8

Mann Whitney U Tests

Index	Abstracts			Introductions			Literature reviews		
	U	Z	p	U	Z	p	U	Z	p
MLS				465	4.496	< .000	245	5.376	< .000
MLT				504	4.199	< .000	267	5.182	< .000
MLC	670	-2.846	.004	1006	-0.377	.704			
C/S	703	2.592	.01						
VP/T	626.5	3.182	.001	318	5.616	< .000			
C/T	674	2.816	.005						
DC/C							266	5.191	< .000
DC/T	663	2.9	.004	510.5	4.15	< .000			
T/S	982	0.44	.66	894	1.23	.219	742	0.985	.322
CT/T							291.5	4.965	< .000
CP/T	681	-2.762	.006	762.5	2.231	.026	262	5.226	< .000
CP/C	606.5	-3.419	.001	1016	0.301	.764	387	4.121	< .000
CN/T				596.5	3.495	.000			
CN/C	1003.5	0.274	.787	1051	0.034	.976			

Discussion

This study investigated variation in syntactic complexity between natural and social science RA corpora. There were statistically significant differences in 10 L2SCA indices between the natural and social science corpora. RAs from the social science corpus contained longer sentences, as demonstrated by the mean length of sentence and mean length of T-unit indices, than RAs from the natural science corpus. This finding is in line with Lu et al. (2021), who found that the introductory sections in the field of anthropology utilized the longest sentences for multiple steps. In addition, sociology texts used longer sentences than texts from the two engineering disciplines examined in Lu et al.'s study. In this study, the RAs from the social science corpus contained significantly more clauses per sentence and more clauses per T-unit than the RAs from the natural science corpus. In addition, the social science research articles contained significantly more verb phrases per T-unit. Regarding subordination, the RAs from the social

science corpus contained significantly more dependent clauses per clause and per T-unit. This finding is also in line with Lu et al., who found that the two social science disciplines used more finite dependent clauses than the two engineering disciplines in the five focal rhetorical move-steps. The relatively low DC/C ad DC/T indices are not surprising, since Biber et al. (2011) claimed that dependent clauses tend to be more frequent in conversation than in academic writing. The difference in subordination implies that social science writers tend to unpack information in finite clausal components more often than natural science writers.

The complex T-unit ratio was also significantly higher in the social science corpus. Furthermore, RAs from the social science corpus contained significantly more coordinate phrases per T-unit. In addition, the RAs from the social science corpus contained significantly more complex nominals per T-unit. The relatively high values in the phrasal complexity indices are not surprising, since Biber et al. (2021) claimed that noun phrase constituents (rather than clause constituents) and complex phrases (rather than clauses) are more frequent in academic writing than in conversation.

The fact that the social science corpus in this study was more complex (with a higher occurrence of longer sentences, complex noun phrases, etc.) might be attributed to the inclusion of RAs from the field of applied linguistics. This was also observed by Casal et al. (2021), who found that applied linguistics RAs were more complex than RAs from the field of economics and psychology. Texts from the field of applied linguistics exhibited a higher frequency of four syntactic complexity indices (i.e., overall T-unit complexity, overall sentence complexity, elaboration at clause level, and noun phrase per clause) compared to the fields of economic and psychology.

Conclusion

This study revealed statistically significant differences in syntactic complexity across two academic disciplines: the natural and social sciences. Variation could be observed between the abstract, introduction, and literature review sections of the RAs analyzed.

There are several limitations to this study. First, the RAs were collected from only three natural science journals and three social science journals. Future studies could build a corpus using a greater number of journals. Second, the RAs were collected from volumes published in 2021; however, there might be variation over time which is worth considering. Therefore, future studies could expand the timeframe to investigate these temporal variations. Third, this study only investigated variation in terms of syntactic complexity; however, greater variations might be found in lexical complexity across different academic disciplines.

There are several implications of this study for instructors at the tertiary level. First, these results indicate significant variations between the academic writing of the natural and social sciences. Instructors of English for specific purposes (ESP) and English for academic purposes (EAP) should be aware of the seemingly subtle variations between these disciplines and instruct the linguistic characteristics of the specific field in which they are teaching. For instance, EAP teachers in the natural sciences should be aware that subordination is not a characteristic feature of RAs in the genre.

References

Alsahlanee, A., & Jaganathan, P. (2023). Phrasal syntactic complexity measures in linguistics research articles written by Iraqi and English L1 writers. *3L: Language, Linguistics, Literature, 29*(1), 154–168. http://dx.doi.org/10.17576/3L-2023-2901-11

Biber, D. (1991). *Variation across speech and writing*. Cambridge University Press.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development. *TESOL Quarterly, 45*(1), 5–35. https://doi.org/10.5054/tq.2011.244483

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21–46). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.32.02bul

Bulté, B., & Housen, A. (2014). Conceptualizing and measuring short-term changes in L2 writing complexity. *Journal of Second Language Writing, 26*, 42–65. https://doi.org/10.1016/j.jslw.2014.09.005

Casal, J. E., Lu, X., Qiu, X., Wang, Y., & Zhang, G. (2021). Syntactic complexity across academic research article part-genres: A cross-disciplinary perspective. *Journal of English for Academic Purposes, 52*, 1–12. https://doi.org/10.1016/j.jeap.2021.100996

Casal, J. E., Shirai, Y., & Lu, X. (2022). English verb-argument construction profiles in a specialized academic corpus: Variation by genre and discipline. *English for Specific Purposes, 66*, 94–107. https://doi.org/10.1016/j.esp.2022.01.004

Chen, D., & Manning, C. D. (2014). *A fast and accurate dependency parser using neural networks*. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) (pp. 740–750).

Davies, M. (2009). The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics*, *14*(2), 159–190. https://doi.org/10.1075/ijcl.14.2.02dav

Davies, M. (2010). The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing, 25*(4), 447–464. http://doi.org/10.1093/llc/fqq018

Davies, M. (2023, June 28). Corpus of Contemporary American English. Help. https://www.english-corpora.org/coca/help/texts.asp

Egbert, J., Biber, D., & Gray, B. (2022). *Designing and evaluating language corpora: A practical framework for corpus representativeness*. Cambridge University Press. https://doi.org/10.1017/9781316584880

Egbert, J., Larson, T., & Biber, D. (2020). *Doing linguistics with a corpus: Methodological considerations for the everyday user*. Cambridge University Press. https://doi.org/10.1017/9781108888790

Ellis, N. C. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, *24*(2), 143–188. https://doi.org/10.1017/S0272263102002024

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Hu, G., & Wang, G. (2014). Disciplinary and ethnolinguistic influences on citation in research articles. *Journal of English for Academic Purposes*, *14*, 14–28. https://doi.org/10.1016/j.jeap.2013.11.001

Hyland, K. (1999). Academic attribution: Citation and the construction of disciplinary knowledge. *Applied Linguistics*, *20*(3), 341–367. https://doi.org/10.1093/applin/20.3.341

Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing*. University of Michigan Press.

Hyland, K. (2008). As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes*, *27*(1), 4–21. https://doi.org/10.1016/j.esp.2007.06.001

Hyland, K., & Tse, P. (2007). Is there an "academic vocabulary". *TESOL Quarterly*, *41*(2), 235–253. https://doi.org/10.1002/j.1545-7249.2007.tb00058.x

JASP Team (2023). JASP (Version 0.17.2) [Computer software].

Klein, D., & Manning, C. D. (2003). *Accurate unlexicalized parsing*. In Proceedings of the 41 Annual Meeting on Association for Computational Linguistics - ACL '03 (Vol. 1, pp. 423–430). http://doi.org/10.3115/1075096.1075150

Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication*. [Unpublished doctoral dissertation]. Georgia State University. https://doi.org/10.57709/8501051

Kyle, K., & Crossley, S. (2017). Assessing syntactic sophistication in L2 writing: A usage-based approach. *Language Testing*, *34*(4), 513–535. https://doi.org/10.1177/0265532217712554

Kyle, K., & Crossley, S. A. (2018). Measuring syntactic complexity in L2 writing using fine-grained clausal and phrasal indices. *The Modern Language Journal*, *102*(2), 333–349. https://doi.org/10.1111/modl.12468

Kyle, K., Crossley, S., & Verspoor, M. (2021). Measuring longitudinal writing development using indices of syntactic complexity and sophistication. *Studies in Second Language Acquisition*, *43*(4), 781–812. https://doi.org/10.1017/S0272263120000546

Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford University Press.

Levy, R., & Andrew, G. (2006). *Tregex and Tsurgeon: Tools for querying and manipulating tree data structures*. Citeseer.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics*, *15*(4), 474–496. https://doi.org/10.1075/ijcl.15.4.02lu

Lu, X. (2011). A Corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly*, *45*(1), 36–62. https://doi.org/10.5054/tq.2011.240859

Lu, X., Casal, J. E., Liu, Y., Kisselev, O., & Yoon, J. (2021). The relationship between syntactic complexity and rhetorical move-steps in research article introductions: Variation among four social science and engineering disciplines. *Journal of English for Academic Purposes*, *52*, 1–13. https://doi.org/10.1016/j.jeap.2021.101006

Mazgutova, D., & Kormos, J. (2015). Syntactic and lexical development in an intensive English for Academic Purposes programme. *Journal of Second Language Writing*, *29*, 3–15. https://doi.org/10.1016/j.jslw.2015.06.004

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, *24*(4), 492–518. https://doi.org/10.1093/applin/24.4.492

Park, S., & Yoon, S. (2021). Syntactic complexity of EFL learners' casual conversation, monologue, and writing. *The Journal of Studies in Language*, *37*(1), 75–89. https://doi.org/10.18627/jslg.37.1.202105.75

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

Thongyoi, K., & Poonpon, K. (2020). Phrasal complexity measures as predictors of EFL university students' English academic writing proficiency. *REFLections*, *27*(1), 44–61. Retrieved from https://doi.org/10.61508/refl.v27i1.241750

Wind, A. M. (2020). A dynamic usage-based approach to the longitudinal development of verb argument constructions in second language writing: Two case studies. In C. Kalman (Ed.), *DEAL 2020: A Snapshot of diversity in applied linguistics* (pp. 39–70). Eötvös University Press. https://www.eltereader.hu/media/2020/12/Kálmán-Csaba_WEB.pdf#page=58

Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity*. Hawaii: University of Hawaii.

This volume is our fourth in line in a series where we seek to connect global issues with local research in the field of applied linguistics. The studies presented here speak for the diversity of research taking place at the Department of English Applied Linguistics, showcasing the curiosity and dedication with which the articles were written. Whether a particular piece of research is theory-based or empirical, the commonality is that they all present implications to further understand and aid the language learning process. Along with our colleagues, some of our PhD students have also been part of this creative process of research and write-up in 2023.

● **ATTILA M. WIND** & **BRIGITTA DÓCZI,** Editors

ELTE | FACULTY OF HUMANITIES

ELTE | EÖTVÖS UNIVERSITY PRESS