# UZRT2016 : empirical studies in applied linguistics

**Edited book / Urednička knjiga**

*Publication status / Verzija rada:* **Published version / Objavljena verzija rada (izdavačev PDF)**

*Publication year / Godina izdavanja:* **2017**

*Permanent link / Trajna poveznica:* https://urn.nsk.hr/urn:nbn:hr:131:854816

https://doi.org/https://doi.org/10.17234/UZRT.2016

*Rights / Prava:* In copyright / Zaštićeno autorskim pravom.

*Download date / Datum preuzimanja:* **2025-03-29**

Sveučilište u Zagrebu
Filozofski fakultet
University of Zagreb
Faculty of Humanities and Social Sciences

*Repository / Repozitorij:*

ODRAZ - open repository of the University of Zagreb
Faculty of Humanities and Social Sciences

DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

# UZRT 2016

## Empirical Studies in Applied Linguistics

Edited by Stela Letica Krevelj
and Renata Geld

FF press

UZRT 2016: Empirical Studies in Applied Linguistics

Edited by Stela Letica Krevelj and Renata Geld

Zagreb, FF press

http://wp.ffzg.unizg.hr/ffpress

ISBN: 978-953-175-662-4

Collection © 2017 FF press
Papers © 2017 The Contributors
Cover image © 2017 Ivana Režek

Layout: Nataša Vuković

All parts of this publication may be printed and stored electronically.

## CONTENTS

Foreword

**Assessing skills, prospects and outcomes**

17 *Judit Dombi*
**Working towards a Model of Intercultural Communicative Competence for Advanced-level EFL Learners**

22 *Katalin Doró*
**Source Distortion in Patchwritten EFL Academic Texts**

31 *Vera Savić*
**Reading Difficulties: What Do We Learn from Young English Language Learners?**

43 *Zoltán Lukácsi*
**The Language Instruction Scheme for Graduates with a Withheld Degree**

52 *Gabriella Lőcsey*
**Keyword Analysis of Short Stories for Young Language Learners**

63 *Darija Omrčen*
**Vocabulary Acquisition Revisited: The Case of Human Movement Science**

81 *Jasenka Čengić, Silvija Hanžić Deda & Ana Gabrijela Blažević*
**Evaluation of the FHSS (Faculty of Humanities and Social Sciences) Doctoral Program in Foreign Language Education: Questionnaire Design**

**Individual differences, diversity and learners with disabilities**

90 *Katalin Piniel & Ágnes Albert*
**L2 Motivation and Self-efficacy's Link to Language Learners' Flow and Anti-flow Experiences in the Classroom**

104 *Robert Märcz*
**The Mean Examiner: How Language Anxiety Affects Performance during an Exam**

115 *Alenka Mikulec & Renata Šamo*
**How Important Is English to Croatian Pre-Service Primary School and Preschool Teachers Not Majoring in English?**

127 *Iva Gugo & Renata Geld*
**Frequency and Type of Gesture in the Blind and the Sighted in L1 and L2**

139 *Agnieszka Kałdonek-Crnjaković & Zrinka Fišer*
**'Dyslexia-friendly' Approaches in the Teaching Practice of Croatian English as a Foreign Language (EFL) TeachersListening Comprehension Tests**

## 1.1. Foreword

This volume stems from University of Zagreb Round Table 2016 (UZRT 2016) conference, a yearly event which is organized alternately by the University of Pecs and the University of Zagreb. The event started as a platform through which younger researchers would be given the opportunity to receive feedback on their research from peers and more experienced colleagues.

The volume comprises 12 papers which cover various empirically studied topics and contribute to the wealth of knowledge in the field of applied linguistics. Some studies propose new theoretical models, and other have a narrower focus and inquire into specific issues of concern to the field of language learning and use. We have broadly categorized them into two areas of interest; studies pertaining to the assessment of skills, prospects and outcomes, and those investigating individual differences, diversity and learners with disabilities.

The first seven papers in this volume address the topic of assessment of either competencies and skills acquired through different educational courses and programs, or pedagogical prospects and value of instructional material and assessment tools. Out of the remaining five papers in the second part of the volume, three papers deal with individual differences, or more specifically, affective aspects in the process of language acquisition and use, and two papers deal with the characteristics and needs of learners with disabilities and difficulties.

Even though most of the papers have the English language as their focus, we believe that the range of educational and sociolinguistic settings as well as different age groups of the participants included in the analyses make the volume appealing to a wider audience of researchers, teachers and students, as well as curriculum designers and developers.

We would like to thank the authors and reviewers for their thoughtful contributions as well as professors Marianne Nikolov and Jelena Mihaljević Djigunović, the initiators of this valuable collaboration between the two universities, for their support to this yearly event.

*The editors*

# Working towards a Model of Intercultural Communicative Competence for Advanced-level EFL Learners

**Judit Dombi**
*Department of English Linguistics, University of Pécs, Hungary*
*dombi.judit@pte.hu*

## 1. Introduction

The idea that foreign language teaching (FLT) should prepare language learners for communication in real life situations is not new: over the past four decades research has focused on how to help language learners function as competent communicators in interactions with partners of various language and cultural backgrounds. For successful communication in the globalized world, it is no longer sufficient to master a linguistic code in order to exchange ideas and negotiate meaning. Language users need to be able to apply their knowledge, attitudes, skills and critical awareness (Byram, 1997) that help them in situations where the other parties are likely to have beliefs, ideologies, values and practices markedly different from their own. This need has been recognized by policy makers: the Common European Framework of Reference for Languages (Council of Europe, 2001, pp. 103-105) includes intercultural awareness and intercultural skills under general competences language learners should attain. Intercultural awareness and intercultural skills are widely referred to in FLT contexts as intercultural communicative competence (ICC) (Byram, 1997), a term emphasizing a link between the construct and established trends in FLT.

Individual differences that characterize language learners play an important role in how learners behave and interact in intercultural encounters. Thus, the complexity of the learner and of the language learning process also need to be taken into consideration. The world is perplexing and ever changing, as well as individuals: the differences characterizing learners and making them distinct from one another add to the intricacy of the issue. The past experiences, fears, expectations, motives, beliefs and attitudes learners bring to intercultural interactions play a role as important and conducive to their success as are their competences.

## 2. An overview of the literature on ICC and individual differences

### 2.1. Intercultural Communicative Competence

The most often quoted model of intercultural communicative competence was proposed by Byram (1997). In his conceptualization, ICC is made up of discourse competence, sociolinguistic competence, strategic competence and intercultural competence. The first three competences were introduced in applied linguistics (Canale, 1983; Canale & Swain, 1980; van Ek, 1986). The fourth constituent, intercultural competence, comprises (1) attitudes, (2) knowledge, (3) skills, and (4) critical cultural awareness (Byram, 1997).

The increased interest in intercultural interactions has resulted in numerous empirical studies on how individuals behave, communicate and eventually, cope in these interactions. Some studies were conducted to address the difficulties participants faced when sojourning (Fantini, 2006; Matsumoto, et al., 2001; Matsumoto, LeRoux, Bernhard, & Gray, 2004), living (Jenkins, 2008; Matsumoto, et al., 2003) or studying (Dombi, 2011; Tran, 2009) abroad.

Other studies were concerned with how educational settings may contribute to the development of ICC (Jenkins, 2008; Kramsch, 2010; Menyhei, 2016; Xiao & Petraki, 2007). All studies analyzed data on ICC; however, there are important differences in the ways they conceptualize and assess ICC. As there are many terminological and conceptual inconsistencies, the focus is on instruments that are relevant to the present study: they use similar constructs, or employ analogous methods or survey similar participants.

Arasaratnam (2009) used an instrument to measure participants' ICC based on Spitzberg and Cupach's (1984) construct of communication behavior that is both effective and appropriate. ICC was measured along three dimensions: cognitive, behavioral and affective. The following items constitute her measure: (1) attitudes towards other cultures (ATOC), (2) ethnocentrism, (3) motivation, (4) interaction involvement, (5) intercultural communication competence (ICC). Participants of the study (N=302) were students of a large Australian university, representing diverse cultures. Data were subjected to regression analysis, factor analysis and correlation analysis. The findings revealed positive relationships between ICC and ATOC, ICC and motivation, and ICC and interaction involvement; a negative correlation was found between ICC and ethnocentrism. The results from the regression indicated that ATOC, motivation, and interaction involvement were all predictors of ICC.

Matsumoto and his associates completed a series of studies on Japanese sojourners to the US (Matsumoto et al., 2001, 2003, 2004) to uncover differences in the potential for intercultural adjustment using their own instrument, the Intercultural Adjustment Potential Scale (ICAPS). The ICAPS is based on the assumption that intercultural conflict is inevitable and adjustment depends on the ability to manage conflicts well (Matsumoto, et al., 2004). The instrument was used to predict how participants would respond to their new environment and how well they could adopt to new circumstances. Their findings were in accordance with theoretical assumptions about some psychological components necessary for successful intercultural adjustment, especially concerning emotion regulation, openness, flexibility and creativity (Matsumoto, et al., 2001, p. 505).

Ying (2002) also studied Asians temporarily residing in North America. Her participants were Taiwanese university students studying in the U.S. She hypothesized that students are more likely to form intercultural relationships if they have (1) a more extroverted personality, (2) a more robust knowledge about the host culture, (3) favorable attitudes towards befriending Americans, and (4) better communication skills in English. Ying measured personality, knowledge, attitudes, communication skills, social environment and social network composition. Results showed that students had some understanding of America, expressed equally positive attitudes towards forming relationship with American and Taiwanese peers, had moderate English competence, and their social networks mainly consisted of Chinese peers. More extroverted students reported more intercultural contacts and students understanding American culture more reported better relationships with Americans and more confidence in interacting.

In sum, personality and other learner characteristics were found to contribute to participants' predisposition to succeed in intercultural encounters.

## 2.2. Individual differences: Affective and communication variables

Applied linguists have long been engaged in mapping the internal factors, which are responsible for differences in individuals' language learning (Clément & Gardner, 2001, p. 21; Cohen & Dörnyei, 2002; Dörnyei, 2005; Ellis, 1994). Individual differences (ID) were defined by Dörnyei as "dimensions of enduring personal characteristics that are assumed to apply to everybody and on which people differ by degree" (2005, p. 4). In this section the most widely included affective and communication variables are discussed which are assumed to contribute to successful intercultural encounters.

Motivation, the force that drives people to pursue different ambitions, has long been in the center of attention (Dörnyei, 1990, 1994; Gardner & Smythe, 1975; Gardner, 1985). Research shows that motivation is the second best predictor of L2 proficiency after aptitude (Dörnyei, 1998, 2001; Gardner, & MacIntyre, 1993). In literature, language learners' motivation is either described along the intrinsic-extrinsic or the instrumental-integrative continuum (Dörnyei, 2005, pp. 73-80; Ellis, 1994, pp. 74-76).

Studies on Hungarian EFL learners' motivation to learn English found that students' main motives were intrinsic and instrumental. These results may be rooted in the specific contexts of these studies. The privileged position of English in Hungary (Nikolov 2002, 2003, 2007; Nikolov & Józsa, 2003) contributes to students' underlying beliefs that it is in their own interest to learn this global language. This accounts for instrumental motivation: in a foreign language (FL) learning context students' main concern is not integrating, but making use of the language.

Attitudes have also been found to be strongly related to language learning success (see Dörnyei, Csizér, & Németh, 2006; Ellis, 1994; Gardner, 1985). Attitudes include learners' beliefs and feelings towards the target language speakers, culture, and the social value of language learning. Increased opportunities for intercultural interactions are often perceived to have beneficial effects on learners' attitudes as they may result in an "enhanced understanding of the target language culture and a more sympathetic attitude to native speakers" (Coleman, 1997, p. 7). Previous research has shown that time spent in an English speaking country (Nagy, 2008) or in a foreign country (Dombi, 2011; Fantini, 2006) significantly enhances students' attitudes and openness towards other cultures, thus fostering ICC.

Ample studies have addressed the effect of intercultural contact on learners' attitudes towards the target culture, motivation and their language proficiency (Csizér & Dörnyei, 2005; Csizér & Kormos, 2008; Szaszkó, 2010). Attitudes towards speakers of a language most often determine attitudes towards the target language (Dörnyei, Csizér, & Németh, 2006), and thus it also has an impact on motivation to learn the language. Studies have also emphasized that extensive intercultural contact raised participants' awareness towards intercultural issues and contributed to the development of favorable attitudes towards other cultures (Csizér & Kormos, 2009; Szaszkó, 2010).

Apart from affective variables, certain communication variables may also contribute to individuals' success in intercultural interactions. The most often studied communication variable is willingness to communicate (WTC). Research on WTC derives from communication research in the native language. The construct proposed by McCroskey (1992) defines WTC as the probability that an individual would initiate a conversation if there were an opportunity. McCroskey and Richmond (1990) suggested that WTC in individuals' mother tongue is likely to be a personality trait, whereas in case of language learners, a more complex understanding of WTC is needed, since it is to a great extent situationally conditioned and involves a wide range of feelings.

Two other communication variables are closely associated with WTC: communication apprehension (CA) and perceived communicative competence (PCC) (MacIntyre, 1994). In McCroskey's definition, CA is the "level of fear or anxiety" that individuals associate with interpersonal communication (1992, p. 1). Thus, conceptually, CA is similar to foreign language anxiety.

The way individuals assess their own L2 communication abilities is also of importance. Findings of previous studies on communication in L2 suggest that students' PCC significantly impacted their communication behavior. Studies have consistently shown a relationship between CA, PCC and WTC both in the mother tongue (McCroskey & Richmond, 1990) and in a L2 (MacIntyre, Baker, Clément, & Donovan, 2003; Nagy, 2008, 2009; Yashima, Zenuk-Nishide & Shimizu, 2004).

In the Hungarian context, Nagy (2009) conducted a study on English majors' WTC in EFL. Her statistical analysis revealed that only learners' self-perceived proficiency influenced their WTC, whereas language anxiety was not directly related to the construct. However, learners' communicational anxiety was significantly related to their perceived competence. WTC had no direct relationship with actual L2 behavior; thus, high level of WTC did not result in actual use of EFL in real life situations.

## 3. A study on EFL learners' ICC

### 3.1. Aim

This study aimed to build and test a model of EFL language learners' intercultural communicative competence in relation to other learner variables in order to find out how certain individual difference variables influence the way language learners behave in intercultural situations.

The study addressed the following research questions:

RQ1: How does ICC relate to participants' affective ID variables?

RQ2: How does ICC relate to other communication variables?

RQ3: How does the frequency of IC contact contribute to ICC?

RQ4: How can ICC be modeled in relation to other learner variables?

### 3.2. Participants

Participants were first-year B.A. students of English Studies at a large Hungarian university. Altogether 117 students filled in the questionnaire. Fifteen participants were excluded for various reasons (i.e. incomplete questionnaire, native speaker of English). Consequently, the final number of participants was 102 (female n=71, male n=31). Background data on participants are presented in Table 1.

Table 1. Background data on participants.

| | N | Mean | Min | Max | SD |
|---|---|---|---|---|---|
| Age | 102 | 20.3 | 19 | 23 | 1.4 |
| Number of years studying English | 102 | 10.5 | 5 | 12 | 3.7 |
| Number of years of intensive English language learning | 102 | 5.9 | 5 | 7 | 3.2 |

As Table 1 presents, participants' average age was 20.3 years, they had been learning English for an average of 10.5 years, out of which 5.9 years were devoted to intensive language learning (more than 5 hours/week).

### 3.3. Instrument and procedures

In order to better understand how individual differences interact with students' ICC, a questionnaire was designed and piloted (Dombi, 2013). The questionnaire comprised five sections, aiming to elicit data on different individual difference and communication variables.

*Section I* was made up of nine items on willingness to communicate (WTC) to which participants needed to answer by giving percentage values, 0 meaning they are absolutely unwilling to act as described by the statement, 100 meaning they are absolutely willing to.

*Section II* included 18 self-assessment items: nine on perceived intercultural communicative competence (PICC) and nine on perceived communicative competence (PCC). Participants needed to give their answers in percentages, indicating how competent they believe they were in the situations described by the statements, 0 meaning absolutely incompetent, 100 meaning absolutely competent.

*Section III* comprised 18 items on communication apprehension (CA). Participants were requested to indicate how much they agreed with the statements on a five-point Likert scale ranging from 5 (strongly agree) to 1 (strongly disagree).

*Section IV* consisted of 56 items on various affective aspects and ICC. This section combined various scales: on motivation (MOT) on perceived L2 competence (PL2), and on intercultural contact (ICO). This section also comprised 23 items on ICC: seven on knowledge, seven on skills and nine on attitudes. Participants indicated how much they agreed with the statements on a 5-point Likert scale (5: strongly agree, 1: strongly disagree).

*Section V* elicited data on participants' language learning background.

Table 2 provides a list acronyms used to refer to the variables.

Table 2. Acronyms used in the study

| Acronym | Variable |
|---|---|
| **WTC** | willingness to communicate |
| **PICC** | perceived intercultural communicative competence |
| **PCC** | perceived communicative competence |
| **CA** | communication apprehension |
| **ICC** | intercultural communicative competence |
| **MOT** | Motivation |
| **PL2** | perceived L2 competence |
| **ICO** | intercultural contact |

The questionnaires were administered in April, 2011 at the university under consideration. Taking part in this study was voluntary, and all first-year students enrolled to a compulsory English language development course gave their informed consent. The filling in the questionnaires lasted 30 minutes and was supervised by the author. For the statistical analysis SPSS 14.0 and AMOS 4.0 (Arbuckle, 1999) were used.

### 3.4. Results

First, principal component analysis was performed. The scales constituted principal components; as confirmed by the reliability checks. Alpha values of the scales are presented together with the results in Table 3.

Table 3. Performance scores on the scales

| | N valid | N missing | Min | Max | Mean | SD |
|---|---|---|---|---|---|---|
| **ICC** (Alpha=70) **5-point Likert-scale** | 102 | 0 | 2.52 | 4.35 | 3.48 | .39 |
| **PICC** (Alpha=.70) 0-100 % | 102 | 0 | 34.11 | 100 | 76.48 | 13.42 |
| **WTC** (Alpha=.85) 0-100 | 102 | 0 | 40.56 | 100 | 78.39 | 16.09 |
| CA (Alpha=.93) **5-point Likert-scale** | 102 | 0 | 1.56 | 4.61 | 2.76 | .79 |
| **MOT** (Alpha=.71) **5-point Likert-scale** | 102 | 0 | 3.36 | 5.00 | 4.44 | .41 |
| **PCC** (Alpha=.90) **0-100 %** | 102 | 0 | 41.11 | 100 | 80.33 | 15.39 |
| **ICO** (Alpha=.74) **5-point Likert-scale** | 102 | 0 | 1.67 | 5.00 | 3.80 | .62 |
| **PL2** (Alpha=.78) **5-point Likert-scale** | 102 | 0 | 2.00 | 4.71 | 3.37 | .66 |

Two scales elicited information on participants' intercultural communicative competence: the ICC scale and the PICC scale. However, in drawing the model one variable on ICC was needed. To have one single variable instead of these two, a linear transformation was used.

Final values for ICC are presented in Table 4.

Table 4. Students' scores on the combined ICC (ICC_C, measure 0-100)

| | N | N missing | Min | Max | Mean | SD |
|---|---|---|---|---|---|---|
| ICC_C | 102 | 0 | 41.0 | 90.20 | 69.34 | 9.71 |

### 3.4.1. RQ1: How does ICC relate to participants' affective ID variables?

To find out more about the relationship between ICC scores and affective individual differences, the Pearson product-moment correlation coefficient was computed for participants' combined ICC scores, their WTC, CA and MOT. The analysis revealed a significant ($p < .01$) correlation between ICC and the affective variables WTC ($r = .529$), CA1 ($r = -.627$) and MOT ($r = .298$). Thus, in line with previous expectations, sound relationships were detected between WTC, CA and ICC, and low, yet still significant relationships between MOT and ICC.

To get a clearer picture on the relationships among these variables, stepwise regression analysis was performed. This type of statistical analysis seeks to explain the variance in the level of one variable on the basis of the level of other variable(s). Hence, the analysis was meant to find out how much variance in individuals' ICC scores can be explained by the variables WTC, CA and MOT. Thus, ICC was entered as a dependent variable, and WTC, CA and MOT were entered as independent variables (predictors). The variables were entered step-wise, in the following order: CA, WTC, and MOT.

Results suggest that the predictors entered this way explained a significant amount of variance in students' ICC scores ($F1,100 = 64.68$; $p < .01$). As for the dependent variables, CA explained 39 percent of variance in participants' ICC scores ($R^2adj = .387$ for CA), WTC explained an additional six percent ($R^2adj = .450$ for CA and WTC), whereas the inclusion of MOT into the model resulted in an additional two percent of explained variance ($R^2adj = .467$ for the three dependent variables).

### 3.4.2. RQ2 and RQ3: How does ICC relate to other communication variables and to the frequency of intercultural contacts (ICO)?

In a second round, similarly to previous steps, correlation analysis and regression analysis were carried out to map the relationship between participants' ICC scores and their communicative profile, i.e. their perceived communicative competence (PCC), their perceived L2 competence (PL2) and the frequency of their intercultural contacts (ICO).

The analysis revealed significant ($p < .01$) relationships between each of the variables and ICC. PCC had the highest correlation ($r = .709$) with ICC, and PL2 also had a rather high value ($r = .610$). The ICO variable had somewhat more modest, yet still significant correlation ($r=.432$) with ICC.

Regression analysis provided an insight into the relationship between these individual difference variables and ICC. Similarly to previous steps, ICC was entered as a dependent variable, whereas the other individual difference variables were entered stepwise as independent variables, first PCC, then PL2 and finally, ICO. However, the analysis excluded ICO and only tested the models with PCC and PL2, revealing that participants' ICO cannot explain any variance in their ICC scores.

Results reveal that PCC and PL2 explain a significant amount of variance in students' ICC scores ($F1,101 = 61.63$; $p < .01$). The dependent variable PCC explained almost 50 percent of variance in students' ICC scores ($R^2adj = .498$ for PCC), and PL2 explained an additional five percent ( $R^2adj = .546$ for PCC and PL2).

### 3.4.3. RQ4: How can ICC be modeled in relation to other learner variables?

Since correlation analysis shows only to what extent certain variables are related, but does not explain causation, a more complex statistical analysis was applied

to get a better picture on participants' ICC in relation to the observed affective and communication variables. To achieve this, first a model of the cause-effect relationships of the observed variables was proposed, which could be tested as a second step. A type of structural equation modeling, path analysis, was used to test the hypothesized model. The proposed model was drawn up based on correlations presented above.

Based on these, a model was constructed (see Figure 1). In this model the affective variables WTC and CA and other communication variables PCC and PL2 were pictured as related; the affective variable MOT and the individual difference variable ICO were also presented as related. As regression analysis indicated that PCC, CA and WTC explained the most variance in ICC scores, a direct path was suggested from these variables to ICC. Even though regression analysis excluded ICO, a path from this variable to ICC was also proposed since the literature suggests that experience with the target culture fosters ICC. Correlation analysis also revealed a significant but moderate relationship between these two values.

As path models follow certain common drawing conventions (Schumacker & Lomax, 2004) these conventions are used in Figures 1 to 4. Observed variables are presented in rectangular boxes. Lines directed from one observed variable to another signal direct effects, whereas curved, double-headed lines denote covariance, meaning that the marked variables are correlated. Each dependent variable also has an error term, indicated by a circle around the error term (Byrne, 2010; Schumacker & Lomax, 2004).

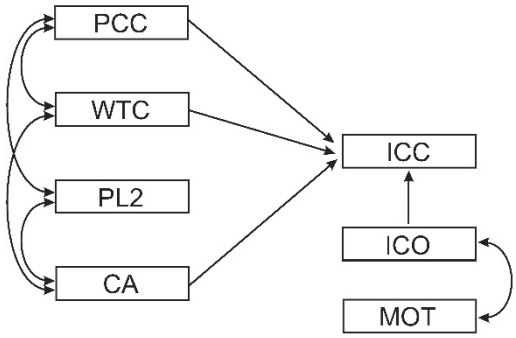


*Figure 1.* Proposed model of ICC

The proposed model's fit to the dataset was tested using AMOS 4.0, a computer program designed for structural equation modeling (SEM) (Arbuckle, 1999). There are numerous criteria for assessing model fit, including (1) $\chi^2$; (2) $\chi^2$ divided by the number of degrees of freedom; (3) goodness-of-fit index (GFI); (4) adjusted goodness-of-fit index (AGFI); (5) Bentler-Bonett normed fit index (NFI); (6) Tucker-Lewis coefficient (TLI); (7) root mean square error of approximation (RMSEA); (8) expected cross-validation index (ECVI); and (9) comparative fit index (CFI) (Loehlin, 2004; Raykov & Marcoluides, 2006; Schumacker & Lomax, 2004). These indices are meant to show how much the model deviates from the null hypothesis of no relationships between the constituents.

As for GFI, AGFI, NFI, TLI, CFI, values may be placed on a 0 –1 scale, 0 indicating no fit, whereas 1 indicates a perfect fit of the model. As for RMSEA, a value less than .05 indicates a good model fit. $\chi^2$/df ratio should not exceed 2 for the model to be accepted (Schumacker & Lomax, 2004, pp. 82-83). The probability (p) value of the $\chi^2$must exceed 05, otherwise the model has to be rejected. All these fit measures are provided

together with the tested and re-tested models. However, taking into consideration the relatively small sample size (N=102), the RMSEA index is of crucial importance, as it is insensitive to sample size (Loehlin, 2004, p. 68). The proposed model was tested with AMOS 4.0 (Arbuckle, 1999), and the results with the standardized path coefficients and the goodness-of-fit measures are shown in Figure 2.

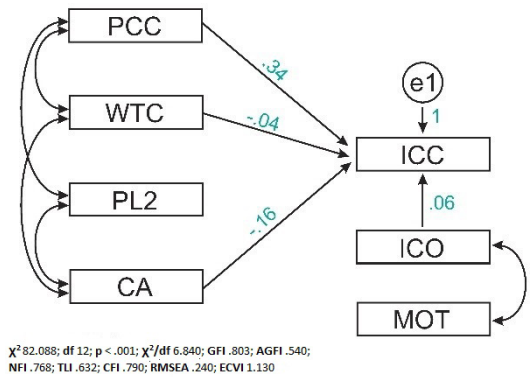


Figure 2. The proposed model tested.

As Figure 2 shows, the $\chi^2$ for the proposed model was 82.08 with 12 degrees of freedom, thus, $\chi^2$/df ratio was 6.84, which was above the level of acceptance. However, not only was this value problematic with the initial model, but the probability did not exceed .001, so that the model significantly differed from the dataset and had to be rejected.

In building the second model, we relied on AMOS' suggestion that the path from WTC to ICC was not significant, and deleted it. Moreover, AMOS also recommended to make all the four dependent variables inter-correlated, so we changed the model to fit these suggestions. Furthermore, the path from ICO to ICC was also found to be insignificant. However, we decided to keep it for the time being, to check how this revised model worked. The revised model was re-submitted to analysis, and its results are shown in Figure 3.

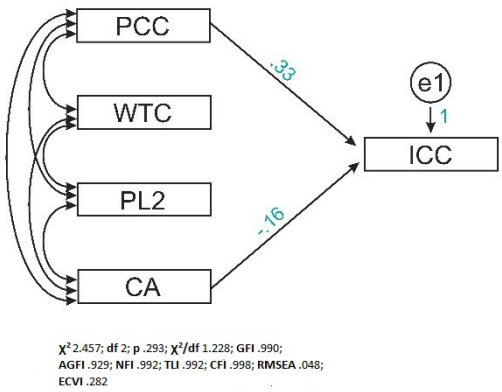


*Figure 3*. Revised model

In the revised model all paths but the one from ICO to ICC were significant, and the correlation values between PCC, WTC, CA and PL2 were convincing (as presented in Figure 3). However, the goodness-of-fit measures were still unacceptable ($\chi^2$/df = 3.75, p = < .001, and RMSEA = .165). Thus, further revision of the model was necessary.

AMOS had previously suggested elimination of the path from ICO to ICC to gain better fit measures. As a result of this, it was concluded that in order to adequately model English majors' ICC, the variables ICO and MOT had to be excluded from the model. The final model consisted of the inter-correlated PCC, PL2, CA and WTC, with significant paths from PCC to ICC, and from CA to ICC (see Figure 4). The model was re-entered for analysis and was found to fit the dataset with very good goodness-of-fit indices: $\chi^2$/df ratio was 1.22 with a probability of .293, and the RMSEA was .048.

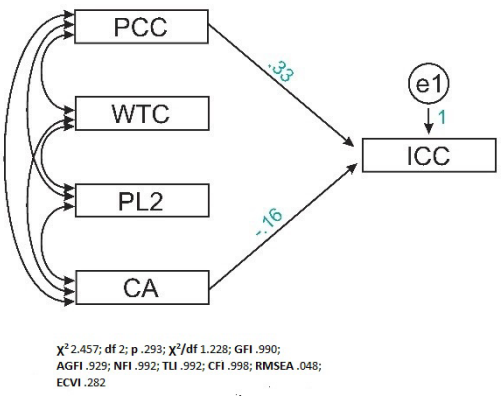


*Figure 4.* The final model.

### 3.5. Discussion

Findings concerning the relationship between affective variables revealed the strongest negative relationship between CA and ICC, indicating that anxiety is most likely to affect performance in intercultural situations. Willingness to communicate in English was also found to significantly correlate with ICC: the more ready students are to engage in communication, the more likely they are to be successful in such interactions. This relationship is rooted in the multi-faceted nature of the ICC construct: someone with high scores on the combined ICC scale has high scores on the ICC Attitudes and the Perceived ICC scales as well, both of which comprise items describing imagined interactions with members of other cultures. Thus, it is likely that the more eager students are to take part in such conversations, the more likely they develop positive attitudes towards others, which in turn, fosters their ICC.

The relationship between motivation and ICC was also significant; however, much weaker than the previous two affective variables. This finding shows similarities with findings of Szaszkó (2010) who found that in the case of Hungarian adult EFL learners, intercultural contact had more impact on their English communicative competence than their language learning motivation.

Regression analysis confirmed the assumption that anxiety is of utmost importance in intercultural encounters: it was found that most variance in students' ICC

scores could be explained by their CA scores, and the addition of WTC and motivation did not explain significantly more variance in ICC. Revisiting the findings of previous studies in similar settings, more empirical evidence supports this result: debilitating anxiety, i.e. the extreme feeling of insecurity is a factor participants frequently mention as the major obstacle to success in intercultural encounters (Dombi, 2013; Menyhei, 2016; Nagy, 2009). Students in these studies reported feeling nervousness, apprehension or even panic when it came to speaking in English with members of other cultures. The findings of the present study confirm that the apprehension experienced when talking in English prevents students from extensive intercultural contacts, negatively influencing both their WTC and their motivation, thus affecting their ICC.

A strong relationship was found between perceived communication competence and ICC. However, the interpretation of this finding requires some caution: since correlation does not show which variable influences the other, this outcome can be interpreted in two different ways and they both make sense. We may assume that the students who thought they had better communication skills had higher ICC, as their better perceived communicative competence made them more self-confident, and thus more likely to take part in intercultural interactions, which, in turn, made them more experienced and competent in such situations. However, this can also work the other way around: higher ICC, i.e. more success in intercultural situations, can also boost students' self-confidence, making them believe they are good at communication in the target language.

Moreover, PCC and ICC were related, as ICC also has to do with communication: students who believe they are more competent communicators in English are likely to project the same competency in intercultural situations in which they communicate in English. This finding is supported by findings in Nagy's (2009) study: PCC was found to be more likely to influence students' communication behavior than their linguistic self-confidence (p. 92).

A more advanced statistical procedure, regression analysis, proved that PCC explained almost 50 percent in the variance of students' ICC scores, thus revealing PCC to be an important predictor of ICC. Students' perceived L2 competence was also found to be strongly related to their ICC. With regard to this variable, the relationship seems more straightforward (still not evident, as correlation alone is insufficient to establish a cause-effect relationship): students who believed they had better English proficiency were likely to score higher on ICC. This may be due to the fact that high level of perceived L2 competence, similarly to good PCC, helps students to be more confident and self-assured in interactions. However, since more students are in the average PL2 category than in the average PCC category, it can be observed that there are students who believe they are good at English, but fewer of them believe they are good at communication in English.

As for the last individual difference variable, ICO, the results showed a relatively high frequency of intercultural contact, which was not surprising, as the university town is known for the high number of foreign students.

Contrary to expectations, a significant, yet surprisingly weak relationship was found between students' ICC and the frequency of their intercultural contact. Moreover, stepwise regression analysis excluded the ICO variable and concluded that it did not explain variance in students' ICC scores. These results were rather unexpected, as one would assume that students' exposure to foreign cultures through contact contributes to their ICC to a great extent. Nagy's findings (2008) on English majors who previously worked as au-pairs in the UK revealed that as a result of their first-hand experiences

with members of other cultures, participants were ready to drop their previous stereotypes, and became more aware of the importance of cultural diversity. Similarly, Fantini (2006) found that exchange students' extended contact with local residents in Ecuador was among the factors mostly contributing to their enhanced ICC. In this study, however, there was no empirical evidence to support the claim that the frequency of intercultural contact had a significant impact on ICC. Since Fantini's (2006) and Nagy's (2008) study surveyed participants who spent time in a foreign country, findings may imply that extensive contact with a foreign culture away from one's home environment could result in these beneficial outcomes, whereas the impact of foreign contact on one's ICC while residing at home may be more limited.

The most important finding of the SEM analysis was that the initially suggested model (Figure 1) of English majors' ICC did not provide acceptable goodness-of-fit indices, and had to be rejected. The underlying problem with the proposed model was that it hypothesized direct relationship between WTC and ICC, based on findings of the correlation and regression analyses. Based on these results it was reasonable to assume that the more willing students are to take part in English conversations, the more likely it is for them to develop their intercultural communicative competence. The analysis, however, rejected this hypothesis, as the goodness-of-fit values gradually improved after deleting the direct path from WTC to ICC.

Willingness to communicate, however, remained a part of the model, although in a slightly modified position. The revised models contain WTC as a part of a chain of inter-correlated variables, including perceived L2 competence, perceived communication competence and anxiety. Based on findings of Nagy (2009) with a very similar population, the original model did not hypothesize a significant relationship between WTC and PL2; however, SEM revealed that a revised model with such correlation fits the dataset to a much greater extent. This result may be construed as students with better perceived L2 competence tend to be more satisfied with themselves and less insecure in communication; thus, they are more likely to engage in interactions in English. This leads to the interpretation that not only the perceived level of proficiency plays a role here, but perceived communication competence does so as well. The better communicators students believe they are, the more likely it is for them to take part in interactions in English; this is a finding in line with Nagy's (2009) results.

Important characteristics of the initial model which were maintained throughout the revisions are the direct paths from perceived communicative competence and communication apprehension to ICC. The fact that these paths were found to be significant confirms the findings of regression analyses: these are the two most important predictors of ICC.

The initial model, as well as the first two revised models suggested a further path that finally proved to be non-significant: the direct path between the frequency of intercultural contact and ICC. As other empirical studies (Dombi, 2011; Matsumoto, et al., 2001, 2003; Nagy, 2008) suggest that acquaintance with members of other cultures makes students more interculturally-minded and successful in intercultural situations, this variable was included in the model, even though the regression analysis did not show significant predictive value for it. Initially, it was proposed that students' motivation to learn English and the frequency of their intercultural contact were related: motivated students were thought to seek intercultural contact for various reasons, such as practicing the language by meeting speakers of English. Frequent intercultural contact, on the other hand, was hypothesized to sustain and further increase students' motivation to learn English. Surprisingly, this link was disproved.

Not only did SEM reveal that these variables were not directly related to ICC, but it was also suggested that they had nothing to do with the construct, and consequently had to be removed from the model.

The final model shows that the two variables directly influencing ICC were (1) perceived communicative competence in a positive and (2) communication apprehension in a negative way. This means that students with a more confident self-image have higher ICC, most probably because this confidence assures them in intercultural situations and helps them overcome difficulties or breakdowns in communication. This finding reveals that learners' self-image is of utmost importance: the PCC scale did not measure actual communication competence, but perceived communication competence. Therefore, there is no evidence that students who believe they are good communicators are actually good or not. However, it seems that the belief of being good at communicating in English allows these students to take advantage of intercultural situations and develop their ICC.

Highly anxious students, on the other hand, fear engaging in intercultural situations, possibly due to their lack of self-confidence. Over-anxious students trying to avoid interactions deprive themselves of the benefits of learning from such encounters and have fewer chances to develop their ICC. As a result, special attention should be paid to help anxious students overcome their fear to communicate in English.

## 4. Conclusion

One of the main aims of the study was to draw a theoretically sound and empirically based model that adequately presents English majors' ICC. This was achieved through performing structural equation modeling on the dataset. The final model presents four inter-correlated individual difference variables, WTC, CA, PL2 and PCC with direct paths from PCC and CA to ICC, suggesting that students' willingness to use English, their ideas about their own performance and their apprehension from communication situations are strongly related, and out of these variables perceived communication competence and communication apprehension directly affect ICC.

This indisputably has a pedagogical implication: knowing the variables influencing ICC may help instructors to tailor their teaching so as to address the individual differences accounting for variation in ICC.

It was shown that students' ICC was mostly affected by their anxiety. This result echoed findings of previous studies conducted with similar English language majors (Nagy, 2009; Tóth, 2007, 2011). It would be crucial to reduce learners' nervousness about speaking in English, since their anxiety has negative effects on their development: it debilitates their performance, and most often prevents them from interacting with others in English. Instructors should pay special attention to reducing learners' anxiety in classrooms, which can be achieved by creating a relaxed and friendly atmosphere free of competition. Raising students' awareness about negative effects of anxiety is also crucial, as it may induce more conscious actions.

Findings also suggest that students' self-image as communicators (their perceived communicative competence and their perceived L2 competence) are of utmost importance: if students believe they are good communicators in English, they are more likely to be self-confident and are more likely to take part in intercultural encounters. Thus, instructors should help students achieve a realistic self-image about their performance in English and support them if they lack self-confidence.

## 5. References


Arbuckle, J. L. (1999). Amos (Version 4.0) [Computer Software]. Chicago: SPSS.

Arasaratnam, L. A. (2009). The development of a new instrument of intercultural communication competence. *Journal of Intercultural Communication*, *20*. Retrieved 21 May 2011 from http://www.immi.se/intercultural/nr20/arasaratnam.htm

Byram, M. (1997). *Teaching and assessing intercultural communicative competence.* Clevedon: Multilingual Matters.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. New York: Routledge.

Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards, & R. W. Schmidt (Eds.), *Language and communication* (pp. 2-27). London: Longman.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1-47.

Clément, R., & Gardner, R. C. (2001). Second language mastery. In H. Giles, & W. P. Robinson (Eds.), *Handbook of language and social psychology* (pp. 489-504). London: Wiley.

Cohen, A. D., & Dörnyei, Z. (2002). Focus on the language learner: Motivation, styles, and strategies. In N. Schmitt (Ed.), *An introduction to applied linguistics* (pp. 170-190). London: Edward Arnold.

Coleman, J. A. (1997). Residence abroad within language study. *Language Teaching*, *30*(1), 1-20.

Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching, assessment.* Cambridge: Cambridge University Press.

Csizér, K., & Kormos, J. (2008). The relationship of inter-cultural contact and language learning motivation among Hungarian students of English and German. *Journal of Multilingual and Multicultural Development, 29,* 30-48.

Csizér, K., & Kormos, J. (2009). Modeling the role of inter-cultural contact in the motivation of learning English as a foreign language. *Applied Linguistics, 30,* 166-185.

Csizér, K., & Dörnyei, Z. (2005). The internal structure of language learning motivation and its relationship with language choice and learning effort. *Modern Language Journal, 89,* 9-36.

Dombi, J. (2011). Exploring Korean EFL students' intercultural experiences. Presented at the XX. International Grastyán Conference, Pécs, Hungary

Dombi, J. (2013). A mixed-methods study on English majors' intercultural communicative competence. (Unpublished doctoral dissertation). University of Pécs. Pécs

Dörnyei, Z. (1990). Conceptualizing motivation in foreign language learning. *Language Learning, 40*, 45-78.

Dörnyei, Z. (1994). Motivation and motivating in the foreign language classroom. *The Modern Language Journal, 78*, 273-283.

Dörnyei, Z. (1998). Motivation in second and foreign language learning. *Language Teaching, 31*, 117-35.

Dörnyei, Z. (2001). *Teaching and researching motivation.* England: Pearson Education Limited.

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

Dörnyei, Z., Csizér, K., & Németh, N. (2006). *Motivation, language attitude and globalisation: A Hungarian perspective*. Clevedon: Multilingual Matters.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Fantini, A. (2006). Exploring and assessing intercultural competence. Final report. Brattleboro, VT: EIL. Retrieved 13th September 2010 from http://www.experiment.org/documents/FinalGSIResearchReport12.06.pdf

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

Gardner, R., & MacIntyre, P. D. (1993). A student's contributions to second-language learning. Part 2: Affective variables. *Language Teaching, 26*, 1-11.

Gardner, R.C., & Smythe, P.C. (1975). Motivation and second language acquisition. *The Canadian Modern Language Review, 31*, 218-230.

Jenkins, S. (2008). Adopting an intercultural approach to teaching English as an international language. *Essential teacher*, *5*(4), 19-21.

Kramsch, C. (2010). The symbolic dimension of the intercultural. Plenary delivered at the Intercultural communication conference, University of Arizona. Manuscript of speech sent via email on 24 January, 2010.

Loehlin, J. C. (2004). *Latent variable models: An introduction to factor, path, and structural analysis*. Hillsdale, NJ: Laurence Erlbaum.

MacIntyre, P. D. (1994). Variables underlying willingness to communicate: A casual analysis. *Communication Research Reports, 11*, 135-142.

MacIntyre, P. D. (1999). Language anxiety: A review of the research for language teachers. In D. J. Young (Ed.), *Affect in foreign language and second language learning: A practical guide to creating a low anxiety classroom atmosphere* (pp. 24-45). Boston, MA: McGraw-Hill College.

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.

MacIntyre, P. D., Baker, S., Clément, R., & Donovan, L. A. (2003). Talking in order to learn: Willingness to communicate and intensive language programs. *The Canadian Modern Language Review, 59*(4), 580-607.

Mackey, A., & Gass, S. (2005). *Second language research: Methodology and design.* Mahvah, NY: Lawrence Erlbaum.

Matsumoto, D., LeRoux, J. A., Ratzlaff, C., Tatani, H., Uchida, H., Kim, C., & Araki, S. (2001). Development and validation of a measure of intercultural adjustment potential in Japanese sojourners: The Intercultural Adjustment Potential Scale (ICAPS). *International Journal of Intercultural Relations*, *25*, 483–510

Matsumoto, D., LeRoux, J. A., Iwamoto, M., Choi, J. W., Rogers, D., Tatani, H., & Uchida, H. (2003). The robustness of the Intercultural Adjustment Potential Scale (ICAPS). *International Journal of Intercultural Relations*, *27*, 543–562.

Matsumoto, D., LeRoux, J. A., Bernhard, R., & Gray, H. (2004). Personality and behavioral correlates of intercultural adjustment potential. *International Journal of Intercultural Relations, 28*, 281-309.

McCroskey, J. C. (1982). *Introduction to rhetorical communication*. Englewood Cliffs, NJ: Prentice-Hall.

McCroskey, J. C. (1992). Reliability and validity of the willingness to communicate scale. *Communication Quarterly, 140*, 16-25.

McCroskey, J. C., & Richmond, V. P. (1987). Willingness to communicate. In J.C. McCroskey, & J.A. Daly (Eds.), *Personality and interpersonal communication* (pp. 129-156). Newbury Park, CA: Sage.

McCroskey, J. C., & Richmond, V. P. (1990). Willingness to communicate: Differing cultural perspectives. *The Southern Communication Journal, 56*, 72-77.

Menyhei, Zs. (2016). Developing English majors' intercultural communicative competence in the classroom: We have learnt (…) how much more we need to learn about Intercultural Communication. (Unpublished doctoral dissertation). University of Pécs, Pécs.

Nagy, B. C. (2008). Linguistic and socio-cultural outcomes of the *Au Pair* experience in the United Kingdom. *Language and Intercultural Communication*, *8*(3), 172-192.

Nagy, B. C. (2009). *'To will or not to will': Exploring advanced EFL learners' willingness to communicate in English.* (Unpublished doctoral dissertation) University of Pécs, Pécs.

Nikolov, M., & Józsa, K. (2003). *Az idegen nyelvi készségek fejlettsége angol és német nyelvből a 6. és 10. évfolyamon a 2002/2003-as tanévben.* [Foreign language competence in English and German in grades 6 and 10 in the 2002/2003 academic year.] Budapest: OKÉV.

Nikolov, M. (2002). *Issues in English language education.* Bern: Peter Lang AG.

Nikolov, M. (2003). Az idegennyelv-tanítás megújulásának hatásai. [Effects of renewing foreign language teaching.] *Új Pedagógiai Szemle*, *53*(3), 46-57.

Nikolov, M. (2007). A magyarországi nyelvoktatás-fejlesztési politika: Nyelvoktatásunk a nemzetközi trendek tükrében.[Hungarian foreign language development policy. Our language teaching and international trends.] In I. Vágó (Ed.), *Fókuszban a nyelvtanulás* (pp. 43-72). Budapest: Oktatáskutató és Fejlesztő Intézet.

Raykov, T., & Marcoulides, G. A. (2006). *A first course in structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum.

Spitzberg, B. H., & Cupach, W. R. (1984). *Interpersonal communication competence*. Beverly Hills, CA: Sage.

Szaszkó, R. (2010). *The effects of intercultural contact on adult Hungarians' motivation for learning English as a foreign language*. (Unpublished doctoral dissertation). Eötvös Loránt University, Budapest.

Tóth, Zs. (2006). First-year English majors' perceptions of the effects of foreign language anxiety. In M. Nikolov, & J. Horváth (Eds.), *UPRT 2006: Empirical studies in English applied linguistics* (pp. 25-37). Pécs: Lingua Franca Csoport.

Tóth, Zs. (2007). *Foreign language anxiety: A study of first-year English majors*. (Unpublished doctoral dissertation). Eötvös Lóránt University, Budapest.

Tóth, Zs. (2011). Foreign language anxiety and advanced EFL learners: An interview study. *WoPaLP, 5*, 39-40.

Tran, L. T. (2009). Making visible hidden intentions and potential choices: International students in intercultural communication. *Language and Intercultural communication*, *9*(4) 271-284.

Van Ek, J. A. (1986). *Objectives of foreign language learning. Volume I: Scope.* Strasbourg: Council of Europe, Publications Section.

Xiao, H., & Petraki, E. (2007). An investigation of Chinese students' difficulties in intercultural communication and its role in ELT. *Journal of Intercultural communication*, *13,* Retrieved 20 May 2010 from http://www.immi.se/intercultural/nr13/petraki.htm

Yashima, T., Zenuk-Nishide, L., & Shimizu, K. (2004). The influence of attitudes and affect on willingness to communicate and second language communication. *Language Learning, 54,* 119-152.

Ying, Y. (2002). Formation of cross-cultural relationships of Taiwanese international students in the United States. *Journal of Community Psychology*, *30*(1), 45-55.

# Source Distortion in Patchwritten EFL Academic Texts

**Katalin Doró**
*University of Szeged, Hungary*
*dorokati@lit.u-szeged.hu*

## 1. Introduction

Learning to develop appropriate source use is an essential part of academic writing. Efficient citation practice shows knowledge of the field, supports the line of argument, credits the author of the source texts and, even more importantly, indicates the place of the author in the argument. Novice writers, especially those using an L2, face considerable difficulties when asked to produce essays, summaries and literature reviews in which they should not only focus on textual mechanisms, but also integrate the ideas of others, evaluate sources and show authorial stance. Texts produced by university students, either in their L1 or L2, include forms of source misuse and inappropriate citation practices to an alarming degree. This is often the case even when students are familiar with general citation conventions and consequences of source misuse. Inappropriate citation can lead to misunderstanding, unclear authorial stance, or even charges of plagiarism. While previous research has extensively studied the types of source misuse, the reasons behind this behaviour, students' and instructors' attitudes and possible forms of intervention, less is known about how the source texts are changed inappropriately or kept untouched when integrated into students' own texts.

This paper analyses the types of source distortion found in patchwritten academic texts (seminar papers, thesis drafts and final versions) written by Hungarian undergraduate and graduate students of English. Compared to the paraphrasing error studies carried out under controlled circumstances with the rewriting of short paragraphs and a given set of sources, the texts under investigation integrate various sources to different degrees in order to form longer pieces of academic texts. Types of source misuse, such as the lack of appropriate and clear source indication, sentence structure errors, semantic distortion and other writing mechanism problems (e.g., numbering and referencing untouched), are reported and discussed.

## 2. Previous research on citation practices

Source text use by foreign or second language (L2) writers has received particular attention in the last few decades. The integration of the textual chunks of other authors has been referred to by different terms, including, but not limited to, *textual borrowing* (Baily & Challen, 2015; Keck, 2014, Petrić, 2012; Shi, 2004), *source text borrowing* (Weigle & Parker, 2012), *copy-paste plagiarism* (Haen & Molnar, 2014; Mozgovoy, Kakkonen & Cosma, 2010), *textual appropriation* (Shi, 2006, 2012), and *patchwriting* (Pecorari 2003; Howard, 1993, 1999; Howard, Serviss & Rodrigue, 2010; Li & Casanave,

2012). Most scholars and higher education instructors agree that inappropriate text reuse is most often not the result of a deliberate choice to deceive the reader. Rather it stems from a number of factors that may co-occur, such as inadequate reading skills, low language proficiency, lack of subject knowledge, misinterpretation of the sources, no clear expectations or misunderstandings about the appropriateness of close copying (Davis, 2013; Doró, 2014; Harwood & Petrić, 2012; Li & Casanave, 2012; McCulloch, 2012, Pecorari, 2003; 2008, Petrić, 2004; 2015; Shi, 2012; Stockall &Villar Cole, 2016). On top of these, the little writing practice accumulated during the few years of secondary and higher education and the limited time put into a writing task that is too difficult push many students to copy from the texts they read. Often task completion is boiled down to the accumulation of the required number of words and textual copying helps to fill the pages more quickly.When the sources are not clearly marked, there is an obscure ratio of what the student author is saying and what comes, often verbatim, from other sources. As Pecorari (2003) puts it,

> the real nature of source use is only known to the writer, who uses conventional metatextual devices (citation, quotation marks, etc.) to signal the relationship. The reader decodes these signals and interprets the relationship between source and citing text. Ordinarily, these signals are the reader's only source of information about source use. If they are not used accurately, then the real nature of the source use is obscured. Transparency, then, means signaling the relationship between source and citing text accurately; its opposite is often termed plagiarism. (p. 324)

Researchers have highlighted different reasons behind textual plagiarism with debates over the role of cultural differences, language proficiency, task type and intentionality (Maxwell, Curtis, & Vardanega, 2008; Pecorari, 2003; Pecorari &Petrić, 2014; Setoodeh, 2015). Most researchers, however, agree that close copying from sources is a natural stage many novice writers go through while learning appropriate source use (McCulloch, 2012, Pecorari, 2003; 2008, Petrić, 2004; 2015; Shi, 2012). Pecorari introduced the term *patchwriting* to refer to the attempt to paraphrase with limited lexical or syntactic changes made to the source text without the intention to plagiarize. Extended patchwriting may also refer to the integration of longer textual chunks into students' writing with no or minimal changes and some connecting elements (e.g., reporting verbs and connectives) resulting in full paragraphs or even longer sections that really look like patches (Doró, forthcoming).

Hyland (2012) rightly points out that acquiring field-specific academic writing is a long process and it involves more than simply learning the mechanisms of text production. Hamilton (2016) in his work in an Australian higher education context found that students are expected too early to apply referencing conventions perfectly, creating frustration both from the students' and the instructors' side. Academic writing courses do help students in learning about field-specific citation practices, but the application of this knowledge is difficult and requires massive reading of source texts to get familiar with the discourse conventions of the chosen research community. Tomaš (2011) concludes that instructors may miss valuable opportunities to address the functions of textual borrowing when they are explaining assignments. Studies have also addressed the positive influence and learning potential of computer-assisted forms of citation practices that may aid face-to-face writing support (Baily & Challen, 2015; Choi, 2012; Stockall &Villar Cole, 2016).

Studies on L2 writers' citation behaviour and source use have used both qualitative and quantitative methods and included participants from different language and educational backgrounds (e.g., Baily & Challen, 2015; Luzón, 2015; Yoke, 2013). The most recent research on textual borrowing has used corpus methods

to balance for the frequent assumptions about the nature and the degree of source misuse students have in their assignments (Keck, 2014). As Keck (2014) points out, larger corpus-based studies can counterbalance for the type of conclusions drawn by earlier research mainly done by hand with a limited number of student texts. While corpus methods give us exact frequency counts of certain types of source use, studies like that of Keck also limit researchers to the analysis of specific task types, such as the summary of assigned sources.

## 3. The study

### 3.1. Rationale and research questions

The present study looked at texts produced by undergraduate and graduate level EFL students to evaluate what types of source distortion happen when they integrate source texts into their own. As the main aim was to list and illustrate main distortion types, a quantitative analysis of source misuse was not included. Instead, a manual checking of student texts and, when available, the source texts was done not only to document the presence and level of patchwriting, but also to see how the original texts are changed when students rely very heavily on textual borrowing.

### 3.2. Participants

Participants were students of English at a Hungarian university where the language of instruction for their English course is English. They did not include first-year students who have minimal or no previous academic writing experience in English, low level English proficiency and often submit heavily plagiarized assignments until they understand the general requirements of source use. The English Study program of the participants includes an introductory writing course in their first year, an academic writing class in their second year and a thesis writing seminar in their third year. Additional academic reading and writing are practiced and discussed in disciplinary seminars in which they receive writing assignments that require source integration. Therefore, the participants had general knowledge about field-specific writing conventions and knew about academic integrity requirements. The applied linguistics texts produced by the students use in-text, integral and end of the paragraph citations all of which should clearly credit the sources used, mark the relationship between source and student texts and show authorial stance.

### 3.3. The corpus

The corpus consists of approximately 200.000 words of academic essays, summaries, comments, literature reviews and source-dependent introductory chapters of BA and MA level theses. All texts were produced on applied linguistics topics as regular course assignments or as a final graduation requirement. Unlike many studies that assign participants a set of sources, these texts included those that students searched for themselves and found relevant for their chosen topic.

# 4. Results and discussion

## 4.1. Incorrect source attribution

Although the participants had had some practice with academic writing and received instruction on source integration, extreme source misuse, such as verbatim copying of large sections, often without final editing, was detected in the corpus. This form of extended patchwriting, fortunately, is rare. However, much more often students identify key phrases, sentences or even larger chunks and partly reorganize them and paraphrase them to form what they consider to be their own texts. Reference to the source in most of these cases happens only at the very beginning or very end of the paragraphs, leaving the reader uncertain about the proportion of original and copied sections. In the examples 1 and 2 the student texts create the impression of a legitimate summary while these passages are, in reality, close copies of the original ones (source texts are provided below the student texts).

> (1) Nation adds that most English content words can change their form by adding prefixes or suffixes. According to him the affixes can be divided into two types such as inflectional and derivational ones.
> Cf. Most of the content words of English can change their form by adding prefixes or suffixes. These affixes are typically divided into two types: inflectional and derivational (Nation, 2001, p. 263).

> (2) One of the most important messages of this study is that it is never too late to pursue an interesting, challenging career based on passion. Unfortunately it is still a problem that people with dyslexia do not get the appropriate help. Naturally, it would be the ideal state if dyslexic students needs were satisfied by the educational system but we could not reach this aim. Fink however draws the attention to the possibility that if individuals with dyslexia lack the necessary background and training but are interested in pursuing challenging fields, career counselors can brainstorm with them to help figure out how to obtain the necessary training (Fink, 2002).
> Cf. It is never too late to imagine and pursue an interesting, challenging career based on passion. If individuals with dyslexia lack the necessary background and training but are interested in pursuing challenging fields, career counselors can brainstorm with them to help figure out how to obtain the necessary training (Fink 2002, p. 7).

It often happens that the sources are distorted when the authors' names are misspelled or wrong publication dates or numbering are provided. In most cases this can be a simple typing error or a sign of a copy-paste writing methods, but may not influence the source text adaptation. The same happens when a source is missing from the reference list, but is indicated in the text. However, if the source is not easily traceable from an in-text reference or references are mixed or randomly provided, we can talk about real source distortion. In example 3 below, one of Nation's 1989 articles is indicated as the source, but the information provided, namely error correction, is not discussed in this Nation 1989 source.

> (3) In dictionary usage words can also be checked if they are existing or not. In dictionaries the difference between the known and the new words can be found, and it is useful in error correction, too… (Nation 1989)

Example 4 illustrates a student's effort to summarize in one single paragraph not only a source article, but also some of the further sources that appear in this one. While the primary source is referenced incorrectly, both in the text and in the reference list, the secondary sources are missing from the reference list. The last sentence in

the passage refers to "the researchers", but it is unclear which study and authors are mentioned here. This is a frequent problem and a serious form of source distortion in which primary and secondary sources are mixed. This practice makes it unclear how many of these sources the student has actually read and to what extent the summary is his or her own integration and interpretation of the referenced readings and not a condensed version of the research background sections of a few key articles.

> (4) Kata Csizér, Judit Kormos and Ágnes Sarkadi also mention in their study that dyslexia as a term is difficult to define for two reasons. ... The authors prefer the definition of the British Dyslexia Association (1998), which is widely accepted in Europe. According to this statement, dyslexia is "a complex neurological condition which is constitutional in origin. The symptoms may affect many areas of learning and function, and may be described as a specific difficulty in reading, spelling and written language." They mention studies conducted by Sparks, Ganschow, and their colleagues, in which they examined students with dyslexia who scored significantly lower than students without dyslexia on all of the components of the Modern Language Aptitude Test, which measures language aptitude, an important determinant of language learning success. Moreover, the researchers found that a number of other linguistic skills considered relevant in L2 learning, such as word recognition, spelling, pseudoword reading, word and nonword repetition, were in dyslexic language learner (Kata Csizér, 2010).

Previous studies have also pointed out the frequent incorrect source attribution of student writers, including the absence or under-use of in-text sources, the overuse of integral citations and excessive quotations (for a review see Luzón, 2015). In our corpus, the focus was on patchwritten texts with excessive textual borrowing. The four examples above show different referencing problems, but interestingly all of them indicate sources in the text, even though these are sometimes missing from the final reference lists. This could be an indicator of an alarming tendency of close-copying source texts not only for the ideas they contain, but also for the references they include and the texts themselves which already draw on a number of sources. This saves both time and effort to search for, read and integrate various sources, when all of these are readily available in some selected articles. If the students are not making very obvious referencing and syntactic errors in their texts, their writing can easily pass as assignments that meet the requirements.

### 4.2. Sentence structure errors and basic grammar problems

Students who draw too closely on source texts, yet try to make an effort to paraphrase to a minimal extent, soon run into the problem of sentence structure errors. Those who have low language proficiency which does not allow them to fully process the source texts, or who do a quick, careless job, may also leave basic grammar mistakes in their writing. Example 5 was most likely the result of the misunderstanding or simple misreading of the source text, which was a lecture handout the student had received. Sg was interpreted as someone instead of something, and the sentence this way makes little sense, which should have been noticed if the text had been proofread.

> (5) Face threatening act (FTA) appears when <u>someone</u> is said that represents a threat to another individual's expectations while face saving act appears when someone is said in order to lessen the possible threat (Brown and Levinson, 1987).
> Cf. Face threatening act (FTA) - when sg is said that represents a threat to another individual's expectations regarding self-image. Face saving act - when sg is said in order to lessen the possible threat (lecture handout).

Students with lower language proficiency level often produce very uneven texts. Borrowed chunks and partially paraphrased parts are mixed and sometimes connected in an ungrammatical way. These cases are relatively easy to notice, unlike the examples in the previous section in which ready-made sentences are lifted verbatim from source. Probably the student author of example 6 did not understand the source text or lacked language skills to paraphrase.

> (6) According to Harmer teaching listening skills shares similarities with reading, but difficulties arouses, most importantly, the text itself is different. While the written text is static and can read againat the speed of the reader, the spoken text being on tapes can be repeated, but only at its speed not the reader's (Harmer, 1991).

### 4.3. Semantic distortion

Various types and degrees of meaning distortion can happen while sources are borrowed, summarized or paraphrased. Reasons may vary including, among others, inadequate reading and note-taking skills, lack of careful proofreading with the sources at hand, low proficiency level and the misinterpretation of the sources used. Very few studies have addressed the semantic distortion of source texts that are borrowed and integrated into students' writing. Also, when these excessive forms of textual borrowing (also called near copies and close paraphrases by Shi [2004]) are detected either in research or in educational settings, little further analysis happens concerning the way the source text is changed. A careful, word-for-word comparison of the source and the student texts can reveal some unexpected solutions students have. Examples 7–9 illustrate the minimal changes the authors made to the source texts. All three students referenced their sources and all seem to be legitimate paraphrases or summaries at first glance. Interestingly, the three examples below exchanged one single word in the original text, resulting in meaning distortion: *ambiguous* is changed to *vague*, *valuable* to *valid* and *fairly* to *very*. We do not know whether these word changes were attempts to paraphrase or whether the students simply felt more comfortable using the new words without noticing the differences in meaning.

> (7) This view is supported by Lance (1977) who states that the meaning of grammar is not vague, but polysemous, having more meanings.
> Cf. When an individual uses the word *grammar*, he/she may be referring to any one (or more) of a variety of theoretical constructs. The term *grammar*, in other words,is polysemous—not ambiguous, but "having more than one basic meaning.… (Lence, 1977, p. 43).

> (8) Mufwene claimed that the subject's grammar supported the structural claims of Bickerton's hypothesis so the UG orientation of Bickerton's LBH seems to be valid.
> Cf. Mufwene asserted that the subject's grammar supported the structural claims of Bickerton's hypothesis, though genetic claims were not similarly supported. The UG orientation of Bickerton's LBH was deemed valuable, though Mufwene suggested that UG features of acquisition are also available to adults and hence would afford them agency in the creolization context (Barikmo, 2007, p. 23).

> (9) First she speaks about *real-life listening* situations such as listening to the news, weather forecast, sport reports, chatting at a party, receiving instructions on how to do something, watching a film or a television programme. Moreover, she follows with hearing a speech or a lecture, attending a formal occasion etc. She adds that this list is naturally rough and incomplete, but very representative.

Cf.Now this list is naturally rough and incomplete, but fairly representative (Ur, 1984, pp. 2-3).

Student in example 9 selectively copied some items from a list of thirty provided in the source, working with a close copy for most of her paragraph. Reference to Ur (1984) is made at the end of the paragraph, in parenthesis. The reporting structures at the end of the paragraphs indicate that the student was familiar with academic writing mechanisms and conventions and the importance of referencing. Nevertheless, most of the text is borrowed.

The above examples also support the view of previous research which concludes that the teaching of the basic rules of academic writing may not produce assignments with legitimate source use. All samples reference their sources and most show the knowledge of field-specific integral citation. Yet, this remains a surface level knowledge if the rest of the text is closely copied from the sources and only a skeleton is provided for the textual borrowing. This is an alarming tendency as students easily deceive their readers, and may also feel they are doing the right job, either intentionally or unintentionally.

## 5. Conclusion

This study reported on some frequent textual borrowing mechanisms of students of English in their applied linguistics assignments. Different levels and forms of source misuse were identified, focusing on cases in which large textual chunks were lifted from the sources. Patchwriting, near copies and close paraphrases were found to be common forms of textual borrowing also in previous research (Shi, 2004; Jahić, 2011). Reasons for borrowing extensively from the sources may vary, but the examples of incorrect source attribution, syntactic and semantic distortions discussed in this study suggest that students with heavy source reliance aim to produce texts that look good on the surface. In-depth interview studies in the future may shed more light on what academic writing strategies push students to adopt the strong textual borrowing reported in this study and whether they are aware of source distortion.

It is important to raise both students' and instructors' awareness to the excessive textual borrowing and source integration problems through the examples of student (peer) texts and highlight the importance of good reading, note-taking skills and writing practice. Through this it is hoped that the frequency of unintentional plagiarism and source misuse can be lowered.

## 6. References

Baily, C., & Challen, R. (2015). Student perceptions of the value of Turnitin text-matching software as a learning tool. *Practitioner Research in Higher Education,9*(1), 38–51.

Barikmo, H. (2007). Perspectives on creole genesis and language acquisition.*Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics, The 2007 APPLE Award Winning M.A. Essay in Applied Linguistics*, *7*(2),1–25.

Brown, P., & Levinson. S. C. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Choi, Y. H. (2012). Paraphrase practices for using sources in L2 academic writing. *English Teaching, 6*(2), 51–79.

Csizér, K., Kormos, J., & Sarkadi, Á. (2010). The dynamics of language learning attitudes and motivation: Lessons from an interview study of dyslexic language learners. *Modern Language Journal, 94*(3), 470–487.

Davis, M. (2013). The development of source use by international postgraduate students. *Journal of English for Academic Purpose*, *12*(2), 125–135.

Doró, K. (2014). Why do students plagiarize? EFL undergraduates' views on the reasons behind plagiarism. *Romanian Journal of English Studies, 11*(1), 255–263.

Doró, K. (forthcoming). Extended patchwriting in EFL academic writing – signs and possible reasons In M. Chitez, C. Doroholschi, O. Kruse, L. Salski, & D. Tucan (Eds.), *University writing in Central and Eastern Europe: Tradition, transition and innovation*. Springer.

Fink, R. P. (2002.). Successful careers: The secrets of adults with dyslexia. *Career Planning and Adult Developmental Journal, 18*(1), 118–135.

Haen, J.J., & Molnar, K.K. (2014). Exploration of the ethical maturity of an undergraduate student cohort. *Journal of Academic and Business Ethics,9*, 1–10.

Hamilton, J. (2016). Attribution, referencing and commencing HE students as novice academic writers: Giving them more time to 'get it'. *Student Success, 7*(2), 43–49.

Harmer, J. (1991). *The practice of English language teaching*. London: Longman.

Harwood, N., & Petrić, B. (2012).Performance in the citing behavior of two student writers. *Written Communication*, *29*(1), 55–103.

Howard, R. M. (1993). A plagiarism pentimento. *Journal of Teaching Writing*, *11*(3), 233–246.

Howard, R. M. (1999). *Standing in the shadow of giants: Plagiarists, authors, collaborators.* Stamford, Conn.: Ablex Pub.

Howard, R. M., Serviss, T., & Rodrigue, T. K. (2010). Writing from sources, writing from sentences. *Writing and Pedagogy, 2*(2), 177–192.

Hyland, K. (2012). *Disciplinary identities: Individuality and community in academic writing*. Cambridge: Cambridge Applied Linguistics.

Jahić, A. (2011). An insight into textual borrowing practices of university-level students in Bosnia and Herzegovina. *Journal of Academic Writing, 1*(1), 173–179.

Keck, C. (2014), Investigating textual borrowing in academic discourse: A need for a corpus-based approach In V. Cortes & E. Csomay (Eds.), *Corpus-based research in applied linguistics: Studies in honor of Doug Biber (pp. 177–196).* Philadelphia: John Benjamins.

Lance, D. (1977). What is "Grammar"? *English Education, 9*(1), 43–49.

Li, Y.,& Casanave, C. P. (2012). Two first-year students' strategies for writing from sources: Patchwriting or plagiarism?*Journal of Second Language Writing*,*21*(2), 165–180.

Luzón, M. J. (2015). An analysis of the citation practices of undergraduate Spanish students. *Journal of Academic Writing, 5*(1), 52–64.

Maxwell, A., Curtis, G.J., & Vardanega, L. (2008). Does culture influence understanding and perceived seriousness of plagiarism? *International Journal for Educational Integrity, 4*(2), 25–40.

McCulloch, S. (2012). Citations in search of a purpose: Source use and authorial voice in L2 student writing'. *International Journal for Educational Integrity, 8*(1), 55–69.

Mozgovoy, M., Kakkonen, T., & Cosma, G. (2010). Automatic student plagiarism detection: Future perspectives. *Journal of Educational Computing Research,43*(4), 511–531.

Nation, I.S.P. (1989) Dictionaries and language learning. In M.L.Tickoo (Ed.), *Learners' Dictionaries: State of the ArtRELC Anthology Series 23* (pp. 65–71).Singapore: SEAMEO Regional Language Centre.

Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Pecorari, D.,&Petrić, B. (2014). Plagiarism in second-language writing. Language Teaching, *47*(3), 269–302.

Pecorari, D. (2003). Good and original: Plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing, 12*(4), 317–345.

Pecorari, D. (2008). *Academic writing and plagiarism: A linguistic analysis*. New York: Continuum.

Petrić, B. (2004). A pedagogical perspective on plagiarism. *NovELTy, 11*(1), 4–18.

Petrić, B. (2012). Legitimate textual borrowing: Direct quotation in L2 students' writing. *Journal of Second Language Writing, 21*, 102–117.

Petrić, B. (2015). What next for research on plagiarism? Continuing the dialogue. *Journal of Second Language Writing, 30*, 107–108.

Setoodeh, K. (2015). The effect of proficiency and task type on the use of paraphrase type in writing among Iranian EFL university students. *Theory and Practice in Language Studies, 5*(12), 2480–2489.

Shi, L. (2004). Textual borrowing in second-language writing. *Written Communication, 21*(2), 171–200.

Shi, L. (2006). Cultural backgrounds and textual appropriation. *Language Awareness, 15*(4), 264–282.

Shi, L. (2012). Textual appropriation and citing behaviors of university undergraduates. *Applied Linguistics*, *31*(1), 1–24.

Stockall, N., & Villar Cole, C. (2016). Hidden voices: L2 students' compensatory writing strategies. *Teaching in Higher Education, 21*(3), 1–14.

Tomaš, Z. (2011). Textual *borrowing across academic assignments: Examining undergraduate second language writers' implementation of writing instruction*. Unpublished doctoral dissertation, University of Utah.

Ur, P. (1984). *Teaching listening comprehension*. Cambridge: Cambridge University Press.

Weigle, S. C., & Parker, K. (2012). Source text borrowing in an integratedreading/ writing assessment. *Journal of Second Language Writing, 21*(2), 118–133.

Yoke, L.L. (2013). Errors in paraphrasing and strategies in overcoming them. *Journal of Creative Practices in Language Learning and Teaching, 1*(1), 4–17.

# Reading Difficulties: What Do We Learn from Young English Language Learners?

**Vera Savić**
*University of Kragujevac, Faculty of Education in Jagodina, Serbia*
*verasavic035@gmail.com*

## 1. Introduction

Research suggests that reading difficulties in English as first (L1) or second language (L2) are rather prominent globally, often affecting negatively learners' self-esteem, motivation, attitudes, confidence, and academic and career prospects negatively (International Reading Association 1999, 2007; Montgomeri, 2007; OECD, 2014; Rasinski, 2013; Westwood, 2008). Although English L2 learners are among the groups that can experience serious failure in reading (Westwood, 2008), there is still limited research in this field, especially in early/beginning reading in English as a foreign language (EFL). Serbia is no exception: since the introduction of English as a compulsory school subject from primary Grade One in 2003, there have been no research studies to verify what can be achieved in early reading skill development, and/or to point to reading difficulties and areas requiring effective action for improving EFL reading (Savić, 2016).

International Reading Association (1999) defines reading as "a complex system of deriving meaning from print" (p. 3), involving motivation to read, strategies to make meaning from print, ability to read fluently and to decode new words, and the knowledge and skills to connect speech sounds to print. Grabe and Stoller (2011) describe the components of reading abilities: lower-level processes and higher-level processes, the former referring to more automatic language processes like lexical access, syntactic parsing and semantic proposition formulation, and the latter describing the process of comprehension and construction of the text and situation model of reader interpretation. Both groups of components of reading abilities are equally difficult to develop (and both of them are aspects of working memory processing), and may be the reasons for reading difficulties: although higher level language skills are important for comprehension, "young children's reading comprehension is strongly predicted by other lower level language skills, such as word reading accuracy and verbal and semantic skills" (Cain, Oakhill & Bryant, 2004, p. 232). Westwood (2008) argues that struggling readers often cannot use the word recognition strategies rapidly, or are not able to decode words with difficult (irregular) spelling patterns due to poor phonic skills and ineffective decoding strategies. If they try decoding letter-by-letter, they become very slow, thus "overloading their working memory and impairing comprehension" (Westwood, 2008, p. 18). As a consequence, beginning readers often depend on top-down processing and on transfer of L1 reading processes.

What is more, research has shown that two different strategies are used in reading, depending on orthographic depth: sublexical (alphabetic) and lexical (whole word) (Perfetti & Dunlap, 2008). Perfetti and Dunlap (2008) contend that two

different types of errors are thus caused: using the lexical strategy in reading "leads readers, when they make errors, to respond with real words based on shared letters or partial visual overlap with the target word, for example, responding 'near' for the word 'never'", while the sublexical strategy "leads to errors with high phoneme overlap with the target word, even when that means producing non-words" at the expense of lexicality (Perfetti & Dunlap, 2008, p. 27). Since Serbian is an alphabetic language with shallow orthography, the tendency to use the sublexical strategy may be more common than the use of the lexical strategy for Serbian readers when reading in Serbian (L1), and may also be transferred to their reading in English (L2). Therefore, the errors that might appear in their reading in English as L2 would result from their tendency to use the same strategy instead of the more appropriate lexical one, i.e. they would probably sound out new words applying the learned letter-sound correspondences, instead of responding with a real word similar to the target one. Moreover, due to the dominant strategy, the reading problems of Serbian beginning readers may involve the rate of reading and accuracy of spelling, rather than phonological awareness or decoding (Perfetti & Dunlap, 2008).

There may be other negative transfer effects, like the influence of L1 syntactic knowledge, of false cognates and orthographic differences, which may cause reading difficulties with beginning readers, such as pronunciation difficulties, slow reading and inaccurate comprehension (Grabe, 1991). Moreover, if a child possesses poor knowledge of vocabulary, or limited knowledge of the world related to the topic of the text, a child will experience difficulties in making necessary inferences (Oakhill, Cain & Elbro, 2015). Apart from vocabulary knowledge, syntactic skills of beginning readers may be the cause of reading comprehension difficulties. Although young learners can make inferences, they sometimes do not do it even upon explicit prompting with questions (Oakhill et al., 2015). Based on their own research, Cain, Oakhill, Barns & Bryant (2001) maintain that inference failure of poor comprehenders may be the result of the readers' difficulty to select the relevant information for inference making, and of their failure to remember the information necessary for making inferences.

More generally, Grabe and Stoller (2011) argue that inefficient operation of lower-level and higher-level comprehension processes is the principal cause of reading difficulties, both with L1 and L2 readers. The reasons may involve text difficulty, inadequate background or linguistic knowledge of the reader, or underdeveloped reading efficiency. The authors explain that in such cases L2 readers may try to translate the text slowly and mechanically in order to understand it, or they may rely on their past experience as a means of comprehending the text. A large number of studies has focused on the importance of bottom-up and top-down processing in L2 reading comprehension (Chamot & El-Dinary, 1999; Macaro & Erler, 2008), concluding that less successful readers rely mostly on phonetic decoding or prior knowledge to make wild guesses, while more successful readers make inferences by combining top-down and bottom-up processes, and by using a variety of strategies and their own world knowledge.

Rasinski (2004a, 2004b, 2013) contends that the lack of proficiency in reading fluency is the most frequent reason of reading comprehension difficulty. Reading fluency "refers to accurate and automatic decoding of the words in the text, along with expressive interpretation of the text, to achieve optimal comprehension" (Rasinski, 2004a, p. 2), which emphasizes the significance of accurate and fast decoding, and appropriate expression and phrasing, for achieving fluency. Research has indicated a strong correlation between reading fluency and comprehension, thus highlighting the

need to include measurements of oral reading fluency in current studies of reading skills development and in assessing the effectiveness of reading and intervention programmes (Fuchs, Fuchs, Hosp & Jenkins, 2001). Accuracy in word decoding, speed and automaticity in word recognition, expressiveness and prosody are three dimensions of reading fluency that are critical for comprehension (Rasinski, 2004a). Comprehension difficulties appear when these three dimensions are absent from a child's reading (Rasinski 2004a, 2004b). Accuracy in word decoding can be measured by counting the number of words read correctly in a minute; when reading in English, Fifth graders should read approximately 100-125 words correct per minute, and their oral reading errors may fall under several categories of miscues, like mispronunciations, substitutions, reversals, omissions, or teacher prompts (Rasinski, 2004a). These quantitative measures can point to the types of decoding difficulties, the kinds of cueing systems (semantic, syntactic, and graphophonic) readers are applying or fail to apply while reading in L2, and to the use of self-corrections and comprehension monitoring skills (Fuchs at al., 2001; Serafini, 2010). What is more, analysis of miscues gives an insight into the reader's "construction of meaning and the process of comprehension" (Goodman & Goodman, 1998, p. 107). High-quality miscues and low quality miscues influence comprehension differently: the former are semantically and syntactically acceptable and do not interfere with comprehension (the sentence makes sense), while the latter are not semantically and syntactically acceptable and impede comprehension (Goodman, 1998). Obviously, reading difficulties at word level inevitably negatively affect reading comprehension skills, causing difficulties at text level.

It is indisputable that comprehension is the essence and "the ultimate goal of reading" (Nation, 2005, p. 248). However, when measuring children's comprehension, scores do not tell the researchers much about these processes: reading tests provide information about the product of reading comprehension, rather than about "the processes (or deficiencies in particular processes) that resulted in the child arriving at that particular score" (Oakhill, Cain & Elbro, 2015, p. 30), or about particular reading difficulties experienced in the reading process. The paper presents the results of a research study that aims to shed more light on reading difficulties of beginning English L2 readers and on the effect reading difficulties have on reading comprehension. The following sections focus on the research study and present its results and suggestions for further investigations in the field.

## 2. The study

### 2.1. Aims

The principal aim of the study was to identify reading difficulties experienced by beginning readers studying English in state school settings in Serbia. More precisely, we defined the following research questions: 1. What difficulties do beginning EFL learners experience in oral reading? Our assumption was that beginning EFL readers would experience decoding difficulties and produce a variety of miscues. 2. What effect does reading accuracy have on reading comprehension? We assumed that reading accuracy and reading comprehension were interdependent. 3. What difficulties do learners experience in comprehending a text? Our assumption was that beginner EFL readers would fail to use effectively the comprehension monitoring skills in reading.

## 2.2. Methodology

Mixed-method approach was applied and both quantitative and qualitative data were collected.

## 2.3. Participants

Twelve (N=12, 7 boys and 5 girls) 11-year-old EFL learners were selected from six state primary schools located in five geographically distant cities in Serbia. The learners formed a sub-sample of a larger study of reading skills of young EFL learners in Serbia (Savić, 2014; Savić, 2016). At the time of the survey, the participants had been learning English formally for four full school years (Grades 1-4) and two months (the beginning of Grade 5), and had started EFL literacy development in Grade 3. Regarding the English language study context, the participants shared most of the conditions: they had followed the English curriculum from Grade 1 with two 45-minute lessons a week; they shared Serbian as a common mother tongue and were literate in it; they were taught by specialist English language teachers holding a bachelor's or a master's degree in teaching English as a foreign language, with teaching experience ranging from nine to twenty-five years. According to Grade 4 curriculum prescribed for English by the Ministry of Education of Serbia, literacy related competence objectives involved reading with understanding a short text (up to 50 words) comprising mainly familiar words, phrases and language structures prescribed by the grade curriculum. All the participants in the study had reported top grades in English awarded at the end of Grade 4, which indicated that they had successfully mastered the required literacy skills. The approach applied in sampling was purposeful non-probabilistic sampling (Patton, 2002), and competent grade-level readers were chosen to provide valuable information for deeper understanding of the process of reading in a foreign language and to allow identification of reading difficulties.

## 2.4. Instruments

The principal instrument was a prompted think-aloud protocol interview, comprising the reading text of the *Cambridge English Young Learners: Young Learners English Tests: Sample Papers – Flyers, 2013* (see the Appendix). The reading text was on a familiar topic (family picnic), but somewhat above the reading competence level of the participants in order to satisfy the demands for a tool for collecting running records for miscue analysis (Goodman, 1998; Serafini, 2010). Also, the text was new to the readers, it was complete (a three-paragraph narrative text with a beginning, middle, and end), and it was "long and challenging enough to produce sufficient numbers of miscues for patterns to appear" (Goodman & Goodman, 1998, p. 103). Face-to-face individual interviews with the participants were conducted in L1 in the form of performance-based assessment, involving open-ended questions as prompts. The interview in the original study (Savić, 2014; Savić, 2016) consisted of four parts; the first two parts aimed to determine the participants' attitudes to reading and their background contexts, while in parts three and four the participants' reading fluency and reading difficulties were determined. The results presented here related only to parts three and four, which aimed to elicit the participants' thoughts/cognitive processes going on in the process of reading, as well as the decoding skills used, the cueing systems applied, and the

reading difficulties experienced. The interviews produced both quantitative (running records) and qualitative data (the participants' answers to open-ended questions).

### 2.5. Procedure

The interviews were carried out face-to-face with individual readers, in regular classes, in quiet premises of the participants' schools. They lasted up to 45 minutes (the regular class period), were audio recorded and later transcribed and analysed. The interview with each participant began with general questions related to a participant's attitudes to and habits of reading, reading in English, and extensive reading. After that the reading task was introduced by the researcher and a participant was invited to look at the picture and try to predict what the story would be about. The participant was then instructed to read the first paragraph aloud, and then his/her comprehension was checked through questions related to the content of the paragraph. After that, the participant was encouraged to express her/his thoughts related to the story and its content, as well as to predict the development of the story in the paragraph that followed. He/She was also encouraged to guess both the meaning and pronunciation of the new words. The same procedure was repeated with the second paragraph, but before the third, the final and the most linguistically challenging paragraph, the participant was instructed to read the paragraph first silently and then aloud. It is important to stress that the participant was never interrupted while reading, nor did he/she receive any assistance with pronunciation or comprehension, or was urged to skip the word/phrase and continue reading when difficulties appeared.

## 3. Results and discussion

### 3.1. Types of miscues and reading comprehension

The types of miscues in oral reading were determined by listening to the recordings made in the interviews and by comparing them to the field notes made by the researcher. Table 1 shows that there were five different miscue types: 1. mispronunciations, producing non-words; 2. substitutions; 3. omissions; 4. insertions; and 5. rereadings.

Table 1. Types of miscues made by the twelve readers

| Type of miscue | Miscue categories |
|---|---|
| mispronunciation (non-words) | - mispronunciation of past tense inflection for regular verbs *-ed* as /ed/ in all cases (non-words): *called, carried, shouted, looked, answered.*<br>- mispronunciation of vowel clusters like *-ai-*, *-au-*, *-ea-*, *-ie-*, *-ou-*, when they appear between consonants (non-words): *near, sweaters, couldn't, because, said, brought, Treasure, course, mouth, found, our.*<br>- mispronunciation of vowels (non-words): *her, takes, few, find, took, other, tell, saw, kind, other, last, carefully, minutes, gave, us.* |

| | |
|---|---|
| substitution | *shout (*for) *shouted*<br>*has* (for) *was*<br>*go*(for) *got* (2x)<br>*balls* (for) *dolls*<br>*play* (for) *place*<br>*give* (for) *gave*<br>*careful* (for) *carefully*<br>*your* (for) *our*<br>*five*(for*) few*<br>*your* (for) *our*<br>*want* (for*) went* (2x)<br>*wants*(for*) wasn't*<br>*on* (for) *one*<br>*in* (for) *it*(3x)<br>*he*(for) *she*<br>*bag* (for) *bad*<br>*give* (for) *gave*<br>*was*(for) *wasn't*<br>*the* (for) *her*<br>*live*(for) *have*<br>*to* (for) *and*<br>*found* (for) *find* |
| omission | -omission of the indefinite article: *a picnic*<br>-omission of a conjunction: *and*<br>-omission of the definite article: *the swings* |
| insertion | - insertion of a preposition:<br>*to (went to on the swings*) (by 5 readers)<br>*to (to near)*<br>- insertion of an article:<br>*the (to the school)*<br>*the (the Emma)* (by 2 readers)<br>*a (a lots of dolls*)<br>- insertion of an auxiliary verb:<br>*is* (*the doll is always sits*) (by 2 readers) |
| rereading | -repetition of individual words |

The running records indicate that the participants' pronunciation difficulties mostly involved mispronunciation that produced non-words (observed in 11 out of 12 participants), mostly due to the wrong pronunciation of past tense affixes, irregular paste tense forms, vowels in conjunctions, adjectives, adverbs, and determiners. Another big group of miscues involved substitutions, mainly of verbs (e.g. infinitive for past tense), pronouns (e.g. prepositions were used instead of pronouns, like *in* instead of *it*,), nouns (replaced by phonetically similar nouns, like *balls* instead of *dolls*, or *play* instead of *place*), and adverbs (adjectives for adverbs, like *careful* instead of *carefully*). Omissions included the omissions of articles and conjunctions, while insertions

involved inserting prepositions and auxiliaries. Some pronunciation difficulties may have resulted from using the sublexical strategy, transferred from L1 reading, by applying the learned letter-sound correspondences (Savić, 2016). This confirms our assumption that the participants would fail to use effectively the comprehension monitoring skills while reading in English, mainly due to the negative transfer of their L1 reading skill.

Although the twelve participants considered themselves to be successful readers, they showed awareness of reading difficulties and were able to describe their own difficulties related to reading in English. The self-reported oral reading difficulties involved pronunciation of new words, pronunciation of some known but difficult words, and pronunciation of inflections. However, these pronunciation problems did not necessarily result in comprehension problems, and mispronunciation did not impede comprehension severely unless it referred to unknown words. Comprehension difficulties were mainly caused by making wrong inferences, by unsuccessful guessing of meanings of new words, and by ineffective use of monitoring strategies. Some examples of unsuccessful guessing included the following words: *lake*, wrongly understood as 'river'; *swings*, wrongly understood as 'river bank'; and *sweaters* wrongly understood as 'small plates' and 'small towels'. Also, wrong inferencing involved attributing the meaning 'swimming' to the unknown word *swings*, relying both on the context of the story in which the characters were having a picnic near the lake (top-down processing), and on orthography, i.e. the beginning syllable '*swi-*' (bottom-up processing) (Savić, 2016). The participants' misinterpretations resulted mainly from their limited knowledge of vocabulary and grammar. Some irregular past tense forms (i.e. *took, couldn't, brought, saw*) were reported to be new to the participants and therefore very difficult to be interpreted.

Some participants were very good at guessing the meanings of unfamiliar words after being prompted. For example, *shouted* was a new word for Participant 1, who managed to guess the meaning after being asked to relate it to the part of the text that immediately preceded this verb (direct speech marked with inverted commas; see the Appendix). Although some participants found it difficult to pronounce regular past tense inflections, they correctly interpreted the past tense verbs (e.g. *called, shouted, looked*). However, the other participants either failed to guess the meanings of unknown words (e.g *lake, sweaters*), or were stuck to their misinterpretations, lacking awareness that their guesses were wrong (e.g. *swings, was, shouted, sweaters, kind, carefully, brought, saw*). When prompted, the participants used their background knowledge and top-down processing skills in interpreting the new vocabulary and grammar.

### 3.2. Reading accuracy and reading comprehension

Table 2 presents the quantitative data related to reading accuracy, determined by taking two measures: the number of miscues was recorded for the participants' reading of the first 100 words, and for the whole text (229 words), so that two accuracy rates were calculated. There were two reasons for taking both measures: firstly, there was the difference in difficulty between the two parts of the text, i.e. the first 100 words and the remaining 129 words, the second part being more difficult in terms of unfamiliar vocabulary and unknown grammar; secondly, taking two measures while reading different paragraphs of the same text was considered necessary for increasing validity and reliability of reading proficiency measures.

Table 2.Reading accuracy levels

| Participant (participant's code) | Reading Accuracy (based on Rasinski, 2004a) (Independent Level: 97-100%; Instructional Level: 90-96%;Frustration Level: < 90%) |
|---|---|
| Participant 1 L-1-1-2-13 | Accuracy (first 100 words): 98% (3 errors: omissions – 1, mispronunciations – 1, self corrections - 1) Total miscues (in 229 word narrative text): 10; accuracy: 96%; Level: instructional |
| Participant2 (L-1-1-1-08) | Accuracy (first 100 words): 99% (1 error: substitutions – 1) Total miscues (in 229 word narrative text):12; accuracy: 95%; Level: instructional |
| Participant 3 L-2-1-2-03 | Accuracy (first 100 words): 94% (6 errors: substitutions – 1, mispronunciations – 5) Total miscues (in 229 word narrative text): 20; accuracy: 91%; Level: instructional |
| Participant 4 L-2-1-1-02 | Accuracy (first 100 words): 99% (1 error: mispronunciation) Total miscues (in 229 word narrative text): 6; accuracy: 97%; Level: independent |
| Participant 5 L-2-2-1-15 | Accuracy (first 100 words): 95% (5 errors: substitutions – 1, mispronunciations – 4) Total miscues (in 229 word narrative text): 23; accuracy: 90%; Level: instructional |
| Participant 6 L-3-1-3-16 | Accuracy (first 100 words): 98% (2 errors: insertions) Total miscues (in 229 word narrative text): 6; accuracy: 97%; Level: independent |
| Participant 7 L-3-1-3-05 | Accuracy (first 100 words): 100% Total miscues (in 229 word narrative text): 1; accuracy: 100%; Level: independent |
| Participant 8 L-3-1-3-21 | Accuracy (first 100 words): 97% (3 errors: insertions – 1, omissions – 2) Total miscues (in 229 word narrative text): 6; accuracy: 97%; Level: independent |
| Participant 9 L-4-1-1-09 | Accuracy (first 100 words): 92% (8 errors: insertions – 1, mispronunciations – 7) Total miscues (in 229 word narrative text): 24; accuracy: 90%; Level: instructional |
| Participant 10 L-4-1-1-10 | Accuracy (first 100 words): 91% (9 errors: substitutions – 4, mispronunciations – 5) Total miscues (in 229 word narrative text):29; accuracy: 87%; Level: frustration |
| Participant 11 L-5-1-2-09 | Accuracy (first 100 words): 99% (1 error: mispronunciations – 1) Total miscues (in 229 word narrative text): 3; accuracy: 99%; Level: independent |
| Participant 12 L-5-1-1-27 | Accuracy (first 100 words): 95% (5 errors: substitutions – 1, insertions – 3, mispronunciations – 1) Total miscues (in 229 word narrative text): 22; accuracy: 90%; Level: instructional |

The results showed that the 12 participants' reading involved all three reading accuracy levels, as follows: 5 participants read at the independent level as they were at least 97% accurate in their oral reading, 6 read at the instructional level as they were at least 90% accurate in their oral reading, and one was at the frustration level as he or she was less than 90% accurate in his/her reading (Savić, 2016). There was a difference both in number and quality of miscues among the three levels: the miscues of the participants who read at frustration and instructional levels involved mostly mispronunciation of a large number of known and unfamiliar words (often not recognized as wrong pronunciations), and substitutions that did not fit the context; on the other hand, the participants who read at an independent level mispronounced few known and/or unfamiliar words (managing to guess the meaning of new words), and the insertions they made (mostly prepositions) did not interfere with their understanding of the text. This confirms our assumption that reading comprehension and reading accuracy are interdependent.

## 4. Conclusion

Our study provides significant insight into the development of beginning reading in English as a foreign language. Listening to the participants reading an unfamiliar text in L2 gave valuable data about their development as readers and about reading difficulties they experienced performing the task. Moreover, the prompted think-aloud protocol yielded significant information related to the participants' background knowledge crucial for understanding the text, to their inference making ability, and the ability to connect ideas in the text and to monitor their own comprehension. Oral reading difficulties recorded and reported in the interviews included mispronunciation of some English phonemes and words, as well as inappropriate reading accuracy. The participants showed metacognitive awareness of the problems they experienced in reading in English as L2, referring to them as their own failure to pronounce new and known words and grammar forms and to interpret them. Still, there were some obvious difficulties that the participants did not recognise or report. Some of these difficulties were related to reading aloud, and were observed in the form of miscues like omissions, repetitions, substitutions, and mispronunciation, which affected the participants' accuracy levels and comprehension. The other difficulties involved making wrong inferences when guessing the meaning of unfamiliar words and grammar forms.

Considering the fact that errors are a necessary part of learning a language, reading difficulties should be seen as components of reading skill development that should be recognised and treated timely and properly. Consequently, development of comprehension monitoring skills and improvement of reading accuracy should be seen as integral parts of teaching beginning reading in English as a foreign language.

## 5. References

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*(1), 31-42.

Cain, K., Oakhill, J., Barns, M., & Bryant, P. (2001).Comprehension skill, inference-making ability and their relation to knowledge. *Memory & Cognition*, *29*(6), 850-859. Psychonomic Society, Inc.

*Cambridge English Young Learners: Young Learners English Tests: Sample Papers – Flyers* (2013).

Retrieved February 10, 2013 from: http://www.cambridgeenglish.org/exams/young-learners-english/flyers/preparation/

Chamot, A. U. & El-Dinary, P. B. (1999). Children's learning strategies in language immersion classrooms. *The Modern Language Journal*, *83*(3) (Autumn, 1999), 319-338. Retrieved March 2, 2013 from: http://www.finchpark.com/courses/grad-dissert/articles/bilingual/immersion-learning-strategies.pdf

Fuchs, L.S., Fuchs, D., Hosp, M.K., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading, 5* (3), 239-256.

Goodman, Y. M. (1998). Miscue analysis for classroom teachers: Some history and some procedures. In C. Weaver (Ed.), *Practising what we know: Informed reading instruction* (pp.226-236). Urbana: National Council of Teachers of English.

Goodman, Y. M. & Goodman, K. (1998). To err is human: Learning about language processes by analyzing miscues. *Reconsidering a balanced approach to reading* (pp. 101-123). Urbana: National Council of Teachers of English.

Grabe, W. (1991). Current developments in second language reading research.*TESOL Quarterly, 25*(3), 375-406.

Grabe, W., & Stoller, F. L. (2011).*Teaching and researching reading* (2nd ed.). Great Britain: Pearson Education Limited.

International Reading Association (1999). Using MULTIPLE METHODS of beginning reading instruction: A position statement.*Newark: International Reading Association*.

International Reading Association (2007). *Teaching reading well: A synthesis of the International Reading Association's research on teacher preparation for reading instruction*. Retrieved June 16, 2013 from http://www.reading.org/Libraries/reports-and-standards/teaching_reading_well.pdf

Macaro, E. & Erler, L. (2008). Raising the achievement of young-beginner readers of French through strategy instruction.*Applied Linguistics, 29*(1), 90-119.

Montgomery, D. (2007). *Spelling, handwriting and dyslexia: Overcoming barriers to learning.* Abington: Routledge.

Nation, K. (2005). Children's reading comprehension difficulties. In M. J.Snowling, & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 248-266). Oxford: Blackwell Publishing Ltd. doi: 10.1002/9780470757642.ch14

Oakhill, J., Cain, K. &Elbro, C. (2015): *Understanding and teaching reading comprehension* (A handbook). London: Routledge.

OECD (2014), *PISA 2012 results: What students know and can do – Student performance in mathematics,reading and science* (Volume I, Revised edition, February 2014), PISA, OECD Publishing.

Retrieved August 16, 2014 from http://dx.doi.org/10.1787/9789264201118-en.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks: Sage Publications.

Perfetti, C. A., & Dunlap, S. (2008). Learning to read: general principles and writing system variations. In K. Koda, & A. M. Zehler (Eds.), *Learning to read across languages* (pp. 13-38). New York: Routledge.

Rasinski, T. (2004a, reprinted in 2011). *Assessing reading fluency*. Honolulu: Pacific Resources for Education and Learning.

Rasinski, T. (2004b). Creating fluent readers.*Educational Leadership*, *61*(6), 46-51. Retrieved Feb. 10, 2013 from http://educationalleader.com/subtopicintro/read/ASCD/ASCD_364_1.pdf

Rasinski, T. (2013). Supportive fluency instruction: The key to reading success (Especially for students who struggle). White Paper. Available at: http://thefutureinreading.myon.com/sites/default/files/attached/White%20Paper%20-%20Supportive%20Fluency%20Instruction%20-%20Dr.%20Rasiniski.pdf

Savić, V. (2014). Investigating reading skills of Serbian young learners learning English as a foreign language. In J. Enever et al. (Eds.), *Conference proceedings from Early Language Learning: Theory and practice 2014* (pp. 108-114). Umea: Umea University.

Savić, V. (2016). Reading difficulties in English as a foreign language [unpublished doctoral dissertation]Novi Sad: University of Novi Sad.

Serafini, F. (2010).*Classroom reading assessments.Classroom reading assessment: more efficient ways to view and evaluate your readers.*Portsmouth:Heinemann.

Westwood, P. (2008). *What teachers need to know about reading and writing difficulties*. Camberwell: Australian Council *for* Educational Research.

## 6. Appendix: Reading text

*Cambridge English Young Learners: Young Learners English Tests: Sample Papers – Flyers* (2013).

My name is Betty and I have a little sister called Emma. She has lots of dolls, but her favourite one is called Daisy. Mum and Dad gave it to her when she was a baby and she takes it everywhere with her. She takes it to school and to her bedroom and when we sit down to eat, the doll always sits next to Emma.

Last Sunday, our family went to the park to have a picnic. We took our dog, Treasure, with us and of course, Emma took Daisy too. There were a lot of people in the park because it was sunny. We found a place near the lake to have our picnic. After lunch, Emma and I went on the swings. After a few minutes, Emma said to me, "Betty, I want Daisy on the swing with me. Can you go and get her for me?" "OK!" I answered.

But when I went back to our picnic, Daisy wasn't there. "Mum!" I shouted, "we've lost Daisy!" Dad looked in all the bags and Mum and I looked under our sweaters and other things, but we couldn't find her. I went to tell Emma the bad news, but when I got there, I saw Treasure. He carried Daisy carefully in his mouth. "Look!" said Emma, "Treasure has brought Daisy to play with me. He's very kind."

# The Language Instruction Scheme for Graduates with a Withheld Degree

**Zoltán Lukácsi**
*Euroexam International, Budapest, Hungary*
*Budapest Business School, Budapest, Hungary*
*lukacsi74@gmail.com*

## 1. Introduction

University undergraduates in Hungary are required by law to have at least one state-accredited foreign language examination certificate at level B2 if they are to gain their degrees (Ministry of Human Capacities Decree 18.). Over the past decade and due to excessive standardization, language instruction at the tertiary level has undergone unfavourable curricular changes. A highly contentious issue is the heavy reduction in the time spent on teaching and learning languages. Degree courses at the Budapest Business School schedule 156 contact lessons, whereas Corvinus University offers 104, and some vocational further education programs at the same institutions only provide 91.As a rule, the cost of language instruction is covered by the tuition fee. However, as the student is held responsible for meeting the academic requirements, most universities are reluctant to invest into language teaching and offer language classes as extra-curricular activities that incur additional expenses.

Input features high on the comprehensive lists of factors that influence L2 learning and the rate of acquisition (Ellis, 2004; Littlewood, 2004). One, albeit crude, way of quantifying input is to express it in terms of guided learning hours, i.e. tutor-led contact lessons. Even though de Jong (2009) harshly criticized the Cambridge English Language Assessment webpage (http://www.cambridgeesol.org/exams/exams-info/cefr.html) for the overly simplistic view of L2 learning pace, it still remains the only comparative measure to guide curriculum design. However basic the recommended 150-200 hours might be to reach level B2 from B1, even the most robust instruction program in Hungarian higher education falls behind these recommendations.

Besides, without an entrance examination or a filter test, universities have no information available about students' L2 level at the outset of teaching. Previous research conducted locally at the Budapest Business School (Lukácsi, 2015, 2016) shows that the curriculum cannot adequately cater for the L2 needs of the majority of university undergraduates, and roughly a third of the freshmen will be unable to meet the English language certificate requirement. Indeed, on the country level in 2014, some 50,000 individuals could not collect their degrees through lack of a language exam (Soós, 2015).

## 2. The language instruction scheme

In 2014, the Ministry of Human Capacities commissioned the National Employment Coordination Agency (OFA NKft.) to develop, organize, and control a language

instruction scheme for university graduates with a withheld degree. Private language schools were invited to apply for funding from the 10,000,000€ budget to run free classes in English (72%), German (25%), and French (2%). The intensive courses were arranged into 120, 180, 240, and 360 hour blocks with 6 to 20 contact lessons a week ("Így jelentkezhettek," 2015). Applicants were expected to (a) have completed their university education, (b) have successfully passed their final examination formerly known as the state examination, and (c) be at level B1 in the language in question. Under the terms of their contract, the language learners agreed to pass a B2 exam by the end of 2016 in return for the free instruction.

Thus far, despite the volume of the venture, no scientifically sound evaluation of the language instruction scheme has been conducted. Publicists and government officials depict the project as a total flop with a pass rate of 15% (G. Tóth, 2015) or as a triumphant success with that of 90% (Soós, 2015; Tarnai, 2015). These articles and internal reports are in stark contrast to viable psychometric expectations and current state language examination accreditation requirements alike. Dávid (2014) claims that a realistic pass rate in a language exam should fall between 35 and 65 per cent. The Accreditation Manual (2016) rules that item difficulty must be between 30 and 70 per cent, which implies that a test paper will also lie within these extremes.

# 3. The study

## 3.1. Aims and research questions

The study reported here aimed to answer three research questions.
RQ 01: What is the pass rate among candidates with a withheld degree?
RQ 02: Do candidates with a withheld degree produce the same pass rate as the rest of the candidature?
RQ 03: How to explain potentially different degrees of success between candidates with a withheld degree and the rest of the candidature?

## 3.2. Method

The validity of such a comparative analysis hinges on the existence of two distinct groups of test takers: (a) regular candidates and (b) candidates with a withheld degree. In the rest of the paper, the term "regular candidate" will be used to describe test takers who were not enrolled in the language instruction scheme. As a preliminary requirement to any meaningful comparison, live examination data were first screened to see if a longitudinal analysis was even possible between pre-2015 and 2015 candidates. Therefore, evidence needed to be collected to support the claim that candidates with a withheld degree were altogether new to the examination system. Given that they had to indicate in a checkbox on the language exam application form if they had enrolled in the language instruction scheme, a simple numerical comparison of annual candidature together with identity checks were sufficient. As a result, the validity of the claim that candidates with a withheld degree were different from regular Euro candidates was established.

### 3.3. Participants

The participants were 12,315 candidates at level B2 who sat for the Euro examination in 2015. The vast majority took the monolingual examination, i.e. tests in listening, reading, speaking, and writing (N = 11,748) without the mediation paper. The sample consisted of two subsamples: (a) candidates with a withheld degree (N = 1,920) and (b) regular examinees (N = 10,395).

### 3.4. Design and instruments

Euroexam International is one of the few certified L2 centres in Hungary that use a data collection design which enables direct comparison of test and examinee statistics longitudinally. In a non-equivalent groups anchor test (NEAT) design (Kolen, 2007), each test administration contains a number of repeated items with known characteristics and, further, a stable ability standard is applied to ensure that the same amount of the latent trait is required in consecutive test periods. In essence, it means that regardless of when the candidate sits for the exam and which tasks they have to complete, their results will be fair and directly comparable. For the estimation of item difficulty and language ability, the item responses were processed in the One-parameter Logistic Model: OPLM (Verhelst, Glas, & Verstralen, 1995). To answer the research question about differential degrees of success, Profile-g (Verhelst, 2012) was used to model the statistical profiles of the two cohorts.

### 3.5. Results and discussion

In terms of the results, reported score means are compared first in the four test parts: listening, reading, speaking, and writing. Next, the pass rates in each part will be presented. Since the Hungarian system of accreditation uses a mixture of the conjunctive and the compensatory approaches (Kaftandjieva, 2004) when combining the test papers into the final examination result, overall achievement and success will be reported separately.

### 3.6. Percentage scores and pass rates

Test paper means are calculated as simple arithmetic averages of the reported scores in the two groups of test takers (Table 1). Officially, these are expressed as truncated percentage values even though the scores do not constitute a ratio scale, and the extremes do not mean total lack or possession of the latent trait in question.

Table 1 Mean Test Paper Scores of Regular and Withheld Degree Candidates

| Candidates | Listening | Reading | Speaking | Writing |
|---|---|---|---|---|
| Regular | 60.5097 | 62.2977 | 64.9267 | 59.4248 |
| Withheld | 51.6312 | 60.2512 | 56.5986 | 55.3920 |
| Grand mean | 59.0627 | 61.9612 | 63.5696 | 58.7617 |

Despite the methodological discrepancies discussed previously, Table 1 demonstrates the difference in attainment between regular test takers and those with a withheld degree. In 2015, candidates with a withheld degree fell behind the grand mean, but particularly the regular examinees in every respect. Listening results showed the biggest difference, followed by speaking, writing, and then reading. A viable interpretation of these figures is to claim that the language instruction scheme was most efficient in teaching language learners to read.

An analysis of the pass rate yields a more criterion-referenced overview of the examination results. In this comparison, the groups are subdivided into two categories according to success on the test paper (Table 2). The pass rate then is the percentage of successful candidates in each group.

Table 2 Pass Rate (%) of Regular and Withheld Degree Candidates

| Candidates | Listening | Reading | Speaking | Writing |
|---|---|---|---|---|
| Regular | 54.9312 | 56.8391 | 67.6763 | 49.4340 |
| Withheld | 36.8171 | 51.6667 | 46.7933 | 37.5269 |
| Grand mean | 51.9791 | 55.9887 | 64.2733 | 47.4764 |

The test paper means in Table 1 already described them as underachievers, but the pass rates in Table 2 depict candidates with a withheld degree as substantially weaker than regular candidates. With the exception of reading, they were more likely to fail than to pass, which was particularly prominent in listening and writing. The reason for the increasing difference between the two cohorts resides in the fact that the consecutive administrations are only equated at the standard. The stable pass mark is at 60 points on the reporting scale of 0 to 100. However, no correct responses are always worth 0 and completely flawless responses are always 100 irrespective of the difficulty of the set of tasks. While the same amount of ability is required for a pass from each candidate regardless of exam period, it also implies that nowhere else on the reporting scale will two similar numbers signify the same. As a result, arithmetic averages will distort genuine differences in ability to varying degrees.

As Government Decree 137 / 2008 2.§ (4) b) rules that in terms of type a language exam can be (a) oral, (b) written, or (c) complex if (a) and (b) are combined, the most authoritative criterion when evaluating the language instruction scheme is the overall success (Table 3).

Table 3 Overall Means and Pass Rates of Regular and Withheld Degree Candidates

| Candidates | Mean | Pass rate (%) |
|---|---|---|
| Regular | 60.7989 | 52.9248 |
| Withheld | 56.0775 | 38.2813 |
| Grand mean | 60.0627 | 50.6415 |

As Table 3 shows, collating the test paper results into an overall score distorts and misrepresents individual differences resulting in a reduced distance between the two groups in the mean. By contrast, the pass rate is robust enough to such a distortion and reveals that they were roughly 15% less likely to be successful than

regular candidates. Grounded in empirical evidence, the response to RQ 01 is that 38% of learners from the language instruction scheme obtained a certificate in English from Euroexam International in 2015. This is significantly lower than the 53% pass rate observed among regular candidates.

### 3.7. Reasons for the different degrees of success

In the next part of the study, the statistical profiles of the two contrasted cohorts on the objectively scored test papers will be presented. Profile analysis is similar to differential item functioning (DIF) analysis in that it can detect systematic deviations from model predictions (Verhelst, 2012). However, while DIF works on the item level, profile analysis partitions items into categories, e.g. tasks, and creates deviation profiles accordingly. A deviation profile shows the difference between the expected profile and the observed profile. As such, all deviation profiles will necessarily sum to zero.

Instead of reporting on the results from each administration separately, the findings will be presented from an aggregate where all the candidates from the same group are combined essentially reducing the method to a pairwise comparison. Of particular importance are the deviation profiles of candidates with a withheld degree to see how balanced their performance was and, eventually, how to explain their poor pass rate. In essence, such an analysis can reveal if a group performed unexpectedly well (or poorly) on a task type when compared to what would be predicted based on the total score on the test paper.

Following the Detailed Specifications (Euro Examinations, 2009) as a blueprint, the listening paper contained three task types: (a) short conversations, (b) making notes, and (c) radio programme. A thorough discussion of what each item in the three test tasks aims to measure is impossible in a publication such as this one, but the general principle is to proceed from a global comprehension to understanding details including stance and attitude, and from short, simple passages towards longer, more complex texts. None of the tasks tapping receptive skills require production of language; therefore, guessing could always play a part, especially so in the radio programme, a three-option multiple choice task.

Test paper reliability was always sufficiently high ($\alpha \geq .75$), and DIF was not present. The data were processed in the OPLM-model with the geometric mean of discrimination indices set to $g = 3$. Model fit was always checked and deemed acceptable for practical purposes. Mean person parameter estimates quite bluntly depicted candidates from the language instruction scheme as less able. Figure 1 presents the statistical profiles of withheld degree test takers and regular candidates broken down to task type.

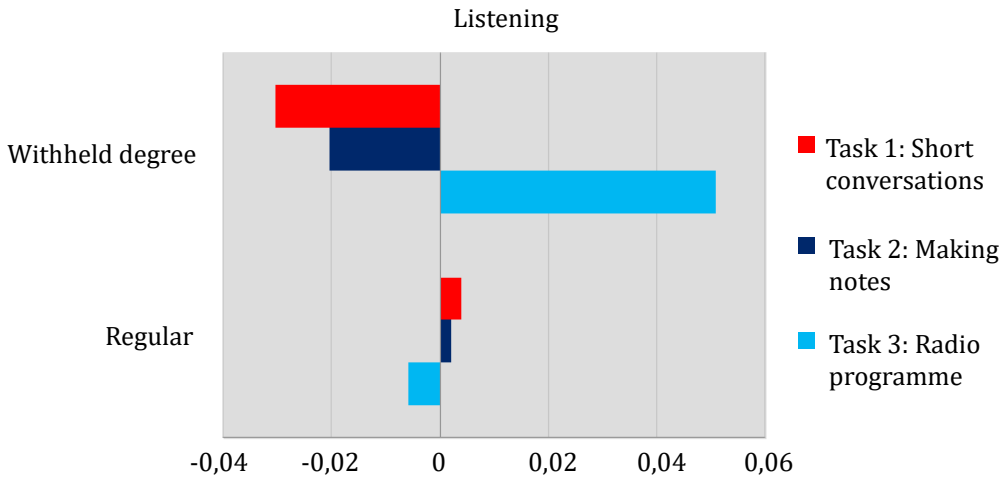


Figure 1. Statistical profiles on the listening paper

In Figure 1, the bars associated with regular examinees are much shorter than in the other group primarily because the sample sizes were not equal. The probability of scoring higher than expected is particularly conspicuous in the withheld degree group on Task 3. On the other hand, the same candidates were more likely to score low on Task 1, a task focusing on more detailed, global comprehension. As model fit was acceptable, the variance in the scores was attributed to differences in the latent trait rather than guessing, test-wiseness or other sources of noise. Arguably, the language instruction scheme successfully prepared these candidates for finding specific, predictable information without understanding longer stretches of spoken text. The two task types where withheld degree test takers underperformed both required comprehension and processing of longer passages and, as such, also placed a heavier burden on working memory capacity.

The reading paper also contained three task types: (a) paragraph headings, (b) scanning, and (c) multiple choice reading. The first two tasks were matching, whereas the last one was a four-option multiple choice task. The psychometric properties of this test paper were similar to the listening part. Figure 2 displays the reading profiles of withheld degree test takers and regular candidates.

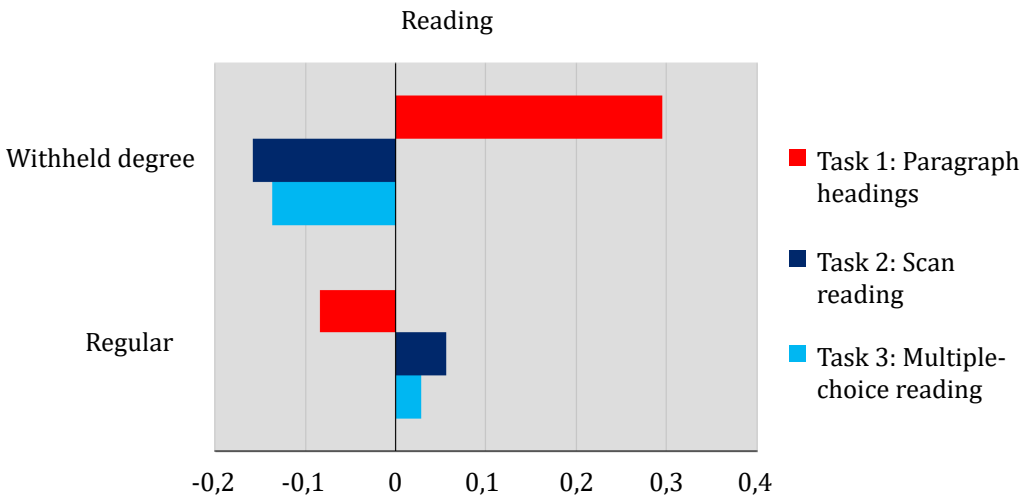


Figure 2. Statistical profiles on the reading paper.

As Figure 2 shows, similarly to the listening profiles, unexpected success on the reading paper was also largely attributed to a single task. This time Task 1: Paragraph headings yielded much better scores than anticipated on the basis of total scores. In this task type, the candidate is expected to choose a heading from a given pool of eight options to six short paragraphs of about 70 words each. The language instruction scheme prepared these language learners for the surface processing of short printed texts. Scan reading and the multiple choice task type both require processing longer passages, which goes beyond local comprehension. It is also important to point out that paradoxically Task 2 is not scanning as the name would suggest but rather search reading (Weir & Khalifa, 2008). While both are classified as expeditious reading, scanning is selective reading at the word level for specific items, whereas search reading never aims for exact word matches.

## 4. Conclusion and implications

Contrary to the enormous stakes of foreign language examinations in Hungary, the role of L2 teaching and learning in higher education is secondary at best. A number of institutions delegate the responsibility of mastering an L2 to the student and for financial reasons refrain from providing L2 education as part of the curriculum. An unwelcome result of this educational policy is reflected in the gradually increasing number of withheld university degrees. By 2014, the number of people unable to collect their degrees had reached 50,000. In response to the growing need for L2 instruction, the Ministry of Human Capacities launched the language instruction scheme providing ample financial support for private institutions to run intensive exam preparation courses. Despite the volume of the budget, no scientifically sound investigation has been launched into evaluating the project.

In this comparative analysis, live examination data from 2015 were reviewed. Even though the sampling applied was comprehensive in the sense that it included the entire population of test takers in Euroexam International, the results only generalize to other certified language centres as far as the efficiency of the language instruction scheme and the level standards in the various language testing systems can be conceived of as constant. With these contentions, the study found that the language instruction scheme helped 38.2813% of the graduates with a withheld degree to their L2 certificates and, more importantly, to their university qualifications.

Any interpretation of such a low pass rate will require further judgement. The language instruction scheme was undoubtedly a success story if the position taken is that each additional degree is a positive result. However, considering the vast amount of resources, the number and intensity of the classes especially in relation to university L2 instruction, and the quantity of the remaining withheld degrees, the language instruction scheme failed to reach its major goal. The language school courses were successful in cramming learners for exam tasks where local comprehension was the focus. The study also found that withheld degree candidates found it hard to deal with longer passages, particularly when global understanding of deep text structures was targeted.

The findings from this study provide empirical evidence in support of previously published recommendations. Nikolov (2011) pointed out that universities should systematically develop students' L2 and suggested that content and language integrated learning should be introduced. Studying a subject per semester in an L2

wouldresult in increasing levels of motivation. Further, such authentic language use would help the language learner acquire expert knowledge and skills eventually leading to learner autonomy. Unless educational policymakers invite professionals in curriculum development, syllabus design and language planning to assist in redesigning the current system of L2 instruction, universities will continue to produce graduates with withheld degrees.

## 5. References


*Akkreditációs kézikönyv* [Accreditation manual]. (2016). OH: NYAK. Retrieved November 24, 2016, from http://www.nyak.hu/nyat/doc/ak2016/ak2016.htm

Dávid, G. (2014). Mitől lesz jó egy nyelvvizsga? [What makes a good language exam?]. *Iskolakultúra, 24*(4), 3-25.

De Jong, J. (2009, June). *Unwarranted claims about CEF alignment of some international English language tests*. Paper presented at the meeting of EALTA, Turku, Finland.

Ellis, R. (2004). Individual differences in second language learning. In A. Davies &C. Elder(Eds.),*The handbook of applied linguistics* (pp. 515-552). Oxford: Blackwell.

Euro Examinations. (2009). Euro detailed specifications. Budapest: EuroExaminations.

G. Tóth, I. (2015). Nyelvi káosz: Botrányba fulladt a diplomamentő akció [Language chaos: The language instruction scheme ends in scandal]. *HVG*. RetrievedNovember 28, 2015, from http://hvg.hu/itthon/201548_botranyba_fulladt_diplomamento_akcio_nyelvl

Government Decree 137. (2008). RetrievedDecember 11, 2015 from http://www.nyak.hu/doc/jogszab/137-2008_EN.pdf

Így jelentkezhettek az ingyenes nyelvtanfolyamra [This is how you can apply for admission to the free language courses]. (2015, July). *Eduline*. Retrieved July 31, 2015, from http://eduline.hu/nyelvtanulas/2015/7/31/diplomamento_kerdesek_XJ5ONM

Kaftandjieva, F. (2004). Section B: Standard Setting. *Reference Supplement to the Preliminary Pilot Version of the Manual for Relating Language Examinations to the CEFR for Languages: Learning, Teaching, Assessment*. Strasbourg: Language Policy Division.

Kolen, M. J. (2007). Data collection designs and linking procedures. In N. J. Dorans, M. Pommerich, & P. W. Holland (Eds.), *Linking and aligning scores and scales*(pp. 31-55). New York, NY: Springer.

Littlewood, W. (2004). Second language learning. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 501-525). Oxford: Blackwell.

Lukácsi, Z. (2015). *How a CEFR-linked test of ESP can inform educational policy*. Paper presented at the meeting of EALTA, Copenhagen, Denmark.

Lukácsi, Z. (2016). Értékteremtés a turizmus-vendéglátás szakos hallgatók szaknyelvi képzésében – avagy: hogyan mérjünk az oktatás segítésére [Added value in the foreign language instruction program for tourism and catering majors – or: how to test to assist education]. *Nyelvvilág*, *17*, 5-18.

Ministry of Human Capacities Decree 18. (2016). Retrieved December 11, 2016 from http://net.jogtar.hu/jr/gen/hjegy_doc.cgi?docid=A1600018.EMM&timeshift=fffffff4&txtreferer=00000001.TXT

Nikolov, M. (2011). Az idegen nyelvek tanulása és a nyelvtudás. *Magyar Tudomány,*

*172*(9), 1048-1057.
Soós, A. (2015, November). *A diplomamentő programról* [About the language instruction scheme for graduates with a withheld degree]. Paper presented at the meeting of Certified Foreign Language Examination Centres, Budapest, Hungary.
Tarnai, R. (2015, December). Sikeres volt a diplomamentő program [The language instruction scheme was successful]. *Pest Megyei Kormányhivatal.* Retrieved December 11, 2015, from http://www.kormanyhivatal.hu/hu/pest/hirek/sikeres-volt-a-diplomamento-program.
Verhelst, N. D. (2012). *Profile-g*: *A computer program for profile analysis*. Tiel: Eurometrics.
Verhelst, N. D., Glas, C. A. W., & Verstralen, H. H. F. M. (1995). *One parameter logistic model: OPLM*. Arnhem: Cito.
Weir, C., & Khalifa, H. (2008). A cognitive processing approach towards defining reading comprehension. *Research Notes, 31*, 2-10.

# Keyword Analysis of Short Stories for Young Language Learners

**Gabriella Lőcsey**
*University of Pécs, Hungary*
*locseyg@gmail.com*

## 1. Introduction

As in recent years a great deal of attention has been given to discourse analysis, the goal of this study is to find out to what extent discourse analysis can be applied in foreign language education. Researchers identify lexical patterns with the help of corpus-based analysis. These lexical patterns determine both the content and the structure of texts (Fischer-Starke, 2010). The identification of these patterns used to be carried out intuitively, but nowadays this is done by computer programs. While researchers used to concentrate on finding and analyzing keywords to recognize and explain essential cultural and societal concepts, an alternative way of keyword analysis has emerged as a result of development of modern technology. *WordSmith Tools* (1996-2008), Scott's software, made it possible to use a quantitative approach to calculating keywords in a sample of texts or a corpus and to compare them to a reference corpus. This software allows for any word to become key if it occurs frequently enough in the examined text when compared to a reference corpus (Baker, 2004). These keywords indicate the content; this way they are also important from the educational perspective as well. According to Nation (2006), the knowledge of more than 95% of the written and spoken texts is needed for learners to be able to understand the meaning of a text. The more keywords of a text a language learner knows, the better the understanding of the content of the text. Analyzing keywords of various texts on different language levels might provide a keyword list that could be applied and would be advisable to use in foreign language education to make the instruction more effective. In this pilot study, texts of short stories for young learners were analyzed to find out what keywords and how many word-families the learners need to know to cope with tales.

## 2. Theoretical Background

### 2.1. Keyword and concepts

What does keyword mean? According to Pierre Guiraud, (1954) who first used this concept, keyword ('mots-clés') means simply a statistically significant lexical item. Enkvist (1964) and Gray (1964) referred to keywords as style markers whose frequencies differ significantly from their frequencies in a norm. Style markers are contextually bound linguistic elements. "Repetition is the notion underlying both style markers and hence keywords, but not all repetition, only repetition that statistically deviates from the pattern formed by that item in another context" (Culpeper, 2009, p.33).

Stubbs (2010) introduced three different senses of the term 'keywords'. Sense 1 is cultural and as Wierzbicka (1997, p.156) states, keyword is a 'focal point around which entire cultural domains are organized'. Firth (1957) and Williams (1967/1983) are the best examples of sense 1, as they found the most salient words intuitively describe culture in the finest way. However, Williams' work was grounded in a cultural studies: his list cannot be a good ground to define a 'general theory' (Stubbs, 2010, p.25). Stubbs (2010) argued that Williams had no theory of how the vocabulary of a language was structured and what the connection between the words, texts and text-types was. Sense 2 comes from comparative quantitative corpus analysis and it is statistical. 'Keywords are words which are significantly more frequent in a sample of text than would be expected, given their frequency in a large general reference corpus' (Bondi, 2010 p.25). Mike Scott's keyword software, *WordSmith Tools* (Scott, 1998) made it easy and fast to carry out corpus analysis. Baker (2004) highlighted the benefit of quantitative keywords and claimed that they guide the researcher to discover salient concepts in a text since the lexical dissimilarities of various texts are in focus. Sense 3 relates to Francis' approach and it is corpus-driven as it scrutinizes alterations in vocabulary and grammar use. Francis (1993) examined how people express their shared values and how the meaning is conveyed by a changeable lexico-grammar pattern.

The advent of computer programs such as *WordSmith Tools* (Scott,1999) and *W Matrix* (Rayson 2005, 2008) has facilitated both grammatical and semantical analyses of texts and allowed researchers to identify keywords easily and rapidly (Culpeper, 2009). Any word can become a keyword if it appears frequently enough in the examined text when compared to a reference corpus. Scott distinguishes three types of keywords: proper nouns, content keywords and high frequency words. Content words are words that people would recognize as key and they are markers of 'aboutness' of a certain text. High-frequency words are for instance *because*, *shall* or *already* which may be style markers (Baker, 2004). However, keywords do not need to consist of single words. Keyword lists consist of words of two-, three- and four- word 'clusters' (Scott, 1999) or lexical bundles (Biber, Johansson, Leech, Conrad and Finegan 1999, p.990).

### 2.2. Previous research

More and more studies have been conducted on keywords of various genres: romantic fiction (Tribble, 2000), political correctness in newspapers (Johnson, Culpeper and Suhr, 2003). Xiao and McEnery (2005) were interested in spoken and written discourse and Culpeper (2009) studied key part-of-speech and key semantic domains in addition to keywords in Shakespeare's Romeo and Juliet.

Fischer-Starke (2010), in her book, provides a detailed summary of previous research on keywords by highlighting the useful and successful combinations of quantitative data and qualitative analysis in stylistics in Toolan's (2004) and Culpeper's (2002) work. The aim of using mixed methods is to encode literary meanings in the data by looking at collocations of the keywords in their concordance lines.

### 2.3. Issues in keyword analysis

The selection of an appropriate reference corpus in keyword analysis is essential. Some researchers (Scott & Tribble, 2006) argue that a large sample is advisable, while

others (e.g., Xiao & McEnery, 2005) did not find the size of the datasets for comparison very important in making a keyword list. The scope of the reference corpus is not as important as its relevance to the corpus being examined. It is precisely the genre of the reference corpus that prominently influences the keywords identified.

An essential parameter set in a Keyword program is the minimum frequency cut-off point. It means that the program 'excludes the words that will be identified as unusual simply because they happen not to have occurred or to have occurred very infrequently in the dataset of the reference corpus' (Culpeper, 2009, p.36). The choice of minimum frequency of cut-off often depends on the size of the data set. In this study the Keynes cut-off point is 25, originally set in the Keyword analysis program.

The other important parameter is the test for statistical significance. This figure 'calculates the significance of the unusualness of the keyword' (Culpeper, 2009, p. 36). A text with a keyword ratio of 0.01 is an extremely high keyword text and it means that it contains many words distinct to this text. In contrast, a low keyword text with a ratio of 0.009 uses mainly general words.

## 3. The Study

### 3.1. The aims of the pilot study

The aim of this pilot study was twofold. It attempted to explore the keywords in a selected corpus of 48 online short stories; how many word-families short stories consisted of and how many of these word-families young learners need to be familiar with to cope with the text confidently.

### 3.2. Research questions

This study was conducted to provide answers to the following research questions:
1. What are the keywords of 48 short stories?
2. How many word families are there in the whole texts?
3. How many word-families do young learners need to be familiar with in order to read these short stories?

### 3.3. Data analysis and procedures

Forty-eight short stories were chosen as corpus data out of 60 short stories for young language learners provided by The British Council on its website (https://learnenglishkids.britishcouncil.org/en/shortstories?page=0%2C0%2C0%2C0%2C0%2C0 ). The corpus consisted of 10,087 words. These tales are recommended to children of three age ranges; 1-5, 6-8 and 9-11. The titles of the stories are listed in the Appendix in the order of their appearance on the website. The selected reference corpus was the 14-million-word mixed written-spoken, US-UK corpus developed by Paul Nation (2007), as a basis for the first 2k of the British National Corpus- Corpus of Contemporary American English lists (BNC-Coca list).

Four procedures were applied in order to answer the research questions. First, keyword analysis was performed with the help of the Keywords Extractor (Cobb, 2007)

which is a part of the *Compleat Lexical Tutor* (Cobb, 2007).' This program determines the defining lexis in a specialized text or corpus by comparing the frequency of its words to the frequency in a more general reference corpus' (http://www.lextutor.ca/cgi-bin/range/texts/index.pl). Then the concordance plots of some words were investigated by AntConc Computer Software (Version 3.4.3). Vocabulary Profilers (Cobb, 2007) were also applied to find out which frequency band each word of the texts belong to; and finally, a special version of Vocabulary Profilers, VP-kids (Cobb, 2007) was used to look at how many words the texts contain from 10 frequency bands of roughly 250 families. The Keynes cut-off point is 25 in this study which was originally set in the Keyword analysis program.

## 3.4. Results and discussion

### 3.4.1. Keywords analysis of short story corpus

The keyword list (Table 1) shows all the words in the short story corpus that are at least 25 times more frequent in the text than in the reference corpus. The number preceding each word is the number of times more recurrent this word is in the short stories than it is in the corpus BNC-COCA list. For example, in the output the 27, 516 princess (Table 1) means that the word princess is 27,516 times more frequent in the short story corpus than it is in the reference corpus. This possibly means that 'princess' word has an important role (key) in the analyzed corpus.

Table 1 Keywords analysis of short story corpus

| | | | | | |
|---|---|---|---|---|---|
| (1) | 27,516.00 | princess | (36) | 79.00 | tummy |
| (2) | 18,827.00 | okay | (37) | 66.21 | cave |
| (3) | 17,379.00 | clothes | (38) | 63.24 | monster |
| (4) | 11,586.00 | computer | (39) | 62.97 | surf |
| (5) | 10,138.00 | quinine | (40) | 62.07 | lantern |
| (6) | 8,689.00 | caliph | (41) | 62.06 | circus |
| (7) | 5,793.00 | especially | (42) | 61.19 | bean |
| (8) | 4,055.00 | splodge | (43) | 59.66 | Santa |
| (9) | 3,258.50 | Persia | (44) | 54.59 | giant |
| (10) | 2,896.33 | ghoul | (45) | 54.31 | clown |
| (11) | 1,448.17 | ogre | (46) | 50.37 | tower |
| (12) | 579.27 | snowman | (47) | 49.51 | eagle |
| (13) | 452.56 | harp | (48) | 47.95 | mouse |
| (14) | 440.78 | pyramid | (49) | 47.78 | lion |
| (15) | 434.50 | rainforest | (50) | 47.10 | torch |

| | | | | | |
|---|---|---|---|---|---|
| (16) | 413.79 | malaria | (51) | 47.05 | roar |
| (17) | 386.20 | hippo | (52) | 44.56 | jungle |
| (18) | 382.55 | ouch | (53) | 44.54 | jack |
| (19) | 310.36 | Ramadan | (54) | 39.04 | seed |
| (20) | 296.23 | dinosaur | (55) | 38.79 | emperor |
| (21) | 271.56 | lonely | (56) | 37.78 | snake |
| (22) | 241.39 | shush | (57) | 36.82 | elephant |
| (23) | 206.90 | tweet | (58) | 33.55 | asleep |
| (24) | 169.27 | dragon | (59) | 31.48 | shout |
| (25) | 163.95 | Greece | (60) | 31.26 | knight |
| (26) | 160.92 | porridge | (61) | 30.94 | clue |
| (27) | 149.81 | pigeon | (62) | 29.48 | castle |
| (28) | 129.51 | buzz | (63) | 28.40 | medicine |
| (29) | 125.93 | icon | (64) | 27.50 | fame |
| (30) | 120.69 | marathon | (65) | 26.49 | tunnel |
| (31) | 117.43 | tomb | (66) | 26.16 | animal |
| (32) | 92.83 | teddy | (67) | 25.83 | climb |
| (33) | 91.06 | planet | (68) | 25.32 | ocean |
| (34) | 85.20 | hooray | (69) | 25,19 | bounce |
| (35) | 81.59 | parrot | | | |

With the help of the Concordance plot the frequency and the places of the words' occurrence in the corpus can be investigated. The more common a word is in the corpus, the darker it is in that part of the concordance plot. According to the Keyword list, 'princess' was the most frequent word; thus, a question emerged whether this word appeared in different parts of the corpus or can it be densely found in one particular section. As it can be clearly seen from the concordance plot (Figures 1-3) 'princess' appears 19 times in the texts (Figure 1) and this word is seemingly repeated in one or two stories. The first 20 most frequent keywords are generally content words and nouns and this is also the characteristic feature of the rest of the list. However, there are two exceptions in the first twenty words: *okay* (5 times in Figure 2) which 'serves as a routine compliant response' (Biber et al, 1999, p. 1090) and *ouch* (14 times in Figure 3) which is 'an interjection' and has 'an exclamatory function' (p.1083) expressing sudden physical pain.

The word 'okay' can be found in different part of the corpus. The title of the last short story is 'I'm too ill' and this must be the reason why 'ouch' appears so frequent at the end of the corpus.

Figure 1. Concordance plot of 'princess' (AntConc (Version 3.4.3) [Computer Software].

Figure 2. Concordance plot of 'okay' (AntConc (Version 3.4.3) [Computer Software].

Figure 3. . Concordance plot of 'ouch' (AntConc (Version 3.4.3) [Computer Software].

Another finding of interest concerns the appearance of words such as quinine, caliph, Persia, ghoul, pyramid, malaria, Ramadan, Greece and marathon. No doubt that these words are mostly culturally dependent words and may not be a part of from the general knowledge of young language learners who are not part of the culture or tradition which they refer to. Short stories of different cultures have different keywords, as earlier Wierzbicka (1997) pointed out the relationship between keyword and culture. The selection of various expressions and traditional stories from the different parts of the world is welcome and can be a good start of awareness raising as it may make young learners and their social context more understanding and tolerant towards people from other cultures. Nowadays, it is essential because most of us live in a multi-cultural environment.

### 3.4.2. Lexical Vocabulary Profile

The lexical frequency profile method was developed by Laufer & Nation (1995). This procedure divides the words of texts according to which frequency band each word belongs to: first 1,000 most-frequent, second 1,000 most-frequent, the most-frequent 'academic' words not in either of the other two lists and the remainder or 'off list'. Vocabulary Profilers were adapted for Web by Tom Cobb (2007) (http://www.lextutor.ca/cgi-bin/range/texts/index.pl). This program shows the numbers and percentages of words and word families in the scrutinized English text coming from each of the three word lists and those which are not recognized.

Table 2 shows the number of word families, types, tokens and percentages at each level. The texts of these 48 short stories offered to young learners a total of 10,087 words. The first row of the Table 2 shows that 7,937 tokens belong to the K1 frequency band (first 1,000 most-frequent words) in the texts of the short stories. This amounts to 78.69% of the total words and it consists of 842 types. More than half (4,678) of these words are function words and 3,259 are content words. The second 1,000 most frequent words (K2 words) account for 913 tokens of 361 word types and 273 word families. Altogether 87.74% of the total words belong to K1 and K2 frequency band (K1+K2). Another finding of interest is that academic words also appear in these tales:

42 tokens of 16 different types and these types make up 14 word-families. The number of off-list words is extremely high, 1,195 words of 469 types. The appearance of off-list words has a higher percentage (11.82%) than the K2 words and academic words together (9.47%).

Table 2 Families, types, tokens at each word level in 48 short stories

| Levels | Families | Types | Tokens | Percent |
|---|---|---|---|---|
| K1 Words (1-1000): | 549 | 842 | 7,937 | 78.69% |
| Function: | ... | ... | (4,678) | (46.38%) |
| Content: | ... | ... | (3,259) | (32.31%) |
| K2 Words (1001-2000): | 273 | 361 | 913 | 9.05% |
| K1+K2 | | ... | ... | (87.74%) |
| AWL Words (academic): | 14 | 16 | 42 | 0.42% |
| Off-List Words: | ? | 469 | 1,195 | 11.85% |
| Total | 836+? | 1,688 | 10,087 | 100% |

Table 3 shows the summary of the data analysis of Vocabulary Profilers. It shows that the texts are built of 10,087 tokens and 1,688 types. The lexical density of the texts, the ratio of the content and total words (0.54%) shows that slightly more than half of the words are content words in these short stories.

Table 3 Summary of the data analysis of Vocabulary Profilers

| | |
|---|---|
| Words in text (tokens): | 10,087 |
| Different words (types): | 1,688 |
| Type-token ratio: | 0.17 |
| Tokens per type: | 5.98 |
| Lex density (content words/total) | 0.54 |

### 3.4.3. Vocabulary Profilers for kids (VP-kids)

VP-kids matches the text of the short stories against 10 modified 250-word lists generalized from several studies of children's oral productions by Stemach and Williams (Cobb, 2007). Table 4 lists the classification of word families, types, tokens and the off-list words across the ten frequency levels. The most tokens are from the KID250-1 level (1 is the most frequent) and on the second most frequent level, KID250-2 there are 1,128 words. Furthermore, the numbers of word families, types and tokens are higher on the more frequent level. The number of words in each category gradually decreases except for the tokens of the first two levels where this reduction is radical. The most interesting outcome is that there are numerous off-list words (910 tokens, 409 types) in the texts and their number is close to the number of word tokens on the second level (1,128).

Table 4 Number of families, types and tokens

| Freq. Level | Families | Types | Tokens | Coverage% | Cum% |
|---|---|---|---|---|---|
| **Kid250 - 1:** | 182 | 338 | 6,331 | 62.78 | 62.78% |
| **Kid250 - 2:** | 173 | 270 | 1,128 | 11.18 | 73.96% |
| **Kid250 - 3:** | 137 | 191 | 517 | 5.13 | 79.09% |
| **Kid250 - 4:** | 94 | 116 | 350 | 3.47 | 82.56% |
| **Kid250 - 5:** | 71 | 95 | 268 | 2.66 | 85.22% |
| **Kid250 - 6:** | 66 | 80 | 163 | 1.62 | 86.84% |
| **Kid250 - 7:** | 58 | 67 | 141 | 1.40 | 88.24% |
| **Kid250 - 8:** | 45 | 55 | 122 | 1.21 | 89.45% |
| **Kid250 - 9:** | 45 | 48 | 115 | 1.14 | 90.59% |
| **Kid250 - 10:** | 20 | 22 | 40 | 0.40 | 90.99% |
| **Off-List known:** | 237 | 259 | 457 | 4.53 | 95.52% |
| **Off-List unknown:** | ? | 150 | 453 | 4.49 | 100.00% |
| Total | 1,128+? | 1,691 | 10,085 | 100% | 100% |

I looked for the first ten words of the keywords list in Table 2 in the 10 modified 250-word lists of Table 4 and the outcome is astonishing. Six out of ten, *quinine, caliph, especially, splodge, Persian and ghoul* are among the words on off-list but known. The other words such as *princess* can be found in the word group 6, *okay* in group 2, *clothes* in group 3 and the word *computer* in group 5.

## 4. Conclusion

The aim of this study was to analyze the keywords of 48 online short stories offered to early English learners in order to investigate how many word families can be found in the short stories altogether. Nation (2006) stated that 98% coverage is ideal, to entirely comprehend written and spoken texts. According to Table 4, in the text of these short stories the 98% coverage can be reached only if young learners also have the knowledge of some of the off-list words besides the other words on the frequency list. The different themes of these short stories are also thought-provoking and the high occurrence of culturally dependent keywords raises the issue of how much time a word needs to step from the off-list a bit higher at least to KID250 -10. Further research is necessary because of the limitations of the study. Only 48 stories were analyzed in this study and the analyses only offer insights into these short texts. It would be good to find out about many more stories. Keyword analysis can be of great help in language teaching as it can have various meaningful results. It can help teachers and learners comprehend and focus on the actual content and can give a kind of guidance on what words or word-families teachers need to teach to make young language learners understand this level of written and spoken English.

## 5. References


Anthony, L. 2014. AntConc (Version 3.4.3) [Computer Software]. Tokyo, Japan: Waseda University. Available from http://www.laurenceanthony.net

Baker, P. 2004. Querying keywords: Questions of difference, frequency and sense in keywords analysis. *Journal of English Linguistics*, *32*(4), 346-359

Biber, D., Johansson, S., Leech, G., Conrad, S. &Finegan, E., (1999). *Longman grammar of spoken and written English*. London: Longman.

Bondi, M. 2010. Perspectives on keywords and keyness: An Introduction. In *Keyness in Texts,* M. Bondi & M. Scott (Eds), 1-21.Amsterdam: Benjamins.

Cobb, T. 2007. Keywords Extractor v.2 [computer program].

http://www.lextutor.ca/cgi-bin/range/texts/index.pl. Cobb, T. 2007. VP-Kids v.9 [computer program].

http://www.lextutor.ca/cgi-bin/range/texts/index.pl.

Cobb, T. Web VP Classic v.4 [computer program]. Accessed 3 Jan 2016 from

http://www.lextutor.ca/cgi-bin/range/texts/index.pl.

Culpeper, J. 2002. Computers, language and characterization: An analysis of six characters in *Romeo and Juliet*. In U. Melander-Marttala, C. Ostman & M. Kytö (Eds.), *Conversation in Life and in Literature: Papers from the ASLA Symposium*, *Association Suédoise de Linguistique Appliqué (ASLA),* 15. (pp.11-30). Uppsala: Universitetstryckeriet. (available at: *h*ttp://www.lexically.net/wordsmith/ corpus_linguistics_links/papers_using_wordsmith.htm)

Culpeper, J. 2009. Keyness. Words, parts-of-speech and semantic categories in the character-talk of Shakespeare's Romeo and Juliet. *International Journal of Corpus Linguistics 14*(1), 29- 59.

Enkvist, N. E. 1964. "On defining style". In N. E. Enkvist, J. Spencer & M. Gregory (Eds.), *Linguistics and style*, (pp. 1–56). Oxford: Oxford University Press.

Firth, J. 1957. *Papers in Linguistics* 1934-1951. London: Oxford University Press.

Francis, G. 1993. A corpus-driven approach to grammar: principles, methods and examples.

In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds), *Text and Technology* (pp.137-56). Amsterdam: Benjamins.

Fischer-Starke, B. 2010. *Corpus Linguistics in Literary Analysis: Jane Austen and her Contemporaries.* London: Continuum. DOI: 101515/cast-2011-0010.

Gray, H. 1964. *Anatomy* or the London Telephone Directory of 1960.

Guiraud, P. 1954 [1970]. *Les Caractéres statistiques du vocabulaire*. (pp. 64–67) reprinted In: P. Guiraud & P. Kuentz (Eds.), *La Stylistique Lectures*, (pp.222–224). Paris: Klincksieck.

Johnson, S., Culpeper, J.& Suhr, S. (2003) From 'politically correct councillors' to 'Blairite nonsense': discourses of 'Political Correctness' in three British newspapers. *Discourse and Society* 14(1). 29-47.

Nation, I.S.P. 2006. How large a Vocabulary is needed For Reading and Listening? *The Canadian Modern Language Review, La Revue canadienne des langues vivantes, 63*(1), 59-82.

Nation, I.S.P. & Beglar, D. 2007. A vocabulary size test. *The Language Teacher, 31*(7), 9-13. Rayson, P. 2005. *WMatrix: A Web-based Corpus Processing Environment*. Computing Department, Lancaster: Lancaster University. http://www.comp. lancs.ac.uk/ucrel/wmatrix/. Rayson, P. 2008. From key words to key semantic domains". *International Journal of Corpus Linguistics*, *13* (4), 519–549.

Scott, M. 1998. *WordSmithTools*. Version 3. Oxford: Oxford University Press.

Scott, M. R. & Tribble C. 2006. *Keywords and corpus analysis in language education*. Amsterdam & Philadelphia: John Benjamins.
Stubbs, M. 1996. *Text and corpus analysis*. Oxford: Blackwell.
Stubbs, M. 2010. Three concepts of keywords. In *Keyness in Texts*, M. Bondi & M. Scott (Eds), 21-43. Amsterdam: Benjamins.
Toolan, M. 2004.Values are descriptions; Or, from literature to Linguistics and back again by way of keywords*. Belgian Journal of English Language and Literatures (BELL New Series 2*): 11-30.
Wierzbicka, A. 1997. *Understanding cultures through their key-words*. Oxford: Oxford University Press.
Williams, R. 1976/1983. *Keywords. A Vocabulary of Culture and Society*. London: Fontana.
Xiao, R. & McEnery, T. 2005. Two approaches to genre analysis: Three genres in Modern American English. *Journal of English Linguistics*, *33*(1), 62–82.

## 6. Appendix

**Titles of 48 stories**

1. My Secret Team
2. A Dog's life
3. Our colorful world
4. Ratty robs a bank
5. The Great Race
6. Why Anansi has thin legs?
7. The princess and the dragon
8. Dinosaur Dig
9. Spycat
10. The Lucky Envelope
11. The Mummy
12. Eric the Engine
13. The voyage of the animal
14. The Lucky Seed
15. My Favourite Day: Chinese New Year
16. Superhero High
17. The First Marathon
18. Santa's Little Helper
19. The story of quinine
20. No dogs
21. The lion and the mouse
22. The treasure map
23. The bird king
24. Teddy's adventure
25. My favourite day - Eid al Fitr
26. Ali and the magic carpet
27. Monster shopping trip
28. The hungry dragon
29. The cold planet
30. Buzz and Bob's big

31. The lazy bear
32. Jack and the beanstalk
33. George and the dragon
34. Goldilocks and the three bears
35. Pyramids in Paris
36. My favourite clothes
37. What's that noise?
38. The greedy hippo
39. The animal shelter
40. The haunted house
41. Twin's week
42. ABC Zoo
43. Circus escape
44. Planet Earth
45. I couldn't believe my eyes
46. The snowman
47. The lantern(A Ramadan story)
48. I'm too ill

# Vocabulary Acquisition Revisited: The Case of Human Movement Science

**Darija Omrčen**
*Faculty of Kinesiology University of Zagreb, Zagreb, Croatia*
*darija.omrcen@kif.hr*

## 1. Introduction

That women excel men in language ability and skills (Gu, 2002; Kimura, 1999; Moir & Jessel, 1992; Wilson, 1975) is a continuously challenged stereotype whose refutation seems to be corroborated time and again. Some studies found no gender differences in this respect (Agustín Llach & Gallego, 2012; Brantmeier, 2003; Morris, 1998; Voyer & Voyer, 2014). The results of some other studies showed that men bettered women (Boyle, 1987; Scarcella & Zimmerman, 1998), whereas in some women did better in one and men in the other segment of a study (Newman, Groom, Handelman, & Pennebaker, 2008; Lin, 2011; Lin & Wu, 2004; Xia, 2013). In some cases, results regarding the absence/presence of differences between men and women depended on the methodology and experimental design of research (Ali, 2016; Pahom, Farley, & Ramonda, 2015). This inconsistency in findings holds true both for language ability/skills in general and for L2 language acquisition in many of its possible aspects.

Research into English for Specific Purposes (ESP) vocabulary acquisition in kinesiology students has adduced the absence of differences between men and women in a written multi-task ESP test, dictation and English into Croatian open-ended translation (Omrčen & Bosnar, 2008). No differences between men and women were found in the knowledge of selected English grammatical constructions either (Omrčen, 2011a). Another study showed that men achieved better results in translating higher technicity-level terms into Croatian, whereas women were better in translating terms of a lower technicity level (Omrčen & Bosnar, 2010). Research into the knowledge of German kinesiological vocabulary, carried out using German into Croatian open-ended translation tests, has also confirmed the lack of differences between male and female students of kinesiology (Omrčen, 2010, 2011b)

### 1.1. Vocabulary assessment

Language ability is an intricate concept that is commonly measured by various tests of two spoken and two written language skills – listening and speaking on the one hand, and reading and writing on the other. Vocabulary knowledge, accompanied by the choice of words and style (Powers, 2010, p. 3), is said to be a subskill that lies in the background of all four language skills.

The choice of language assessment tools varies in compliance with its purpose. The plethora of omnifarious tests makes this choice easier, however, the selection of tests should be done by both clearly pinpointing the aim of testing and by structuring the assessment tool with caution. To test vocabulary knowledge various tests have been applied varying with regard to the research aims, participants, etc.

Research addressing contextual and decontextual vocabulary knowledge yielded disparate results. Van Zeeland (2013) found that learners' contextual and decontextual vocabulary knowledge agree in 65% of the cases. Kelly (1990) regarded contextual clues to vary as for their relevance in learning a word, and Gu (2015, p. 4) considered some contextual clues as helpful, and others as useless or even confusing.

Multiple-choice tests (termed *multiple-choice tasks* in this text) (MCT), frequently used to test L2 acquisition, have been subject to analyses commencing from different points of view. In general, validity and reliability, as well as design (Considine, Botti, & Tomas, 2005) of MCTs and guidelines for their construction (Haladyna, Downing, & Rodriguez, 2002; Moreno, Martínez, & Muñiz, 2006) have long been in the focus of research. As for L2 acquisition, MCTs have been used to test incidental vocabulary learning (Laufer & Hulstijn, 2001; Waring & Nation, 2004; Yoshii, 2006), comprehension-focused extensive reading (Horst, Cobb, & Meara, 1998), second language vocabulary acquisition through reading (Pulido, 2003), vocabulary size (Wood, 1999; Laufer & Nation, 1999), etc. Although MCTs are believed to strongly encourage guessing (Walstad & Becker, 1994, p. 193) and to imply only the simple recognition of facts (Bennett, Rock, & Wang, 1991), MCTs offer an opportunity for each item in it to reflect specific content (Haladyna et al., 2002, p. 312). Since in a MCT choices may be kept independent (Haladyna et al., 2002, p. 312), inclusion both of a single correct choice (a key) and of students' typical errors as distractors into choices (Haladyna et al., 2002, p. 312; McNamara, 2005, p. 6) increases the accuracy-related requirements crucial in language for specific purposes. To be able to include the best possible distractors in the choice set, a test maker should be familiar with errors that test takers most frequently make with regard to selected terms. Ultimately, MCTs are considered to possess a high degree of reliability, which is the result of the scoring procedure that is objective and unbiased (Haladyna, 1999; Douglas, 2000, p. 225).

### 1.2. Agents playing a crucial role in vocabulary acquisition

Vocabulary acquisition in an L2 is not, and cannot be, devoid of both external and internal influences on a learner. Those influences are numerous, and difficult to control. An array of agents playing an important role in L2 (vocabulary) acquisition is comprised of age (DeKeyser, 2000; Birdsong, 2006; Mihaljević Djigunović, 2014), motivation (Dörnyei, 1998), motivational strategies (Cheng & Dörnyei, 2007), attitudes and motivation (Mihaljević Djigunović & Bagarić, 2007), affect (Mihaljević Djigunović, 2007; Qin, 2007), learning strategies and styles (Ehrman & Oxford, 1995), gender (Khamkhien, 2010), attention (Schmidt, 2010), use of web technologies (Wang & Vásquez, 2012), etc. Other aspects affecting foreign language for specific purposes acquisition range from developing teaching materials for language for specific purposes learners (Candlin, Bhatia, & Jensen, 2002), to content (disciplinary, specialist, professional) knowledge (Smit & Dafouz, 2012), approach to teaching a foreign language for specific purposes (Zhang, 2007), to research-based language education (Hyland, 2002), etc.

## 2. Study

To realize the aims of research presented in this paper an action research model was adopted to provide additional perceptions regarding English kinesiological terminology acquisition-related issues, which would consequently contribute to

more efficient teaching-oriented guidelines. In compliance with the action research model paradigm, a four stage approach to the matter, as specified by Sagor (2005) was applied. In other words, first, the research targets were set. Secondly, the best- suited approaches to obtain the desired data were prepared. Thirdly, data were collected, and finally, the yielded results were analysed to be able to draw on the possible practical implications.

### 2.1. Aim

The aim of the research was to identify any statistically significant differences in terms of English kinesiological vocabulary knowledge between male and female students of kinesiology, as well as to identify which of the five selected agents – four subjects from the university study of kinesiology (*Systematic Kinesiology, Biomechanics, Basic Kinesiological Transformations* and *Advanced English Usage in Kinesiology*) and the number of years of learning English – played a decisive role in the acquisition of kinesiology-specific vocabulary. Participants' exam grades in the four aforementioned subjects were regarded as indicators of their academic achievement.

The selection of the four subjects was guided by the contents they cover. *Systematic Kinesiology* provides students with the knowledge of general principles of human movement, general principles of managing the process of physical exercise, and with the knowledge of consequences that these processes have on human organism. In *Biomechanics* the students are taught the basic laws of biomechanics of human movement, and the key aspect in *Basic Kinesiological Transformations* is the knowledge of morphological characteristics, cardiorespiratory endurance, motor abilities and skills and their possible transformations. In all, these three subjects rely on the basic concepts of kinesiology. All three subjects have Croatian as the language of instruction. The aim of the subject *Advanced English Usage in Kinesiology* is to teach, among other topics, the key terms in English that are connected with selected anthropological characteristics, e.g. morphological characteristics, energy production in human body, motor abilities, but also the training process connected with their development (transformations). In other words, one part of the syllabus is based on the selection of terminology from the contents of the aforementioned kinesiology-specific subjects.

To achieve the research aims, an assessment tool was created which would make it possible to collect the necessary data and to conduct a research among kinesiology students. The sample was comprised of 58 male and 24 female participants, students from the Faculty of Kinesiology at the University of Zagreb (Croatia) who all attended the classes in the subject *Advanced English Usage in Kinesiology* within the university study of kinesiology. The ratio of men and women corresponded to the ratio of male and female students at the Faculty of Kinesiology. Although the sample was a convenience one, it best suited the aims of the research in that the results were expected to yield some practical implications for implementation in the syllabus of the previously mentioned subject in which English kinesiological terminology is taught. Correspondingly, the following research questions were formulated:

RQ1: Are there any differences between male and female kinesiology students in their knowledge of a selected set of English kinesiology-related terms?

RQ2: Which of the agents (exam marks in the subjects *Systematic Kinesiology, Biomechanics, Basic Kinesiological Transformations* and *Advanced English Usage in Kinesiology*, as well as the number of years of learning English) have greater relative effect on the knowledge of a selected set of English kinesiology-related terms in male and female kinesiology students?

## 2.2. Instruments and procedure

An assessment tool consisting of five tasks was used: i. a five-item English into Croatian term translation multiple choice task; ii. a five-item Croatian into English term translation multiple choice task; iii. a five-item English by English term denotation multiple choice task; iv. a ten-item open-ended Croatian into English term translation task; and v. a ten-item true or false task. The focus of the constructional knowledge in an L2 for specific purposes is primarily at a lexical level, i.e. on multi-word terms (although multi-word expressions in a broader sense, e.g. *in contrast to, with regard to*, are not negligible), and less a syntactical one. Further, the selection of context-free tasks – multiple-choice and open-ended term translation task – was made in compliance with the fact that the context at sentence and text level, again in contrast to general language, is not necessarily helpful in understanding a term; on the contrary, it might even be confusing, as previously pointed out (Gu, 2015, p. 4). Since the scope of vocabulary testing was not on the contextual knowledge, the decontextualized type of MCTs was selected as a dominant format for this research. Due to the fact that there were cases in which past research showed that context was redundant for vocabulary testing on the one hand, and since language for specific purposes demands terminological accuracy which is not context-dependent on the other, textual context was omitted in three MCTs and one open-ended task in this research. An additional reason for omitting context for each of the terms in the MCTs was the fact that the broader frame of kinesiology-related contextual knowledge was known, which provided a wide clue for vocabulary knowledge.

Further, since the testing was done for a specific population and for a specific subsystem of language, i.e. for language for specific purposes, the accuracy regarding two aspects was critical. The first aspect involved domain knowledge and the accuracy of knowledge of terminology in a mother tongue (Croatian in this case). The second aspect focused on the accuracy of the knowledge of translation equivalence of terms that possess the same communicative value in a foreign language (English in this case). Hence, the selection of translation-type tasks was built on the fact that acquisition of an L2 as a language for specific purposes relies primarily on the knowledge of terminology and translational equivalence of terms between two (or more) languages.

As for the English by English term denotation task, the reasoning behind including it into the assessment tool used lies in synonymy – a crucial semantic relation in the mental lexicon, both in L1 and in L2. Although synonymy is not a desirable feature in language for specific purposes, which strives for terminological accuracy and clarity, it still exists and is frequently inversely proportional to a term's degree of technicality. Since language for specific purposes operates with terms of various technicality degrees, an attempt was made to include precisely such terms whose degree of technicality is not the same. Further and according to Turney, Littman, Bigham and Shnayder (2003), "multiple choice synonym questions can be used to determine the semantic orientation of a word".

According to Nation (2001, p. 359), translations both from L2 into L1 and vice versa involve recall. The open-ended translation format from L1 into L2, as a measure of the learners' productive vocabulary knowledge, was therefore selected as a further vocabulary testing technique.

Not to omit contextual testing at text level completely, a context-based true or false task was included in the research. This task has been devised by taking into account one of its critical features, i.e. its advantageous nature which provides information on whether something has been understood or not (Brown & Hudson, 2002, p. 66).

Finally, the students who participated in this study were familiar with all types of tasks included in the assessment tool, since these types of tasks were previously used in the classes.

### 2.3. Length of the assessment tool and the number of choices

The length of the assessment tool applied in this research (the total number of multiple-choice questions = 15) was within the recommended length of five to 30 multiple-choice questions (Considine et al., 2005, p. 23). In contrast to general language in which various vocabulary teaching and learning strategies may be employed, e.g. paraphrasing, using synonymous words or expressions, etc., language for specific purposes is more demanding in this respect and requires both conceptual and terminological precision.

Although research commonly speaks in favour of a three-option MCT format (Dehnad, Nasser, & Hosseini, 2014, p. 399), researchers opt for the four-option format due to the fact that the decrease of choices increases guessing (Dehnad et al., p. 399). Vyas and Supe (2008) conducted a survey of studies done on the number of choices per item on a multiple-choice test and found that three-option tests were of a similar quality as four- and five-option ones. Still, there are authors who recommend that as many plausible distractors as possible be provided in the choice set (Haladyna et al., 2002). In compliance with this recommendation, eight choices were given for each item in all three MCTs with the intent to increase the difficulty of the three tasks.

### 2.4. Selection of terms and text

The terms selected for all three MCTs and for the open-ended L1 into L2 term translation task were all extracted from the syllabus of the subject *Advanced English Usage in Kinesiology*. Both the English terms and their Croatian translation equivalents have been discussed with the students in these classes. The terms tested in the MCTs were extracted from three specialized areas: a domain that is a common denominator of all scientific disciplines, i.e. from research methodology (e.g., *incidence, perpendicular*); domains of general kinesiology (e.g., *physique, divison, superposition, motor ability, body fat)*; and a sport-specific domain (e.g., *violation, dolphin kick*).

Similarly, the topic of the true or false task was also a topic analysed and discussed with the students both in the classes of the specialized subject of kinesiology and in the subject in which English kinesiological vocabulary is taught, so that the students were familiar both with the domain context and topical knowledge as well as with vocabulary used in it, be it in L1 or in L2. A 200-word long text dealt with components of sports training programmes and the denotation of the terms *volume of training*, *intensity of training* and *density of training*. Although translation of two of them into the Croatian language might seem easy (English *volume of training* → Croatian *volumen treninga*; English *intensity of training* → Croatian *intenzitet treninga*), what perplexes their understanding in the two languages are their divergent semantic fields, i.e. the terms *volume of training* and *intensity of training* are false cognates.

### 2.5. Statistical methods

The sums of scores per task were used as indices of the English kinesiological vocabulary knowledge and served as dependent variables in *t*-tests for independent

samples, gender being a categorical factor. Further, a series of regression analyses was performed for men and women separately having each time one of the five vocabulary knowledge variables as a dependent one. Exam grades in the four selected subjects from the university study of kinesiology – *Systematic Kinesiology, Biomechanics, Basic Kinesiological Transformations* and *Advanced English Usage in Kinesiology* – and the number of years of learning English as a foreign language comprised a set of independent variables that served to identify agents that best predicted the success in each of the five tasks, i.e. that played a decisive role in the acquisition of English kinesiological vocabulary. Statistical package Dell Statistica (data analysis software system), version 12 (Dell Inc., 2015) was used to process the collected data.

## 3. Results and discussion

### 3.1. Gender differences

#### 3.1.1. Open-ended L1 into L2 translation

The yielded results continued to follow the trends outlined by previous analyses of the differences between male and female kinesiologists as regards the knowledge of kinesiological vocabulary in L2. The differences were small but significant, and when significant, this has proven to be in favour of men. Still, even the analysis of items from tasks in which men and women did not differ significantly pointed to some instructive pieces of information that might help to design a syllabus for teaching English as an L2 for specific kinesiological purposes. Overall, there was a significant difference between men (M=18.8; SD=5.7) and women (M=16.1; SD=5.4) in the Croatian into English open-ended term translation test – $t(80)=1.98$, $p \leq 0.05$.

The three most frequently correctly translated terms, by both men and women, were *aerobna izdržljivost* (English *aerobic endurance*), *repetitivna snaga* (English *muscular endurance*) and *anaerobna izdržljivost* (English *anaerobic endurance*) (Table 1). This could be justified by a well-known perception that antonymy, implied in the terms *aerobic* and *anaerobic endurance*, belongs to the set of crucial lexical semantic relations in the mental lexicon, and its uniqueness lies in that it "requires one-to-one relations" (Jones, Murphy, Paradis, & Willners, 2012, p. 1), which draws on its simplicity. Moreover, *aerobic* and *anaerobic* are adjectives, and adjectives possess marked affinity for expressing oppositeness (Jones et al., 2012, p. 4). The reason for the term *repetitivna snaga* (English *muscular endurance*) to be among the three most frequently correctly translated terms together with the terms *aerobic* and *anaerobic endurance* is the notion of *endurance* that conjoins them both conceptually in L1 and L2, and intra-linguistically in terms of its similarity in form in L2 with the terms *aerobic* and *anaerobic endurance*.

Table 1. Percent of incorrect, partially correct and correct translations in the open-ended translation MCT by gender

| CROATIAN TERM | CORRECT ENGLISH TRANSLATION | INCORRECT (%) | | PARTIALLY CORRECT (%) | | CORRECT (%) | |
|---|---|---|---|---|---|---|---|
| | | MEN | WOMEN | MEN | WOMEN | MEN | WOMEN |
| *jakost* | *strength* | 36.2 | 50.0 | 63.8 | 41.7 | - | 8.3 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| *snaga* | *power* | 43.1 | 66.7 | 55.2 | 33.3 | 1.7 | - |
| *repetitivna snaga* | *muscular endurance* | 77.6 | 87.5 | 3.4 | - | 19.0 | 12.5 |
| *mišićna izdržljivost* | *muscular endurance* | 41.4 | 50.0 | 53.4 | 45.8 | 5.2 | 4.2 |
| *aerobna izdržljivost* | *aerobic endurance** | 44.8 | 54.2 | 34.5 | 29.2 | 20.7 | 16.6 |
| *opća izdržljivost* | *stamina* | 63.8 | 54.2 | 29.3 | 37.5 | 6.9 | 8.3 |
| *jakosna izdržljivost* | *strength endurance* | 55.2 | 79.2 | 39.6 | 16.7 | 5.2 | 4.1 |
| *izdržljivost u brzini* | *speed endurance* | 55.2 | 75.0 | 44.8 | 25.0 | - | - |
| *silina* | *force* | 58.6 | 75.0 | 39.7 | 20.8 | 1.7 | 4.2 |
| *anaerobna izdržljivost* | *anaerobic endurance* | 43.1 | 66.7 | 41.4 | 20.8 | 15.5 | 12.5 |

**Aerobic endurance is also termed aerobic fitness, cardiovascular endurance/fitness, cardiopulmonary endurance/fitness. Hence, these were also possible correct translations.*

The three terms most frequently partially correctly translated from L1 into L2 by both men and women were *jakost* (English *strength*), *snaga* (English *power*) and *mišićna izdržljivost* (English *muscular endurance*). However, the survey of data in Table 1 also shows that men translated all the terms but one partially correctly more frequently than women. This partial correctness might point to the fact that, although not being completely sure as regards the correct translation of terms, men dared take a risk and tried to guess them, consequently achieving an overall better result than their female peers. Such a result is to some extent in compliance with past research which pointed to males being more prone to guessing than females (Bolger & Kellaghan, 1990). Although Bolger and Kellaghan (1990) dealt with guessing on MCTs, their result might be extended to open-ended term translations. Additionally, this task requires an analytical approach to learning because on the one hand it requires the knowledge in L1 of the distinction between the terms, included in this task, and concepts that they designate, and their translation equivalency in L2 on the other. This consequently means that in contrast to the perception that women learn more analytically than men, male kinesiologists, at least as far as this task is concerned, appeared to be an exception to this rule. Still, impulsiveness, a feature of men's learning style, might be expressed through guessing. What the result regarding the existence of a significant difference between men and women in this type of task also implies is that men exceeded women in one of the two most difficult tasks (the second most difficult task being the true or false task) in this analysis, i.e. in one segment of the productive aspect of L2 vocabulary knowledge.

### 3.1.2. Multiple-choice tasks

When constructed cautiously, MCTs may provide a rather informative insight into knowledge structure in general, thus also into vocabulary knowledge in a foreign language. The absence of differences between men and women on language-related MCTs in this study is in compliance with results of other researchers' work into the same subject matter, although past research reveals inconsistent data. Cole (1997) and

Grace (2000) found only very small differences between men and women, whereas Bolger and Kellaghan (1990) as well as Hellekant (1994) found no differences between the two genders in this type of task. Further, scarce data from past research into the connectedness between MCTs and guessing – MCTs being believed to encourage it (Walstad & Becker, 1994, p. 193) – reveal that only small differences between men and women were found hereof (Ben-Shakhar & Sinai, 1991).

Further, although no statistically significant differences were obtained between men and women on any of the MCTs, some valuable information displayed in Table 2 might be useful for further application in the process of acquiring kinesiological English terms.

One of those findings confirms a reliance on the knowledge of general English which was evident in the English into Croatian translation MCT. More men than women translated the term *violation* correctly into L1. However, when translated incorrectly, either by men or by women, it was understood as *nasilje*, i.e. *violence*.

Table 2. Frequency and percent of correct translations in three multiple choice tasks

| MULTIPLE CHOICE TASK | ITEM | CORRECT TRANSLATION | MEN | | WOMEN | |
|---|---|---|---|---|---|---|
| | | | COUNT | PERCENT | COUNT | PERCENT |
| English into Croatian translation | *incidence* | *učestalost* (English *occurrence, frequency*) | 12 | 20.7 | 3 | 12.5 |
| | *physique* | *oblik tijela* (English *shape of the body*) | 28 | 48.3 | 10 | 41.7 |
| | *division (in muscle structure)* | *dio mišića* (English *a part of a muscle*) | 19 | 32.8 | 8 | 33.3 |
| | superposition | položaj jedan iznad drugoga* | 36 | 62.1 | 19 | 79.2 |
| | violation (in team sports) | povreda pravila igre isključujući prekršaje na protivničkome igraču i pravila ponašanja u igri** | 10 | 17.2 | 2 | 8.3 |
| Croatian into English translation | *rad nogu (in a stroke termed dolphin technique in Croatian)**** | *dolphin kick* | 25 | 43.1 | 9 | 37.5 |
| | građa tijela | body type | 14 | 24.1 | 12 | 50.0 |
| | motorička sposobnost | motor ability | 33 | 56.9 | 13 | 54.2 |
| | motoričko znanje | motor skill | 25 | 43.1 | 10 | 41.7 |
| | veličina | magnitude | 19 | 32.8 | 11 | 45.8 |
| English by English denotation | *rupture* | *breaking* | 38 | 65.5 | 15 | 62.5 |
| | body fat | all body fat | 43 | 74.1 | 21 | 87.5 |
| | intact | not damaged | 32 | 55.2 | 15 | 62.5 |
| | duct | passage | 26 | 44.8 | 13 | 54.2 |
| | perpendicular | orthogonal | 23 | 39.7 | 9 | 37.5 |

**Position of one above the other.*
***Infringement of the rules of the game excluding fouls committed on an opposing player and the fouls regarding unsportspersonlike conduct.*
****Kicking action in butterfly stroke which is in Croatian termed* leptir/dupin tehnika.

With reference to the result that more men than women knew the meaning of the term *violation* and in spite of its low correct translation rate, it might be supposed that, globally, more men than women participate in team sports, and are hence more familiar with the terminology used in sporting games. Additionally, according to the research by Koivula (2001), many team sports such as soccer, football, handball, ice-hockey, baseball and rugby are regarded as masculine sports, which might also be regarded as a reason why more men are interested in these team sports.

Further, men correctly translated the term *incidence* more frequently than women. The reasons, however, for such a finding are to be sought elsewhere, not in the masculine, feminine or gender-neutral nature of various sports, since this term belongs to an area common to all scientific disciplines, i.e. to the terminology of research methodology. *Incidence* (Croatian *učestalost*) and *incident* (Croatian *slučaj*) are two terms that are strikingly similar in form, so that consequently most participants (both men and women) mistook the former for the latter – the latter again being known to them from general English. However, an additional point emerges here, namely, no such item as *slučaj* was to be found among the possible solution choices. What was offered among the eight items (Figure 1) was the word ***slučaj****nost*. The similarity in the form of the terms *incidence* and *incident* in L2, and partial similarity of the Croatian into English translation equivalent of the L1 term *incident*, might be presumed to have been prioritized by the participants as indicators considered to supposedly link the form of the word/term to its meaning.

| CROATIAN CHOICES | ENGLISH TRANSLATIONS OF CHOICES |
|---|---|
| slučajnost | *chance, mere chance* |
| vjerojatnost | *probability, likelihood* |
| valjanost | *validity* |
| pouzdanost | *reliability* |
| učestalost | *frequency* |
| raspršenost | *dispersion* |
| raspodjela | *distribution* |
| vrijednost | *value* |

Figure 1. The Croatian choices provided in the task for the term *incidence* in the English into Croatian term translation task.

This has confirmed the well-known principle that in terms of cognitive processes, similarity, e.g. in form, appears to be one of the ways in which people categorize things most frequently, either in language or in general (Escribano, 2004). Subsequently, a conclusion may be drawn that in some instances in this study women seemed to have relied more than men on the form of a term both in L1 and in L2 when trying to infer its meaning.

The correct translation from L2 into L1 of the term *physique* was *oblik tijela*, i.e. *shape of the body* (also termed *građa tijela* and *somatotip* in Croatian, and *body type* and *somatotype* in English. More men than women translated the term *physique* from L2 into L1 correctly, which is contradictory to the result obtained for a different term designating the same concept (Croatian *građa tijela*; English *body type*) later on in the L1 into L2 translation task. Hence, further research focusing on this particular subject matter seems to be necessary to make more accurate conclusions in this respect possible.

In the Croatian into English term translation MCT, women were more efficient than men in translating the term *građa tijela* (English *body type, somatotype, body build, physique*). Even more, the poorest results, with respect to other results achieved by men in this task as well as with respect to other results achieved by both men and women, were found in translating this particular term *građa tijela* into L2. A more detailed scrutiny allows for the identification of a probable cause of such a result. First of all, three concepts – *body type* (Croatian: *građa tijela*), *body structure* (Croatian: *struktura tijela*) and *body composition* (Croatian: *sastav tijela*) – are sometimes incorrectly and interchangeably used in Croatian, i.e. in L1, wherefore their acquisition in L2 requires an analytical approach. Further, although synonymous terms, as previously stated, are not desirable in the language for specific purposes because it strives towards the usage of uniform terms to avoid or, at least, reduce misunderstanding and to facilitate accurate communication in scientific and occupational expert communities, terms with the same or very similar meaning do exist. Hence, two terms exist in the Croatian language – *somatotip* and *građa tijela* to designate *body type*. In English any of as many as four synonymous translation equivalents may be used – *body type, somatotype, body build* and *physique*. A perplexing situation caused by failing to grasp a concept designated by one or even more than one term in one language might incur the lack of its comprehension in another language, and the situation may be exacerbated even further if the same communicative value in another language is also conveyed by more than one term. Such concatenation of possible interferences in concept representation and storage in memory significantly reduces conceptual knowledge in the mother tongue which then affects the knowledge of the term in a foreign language. Subsequently, the knowledge of a term and its meaning in the mother tongue need meticulous scrutiny prior to the presentation of its translation counterpart in a foreign language. In contrast to the result in the open-ended term translation task in which men appeared to have expressed a more analytical approach to learning than women, women achieved a better result when translating the term *građa tijela* into L2 because the acquisition of this term, as is evident from previous discussion on it, requires a more analytical approach to learning, which is the style of learning more frequently applied by women.

Another term that women proved to have better knowledge of was the term *veličina* (English *magnitude*), i.e. its translation equivalents in L2. The Croatian term *veličina* has several possible translation equivalents in English (e.g., *size*, *dimension, extent, measure, magnitude, quantity, volume*) out of which only one (*magnitude*) was to be found among the choices. All terms included in the choices referred to a kind of measure, be it a measure of length (*longitude*), width (*latitude*), height (*altitude*), number (*multitude*), extent (*amplitude*), quality or state (*platitude*), greatness of size or amount (*magnitude*) and finiteness (*infinitud*e), and accuracy in the knowledge of the meaning of all the given terms was necessary to select the correct one. Again, a more analytical approach, typically employed by women, in order to acquire the meaning of the term in L2 was necessary.

In the English by English denotation MCT correct translation rate, both for men and women, exceeded 40% in all but one item – the term *perpendicular*. As for this term, the selection of translation choices that the participants had at their disposal did not contain the term *vertical* (*slanting, tilted, orthogonal, parallel, transverse, oblique, skewed, diagonal*) which is usually the first choice by many when it comes to possible synonyms of the term *perpendicular*. Instead, the term *orthogonal* was used, which the students have learned in statistics, which has Croatian as the language of instruction, e.g. in factor analysis and the rotation of factors. Consequently, the participants were

(or should have been) familiar with the term. Further, the topic of statistics-related terminology used in the English language had been discussed with them in the subject *Advanced English Usage in Kinesiology*. Moreover, both the term *vertical* and the term *orthogonal* are also used in Croatian (*vertikalan* and *ortogonalan*) so that the correct solution rate was expected to be higher. Obviously, what hindered a higher correct solution rate was the lack of knowledge of the term *perpendicular* used in English. Whereas two stimuli might have conditioned the knowledge of the term *orthogonal*, only one stimulus – within the context of the university study of kinesiology – could have existed for storing the term *perpendicular* into the memory: the subject in which English kinesiological terminology is taught, i.e. the source which was not preceded by a concept representation stimulus in the mother tongue and within the creation of content knowledge. In addition, a term similar in form to the term *perpendicular* does not exist in Croatian, so that reference to the same concept based on form similarity of its name in the mother tongue was not possible.

### 3.1.3. True or false task

The number of correct solutions, by both men and women, in the true or false task varied (Table 3); however, item 1 stood out as for its highest rate of correct solutions to the true/false choice. The correct marking of the statement as being true when compared to the text depended on the participants' knowledge of the word *gradually*, which belongs to general language and has a low degree of technicality, and its synonymous expression *step by step* in the statement in item 1.

Table 3. Frequency and percent of correct solutions on the T/F task separately for men and women

| No. | T/F ITEM | MEN | | WOMEN | |
|---|---|---|---|---|---|
| | | COUNT | PERCENT | COUNT | PERCENT |
| 1 | The desired fitness level of an athlete must be achieved step by step. | 58 | 100.0 | 23 | 95.8 |
| 2 | The selection of objectives is non-sport-dependent. | 49 | 84.5 | 20 | 83.3 |
| 3 | Combat sports require team capability. | 51 | 87.9 | 18 | 75.0 |
| 4 | The selection of goals to be met is strict and obligatory. | 31 | 53.4 | 13 | 54.2 |
| 5 | Density and intensity of training are synonymous terms. | 54 | 93.1 | 21 | 87.5 |
| 6 | The time aspect of training belongs to the concept of the intensity of training. | 41 | 70.7 | 19 | 79.2 |
| 7 | The time aspect of training belongs exclusively to the concept of the density of training. | 29 | 50.0 | 10 | 41.7 |
| 8 | The density of training refers to time periods between the phases of exertion and the phases of rest. | 49 | 84.5 | 20 | 83.3 |
| 9 | The duration of training is an aspect of the density of training. | 26 | 44.8 | 16 | 66.7 |
| 10 | The frequency of repeating an effort is an aspect of the density of training. | 38 | 65.5 | 16 | 66.7 |

Another high rate of correct solutions by both men and women was obtained for item 5 and the statement that two focal terms – *density* and *intensity of training* – are not synonymous. Such a high correct solution rate was the result of the fact that both in English and in Croatian literature the distinction between the two concepts and the terms designating them is clear. However, further specification, by introducing the aspect of time, of the density and intensity of training as expressed in the underlying text was the reason for the decrease in the rate of correct solutions for items 6 and 7 in both genders. The time aspect continued to perplex men more than women in item 9. Still, the differences between the two genders were not such as to be identified as statistically significant.

### 3.1.4. Effects of exams grades in four subjects from university study of kinesiology and the number of years of learning English on vocabulary acquisition

Although no statistically significant differences were found between the two genders as for the exam grades obtained in the four subjects of university study of kinesiology, women received more excellent grades in three subjects, except for the subject in which English kinesiological terminology is taught. This is a finding that speaks in favour of men excelling women in at least some language aspects. Overall, the results from Table 4 point to rather evenly spread achievements of men and women in three crucial kinesiological subjects and in one technical language-related one. The absence of differences between the genders is informative as for the academic achievement in domain-specific subjects and a subject which combines linguistic and domain knowledge.

Table 4. Examination grades in the selected subjects from the university study of kinesiology (University of Zagreb, Faculty of Kinesiology)

| SUBJECT | VALUE | EXAMINATION GRADE | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | SUFFICIENT | | GOOD | | VERY GOOD | | EXCELLENT | |
| | | MEN | WOMEN | MEN | WOMEN | MEN | WOMEN | MEN | WOMEN |
| Systematic Kinesiology | Count | 19 | 3 | 16 | 5 | 16 | 9 | 7 | 7 |
| | Percentage | 32.8 | 12.5 | 27.6 | 20.8 | 27.6 | 37.5 | 12.1 | 29.2 |
| Biomechanics | Count | 11 | 3 | 14 | 4 | 23 | 9 | 10 | 8 |
| | Percentage | 19.0 | 12.5 | 24.1 | 16.7 | 39.7 | 37.5 | 17.2 | 33.3 |
| Basic Kinesiological Transformations | Count | 3 | 2 | 5 | 5 | 28 | 7 | 22 | 10 |
| | Percentage | 5.2 | 8.3 | 8.6 | 20.8 | 48.3 | 29.2 | 37.9 | 41.7 |
| Advanced English Usage in Kinesiology | Count | - | - | 5 | 5 | 14 | 4 | 39 | 15 |
| | Percentage | - | - | 8.6 | 20.8 | 24.1 | 16.7 | 67.2 | 62.5 |

As for the duration of systematic learning of the L2, men have been learning English for 10.31 years on average and women for 10.96 years. However, this minor difference in the length of learning English was not significant ($p \leq 0.05$).

The series of regression analyses has shown that, as for men, the subject *Advanced English Usage in Kinesiology* was the significant agent that best predicted participants' accomplishment in two out of five vocabulary knowledge tasks. The exam grades in this subject were found to explain a significant proportion of variance in the true or false task ($R^2$=0.199, F(5, 52)=2.595, p=0.036) and to significantly predict the results in it (β=0.412, t(52)=3.148, p=0.003). A significant proportion of variance in the English by English denotation MCT was explained ($R^2$=0.307, F(5, 52)=4.609, p=0.001) and success in this task was significantly predicted (β=0.453, t(52)=3.721, p=0.000) by the academic achievement in the subject *Advanced English Usage in Kinesiology*.

As for women, the results of regression analyses have shown that academic achievement in the subject *Advanced English Usage in Kinesiology* explained a significant proportion of variance in the open-ended Croatian into English term translation task ($R^2$=0.595, F(5, 18)=5.285, p=0.004), and it significantly predicted the female students' success in this task (β=0.401, t(18)=2.638, p=0.017). Interestingly, academic achievement in *Biomechanics* was influential in the English into Croatian multiple-choice translation task where it explained a significant proportion of variance ($R^2$=0.439, F(5, 18)=2.828, p=0.047) and significantly predicted the female participants' success (β=0.576, t(18)=2.416, p=0.026) in this task. As for the success in the English by English multiple choice denotation task, for which the multiple regression equation was statistically significant and explained 53.6% of the variance ($R^2$=0.536, F(5, 18)=4.154, p=0.011), it was significantly predicted by two independent variables – exam grades in *Biomechanics* (β=0.630, t(18)=2.902, p=0.009) and exam grades in *Advanced English Usage in Kinesiology* (β=0.409, t(18)=2.516, p=0.021). In other words, the knowledge obtained in biomechanics as well as the knowledge of English kinesiological terminology learned in the subject *Advanced English Usage in Kinesiology*, the former to a somewhat greater extent than the latter, proved to be critical for the scores in this task.

Overall, these results have shown that, in men, although believed to learn impulsively, the knowledge of vocabulary in L2 is affected by the actual instruction in the subject whose focus lies on teaching English kinesiological terminology. As regards future female kinesiologists, they combined their knowledge acquired both in the L2 instruction and in a kinesiology-specific domain subject in which instruction attention is aimed at physical laws of mechanics on which the science of human movement relies.

These results point to the fact that many kinesiology students ignore the importance of learning English kinesiological terminology and rely on their knowledge of general English, as well as on within-language and inter-language similarity. Still, one of the most compelling results was that the length of learning English as an L2 was not a decisive agent, either for men or for women, for the achievement in the tasks included in this research. Additionally, the Croatian kinesiological terminology has not been standardized yet, so that there are cases when some terms are not uniformly used and interpreted by all kinesiologists. This is then consequently reflected in the mistakes made by students when acquiring English translation equivalents of those terms. Consequently, although instruction in the subject *Advanced English Usage in Kinesiology* strongly focuses on correct terminology both in English and in Croatian, further steps seem to be necessary to standardize Croatian kinesiological terminology.

## 4. Conclusion

Firstly, although men scored better than women in one task, a conclusion might be drawn that actually there is little difference in language ability and skills between male

and female kinesiologists. The results in this research have once again confirmed that the dominance of women in language-related skills has not only decreased but seems to increasingly speak in favour of men, which is in contrast with the long-existing attitude according to which women better men in language skills. In spite of the fact that women are believed to learn more analytically than men, whereas men learn more impulsively than women, male kinesiologists bettered their female peers in the second most difficult subtest out of the total of five. Secondly, the analysis has shown that the subject in which students are taught language for specific purposes as a foreign language fulfils its purpose. Although it might be argued that such a conclusion is only natural and that the obtained result should go without saying, this need not always be the case, since many students rely on their previous knowledge of general English and fail to put time and effort into learning English kinesiological terminology necessary to understand the concepts that are in the basis of theories constituting the scientific discipline in question.

As for the practical implications of this action research and regarding the future structuring of the teaching process in which technical kinesiological English is taught at tertiary level in the domain of the science of human movement, several possible guidelines may be identified.

- When learning English kinesiology-specific terminology, kinesiology students should be advised not to rely on form similarity of terms between L1 and L2, since this might result in an incorrect understanding of those terms.
- Female kinesiology students should be instructed to pay more attention to the productive aspect of the knowledge of English kinesiological terminology, i.e. they should try to connect a Croatian term with its English translation equivalent. Accordingly, more exercises seem to be necessary during the class in which Croatian terms should be the basis on which English terminology would then be practised.
- Both male and female kinesiology students should be encouraged to apply a more analytical approach to learning terminology. However, to be able to do so, even more emphasis than usual during the English classes should be put on using accurate Croatian terms and on clearly describing the concepts designated by those terms.
- More exercises based on semantic relations such as synonymy and antonymy should be introduced, which would then in turn contribute to the overall analytical teaching and learning style.
- More teaching- and terminology-oriented cooperation between subject specialists, i.e. kinesiologists, and language teachers might contribute to a better structuring of syllabi of both kinesiology-specific and L2-specific subjects.

Ultimately, this action research provides some valuable pieces of information regarding some possible future improvements that could help plan beneficial changes to be incorporated into the teaching process. Due to some limitations of the study, continuous future research seems to be necessary to be able to continue fostering the development of syllabi.

## 5. References


Agustín Llach, M.P., & Gallego, M.T. (2012). Vocabulary knowledge development and gender differences in a second language. *Estudios de lingüística inglesa aplicada*, *12*, 45-75. Retrieved September 10, 2016, from http://institucional.us.es/revistas/elia/12/art_3.pdf

Ali, H.O. (2016). Gender differences in using language in the EFL classes: From teachers' views. *International Journal of Humanities and Cultural Studies*, *2*(4), 73-91. Retrieved September 10, 2016, from http://www.ijhcs.com/index.php/ijhcs/index

Bennett, R. E., Rock, D. A., & Wang, M. (1991). Equivalence of free-response and multiple-choice items. *Journal of Educational Measurement*, *28*(1), 77-92.

Ben-Shakhar, G., & Sinai, Y. (1991). Gender differences in multiple-choice tests: The role of differential guessing tendencies. *Journal of Educational Measurement*, *28*(1), 23-35.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, *56*(1), 9-49. doi: 10.1111/j.1467-9922.2006.00353.x

Bolger, N., & Kellaghan, T. (1990). Method of measurement and gender differences in scholastic achievement. *Journal of Educational Measurement*, *27*(2), 165-174.

Boyle, J.P. (1987). Sex differences in listening vocabulary. *Language Learning*, *37*(2), 273-284.

Brantmeier, C. (2003). Does gender make a difference? Passage content and comprehension in second language reading. *Reading in a Foreign Language*, *15*(1). Retrieved September 10, 2016, from http://nflrc.hawaii.edu/rfl/April2003/brantmeier/brantmeier.html

Brown, D., & Hudson, T. (2002). *Criterion-referenced language testing*. Cambridge: Cambridge University Press.

Cole, N.S. (1997). The ETS gender study: How females and males perform in educational settings. *Educational Testing Service Research Report* (143), 5-35.

Candlin, C.N., Bhatia, V.K., & Jensen, C.H. (2002). Developing legal writing materials for English second language learners: Problems and perspectives. *English for Specific Purposes*, *21*(4), 299-320.

Cheng, H.-F., & Dörnyei, Z. (2007). The use of motivational strategies in language instruction: The case of EFL teaching in Taiwan. *Innovation in Language Learning and Teaching*, *1*(1), 153-174. doi: 10.2167/illt048.0

Considine, J., Botti, M., & Tomas, S. (2005). Design, format, validity and reliability of multiple choice questions for use in nursing research and education. *Collegian*, *12*(1), 19-24.

Dehnad, A., Nasser, H., & Hosseini, A.F. (2014). A comparison between three-and four-option multiple choice questions. *Procedia – Social and Behavioral Sciences*, *98*, 398-403.

DeKeyser, R.M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*(4), 499-533.

Dell Inc. (2015). Dell Statistica (data analysis software system), version 12. software.dell.com.

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

Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge: Cambridge University Press.

Ehrman, M.E., & Oxford, R.L. (1995). Cognition plus: Correlates of language learning success. *The Modern Language Journal*, *79*(1), 67-89.

Escribano, P. D. (2004). Exploring cognition processes in second language acquisition: The case of cognates and false-friends in EST. *Ibérica*, *7*, 87-106.

Grace, C.A. (2000). Gender differences: Vocabulary retention and access to translations for beginning language learners in CALL. *The Modern Language Journal*, *84*(2), 214-224.

Gu, Y. (2002). Gender, academic major, and vocabulary strategies of Chinese EFL learners. *RECL Journal*, *33*(1), 35-54.

Gu, Y. (2015). Vocabulary guessing strategies. In C.A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-7). Oxford, UK: Wiley-Blackwell. doi: 10.1002/9781405198431.wbeal1469

Haladyna, T.M. (1999). *Developing and validating multiple-choice test items.* Mahwah, NJ*:* Lawrence Erlbaum Associates, Inc.

Haladyna, T.M., Downing, S.M., & Rodriguez, M.C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, *15*(3), 309-334.

Hellekant, J. (1994). Are multiple-choice tests unfair to girls? *System*, *22*(3), 349-352.

Horst, M., Cobb, T., & Meara, P. (1998). Beyond a clockwork orange: Acquiring second language vocabulary through reading. *Reading in a Foreign Language, 11*(2), 207-223.

Hyland, K. (2002). Specificity revisited: How far should we go now? *English for Specific Purposes*, *21*(4), 385-395.

Jones, S., Murphy, M.L., Paradis, C., & Willners, C. (2012). *Antonyms in English: Construals, constructions and canonicity*. Cambridge: Cambridge University Press.

Kelly, P. (1990). Guessing: No substitute for systematic learning of lexis. *System*, *18*, 199-207.

Khamkhien, A. (2010). Factors affecting language learning strategy reported usage by Thai and Vietnamese EFL learners. *Electronic Journal of Foreign Language Teaching*, *7*(1), 66-85.

Kimura, D. (1999). *Sex and Cognition*. Cambridge, Massachusetts: MIT Press.

Koivula, N. (2001). Perceived characteristics of sports categorized as gender-neutral, feminine and masculine. *Journal of Sport Behavior*, *24*(4), 377-393).

Laufer, B., & Nation, P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, *16*(1), 33-51. doi: 10.1177/026553229901600103

Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, *22*(1), 1-26.

Lin, J., & Wu, F. (2004, April). *Differential performance by gender in foreign language testing*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, IL.

Lin, L.-F. (2011). Gender differences in L2 comprehension and vocabulary learning in the video-based CALL program. *Journal of Language Teaching and Research*, *2*(2), 295-301. doi:10.4304/jltr.2.2.295-301

McNamara, T. (2005). *Language testing*. Oxford: Oxford University Press.

Mihaljević Djigunović, J. (2007). Role of affective factors in the development of productive skills. In M. Nikolov & J. Horvath (Eds.) *Empirical studies in English applied linguistics* (pp. 9-23). Pecs: Lingua Franca Csoport.

Mihaljević Djigunović, J., & Bagarić, V. (2007). A comparative study of attitudes and motivation of Croatian learners of English and German. *Studia Romanica et Anglica Zagrabiensia*, *52*, 259-281.

Mihaljević Djigunović, J. (2014). L2 learner age from a contextualised perspective. *Studies in Second Language Learning and Teaching*, *4*(3), 419-441. doi: http://dx.doi.org/10.14746/ssllt.2014.4.3.3

Moir, A., & Jessel, D. (1992). *Brain sex*: *The real difference between men and women.* New York: Dell Publishing.

Moreno, R., Martínez, R.J., & Muñiz, J. (2006). New guidelines for developing multiple-choice items. *Methodology. 2*(2), 65-72. doi 10.1027/1614-1881.2.2.65

Morris, L.A. (1998). Differences in men's and women's ESL writing at the junior college level: Consequences for research on feedback. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes*, *55*(2), 219-38.

Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Newman, M.L., Groom, C.J., Handelman, L.D., & Pennebaker, J.W. (2008). Gender differences in language use: An analysis of 14,000 text samples. *Discourse Processes, 45*(3), 211–236. doi: 10.1080/01638530802073712

Omrčen, D., & Bosnar, K. (2008). Assessing ESP in kinesiology students. In D. Milanović & F. Prot (Eds.), *Proceedings of the 5th international scientific conference on kinesiology Zagreb, 2008 'Kinesiology research trends and applications'* (pp. 723-727). Zagreb: University of Zagreb, Faculty of Kinesiology.

Omrčen, D. (2010). Wie erfolgreich lernen Studierende die Fachtermini. *KDVInfo*, *19*(36/37), 76-82.

Omrčen, D., & Bosnar, K. (2010). Gender stereotyping bias – assessment of the swimming and underwater diving vocabulary knowledge in English as a foreign language in kinesiology. In P. Vičić & S. Orthaber (Eds.), *Proceedings CD of the international language conference, Celje, 2010, on 'The importance of learning professional foreign languages for communication between cultures'* (pp. 1-6). Celje: Faculty of Logistics, University of Maribor.

Omrčen, D. (2011a). Gramatičko znanje engleskoga kao stranoga jezika u studenata i studentica kineziologije. *Strani jezici*, *40*(2), 145-157.

Omrčen, D. (2011b). Fachvokabelkenntnisse im DaF-Unterricht: Geschlecht und Sportpräferenz. In P. Vičić, S. Orthaber & V. M. Ipavec (Eds.), *Proceedings of the fourth international language conference, Celje, 2011, on 'The importance of learning professional foreign languages for communication between cultures'* (pp. 175-181). Celje: Faculty of Logistics.

Pahom, O., Farley, A., & Ramonda, K. (2015). Are the best language learners from Mars or from Venus? Gender and vocabulary acquisition in the L2 Spanish classroom. *Reading Matrix: An International Online Journal*, *15*(1), 158-172.

Powers, D.E. (2010). The case for a comprehensive, four-skills assessment of English language proficiency. *TOEIC Compendium Study*. Retrieved September 10, 2016, from https://www.ets.org/Media/Research/pdf/TC-10-12.pdf

Pulido, D. (2003). Modeling the role of second language proficiency and topic familiarity in second language incidental vocabulary acquisition through reading. *Language Learning*, *53*(2), 233-284.

Qin, L. (2007). EFL teachers' factors and students' affect. *US-China Education Review*, *4*(3), 60-67. Retrieved September 10, 2016, from http://files.eric.ed.gov/fulltext/ED497391.pdf

Sagor, R. (2005). *The action research guidebook. A four-step process for educators and school teams.* Thousand Oaks, CA: Corwin.

Scarcella, R., & Zimmerman, C. (1998). Academic words and gender: ESL student performance on a test of academic lexicon. *Studies in Second Language Acquisition*, *20*(1), 27-49.

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan & I. Walker (Eds.), *Proceedings of CLaSIC 2010*, Singapore, 2010 (pp. 721-737). Singapore: National University of Singapore, Centre for Language Studies.

Smit, U., & Dafouz, E. (2012). Integrating content and language in higher education. An introduction to English-medium policies, conceptual issues and research practices across Europe. *AILA Review*, *25*(1), 1-12.

Turney, P.D., Littman, M.L., Bigham, J., & Shnayder, V. (2003). Combining independent modules to solve multiple-choice synonym and analogy problems. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP-03)* (pp. 482-489). Borovets, Bulgaria.

Van Zeeland, H. (2013). L2 vocabulary knowledge in and out of context. *Australian Review of Applied Linguistics*, *36*(1), 52-70.

Vyas, R., & Supe, A. (2008). Multiple choice questions: A literature review on the optimal number of options. *The National Medical Journal of India*, *21*(3), 130-133.

Voyer, D., & Voyer, S.D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174-1204.

Walstad, W.B., & Becker, W.E. (1994). Achievement differences on multiple-choice and essay tests in economics. *American Economic Review*, *84*(2), 193-196.

Wang, S., & Vásquez, C. (2012). Web 2.0 and second language learning: What does the research tell us? *CALICO Journal, 29*(3), 412-430.

Waring, R., & Nation, P. (2004). Second language reading and incidental vocabulary learning. *Angles on the English-speaking World*, *4*, 11-23.

Wilson, E.O. (1975). *Sociobiology: The new synthesis*. Cambridge, Massachusetts: Harvard University Press.

Wood, J. (1999). How large a vocabulary do EAP engineering students need? *Reading in a Foreign Language*, *12*(2), 309-323.

Xia, X. (2013). Gender differences in using language. *Theory and Practice in Language Studies*, *3*(8), 1485-1489. doi:10.4304/tpls.3.8.1485-1489

Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology*, *10*(3), 85-101. Retrieved September 10, 2016, from http://llt.msu.edu/vol10num3/yoshii/

Zhang, Z. (2007). Towards an integrated approach to teaching Business English: A Chinese experience. *English for Specific Purposes*, *26*(4) 399-410.

# Evaluation of the Faculty of Humanities and Social Sciences (FHSS) Doctoral Program in Foreign Language Education: Questionnaire Design

**Jasenka Čengić**
*Faculty of Humanities and Social Sciences, Zagreb, Croatia*
*jasenka.cengic@gmail.com*

**Silvija Hanžić Deda**
*Faculty of Humanities and Social Sciences, Zagreb, Croatia*
*silvija.hanzic@gmail.com*

**Ana Gabrijela Blažević**
*Faculty of Humanities and Social Sciences, Zagreb, Croatia*
*anablazevic@gmail.com*

## 1. Introduction

Evaluating academic programs has become an integral part of developing and improving various programs at undergraduate as well as graduate and postgraduate university levels. One of the problems in assessing the quality of university level programs is the fact that most assessment programs rely on student satisfaction surveys only (Hurt, 2004). However, academic programs involve some other dimensions which should be considered in the evaluation process with student satisfaction being only one of them. Other dimensions include meeting the students' desired outcomes, their relationship with the advisors (Szymanska, 2011) and the variable of the course instructors' as well as the thesis advisors' experience (Cuseo, 2003).

Furthermore, a comprehensive assessment of academic programs should include a variety of qualitative and quantitative assessment tools. The use of open-ended questions and focus groups are the recommended procedures for general questionnaire design proposed by Brown (2001), Dörnyei (2003) and Demetriou (2005).

Ready-made instruments cannot always serve the purpose of evaluating specific academic programs. The necessary steps in academic program evaluation, therefore, involve designing instruments that provide enough information for program improvement.

From 2006 to 2014 the Doctoral program in foreign language education (FLE) at the FHSS (Faculty of Humanities and Social Sciences), University of Zagreb (henceforth the Program) was evaluated with the instruments designed by the Program directors. Although the previously used questionnaires managed to reflect the complex picture of the multitude of dimensions encompassed by the Program, they were not validated instruments.

Therefore, in 2014 the evaluation underwent considerable changes. The idea was to recruit the students themselves in the process of collecting the necessary information about what needs to be included in the new assessment protocol and instrument design. After a period of brainstorming sessions and collaborative meetings

between students and the faculty members involved in the Program, a decision was made to design a new questionnaire for evaluating the Program. The aim of this paper is to describe the process of designing the above-mentioned questionnaire.

## 2. The FHSS Doctoral Program in Foreign Language Education

In this section, the Program will be described along with the evaluation procedures in use prior to the Program change in 2014.

### 2.1. The description of the Program

The Program consists of two parts pertaining to the activities designed for doctoral students: a) coursework based on compulsory and elective courses, and b) students' independent research work.

The coursework part consists of four modules and students' independent research work covers three modules (see Table 1).

Table 1. The structure of the Program

| The FHSS Doctoral Program in FLE | | | | | | |
|---|---|---|---|---|---|---|
| Coursework | | | | Independent Research | | |
| Applied linguistics courses | Courses about research methods | FLE themed courses | Various elective courses | Students' current research work presentation | Selecting the area of doctoral research | Dissertation work |

#### 2.1.1. Coursework

*Module 1* comprises six courses (see Table 2). In all the tables that follow below, the first column shows the course titles, the second column shows their status (compulsory or elective), and the third column shows the number of ECTS credits per course.

Table 2. Courses in Module 1

| Course | Status | Credits |
|---|---|---|
| Second Language Acquisition | compulsory | 4 |
| First Language Acquisition | elective | 2 |
| Bilingualism and Multilingualism | elective | 2 |
| The Age Factor | elective | 2 |
| Cognitive Grammar in Foreign Language Learning and Teaching | elective | 5 |
| Language for Specific Purposes | elective | 2 |

*Module 2* includes courses pertaining to research methodology (see Table 3).

Table 3. Courses in Module 2

| Course | Status | Credits |
|---|---|---|
| Research Design and Introduction to Statistics | compulsory | 3 |
| Methods in Researching Foreign Language Education | compulsory | 3 |
| Academic Writing and Oral Presentation Skills | compulsory | 3 |
| Language for Specific Purposes and Language Corpora | elective | 2 |
| Modern Technology and Foreign Language Teaching | elective | 5 |
| Computational Linguistics and Language Technologies | elective | 2 |
| Research Areas in Foreign Language Acquisition | elective | 2 |

*Module 3* consists of topics in foreign language education, and it offers fourteen courses shown in Table 4.

Table 4. Courses in Module 3

| Course | Status | Credits |
|---|---|---|
| Theories and Methods in Foreign Language Teaching | elective | 3 |
| Discourse Analysis and Classroom Discourse | elective | 3 |
| Reading in a Second Language | elective | 3 |
| Communicative and Other Competences in Foreign Language Teaching | elective | 3 |
| Written Discourse in Foreign Language Education | elective | 3 |
| Literature in Foreign Language Education | elective | 3 |
| Cultural Elements in Modern Foreign Language Teaching | elective | 3 |
| Individual Learner Differences and Foreign Language Teaching | elective | 3 |
| Language Learning Strategies | elective | 3 |
| Learner Autonomy | elective | 3 |
| Reception Approach in Foreign Language Learning | elective | 3 |
| Analysis and Evaluation of Teaching Materials | elective | 3 |
| Assessing Communicative Competence in a Foreign Language | elective | 3 |
| Designing Language Tests | elective | 3 |

*Module 4* contains eleven elective courses (see Table 5).

Table 5. Courses in Module 4

| Course | Status | Credits |
|---|---|---|
| Cognitive Linguistics | elective | 2 |
| Neurolinguistics | elective | 2 |
| Developmental Psychology | elective | 2 |
| Mental Lexicon – Theories and Models | elective | 2 |

| | | |
|---|---|---|
| Language Learning Impairments in Monolinguals and Bilinguals | elective | 2 |
| Psychology of Children with Special Needs | elective | 2 |
| Blindness and Foreign Language Learning | elective | 3 |
| The Modern School in the Educational System | elective | 2 |
| Curriculum Theory and Models | elective | 2 |
| Education Strategies and Innovation | elective | 2 |
| Interculturalism and Education | elective | 2 |

### 2.1.2. Independent Research Work

The students' independent research work covers three modules (Modules 5, 6 and 7). The doctoral students set their own goals and schedule their research activities, that is, they work independently towards reaching the required number of credits.

*Module 5* involves collecting credits for independent research work through publishing, attending conferences and presenting research findings, etc. The students organize their work autonomously, or with the assistance of their advisors, faculty members involved in the Program and other experts in the field.

*Module 6* includes the necessary stages in the process of determining and selecting the thesis topic and preparatory work for drafting research proposal and the official research proposal public defense.

*Module 7* is dedicated to planning the research, reporting on the relevant stages of the accomplished tasks, public presentation of preliminary results and key findings, and the actual dissertation writing. Finally, the module is completed with the official evaluation of the thesis and its public defense.

Each doctoral student receives the so called Gloma file that contains personal information, list of courses, and the program schedule. Students' responsibility is to make timely entries, indicating progress (credits earned, exams passed, conferences attended, papers published, etc.).

## 2.2. Previously used questionnaires for evaluation of the Program

From 2006 to 2014 three questionnaires were designed for evaluation of the Program. The evaluation was conducted after each semester and at the end of the Program. In the questionnaires, the following categories were evaluated: the structure of the courses, the relevance of the courses, course instructors' performance and course requirements. Semester evaluation was conducted using a questionnaire containing the list of both compulsory and elective courses which were evaluated through the above-mentioned categories. The students were asked to evaluate each course along the four listed categories using a 5-point scale. In addition to that, there was a text box intended for comments and suggestions regarding individual courses as well as the overall program.

The evaluation of the overall program consisted of three parts: first, some general information about the students was required, such as previous education, work experience and the hours spent in completing the Program requirements. Next, the students were asked to provide an overall evaluation of the program on a 5-point

scale, and the subsequent 30 statements addressed the Program administration, the Program director's availability and the time provided for students' needs, tuition fees, assignment deadlines and reading materials, exam dates and requirements, teachers' expectations, the balance between theory and research, collaboration with thesis advisors, and overall study program organization. The final part of this questionnaire comprised three open-ended questions related to: a) three strongest aspects of the Program, b) three possible improvements, and c) ideas pertaining to improving the skills and competences acquired in the Program.

In the section that follows we are going to describe the concepts and instruments that served as the basis for the new questionnaire design.

## 3. New Questionnaire Design

The idea was to construct a questionnaire that would encompass the multidimensionality of the Program structure and the complexity of the motives and goals of the students enrolling in the Program. More specifically, we examined the existing instruments used prior to 2014, in terms of their applicability to the aims of evaluation of the Program. In addition to the previously used questionnaires, we explored instruments used in other universities. The instrument that we found most useful for constructing a new questionnaire was The Students' Evaluation of Educational Quality (SEEQ) instrument (Marsh, 1983, 1984, 1987; Marsh & Dunkin, 1992).

### 3.1. SEEQ

The SEEQ instrument (Marsh, 1983, 1984, 1987; Marsh & Dunkin, 1992) is frequently used by researchers at the university level (Richardson, 2005). It comprises a number of the so-called teaching dimensions (Marsh, 2007).

The SEEQ is a type of the students' evaluation of teaching effectiveness (SETs). According to Marsh (2007), SETs is the most commonly used instrument for the students' assessment of their study programs. The SETs has an established construct validity (e.g., Cashin, 1988; Howard, Conway, & Maxwell, 1985) and that was the main motive for inserting it in the new questionnaire we were trying to design. SETs aims to identify various teacher profiles.

The standardized SEEQ is designed to evaluate one course at a time, and it is most commonly administered upon completing the course. It consists of 34 items distributed in ten unequal parts: nine parts use six-point Likert scales, and the last part is an open-ended question eliciting additional comments or feedback. The SEEQ questionnaire sections address the following teaching dimensions: the learning process, enthusiasm, organization, group interaction, individual rapport, breadth of coverage, examinations, assignments, overall program evaluation and feedback.

The SEEQ identifies nine factors relevant for the teacher profiles: Learning/ Value, Instructor Enthusiasm, Organization/Clarity, Group Interaction, Individual Rapport, Breadth of Coverage, Examinations/Grading, Assignments/Readings, and Workload/Difficulty, the combinations of which form various teacher profiles which have been proven constant in relation to time, courses and levels of teaching. The association between those profiles and teaching efficacy have not been widely explored (Marsh, 2007). This relationship between the teachers' profiles and their teaching efficacy was to be one of the crucial segments incorporated in our new questionnaire.

More specifically, we wished to examine the interrelation between teachers' profiles, their efficacy in teaching and the students' profiles (see section 3.3 for details about students' profiles).

### 3.2. Questionnaire design process

In this section, we discuss the steps taken during the process of questionnaire design. The questionnaire items were generated from 1) informal individual interviews with the directors of the Program 2) focus group interview and 3) questionnaires previously used for the evaluation of the Program.

Apart from the current Program director, one of the preceding directors of the FHSS doctoral program in FLE was interviewed as well. The aim of the interviews was learning about the Program structure and the changes after 2014 and, more importantly, to determine to what extent the feedback obtained from the previous questionnaires affected and/or initiated the Program change.

The first focus group consisted of 11 doctoral students who formed part of the 2014 cohort. The goal of the focus group was to articulate the issues concerning the structure and content of the Program. The focus group was held after the doctoral students had already attended most of the courses planned for that cohort. The students were asked to comment on the items included in a previously used questionnaire and to say if they thought the items covered the issues related to the students' satisfaction with the Program. Having compared the issues that arose during the focus group discussion to the items listed in the questionnaire previously used for the evaluation of the Program, it was clear that the old questionnaires used for the evaluation purposes of the Program failed to address most of the issued which emerged during the focus group discussion.

There were several important issues that surfaced from the data collected in the focus group. First, although the old questionnaires contained open-ended questions for additional comments, it was concluded that the doctoral students had to be provided with the opportunity to give much more detailed opinions about individual courses and teachers. This would ensure that the comments collected would be useful for gathering information about specific courses and their requirements. Second, the doctoral students' perception of course quality seemed to differ greatly from one another due to their previous education, teaching experience and motivation for enrolling in the Program. In other words, there were two basic ideas that emerged during this focus group: 1) the questionnaire for the evaluation of the program had to be more detailed in terms of addressing individual courses and teachers, and 2) the evaluation of the success of the courses and teachers' profiles should be matched with what we decided to call doctoral students' profiles. These profiles should include their educational background, place of work, and their primary motivation for starting the program.

### 3.3. The first draft of the questionnaire

The new questionnaire [3]consisted of seven parts that had to be administered at different times during the course of the program. The first two parts aimed at collecting

[3] Contact the authors for the full version of the new questionnaire in Croatian and the English.

information about the doctoral students' profiles. The rest of the questionnaire addressed the program itself – the courses available, exams and details about potential thesis advisors. Let us take a closer look at the structure of the questionnaire:

1. The first set of items gathered general information about the students, such as age, gender and previous education. In addition to that, it collected data on individual student's work experience irrespective of the type of employment. Finally, it addressed the question of motivation to enroll in this doctoral program. The items were either multiple choice or filling in the gaps.
2. The second set of items related to the profiles of doctoral students as language teachers. This set is divided into three subsets. The first subset presents nine types of learners that can be found in a language classroom and examines with which of these types of learners are the doctoral students likely to collaborate most effectively. The second subset examined potential "compatibility" between a certain learner profile and the doctoral students' teaching style. The last subset enquired about the doctoral students' perception of the "ideal" language learner.
3. The third set was intended for individual course evaluation and it consisted of five subsets. The first one related to the interest in the course before and after attending the course. The second subset related to the content and the schedule of the course. The third part related to the assigning a grade to the overall course and course teacher's performance. The final two subsets were open-ended questions enquiring about the strong points of the course, as well as the suggestions about what could be improved.
4. The fourth set examined the exam process. This part of the questionnaire was administered after each semester, and included all the exams passed during that period. The first subset examined the clarity of course requirements in relation to the final exam, the connection between learning outcomes and assessment, availability of reading materials and course instructor's feedback. The second subset related to the amount of time invested in completing the course requirements and the number of ECTS credits earned in relation to the effort invested in passing the exam.
5. The fifth set of the questionnaire dealt with the preferences of doctoral students, that is, they were asked to mark the level of agreement with the statements that described the types of teachers they preferred. For instance, the teachers who presented topics systematically and clearly, the ones who used real life situations to present research results, the teachers who were very objective and demonstrated no personal opinions, the teachers who respected the variety of theories and approaches even though they might not have agreed with them, the ones who expressed their opinions clearly, the teachers whose teaching was interactive, the ones who allowed interruptions, the teachers who encouraged independent student work, and the ones who accepted their teaching input to be guided by their students' needs and interests.
6. The sixth set consisted of six subsets. The first subset examined the administration efficiency, the second subset related to the courses schedule and organization and the third subsection was intended for the overall program evaluation. The last three subsets were open-ended questions examining the strong points of the doctoral program's organization, the suggestions for improvements, and the recommendations for expanding the knowledge and skills incorporated in this study program. This part of the questionnaire would be administered at the end of the doctoral studies.

7. The last, seventh set, of this questionnaire examined the quality of the doctoral students' collaboration with their thesis advisor(s). This section ended with an open-ended question calling for further comments. This part of the questionnaire would be administered after the doctoral thesis is defended.

### 3.4. Piloting the first draft of the new questionnaire

The first draft of the new questionnaire was sent to 9 doctoral students in several doctoral programs at the FHSS. Our aim was to obtain feedback about the structure and comprehensibility of the questionnaire. More specifically, they were asked to complete the questionnaire and comment on the following points: the wording of the items, the organization of the questionnaire, and the time frame necessary for completion. The comments obtained were used to improve the clarity and comprehensiveness of the questionnaire items. The result was the version of the questionnaire that was piloted further a few months later.

The participants were recruited from all the cohorts of the Program since its founding in 2006, including the 2014 cohort. Altogether there were 38 participants, out of which 12 were members of the 2014 cohort, and the remaining 26 were students belonging to all the other cohorts of the Program since its founding. The questionnaire was administered online. It is important to note that this final version of the questionnaire was used selectively in the sense that we had to find the elements that could be relevant for all the cohorts. Thus, the participants were asked to complete only those questions that relate to two out of four compulsory courses that were in common to all the cohorts. The items included in the sixth and seventh set of the questionnaire were not included in the piloting. As stated earlier those sets would be administered at the end of the doctoral study, which is a stage the majority of the doctoral students had not reached at the point the pilot study was conducted.

Our next step is to test the sixth and seventh module of the questionnaire on the doctoral students of the 2014 cohort so that the questionnaire could be used in its entirety with the 2016 cohort. One of the future steps will be designing an additional part of the questionnaire which would require feedback from the course teachers as well as thesis advisors to allow for even deeper insight into the needs of the Program.

## 4. Instead of a conclusion

After all the changes introduced at the university level in 2014, the FHSS Doctoral Program in Foreign Language Education, faced the need for changing the Program evaluation scheme. That resulted in constructing a set of questionnaires which, put together, make a new instrument for the integral evaluation of the Program. Such a questionnaire needed to encompass the enrolled students' profiles, both as teachers and as students, and had to cover the academic program in its entirety.

The novelties of the new questionnaire design are: 1) the opportunity for the students to give direct feedback on each individual course, including the course teacher performance, and their exam experiences immediately upon completion, 2) items designed to measure the students' motivation to join the Program and their goals, i.e., what they are aspiring to upon completing their doctoral education, 3) items designed to create a learner profile and a teacher profile of a particular doctoral student.

Collecting relevant data at regular intervals, from every cohort, enables the doctoral program director to adapt quickly and to meet the doctoral students' needs. In

other words, on the account of timely feedback, the study program becomes dynamic and remains up to date.

Apart from timely feedback and motivation for enrolling in the Program, the new questionnaire contains two distinctive subsets that collect information about each doctoral student from two different perspectives: those of doctoral students as teachers and as learners. Establishing potential connections between doctoral students' profiles as teachers and as students is seen as a way of grasping the complexity of the nature of the cohorts of students joining the Program. Both subsets of the questionnaire contain the items examining various cognitive styles and personality traits, so they provide valuable information that can clarify the reasons why individual students evaluated a certain course, course teacher, and other components of the doctoral program the way they did.

Having considered all the aspects of the Program requirements and the enrolled students' profiles, two new aspects were merged into a salient new variable that could influence the students' evaluation of the courses: the students' motivation to join the Program and their goals, i.e., what they are aspiring to upon completing their doctoral education. This new variable has the potential of explaining the variability in individual student's evaluation.

In sum, the data collected by means of the new questionnaire may carry great potential in providing accurate feedback to the directors of the Program. The ultimate value of such information can be twofold: it can be used for improvements in future planning and curriculum design, and in understanding the needs of potential future candidates, based on their backgrounds and interests.

## 5. References


Brown, J. D. (2001). *Using surveys in language programs.* Cambridge, UK: Cambridge University Press.

Cuseo, J. (2003). Assessment of academic advisors and academic advising programs. Retrieved May 5, 2017, from http://www.nacada.ksu.edu/Clearinghouse/Links/assessment.htm

Demetriou, C. (2005). Beyond Student Satisfaction Surveys: Designing a Student Focus Group to Assess Academic Advising. The Mentor: An Academic Advising Journal. Retrieved May 5, 2017, from http://www.psu.edu/dus/mentor/050211cd.htm

Dörnyei, Z. (2003). *Questionnaires in Second Language Research: Construction, Administration, and Processing*. London: LEA Publishers

Hurt, B. (2004). Using the balanced scorecard approach for program assessment of faculty advising. NACADA Journal, 24 (1 & 2), 124–127.

Marsh, H. W. (2007). Students' evaluations of university teaching: dimensionality, reliability, validity, potential biases and usefulness. In R.P. Perry and J.C. Smart (eds.). *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective*, Dordrecht; [London]: Springer, 2007, 319–383.

Richardson, J.T.E. (2005). Instruments for obtaining student feedback: a review of the literature. *Assessment and Evaluation in Higher Education 30*(4): 387–415.

Szymanska, I. (2011). Best Practices for Evaluating Academic Advising http://provost.uncc.edu/sites/provost.uncc.edu/files/media/Best-Practices-Evaluating-AcademicAdvising%2BAttachments.pdf

# L2 Motivation and Self-efficacy's Link to Language Learners' Flow and Anti-flow Experiences in the Classroom

**Katalin Piniel**
*Eötvös Loránd University, Budapest, Hungary*
*brozik-piniel.katalin@btk.elte.hu*

**Ágnes Albert**
*Eötvös Loránd University, Budapest, Hungary*
*albert.agnes@btk.elte.hu*

## 1. Introduction

Although the importance of affect and emotional experiences in language learning has long been acknowledged, in order to assist learners and teachers in the language learning process, studies in the past have primarily focused on negative emotions. While it is quite obvious that negative emotions can only be combated if we understand their impact, it also appears to be logical to assume that positive emotions carry some benefits that are worthy of investigation. Acknowledging this, a recent shift towards researching the effects of positive emotions can be witnessed in psychology, attempting to describe optimal experiences and focusing on "the positive" in general(Seligman & Csíkszentmihályi, 2000). The fact that this trend has also found its way into the field of applied linguistics is well represented by the recent publication of a book on positive psychology in SLA (MacIntyre, Gregersen & Mercer, 2016).The present paper investigates how one type of optimal experience, learners' flow, and its antithesis, anti-flow, are linked to motivated language learning behaviour and self-efficacy beliefs about learning English as a foreign language.

## 2. Background

### 2.1. Flow theory, flow and anti-flow experiences

The flow experience, defined as "the holistic sensation that people feel when they act with total involvement" (Csíkszentmihályi, 1975, p. 36), typically accompanies an activity that people are willing to perform for its own sake, because they enjoy it so much. It is a type of optimal experience with a long research tradition in psychology (Csíkszentmihályi, 1975; Carli, Delle Fave & Massimini, 1988; Csíkszentmihályi, 1997; Nakamura & Csíkszentmihályi, 2002, Csíkszentmihályi, 2014). On the phenomenological level, there seem to be a number of characteristic features of flow. While experiencing flow, people tend to focus on the task so much that they seem to lose self-consciousness for that period; this experience is often labelled as the "merging of action and awareness" (Csíkszentmihályi, 2014, p. 230). They also feel very much in command and are not worried about losing control of the task. An altered sense

of time – time passing very quickly or slowing down – is also characteristic of flow (Csíkszentmihályi, 2014).

Although flow experiences are not very frequent in everyday life, there are three conditions that persistently seem to increase the likelihood of this phenomenon. One such condition is having a clear set of goals that are not distant but immediate (Csíkszentmihályi, 2014). Having concrete goals enhances the probability of clear and immediate feedback (Csíkszentmihályi, 2014), as under these conditions it should be fairly easy to determine whether the goals have been achieved or not. The third condition that should be met in order to experience flow is "a balance between perceived challenges and perceived skills" (Csíkszentmihályi, 2014, p. 232), which can be easily upset by either having to confront challenges that exceed the person's skills (this can lead to anxiety) or by having to embark on tasks that do not pose any challenge for the person (this can lead to apathy or boredom). For this reason, flow can only be described as a temporary state.

Csíkszentmihályi identified different types of anti-flow experiences in various models of flow (Csíkszentmihályi, 1975; Carli, Delle Fave & Massimini, 1988; Csíkszentmihályi, 1997) with apathy and anxietybeing present in all later models. Apathy is conceptualized as the antithesis of flow in the sense that it is characterized by a lack of perceived challenges and abilities. In the case of anxiety, the perceived challenge is present, but its level exceeds the persons' perceived level of abilities. While apathy appears to be an under-researched area within applied linguistics, a number of studies can be found on anxiety related to language learning. Horwitz, Horwitz, and Cope's (1991) widely accepted definition of foreign language classroom anxiety states that it is "a distinct complex of self-perceptions, beliefs, feelings, and behaviours related to classroom language learning arising from the uniqueness of the language learning process" (p. 31). Thus, foreign language classroom anxiety is a situation-specific type of anxiety closely tied to learners' experiences in the foreign language classroom context (MacIntyre, 1999), which can have a detrimental effect on cognition (Eysenck, 1979), including language task performance (Egbert, 2003).

In spite of the relevance and applicability of flow theory in investigating the language learning process, there are relatively few publications available on the issue, perhaps with the exception of Egbert (2003) and Czimmermann and Piniel (2016). In her article, Egbert (2003) proposed a hypothetical model of the relationship between flow and language acquisition. In this model, she emphasizes the *task* (cf. Eysenck, 1979) and its link to the learner's resources (cf. Bandura's (1988, 1997) concept of self-efficacy) available for task completion, in other words, how challenging the task is for the learner. When the learner is presented with a task that matches their skills and provides just the right level of challenge, they are likely to experience the temporary state of flow. This is also echoed by the results of Czimmermann and Piniel (2016) in the university context of English majors. The authors found that learners are more likely to experience flow when the conditions of optimal levels of task difficulty, manageable challenges, and perceived opportunity of control are met. In contrast, when tasks are too difficult/too easy, too challenging/not challenging enough and there seems to be little room for learner control, the learning situation tends to be characterized by students' anti-flow experiences.

### 2.2. Experience and self-efficacy

From the components of flow theory, it becomes clear that a person's perception of their ability to complete a particular task, in other words, how difficult or easy they

find the task, plays a key role. The notion of self-assessment of our abilities is very closely linked to Bandura's concept of self-efficacy beliefs. According to Bandura (1986), self-efficacy can be defined as "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances" (p. 391). Wong (2005) specifies "language self-efficacy beliefs" (p. 248) as beliefs specifically related to language learning (see Horwitz, 1988). Therefore, we can say that in the foreign language classroom self-efficacy beliefs concern not only the resources learners have in general to learn a foreign language but also the resources they believe they have in order to successfully take part in activities and perform tasks in the language classroom (Bandura 1986, 1988).

It has also been theorized that self-efficacy beliefs are strongly rooted in learners' previous experiences, namely, mastery experiences, vicarious experiences, verbal persuasions, and emotions (Bandura, 1997; Mills, 2014). Nevertheless, out of these four factors, the first, mastery experiences, seem to have the greatest effect on self-efficacy beliefs. Therefore, learners' experiences in the language classroom as to how successfully they have performed a task and the feelings that are evoked during this experience will be key determinants of how they think about their own abilities. In turn, self-efficacy beliefs will further influence motivation, more specifically learners' persistence and the effort they intend to invest into learning (Zimmermann, 2000).

### 2.3. Experience and motivation

This takes us to the next individual differences variable, language learning motivation. Besides the motivating effect of experience, which exerts itself through self-efficacy, language learning experience is also a key precursor to motivation as posited by Dörnyei's model of the L2 motivational self-system (Dörnyei& Ryan, 2015). According to this model, a person's effort and persistence in learning a foreign language (referred to as the construct of motivated learning behaviour by,e.g., Kormos & Csizér, 2008; Csizér & Kormos, 2009) largely depends on three elements: the L2 ideal self, which includes future self-guides; the ought-to L2 self, encumbered by perceived expectations; and previous language learning experiences, stemming from past encounters with the L2. In this sense, motivated learning behaviour is influenced by the strength of the ideal L2 self, the ought-to L2-self and the experiences linked to the L2. In the Hungarian secondary school context, Csizér and Kormos (2009) have found language learning experiences to be more influential in determining motivated learning behaviour than the ideal or the ought-to L2 selves. Therefore, it is reasonable to assume that the more positive encounters learners have had with the foreign language, the more likely they will invest energy into developing their language knowledge.

Similarly, Csíkszentmihályi (2014) argues that flow experiences can result in "emergent motivation" (p.234) because of the intrinsically rewarding nature of flow and the resulting increase in skill. Since flow experiences are autotelic, that is, people tend to take part in them because they enjoy them, they urge people to repeat such activities. Repeated practice in turn leads to mastery, which is accompanied by feelings of competence and higher levels of self-efficacy.

On the other hand, intrinsic motivation to learn (in the present case to learn a foreign language) can positively influence the experience of the learning process (Varga & Osvát, 2012). Piniel and Csizér (2013) found support for this claim in the EFL context, where motivated learning behaviour was shown to affect learning experience directly, while experience through self-efficacy and language anxiety indirectly

influenced learners' motivated learning behaviour. From this it seems that motivation is probably a component of a circular relationship structure that includes experience and self-efficacy, and is not merely a source or an end in itself.

## 3. Study

### 3.1. Research question and hypotheses

Based on the above review, it seemed worthwhile to investigate how positive and negative language learning experiences, operationalised as flow and anti-flow, might be linked to motivation and self-efficacy beliefs on a larger sample of students. We hypothesized that a higher level of motivation to learn a foreign language would lead to more experiences of flow (see Figure 1). The positive experiences in the language classroom were expected to increase learners' level of self-efficacy, which in turn might further enhance language learners' motivation. In case of anti-flow experiences (see Figure 2), we hypothesized that learners with lower levels of motivation would be more likely to experience anti-flow, which in turn was expected lead to lower levels of self-efficacy. The lower the levels of self-efficacy learners reported, the lower we expected their levels of motivation to be.

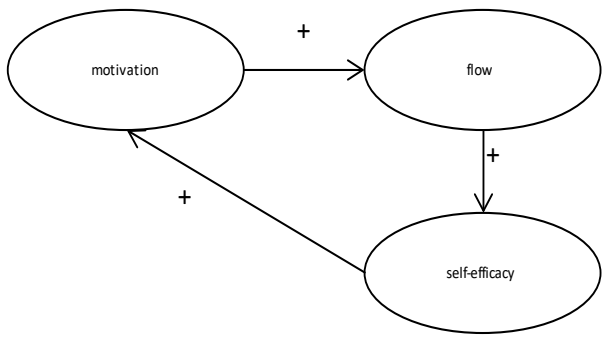


*Figure 1.* Schematic representation of the hypothesized relationship between motivation, self-efficacy, and flow experiences

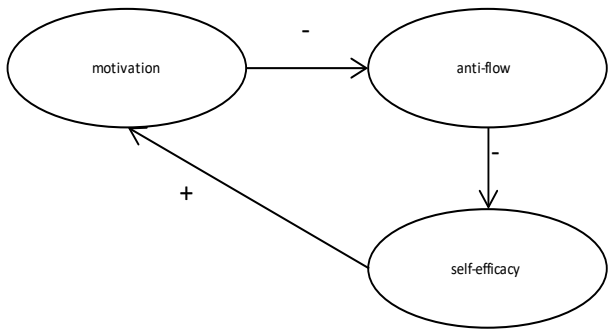


*Figure 2.* Schematic representation of the hypothesized relationship between motivation, self-efficacy, and anti-flow experiences

### 3.2. Method

In order to investigate the relationship of flow and anti-flow experiences with language learning motivation and self-efficacy beliefs, a questionnaire study was designed involving Hungarian high school students. The data was subjected to quantitative analysis using structural equation modelling (SEM)in order to confirm the hypothesized interrelationship of the variables under scrutiny.

#### 3.2.1. Participants

The target population of our study comprised tenth grade high school students in Budapest, Hungary all learning English as a foreign language without a specialized language curriculum which means 3-4 lessons of 45 minutes of language instruction per week. Tenth graders were chosen as the target population in order to control for the novel experience of beginning secondary school in the ninth grade as well as to avoid the influence of the pressure of final examinations in the final grades (grades eleven and twelve). Using this sampling frame, eight schools were selected by way of stratified random sampling. Altogether 214 participants formed the final sample for the present study. The participants consisted of 130 female and 84 male language learners. As specified above, 63.8 % of the students studied English as a Foreign Language at the time of the data collection in three lessons a week and an additional 19.2% in four lessons a week. Moreover, 15.4% also attended extracurricular (non-obligatory) English classes once a week and 1.4% twice a week, organized by their school. Out of the 214 learners, 58% claimed to have been in a country where they have communicated using English, and 41% spent between one and two weeks in the foreign country. Overall, the time students spent abroad ranged from one day to eight years. There were three learners whose stay was over a year long.

#### 3.2.2. Instruments

Seven scales consisting of five-point Likert-scale items were used to collect data on the following constructs: flow experience, apathy (an anti-flow experience), self-efficacy beliefs, classroom anxiety (in terms of fear of negative evaluation of peers, fear of negative evaluation of the teacher, lack of self-confidence), and motivated learning behavior.

Learners' experiences of *flow* in the language classroom were measured using the validated version of Oláh's (2005) Hungarian scale for classroom flow, adapted to the English as a foreign language classroom. This measure was developed for investigating flow experiences in the well-defined situation (see Csíkszentmihályi, 2014) of classroom learning. This was also used by Czimmermann and Piniel (2016) to measure general classroom flow in the English as a Foreign language classroom at university. In their study, classroom flow was found to be related to task-specific flow, but it was a more general measure of flow experiences. The scale included three items referring to general language classroom experiences (example: *I continuously feel that things are going well*). *Apathy* was also measured by Oláh's (2005) instrument which, besides tapping into flow, focuses on anti-flow experiences. The apathy scale consisted of three items (example: *This is unimportant for me*).

*Language classroom anxiety* was measured as a composite of three scales with items adapted from the Foreign Language Classroom Anxiety Scale (Horwitz, et al., 1991; Tóth, 2008): fear of negative evaluation of peers, fear of negative evaluation of the teacher, and lack of self-confidence. The four items measuring the fear of negative evaluation of peers focused on learners' feelings of inhibition associated with the worry about how classmates perceive the learner's performance in a foreign language classroom (example: *I'm afraid the others will laugh at me when I speak English*). The construct of the fear of negative evaluation of the teacher was operationalized with the help of 3 items targeting the negative feelings invoked by the teacher's assessment (example: *I worry that I will not be able to meet the requirements in English class*). The lack of confidence was the final factor operationalized as a constituent of language anxiety. The three items of this scale measured learners' general feelings of uncertainty associated with the foreign language classroom (example*: I am very sure of myself when I speak in English in class*).

The *self-efficacy* scale consisted of four items adapted from the Beliefs About Language Learning questionnaire (Horwitz, 1988; Kuntz, 1996) that operationalized learners' beliefs about their abilities to successfully learn a foreign language (example: *I believe that I have the ability to learn to speak English well*).

Finally, *motivated language learning behavior* was measured with the help of a four-item scale adapted from Kormos and Csizér (2008). This tapped into the extent learners are ready to invest energy and sustain it in the course of foreign language learning (example: *I'm ready to invest considerable effort into learning English*).

### 3.2.3. Procedures

We used think-aloud protocols to pilot the instruments for their use in the foreign language learning context. Three respondents from the population under scrutiny were selected on a voluntary basis to take part in the interviews. After a short orientation on verbalizing their thoughts, the respondents were asked to read the questionnaire items one by one and reformulate them in their own words to check understanding. Based on the interviews, no major modifications were found to be necessary. Then the participant schools were randomly selected from the pool of Hungarian secondary schools in Budapest that matched the criteria outlined above. These schools were then contacted through the principals or vice principals who were informed of the details of the research. Copies of the instruments were made available to the school management. Finally, eight schools agreed to take part in the study, and either the principal or the head of the English Department suggested a class for taking part in the study.

Once the necessary permission was granted, the purpose of the study and its procedures were explained to the learners. The questionnaires were then distributed and filled out. Anonymity was ensured throughout the data collection process. It took students about 35 minutes to respond to all the items.

### 3.2.4. Data analysis

Participants' answers were recorded with the help of the Statistical Package for the Social Sciences (SPSS) 16.0 for Windows. The responses on the five-point Likert scales were assigned numerical values (strongly agree= 5, agree = 4, neither agree nor disagree = 3, disagree = 2 and strongly disagree = 1). In the case of negative items,

the values were assigned in reverse order. The responses to the questionnaire items constitute the observed variables in our study, which were hypothesized to carry information on the latent variables of motivated language learning behaviour, flow/ anti-flow experiences, and self-efficacy respectively.

With the help of Amos 16.0, we used structural equation modelling (SEM) as a technique to analyse our data and the complex relationships among the observed and latent variables in our study. As the first step, the literature suggests that measurement models for each latent variable should be set up, evaluated and modified where necessary (Schumacker & Lomax, 2004). Measurement models provide information about the extent to which the observed variables measure the hypothesized latent variables and are often referred to as confirmatory factor analysis (CFA) models (Schumacker & Lomax, 2004). With CFA models the researcher's goal is to confirm the validity of the theoretical models. In the case of the present study, we drew up and tested measurement models to see whether each latent variable that we intended to tap into (flow, apathy, self-efficacy beliefs and motivated learning behaviour) is indeed identified by the distinct items on the scales (Byrne, 2009; Schumacker & Lomax, 2004). In the case of language classroom anxiety, however, we used a so-called second-order measurement (Byrne, 2009) model where the items comprised aggregates of the following variables: fear of negative peer evaluation, fear of negative teacher evaluation, and lack of self-confidence. Subsequently, based on our hypotheses, all of the above measurement models were combined into distinct structural models with flow and anti-flow experiences respectively.

## 4. Results and discussion

We drew up three different measurement models for the flow and the two anti-flow experiences of apathy and anxiety. As mentioned above, with the help of measurement models, we checked whether the items on the scales adequately measure the latent constructs. In the following figures (Figure 3 through 6), the numbers next to the one-way arrows depict the extent to which the observed variables (shown in the rectangles) carry information about the latent variables (presented in ellipses). Apart from this, we can also see the correlations between pairs of these variables (indicated by the numbers next to the two-headed arrows). The correlation coefficients show that the latent variables are indeed related but at the same time they represent distinguishable constructs.

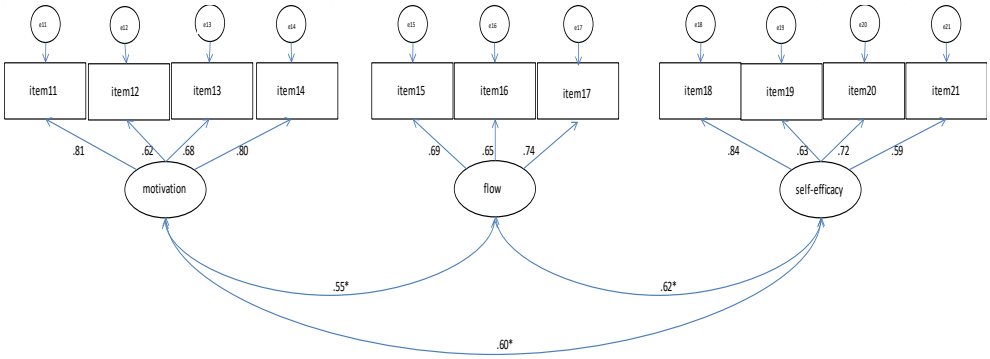


*Figure 3.* Measurement model with flow
*$p$<.05.

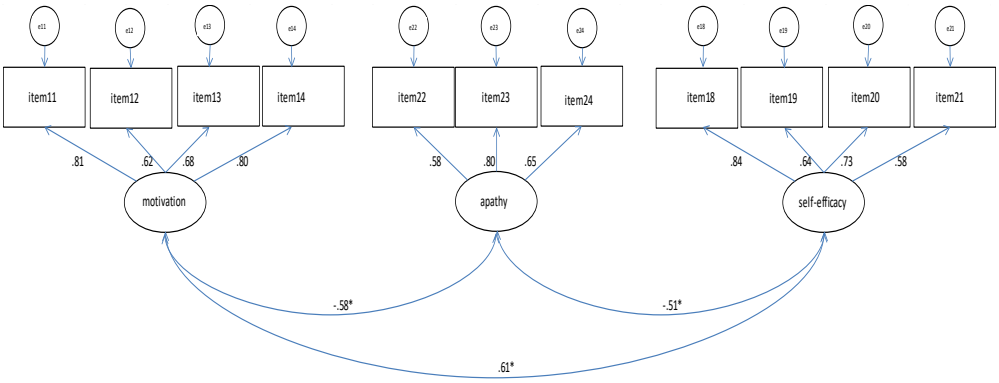


*Figure 4.* Measurement model with apathy
*$p$<.05.

In the case of language classroom anxiety, we used a second-order measurement model because language class anxiety was conceptualized as a composite of fear of negative evaluation of peers, fear of negative evaluation of the teacher, and lack of self-confidence (see Figure 5). This allowed us to use the aggregates of the latent variables in the measurement model with the other variables (Figure 6).

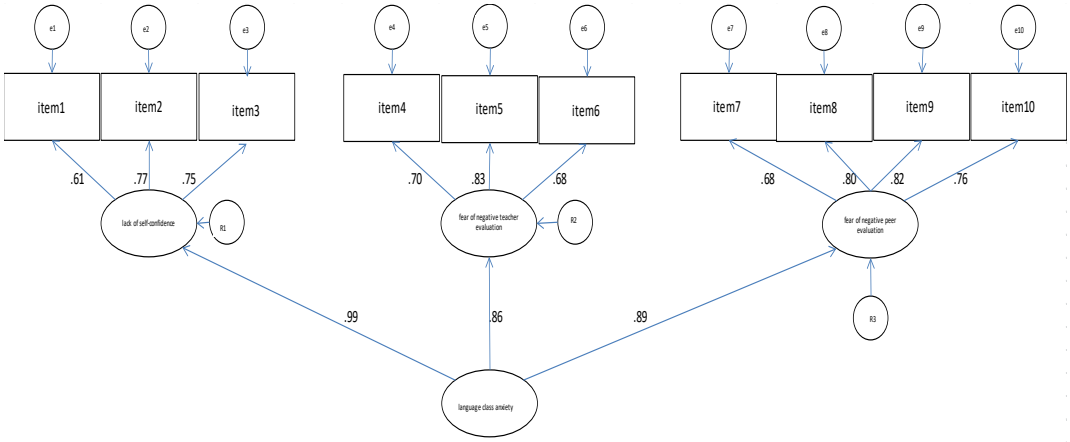


*Figure 5.* Second-order measurement model of language class anxiety

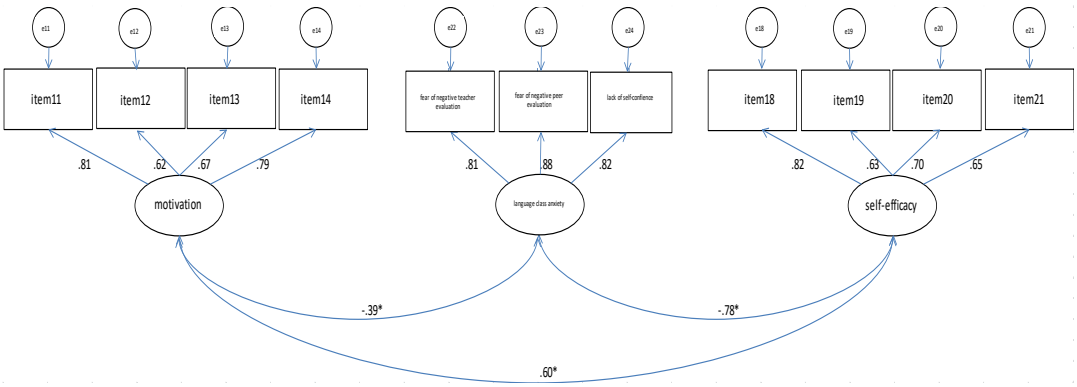


*Figure 6.* Measurement model with language class anxiety
*$p$< .05.

Based on the selected indicators of model fit (see Table 1), there is evidence for the validity argument regarding the questionnaire items and the latent variables they intended to measure. Thus, we could proceed with the descriptive analyses of our scales (see Table 2).

Table 1 Fit statistics of the measurement models

| | CMIN/ df | p | CFI | GFI | NFI | RMSEA | SRMR |
|---|---|---|---|---|---|---|---|
| Measurement model with flow | 1.55 | .007 | .97 | .94 | .92 | .05 | .056 |
| Measurement model with apathy | 1.50 | .012 | .97 | .95 | .92 | .05 | .053 |
| Second-order measurement model of language class anxiety | 2.49 | <.001 | .95 | .93 | .93 | .08 | .049 |
| Measurement model with language class anxiety | 1.95 | <.001 | .94 | .91 | .89 | .07 | .065 |

All the scales included in our instrument also proved to be adequately reliable with Cronbach's alpha values of .72 and higher. Based on the means of five-point scales, we can say that the sample comprised generally motivated learners (*M*=3.85) with moderately high levels of self-efficacy (*M*=3.38) and low levels of language class anxiety (*M*=2.38). It is interesting to note, however, that the average score for flow (*M*=2.88) is lower than expected in light of the level of motivation; on the other hand, the low level of apathy (*M*=2.00) learners seem to experience in the classroom is quite encouraging.

Table 2 Descriptive statistics based on the measurement models

| | | Mean | *SD* | Min. | Max. | Number of items | Reliability ($\alpha$) |
|---|---|---|---|---|---|---|---|
| Language class anxiety subscales | Fear of peer | 2.16 | .92 | 1.88 | 2.34 | 4 | .84 |
| | Fear of teacher | 2.32 | .90 | 2.16 | 2.55 | 3 | .78 |
| | Lack of self-confidence | 2.65 | .93 | 2.37 | 2.88 | 3 | .75 |
| Language class anxiety (aggregate) | | 2.38 | .81 | 2.17 | 2.65 | 3 | .86 |
| Motivated learning behaviour | | 3.85 | .77 | 3.35 | 4.34 | 4 | .81 |
| Self-efficacy | | 3.38 | .82 | 2.81 | 3.73 | 4 | .78 |
| Flow | | 2.88 | .89 | 2.63 | 3.02 | 3 | .73 |
| Apathy | | 2.00 | .79 | 1.86 | 2.22 | 3 | .72 |

### 4.1. Structural models

In the second part of our analyses, we tested the hypothesized relationship between language learners' motivation, their flow and anti-flow experiences and self-efficacy. In the following, we present our three main findings.

First of all, the structural model indicating a positive cyclical relationship among the abovementioned qualities seemed to fit our data well (see Figure 7)(CMIN/df=1.66, CFI=.97, GFI=.94, NFI=.92, RMSEA=.056, SRMR=.057). In line with Litman (2005), our results suggest that the more motivation a language learner has with the urge and curiosity to take part in the learning process, the more likely they will become immersed in language learning tasks and experience flow in the classroom. This kind of experience will provide impetus for the learner to think positively about their abilities to meet challenges in the course of language learning; thus, it will positively influence their self-efficacy beliefs. As the learners see themselvesas someone who is able to meet the challenges in the language classroom, it will prompt them to invest more effort and sustain that effort longer, thus contributing to their motivation.

These findings parallel thoseof Piniel and Csizér (2013) and also reflect the autotelic nature of flow according to which the positive experience of having the skills to meet the challenges of the learning situation further enhances motivation (Csíkszentmihályi, 2014). Moreover, it also echoes the findings of self-efficacy studies where learners with higher levels of self-efficacy have been noted to attribute their successes more to their effort (Hsieh & Kang, 2010). In applied linguistics literature, self-efficacy has also been referred to as a future oriented construct that has been shown to be positively linked with the ideal L2 self, thus prompting the learner to invest more energy into learning in the near future (Iwaniec, 2014).

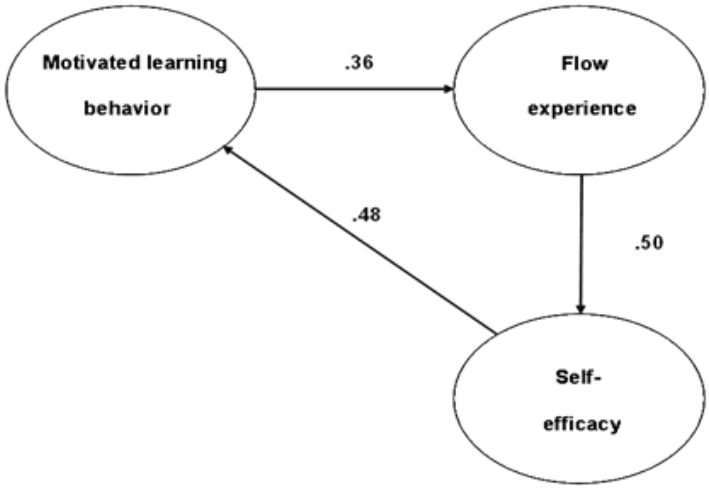


*Figure 7.* The structural model indicating the relationship between motivation, self-efficacy and flow experiences

The second structural model with anti-flow experiences was tested separately with apathy and language class anxiety, as both have been grouped under this umbrella term. Based on acceptable model fit indices (CMIN/df=1.60, CFI=.97, GFI=.95, NFI=.92, RMSEA=.053, SRMR=.055), we can say that in the case of apathy (Figure 8), we found

evidence for a negative cycle, as hypothesized. That is, the lower the motivation to learn a foreign language, the more likely the language learner will experience apathy; in other words, they will find language learning unimportant and will not concern themselves with their skills and the challenges language learning poses. It seems that higher levels of apathy induce lower levels of self-efficacy. Nonetheless, it appears from the data that a sense of self-efficacy will most probably contribute positively to language learning motivation. Therefore, it is perhaps at this point where the vicious circle could be halted by increasing learners' self-efficacy through vicarious and mastery experiences and also associating positive emotions with these experiences.

Finally, in case of the anti-flow experience of language anxiety, the hypothesized model was not confirmed. One of the main reasons for this may be that anxiety is more of an emotional response evoked by the appraisal of an experience (self-efficacy beliefs). Indeed, in psychological literature, anxiety is defined as an affect that is very similar to fear (Németh, 2012). As such, in the present case, it seems to be more of an antecedent to motivation, that is, it plays a key role in determining the energy and persistence with which the action will be carried out.

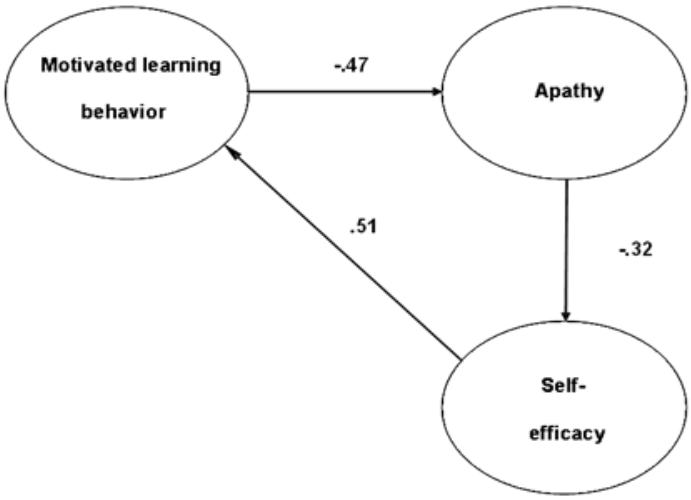


*Figure 8.* Structural model depicting the relationship between motivation, self-efficacy and the anti-flow experience of apathy

## 5. Conclusion

The circular models that have emerged from structural equation modelling suggest three possible points of intervention. Regarding the positive cycle, inducing flow experiences can lead to a higher level of language learning motivation as suggested by Dörnyei (2005) and Csíkszentmihályi (2014) through boosting students' perceived skills. It seems from our results that positive self-efficacy beliefs can also have a direct motivating effect. Moreover, the standardized regression weight in the structural model suggests, in line with the findings of Piniel and Csizér (2013), that there is a moderate link between students' motivation and their flow experiences while learning a language. That is, by increasing learners' motivation, they will more likely experience flow in the classroom.

Unfortunately, studies show that flow experiences tend to be relatively rare in educational contexts (Shernoff & Csíkszentmihályi, 2009).Therefore, the negative cycle where apathy negatively affects self-efficacy beliefs and a lack of motivation leads to apathy is probably also a characteristic scenario in language classrooms. Research suggests that there is a higher likelihood of students experiencing flow in the classroom when learning activities are structured more like extracurricular (Shernoff & Vandell, 2007) and non-academic classes (Shernoff, Csíkszentmihályi, Schneider& Shernoff, 2003), and when there is more room for activity on the students' part. Having a sense of autonomy and the opportunity to interact with others can also result in creating more optimal learning environments (Johnson, 2004). Moreover, virtual learning environments seem to have great potential for inducing flow in learners (Pearce, 2005; Scoresby & Shelton, 2011). Although these findings originate from studies of general education (Shernoff & Csíkszentmihályi, 2009), it might be worthwhile experimenting with them in language classes as well, especially since most of the recommended changes are already present in the repertoire of many language teachers.

## 6. References


Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1988). Self-efficacy conception of anxiety.*Anxiety Research: An International Journal, 1*(2), 77-98.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York, NY: Freeman.

Byrne, B. M. (2009). *Structural equation modelling with AMOS: Basic concepts, applications, and programming.* Mahwah, NJ: Lawrence Erlbaum.

Carli, M., Delle Fave, A., & Massimini, F. (1988). The quality of experience in the flow channels: Comparison of Italian and US students. In M. Csíkszentmihályi& I. Csíkszentmihályi (Eds.), *Optimal experience: Psychological studies of flow in consciousness* (pp. 288-306). Cambridge, UK: Cambridge University Press.

Csíkszentmihályi, M. (1975). *Beyond boredom and anxiety: Experiencing flow in work and play.* San Francisco: CA: Jossey-Bass.

Csíkszentmihályi, M. (1997). *Finding flow: The psychology of engagement with everyday life.* New York, NY: Basic Books.

Csíkszentmihályi, M. (2014). *Flow and the foundations of positive psychology: The collected works of Mihály Csíkszentmihályi.* Dordrecht, Germany: Springer.

Csizér, K., & Kormos, J. (2009). Learning experiences, selves and motivated learning behaviour: A comparative analysis of structural models for Hungarian secondary and university learners of English. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 98-119).Bristol, UK: Multilingual Matters.

Czimmermann, É., & Piniel, K. (2016). Advanced language learners' experiences of flow in the Hungarian EFL classroom. In P. D. MacIntyre, T. Gregersen& S. Mercer (Eds.), *Positive psychology in SLA* (pp. 193-214). Bristol, UK: Multilingual Matters.

Dörnyei. Z., & Ryan, S. (2015). *The psychology of the language learner revisited.* New York, NY: Routledge.

Egbert, J. (2003). A study of flow theory in the foreign language classroom. *TheModern Language Journal, 87,* 499-518.

Eysenck, M. W. (1979). Anxiety, learning, and memory: A reconceptualization. *Journal of Research in Personality, 13,* 363-385.

Hsieh, P. P.-H., & Kang, H.-S. (2010). Attribution and self-efficacy and their interrelationship in the Korean EFL context. *Language Learning, 60,* 606–627.

Horwitz, E. K. (1988). The beliefs about language learning of beginning university foreign language students. *The Modern Language Journal, 72,* 283-294.

Horwitz, E. K., Horwitz, M. B., & Cope, J. A. (1991). Foreign language classroom anxiety. In E. K. Horwitz & D. J. Young (Eds.), *Language anxiety: From theory and research to classroom implications* (pp.27-36). Englewood Cliffs, NJ: Prentice Hall.

Iwaniec, J. (2014). Self-constructs in language learning: What is their role in self-regulation?. In K. Csizér& M. Magid (Eds.), *The impact of self-concept on language learning*(pp. 189-205). Bristol, UK: Multilingual Matters.

Johnson, L. (2008). Relationship of instructional methods to student engagement in two public high schools. *American Secondary Education, 36*(2), 69-87.

Kormos, J., & Csizér, K. (2008). Age-related differences in the motivation of learning English as a foreign language: attitudes, selves and motivated learning behaviour. *Language Learning, 58*, 327-355.

Kuntz, P. S. (1996). *Beliefs about language learning: The Horwitz model.*Retrieved from ERIC database (ED397649).

Litman, J. A. (2005). Curiosity and the pleasures of learning. *Cognition and Emotion, 19,* 793-814.

MacIntyre, P. D. (1999). Language anxiety: A review of the research for language teachers. In D. J. Young (Ed.), *Affect in foreign language and second language learning: A practical guide to creating a low-anxiety classroom atmosphere* (pp.24-45.). Boston, MA: McGraw-Hill.

MacIntyre, P. D., Gregersen, T., & Mercer, S. (Eds.) (2016). *Positive psychology in SLA.* Bristol, UK: Multilingual Matters.

Mills, N. (2014). Self-efficacy in second language acquisition. In S. Mercer, & M. Williams (Eds.), *Multiple perspectives on the self in SLA* (pp. 6-22). Bristol, UK: Multilingual Matters.

Nakamura, J., & Csíkszentmihályi, M. (2002). The concept of Flow. In C. R. Snyder & J. Lopez (Eds.), *Handbook of Positive Psychology* (pp. 89-105). New York, NY: Oxford University Press.

Németh, G. (2012). A menekülő és támadó viselkedés motivációja [Motivation of flight and fight behavior]. In É. Bányai & K. Varga (Eds.), *Affektív pszichológia* (pp.215-235). Budapest, Hungary: Medicina.

Oláh, A. (2005) *Érzelmek, megküzdés,tökéletes élmény*[Emotions, coping, perfect experience]. Budapest, Hungary: Trefort.

Pearce, J. (2005). Engaging the learner: How can the flow experience support E-learning? In G. Richards (Ed.), *Proceedings of E-Learn: World conference on E-Learning in corporate, government, healthcare, and higher education 2005* (pp. 2288-2295). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE).

Piniel, K., & Csizér, K. (2013). L2 motivation, anxiety and self-efficacy: The interrelationship of individual variables in the secondary school context. *Studies in Second Language Learning and Teaching 3*(4), 523-550.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modelling*. Mahwah, NJ: Lawrence Erlbaum Associates.

Scoresby, J., & Shelton, B. E. (2011). Visual perspectives within educational computer games: Effects on presence and flow within virtual immersive learning environments. *Instructional Science, 39*(3), 227-254.

Seligman, M. E., & Csíkszentmihályi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55*(1), 5-14.

Shernoff, D. J., & Csíkszentmihályi, M. (2009). Cultivating engaged learners and optimal learning environments. In R. Gilman, S. E. Huebner & M. J. Furlong (Eds.), *Handbook of positive psychology in schools* (pp. 131-145). New York, NY: Routledge.

Shernoff, D. J., & Vandell, D. L. (2007). Engagement in after-school program activities: Quality of experience from the perspective of participants. *Journal of Youth and Adolescence, 36*(7), 891-903.

Shernoff, D. J., Csíkszentmihályi, M., Schneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly, 18*(2), 158-176.

Tóth, Z. (2008). A foreign language anxiety scale for Hungarian learners of English. *WoPaLP, 2*, 55-78.

Varga, K.,& Osvát, J. (2012). Az ingerszükséglettől a tudászükségletig: Kognitív motiváció [From stimulus needs to knowledge needs: Cogitive motivation]. In É. Bányai & K. Varga (Eds.), *Affektív pszichológia* (pp.237-262). Budapest, Hungary: Medicina.

Wong, M. S. (2005). Language learning strategies and language self-efficacy: Investigating the relationship in Malaysia. *Regional Language Centre Journal, 36(*3), 245-269.

Zimmerman, B.J. (2000) Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25,* 82-91.

# The Mean Examiner:
# How Language Anxiety Affects Performance during an Exam

**Robert Märcz**
*Test Developer, ECL Examinations,*
*PhD student at the Doctoral Programme in English Applied Linguistics and TEFL/TESOL,*
*University of Pécs, Hungary*
*mercrobi@gmail.com*

## 1. Introduction

At the end of the 19th and beginning of the 20th century, psychologists Norman Triplett and Gordon Allport observed that people behave differently when they are in the presence of others. Their findings, however, were contradictory as in some cases the performance of the individuals was increased, while at other times it became worse. This gave birth to the theory of social facilitation, which can also be defined as the notion that people's performance is likely to change when other people are watching them. In his article Zajonc (1965) investigated this phenomenon in the review of studies carried out earlier in the field, and eventually developed his activation theory. This was the first theory attempting to explain why the mere presence of others increases performance sometimes, yet decreases it other times. According to him, the presence of others produces a heightened level of arousal in an individual, which increases the likelihood of doing better on well-learned tasks. However, in the case of complex or unfamiliar tasks, such an increased arousal level decreases performance.

## 2. Previous research

### 2.1. Anxiety and performance

Cottrell (1968) and Henchy and Glass (1968) modified Zajonc's observation suggesting that the source of increased activation in the presence of others was not the mere presence of them, but a feeling of anxiety that we are being judged. This theory, called evaluation apprehension, predicts that when we work surrounded by others, our concern over what they think can enhance or impair our performance. According to these researchers, increased activation only happens when "actors" are afraid of being evaluated. Weiss and Miller (1971) later theorized how activation increases only when the individual perceives negative evaluation.

Zajonc (1980) modified his original theory regarding activation adding that the presence of observers also triggers uncertainty in the actor since they do not know how the other persons will behave. This then causes a state of alertness and this is what increases the level of arousal. Zajonc retained his original idea, in which this heightened arousal level increases performance on simple tasks, but decreases performance of complex tasks.

Guerin (1983) also mentioned uncertainty suggesting that social facilitation effects only occurred when a situation involved uncertainty. In cases when the task or situation is not familiar, when the individual does not know the observer, the uncertainty inherent in such situations increases the arousal level of the individual. Having tested these assumptions, Guerin also observed that such performance differences were only traceable in the case of complex tasks.

### 2.2. Anxiety and foreign language test performance

While participating in an oral language exam, candidates are being watched and evaluated by examiner(s). On the basis of the theories described above, there is reason to believe that the presence and behaviour of the examiners may have a strong impact on the performance of the test takers. We may also assume that, as candidates perform a complex task, that is, speaking in a foreign language, the presence of examiners may have a *negative* effect on their performance.

Regarding the effects the various factors impact upon second language test performance, Bachman (1990) suggested a coherent framework consisting of four factors: (a) communicative language ability, (2) the test taker's personal characteristics, (3) test method factors, and (4) random factors. Within the scope of our research, we are primarily interested in the second set of factors, personal characteristics, which includes the personal attributes of the test taker that potentially affect their test performance. The attributes, which differ throughout each test taker, include their (1) demographic background (e.g., age, gender, socio-economic status and educational background), (2) cognitive characteristics (e.g., aptitude and learning strategies), (3) personality features (self-esteem, anxiety and risk-taking), and (4) socio-psychological features (e.g., attitudes and motivation). Several years following the development of the above framework, Bachman and Palmer (1996) described four sets of individual characteristics affecting test performance: (1) personal characteristics (age, gender, etc.), (2) topical knowledge, (3) affective schemata, and (4) language ability. According to the authors, personal characteristics include "individual attributes that are not part of test takers' language ability but which still may influence performance on language tests" (p. 112). The affective responses of the test takers to the characteristics of the test environment or tasks may also increase or decrease test performance. Test takers' familiarity with the circumstances, under which an exam takes place, may also influence their affective responses to the test task. Anxiety, it seems, is one of the factors which has an effect on foreign language test performance.

Too much anxiety about a test is generally referred to as test anxiety, or test apprehension (Zeidner, 1998). In other words, test anxiety is defined as the tendency to worry about one's own performance (e.g., expectations of failure) and aptitude (e.g., self-deprecatory thoughts) under evaluative or test conditions (Calvo & Carreiras, 1993). Zeidner and Calvo & Carreiras found that anxiety might restrain test takers' ability to retain and recall information. Test anxiety has been defined as an element of general anxiety, including cognitive attentional processes, which interfere with competent performance in academic or testing situations (Spielberger & Vagg, 1995). Anxiety produces a "noise" or interference within our brain and this hinders our ability to retain and recall what we have stored in our memory and also has a negative influence on our ability to comprehend and reason. Sarason (1984) connects test anxiety to the cognitive thought processes that the learner goes through while undergoing a test. Although some degree of arousal is necessary for optimal performance, when the

intensity is too high, we begin to feel nervous and experience anxiety. At this level, anxiety becomes distracting causing our performance to decline. Test anxiety may also be generated when test takers, who have performed poorly in the past, form negative or irrelevant thoughts during an evaluative situation. According to Shi (2012) a diabolic circle may be set in motion as the fear of negative evaluation generates anxiety, which affects performance, and poor performance, in turn, decreases confidence and self-efficacy.

When looking at the connection between anxiety and the various language skills (speaking, reading, writing and listening), it appears that speaking in a foreign language produces a high level of anxiety (Horwitz, Horwitz, & Cope, 1986; MacIntyre & Gardner, 1994). The reason behind this phenomenon implies that language learners may not be able to express themselves clearly, which then leads to a decrease in their confidence and an increased level of anxiety. According to Horwitz et al. (1986) there is "a type of performance anxiety stemming from a fear of failure" (p. 127). Phillips (1992) carried out a study, involving 44 French students, to determine what influence anxiety has on students' performance during an oral test. The findings of her study indicated moderately negative correlations between anxiety and performance. She emphasized that language anxiety has a definite influence on performance, however, it is difficult determine its extent. Horwitz et al. (1986) suggest that anxiety soars to the highest levels when test takers communicate orally during an oral test, which has the potential in provoking both test and oral communication anxiety, simultaneously. Han & Li (2005) propose that oral English test anxiety is related to such elements as test anxiety, communicative apprehension and face anxiety (i.e., the fear of the direct threat to one's self-esteem).

The thorough investigation in the causes of oral test anxiety reveals that factors such as testing techniques, test format, time available, testing environment and clarity of test instructions may all generate test anxiety to a certain extent (Young, 1991). Kessler's study (2010) investigated the characteristics of fluency in students' audio journals recorded in a laboratory setting using mobile audio devices. The frequency of their preferred recording environment was observed and raters analyzed the students' volume, pausing and utterance length in order to determine, whether anxiety had an influence on their fluency and whether the data gained can be connected to the two different environments. Kessler pointed out that the influence of affective factors in a language classroom is obvious. He suggested that this influence was produced through the risk of embarrassment or humiliation the student may face when speaking a foreign language.

Finally, it must be noted that within our competitive society, language tests have become a vital tool for making important decisions (granting entrance to higher education, employment, etc.). Consequently, their stake has increased which, in turn, may increase the level of anxiety candidates feel being administered a language exam.

### 2.3. Features of the Hungarian classroom

The Hungarian education system has undergone many changes in the past 25 years. However, the influence of the old-fashioned, "Prussian" model, which became the norm in Hungary in 1869 when the first education law was drafted, can still be felt in the everyday practice of public education. This model emphasizes factual learning and discipline. The research of Antalné Szabó Ágnes (2006) on teacher talk supports this premise. In analyzing the communication of video-recorded lessons, she discovered

that 94% of the time it was the teacher who initiated communication, constantly reinforcing herself as the primary source of knowledge. On average, students initiated communication only 6% of the time. Regarding the ongoing oral interactions, teachers talked 78% of the time. These patterns confirm the hierarchical nature of the Hungarian classroom, in which it is the teacher who dominates and determines the flow of communication.

Speaking in a foreign language produces a high level of anxiety in itself (Horwitz, Horwitz, & Cope, 1986; MacIntyre& Gardner, 1994) and a situation, characterized by strong hierarchy, may further increase the general anxiety level of students. The observation study of Nikolov (1999) investigated the general teaching conditions in English lessons in secondary schools. She also found that most of the classes were teacher-fronted, that monotony and boredom characterized the lessons in which the most frequent activity was answering questions. When asked to identify students' strengths and weaknesses, "... a few teachers simply left the questions unanswered and most often them tended to identify more weaknesses than strengths" (p. 230). Many of the teachers complained about their students' aptitude, lack of motivation and willingness to communicate adding that they "(...) did not put enough effort and interest into language studies," (p. 231). We have seen in Kessler's study that the influence of affective factors in a language classroom is obvious and that this influence is produced through the risk of embarrassment or humiliation the student may face, when they speak in the foreign language. Once teachers develop a rather negative opinion on their students, chances are greater that they might be embarrassed and/ or humiliated. Teachers' expectations are crucial, as they often work as self-fulfilling prophecies exerting a strong influence upon the self-image and motivation of students (Csapó, 1998; Józsa & Fejes, 2010; Szenczi, 2008). Investigating the potential existence of various emotions related to classroom atmosphere, Imre (2002) and Oláh (1999, 2005) discovered that according to Hungarian students, only 32% of their school activities create joy, occupy their attention and challenge their abilities. Most of the boredom, apathy and anxiety students experience in their life is directly connected to school. In her questionnaire study Imre (2002) found that when classes are not interesting to students, they tend to become less motivated which, in turn, makes teachers resort to disciplinary measures more frequently.

## 3. The present study

### 3.1. Participants and socio-educational context

In summarizing the results of the studies carried out in the Hungarian context, the following may be stated: (1) there is a strong hierarchy in the lessons in which teachers portray a dominant role (Antalné, 2006); (2) most classes are held in a teacher-fronted way and (3) teachers typically possess a low opinion of their students (Nikolov, 1999); (4) two-thirds of the feelings students have towards school is negative (boredom, apathy and anxiety); (5) in most cases, students do not find school/lessons interesting; (6) in such classes teachers more frequently resort to using disciplinary measures (Imre, 2002 and Oláh, 1999 and 2005). Consequently, the Hungarian language classroom appears to be a place exhibiting several contextual factors (1), (2), (3) and (6) that all have the potential to develop anxiety in students.

At the same time, however, possessing a language certificate has high value throughout Hungary. Entrance into Hungarian public higher education institutions is

granted on a credit basis and applicants who possess a language certificate receive extra credits. From 2020, only those students may enter higher education who have earned a language certificate. This implies that taking a language exam can be regarded as a high stakes situation. Typically, it is secondary school students who submit to language exams. There are several language examination systems throughout the country, and one of the leading ones is ECL Examinations. The ECL Consortium (European Consortium for the Certificate of Attainment in Modern Languages) is an association of institutions providing unified, reliable and valid standards for the assessment of attained language knowledge. ECL exams test oral and written ability to effectively communicate through the use of the language of everyday discourse on practical, professional and personal topics at varying degrees of complexity. The ECL oral exam has a paired format; the examinees are required to converse with one another, in the form of an intelligent and meaningful discussion.

Our study included four participants: two pairs of examinees at two different levels of language knowledge. One of the pairs, two secondary school students (hereinafter referred to as teenage boy and girl), participated in the oral exam on level B2, while the other pair, two university students (hereinafter referred to as male and female university students) participated on level C1. Both pairs included a male and a female participant to determine whether gender is important regarding the focus of our study. We involved two secondary school students, since it is typically this age group who actively pursue language exams at level B2. The proportion of university students already possessing a language certificate is higher, which explains why we had one participant with and another one without a language certificate, in the case of the level C1 participants. The other reason for having at least one participant with a language certificate ensures that in this particular case, we had the chance to observe whether his previous exam experience had any influence on his performance.

When candidates subject themselves to an oral language exam, they are being watched and evaluated by examiner(s). Furthermore, in the case of a paired language exam, there is another person, the other candidate. On the basis of the theories mentioned above, we have reason to suggest that the (mere) presence of others may affect the performance of the test takers. Since the candidates perform a complex task, speaking in a foreign language, the presence and behaviour of the examiners may have a detrimental effect on their performance. Additionally, if the examiner behaves in a way which is conducive towards generating anxiety, we may also assume that this can and will enhance these negative effects. On the basis of the literature reviewed above, test takers' level of anxiety may be increased if they feel embarrassed and/or humiliated, if their self-confidence is undermined and if they feel the situation implies a direct threat to their self-esteem. Such feelings may be generated if the examiner (1) gives open and negative evaluation (especially in front of other people); (2) corrects the mistake test takers make; (3) displays an unpredictable behaviour.

The aim of our small-scale qualitative study was to determine whether anxiety, generated under experimental conditions, in the form of mock oral exams, has visible and measurable influence on the performance of the participants taking an oral language exam in English. Consequently, our research questions were as follows:

- Does the anxiety generating behaviour of the examiner (correcting the mistakes of the participants, making derogatory comments and taking notes in front of them) have an influence on the performance of the candidates?
- In what way does the behaviour of the examiner affect the performance of candidates?

### 3.2. Procedure

The ECL oral language exam consists of three parts. The first part includes an introduction, which prepares and warms the candidates up prior to the actual exam. The second part of the exam is a guided conversation in which the examiner determines a certain topic and provides the candidates with some guided questions so they can develop an intelligent discussion with respect to the topic. The third part includes a monologue in which the candidates are asked to speak individually about a topic represented by images and pictures. The participants of our study were asked to take part in three ECL mock oral exams. All exams were recorded on videotape. Immediately following the oral exams, the participants were asked to take a listening comprehension test, including two ten-item tasks. During the first two exams, the examiner exhibited neutral behaviour towards the participants. However, during the third mock exam, he consistently corrected the mistakes of the test takers, took notes while they were speaking and spoke to them in a derogatory way. We hypothesized, that although these mock exams are not regarded as high-stakes situations, the changed behaviour of the examiner would nevertheless exert an influence on the performance of the test takers. Following the third mock oral exam, a semi-structured interview was conducted and recorded with the participants. They were asked two questions:

1) What difference did you notice in the behaviour of the examiner during the third session?
2) How did his changed behaviour affect you?

## 4. Results

### 4.1. Quantitative results

Let us take a look at the results of the three listening comprehension tests (see Table 1). Two participants achieved a lower score following the third oral exam, while the two other participants' score did not change compared to their second test.

Table 1. Results of the listening comprehension tests on levels B2 and C1

| Participant | Score on Test 1 | Score on Test 2 | Score on Test 3 |
|---|---|---|---|
| **Level C1** | | | |
| Male university student | 90% | 100% | 80% |
| Female university student | 80% | 60% | 60% |
| **Level B2** | | | |
| Teenage boy | 80% | 70% | 70% |
| Teenage girl | 80% | 80% | 70% |

Since the sessions were recorded, we counted the number of grammatical errors the candidates made during the mock oral exams (Table 2). We decided to disregard the mistakes connected to phonology and pronunciation and focus purely on grammatical errors.

Table 2. Number of grammatical errors made during the mock oral exams

| Participant | No. of errors made at exam 1 | No. of errors made at exam 2 | No. of errors made at exam 3 |
|---|---|---|---|
| **Level C1** | | | |
| *Male university student* | 9 | 9 | 14 |
| *Female university student* | 14 | 14 | 14 |
| **Level B2** | | | |
| *Teenage boy* | 6 | 5 | 8 |
| *Teenage girl* | 8 | 5 | 18 |

The number of errors made increased in the case of three of the participants, while with respect to one participant (female university student), it remained at the same (relatively high) level. These results seem to confirm our prior expectation that the behaviour of the examiner can have an influence on the participants' performance.

### 4.2. Qualitative results

Following the third session, a retrospective interview was conducted with the participants. The aim of the interview was for the participants to recall how they reacted to the changed behaviour of the examiner. Each participant acknowledged that the changed behaviour of the examiner indeed had an influence upon them, despite the fact that these exams had no stake. The two university students described that their level of anxiety and frustration had been increased when the examiner became overly personal, when he corrected their grammatical errors, when he made a derogatory comment (e.g. "you should know this") and when he took notes.

The female university student admitted that being too personal had an emotional effect upon her, as most people do not like discussing private issues in front of strangers. She also remarked that the derogatory comments made by the examiner had produced "bad feelings" in her, adding: "...in a real exam situation it would undoubtedly generate frustration in me, because I am not that confident in my overall knowledge."

Regarding the constant correction of grammatical errors and note taking, the male university student explained, "... this time, when my errors are highlighted, my brain begins to solve the problems I made in the past (sic) so I am unable to think of what I want to say because I am constantly thinking about how not to make another error as it will be corrected... and I start to focus on my pronunciation and it is not the natural pronunciation that I produce... I begin to mispronounce even those words I was sure of."

The female university student recounted the same experience and illustrated it with a concrete example. She mentioned that the examiner's constant error correcting behaviour made her focus too much on the words she was saying, and as a result she managed to confuse *medium* and *media,* however she had known the difference between the two words. "... under normal or calm circumstances I am aware of the different meanings, I know this, but this time, it turned out to be the opposite."

Regarding the same issue, the male university student added that he had had the same experience in a former language exam situation: "... when I made an error, I knew I made it because the pen started moving and afterwards, I said 'Oh my God', I made an error, but what was it?" He explained how this haunting memory from his past was revived in his mind producing the same feeling as back then.

The two teenagers had different views on the effects of the examiner's behaviour. The boy stated that this changed behaviour had no effect on him, while the girl acknowledged that she had felt anxiety and anger and, as a result, grew distant and reserved.

At the beginning, both mentioned that the error correcting and note taking activities made the exam more personal, because "... in school, our errors are always corrected," suggesting that they were used to this situation. The teenage girl admitted, it was the note taking that had made her anxiety level increase. She said: "I tried not to look at the paper, where my errors were listed, I felt their number was just growing and growing."

We also asked the teenage boy about his reaction to the note taking of the examiner. He responded, "... if it had been a high stakes situation... it would have been more unpleasant...". Elaborating on what she felt due to the note taking, the teenage girl added, "... by the end, it started to annoy me... I became more reserved so as not to make another mistake." "I thought I would not do well. Perhaps if I had not seen the paper it would have been better, but it was almost in my face." She also added, "I let him (the teenage boy) do the talking, and in the meantime I quickly summarized what I wanted to say, so I would not make an error which then would be corrected. By the end I tried to simplify what I said so as no one would find faults with my statements."

When asked what and how they thought about their mistakes being consistently corrected and notes being taken, the teenage boy said, he had not been bothered by the note taking since, as he put it, "actually we can expect an examiner to be condescending." When asked why, he added "I don't know, if someone already works as an examiner I think it's not that big of a problem if he is a bit disrespectful, and anyway, we are, to a certain extent, used to such a situation..." To the question whether he considered himself self-confident, the teenage boy replied, „If I am adequately prepared for the exam, then perhaps, yes, but if I am not well prepared, then not at all."

## 5. Discussion

Considering what the participants mentioned in the interviews, the role of individual differences is obvious. The results of our small scale study confirm the results of researchers who found that factors such as the testing techniques and the testing environment, including the behaviour of the examiner, may generate test anxiety to a certain extent (Han & Li, 2002; Kessler, 2010; Sarason, 1984; Shi, 2012; Young, 1991).

The interviews corroborated the theory of evaluation apprehension, that we tend to become anxious when being judged (Cottrell, 1968; Henchy and Glass, 1968), and the role of uncertainty in generating social facilitation effects. Three out of four participants mentioned that the concern over what the examiner thought affected them. Reflecting upon the error correcting and note taking activities, they remarked that these had annoyed them and made them focus on trying not to make another mistake, which, in turn, diverted their attention from what they really wanted to say. The female university student admitted that her feelings had restrained her ability to retain and recall information adding, that under normal circumstances, she would

have been able to activate her existing knowledge regarding those particular English vocabulary items (*media* and *medium*), in contrast to what actually happened. This is a concrete example for the findings of Zeidner and Calvo & Carreiras that anxiety might hinder test takers' ability to retain and recall information (Calvo & Carreiras, 1993; Zeidner, 1998).

Speaking in a foreign language seems to produce a high level of anxiety because language learners may not be able to express themselves clearly, and it may decrease their confidence and increase their level of anxiety. During an oral exam anxiety may become even more intensive, as in this case both test and oral communication anxiety may become activated simultaneously (Horwitz et al. 1986; MacIntyre& Gardner, 1994). One way to cope with such unpleasant feelings was visible in the case of the teenage girl participant, who described how she had tried not to look at the notes being taken, grew more reserved, let her partner talk and simplified what she wanted to say. This led her to say less than what she had intended to, so the quality of her oral performance was reduced. As in an oral exam situation it is the oral communication of the candidates which is assessed, we can reason that such a strategy may lead to a lower number of points awarded for the performance.

The male university student displayed a specific example of the phenomenon, described by Sarason (1984), in which test anxiety may increase when test takers who have performed poorly in the past form negative or irrelevant thoughts during an evaluative situation. He said that he had already had such an experience in the past, and this event was recalled in his mind as a result of the examiner's attitude. We can also see examples for how the diabolic circle, mentioned by Shi (2012), works: the fear of negative evaluation generates anxiety, which affects performance, and poor performance, in turn, decreases confidence and self-efficacy. In both test taker pairs, the female candidates referred to this: "in a real exam situation, it must have generated frustration, because I am not that confident in my overall knowledge" (female university student) and "... by the end it (the note taking) started to annoy me... I became more distant and reserved so as not to make another mistake..." (teenage girl).

The teenage participants' comments shed light on the current realities of the Hungarian classroom. They illustrate how the error-correcting and note-taking activities of the examiner made the exam 'more personal', because "... in school our errors are always corrected," suggesting that they are used to this situation. This explanation is in accordance with the findings of Nikolov (1999), that the most frequently used oral activity in an English lesson in Hungary is answering questions in a lockstep fashion, always following the IRF cycle: initiation by the teacher (I), reply by the student (R), and feedback (F), in the form of corrections, by the teacher. The same study (Nikolov, 1999) demonstrated that teachers did not have a high opinion on their students and complained about their aptitude and lack of motivation. The teenage boy remarked that test takers may actually "expect the examiner to be condescending," adding, "...it's not that big of a problem if he is a bit disrespectful". These comments may refer to the fact that students are indeed used to being looked down on and to being expected to perform lower than their actual aptitude would position them.

## 6. Conclusion

Our small-scale qualitative empirical study aimed at revealing whether the examiner's behaviour, having the potential to generate anxiety (correcting mistakes, making derogatory comments and taking notes), has any influence on the performance of the

participants. There were three videotaped mock oral exams, followed immediately by a listening comprehension test. On the first two occasions the examiner demonstrated neutral behaviour towards the participants, while on the third occasion he consistently corrected their mistakes, took notes and talked to them in a derogatory way. The results of the listening comprehension tests showed that two participants achieved a lower score following the third oral exam, while the two other participants' score did not change when compared to their second test. Regarding the oral exam, it was found that the number of grammatical errors made increased in the case of three participants.

We wished to determine how the examiner's behaviour affected the performance of the participants in the retrospective interviews recorded following the third mock exam. Three of the four candidates mentioned concrete examples, admitting that their level of anxiety and frustration had increased when the examiner (1) was too personal, when he (2) corrected their grammatical errors, when he (3) made derogatory comments (e.g. "you should know this") and when he (4) took notes. The concrete comments on the examiner's behaviour illustrate the inverse relationship existing between evaluation apprehension and poorer performance on complex tasks, such as taking an oral language test. It was also visible that error correction has the potential to make a participant feel embarrassed. Under such circumstances, they may feel that their self-esteem and self-confidence is threatened and/or questioned, which may divert their attention away from the task. The same may be stated regarding derogatory comments. The teenage participants' answers reflected the practices being carried out in many English language classes in Hungarian public schools (teachers' constant error correction) and illustrated how students are indeed used to being disdained.

## 7. References


Antalné, Sz. Á. (2006). A tanórai kommunikáció jellemzői. In Antalné, Sz. Á.: *A tanári beszéd empirikus kutatások tükrében*. A Magyar Nyelvtudományi Társaság Kiadványai 226.Magyar Nyelvtudományi Társaság. Budapest. 21–32.

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

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

Calvo, M. G., &Carreiras, M. (1993). Selective influence of test anxiety on reading processes. *British Journal of Psychology, 84,* 375–388.

Cottrell, N. B. (1968). Performance in the presence of other human beings: Mere presence, audience, and affiliation effects. In Simmel, E. C., Hoppe, R. A. & Milton G. A. (Eds.), *Social facilitation and imitative behavior*(pp. 91-110). Boston: Allyn& Bacon.

Csapó, B. (1998). Az iskolai tudás felszíni rétegei: mi ttükröznek az osztályzatok? In Csapó, B. (szerk.): *Az iskolai tudás*. Osiris Kiadó, Budapest. 39–72.

Guerin, B. (1983). Social facilitation and social monitoring: a test of three models. *British Journal of Social Psychology, 22,* 203–214.

Han, W. F., & Li, Y. L. (2005).Anxiety in oral English proficiency test.*Research in Foreign Language and Literature*, *5(1),* 55–62.

Henchy, T., & Glass, D. C. (1968).Evaluation apprehension and the social facilitation of dominant and subordinate responses.*Journal of Personality and Social Psychology, 27,* 74–85.

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

Imre, N. (2002). Öröm, unalom és szorongás a tanórákon. Retrieved from: http://ofi.hu/orom-unalom-es-szorongas-tanorakon

Józsa, K., &Fejes, J. B. (2010). A szociális környezet szerepe a tanulási motiváció alakulásában: a család, az iskola és a kultúra hatása. In Zsolnai, A., &Kasik, L. (szerk.): *A szociális kompetencia fejlesztésének elméleti és gyakorlati alapjai*. (pp 134–162). Budapest: NemzetiTankönyvkiadó

Kessler, G. (2010). Fluency and anxiety in self-access speaking tasks: The influence of environment. *Computer Assisted Language Learning, 23(4),* 361–375.

MacIntyre, P. D., & Gardner, R. C. (1994).The subtle effects of language anxiety on cognitive processing in the second language.*Language Learning, 44*, 283–305.

Nikolov, M. (1999).Classroom observation project. In H. Fekete, É. Major & M. Nikolov (Eds.), *English language education in Hungary: A baseline study* (pp. 221-246). Budapest: British Council.

Oláh, A. (1999). A tökéletes élmény megteremtését serkentő személyiségtényezők serdülőkorban. *Iskolakultúra, 6–7.*15–27.

Oláh, A. (2005). *Érzelmek, megküzdés és optimális élmény: Belső világunk megismerésének módszerei.* Budapest: TrefortKiadó.

Phillips, E. (1992). The effects of language anxiety on students' oral test performance and attitudes.*The Modern Language Journal. 76(1),* 14–26

Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: reactions to tests. *Journal of Personality and Social Psychology, 46,* 929–938.

Spielberger, C. D., &Vagg P. R. (1995). Test anxiety: A transactional process model. In Spielberger C. D. &Vagg P. R. (Eds.), *Test Anxiety: Theory, Assessment and Treatment.* Washington, D. C.: Taylor & Francis.

Szenczi, B. (2008). Énkép és tanulás: Nemzetközi kutatási irányzatok és tendenciák. *Iskolakultúra Online*, *1. 2*. 104–118.

Shi, F. (2012). Exploring Students'Anxiety in Computer-based Oral English Test. *Journal of Language Teaching and Research, 3.( 3),* 446–451

Young, D. J. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *Modern Language Journal*, *75(4),* 427–439.

Weiss, R. F., & Miller, F. G. (1971).The drive theory of social facilitation.*Psychological Review, 78,* 44–57.

Zajonc, R. B. (1965). Social facilitation.*Science, 149,* 269–274.

Zajonc, R. B. (1980). Corn presence. In P. B. Paulus, *Psychology of group influence* (pp. 35–60). Hillsdale. NJ: Erlbaum.

Zeidner, M. (1998) *Test anxiety: The state of the art*. New York: Plenum.

# How Important Is English to Croatian Pre-Service Primary School and Preschool Teachers Not Majoring in English?

**Alenka Mikulec**
*Faculty of Teacher Education*
*University of Zagreb*
*alenka.mikulec@ufzg.hr*

**Renata Šamo**
*Faculty of Teacher Education*
*University of Zagreb*
*renata.samo@ufzg.hr*

## 1. Introduction

There is no doubt that English has spread enormously around the world and the trend is likely to continue. This global status can also be supported by the fact that English is given priority in foreign language teaching. According to Crystal (2012), in more than 100 countries around the world (mainly in Europe, Asia, North Africa and Latin America) English has the status of a foreign language, and in most of them it is the main foreign language taught in schools or at least the language which is most likely to be introduced in schools. Many people are thus exposed to and surrounded by English in different contexts. Today, the internet, with English as the dominant language, has turned into one of the most relevant communication channels and information sources, so it may be presumed that the greatest exposure to English would be on the internet.

In view of its status as a *lingua franca*, English has also been identified by Croatian urban adults as the most desirable and important foreign language, both in private and in professional life (Cindrić & Narančić Kovač, 2005). It is not only the preferred foreign language for the majority of Croatian speakers, but it is also a language in which Croatian speakers strive to achieve high proficiency. This is particularly true when other speakers of English as a foreign language (EFL) are compared with Croatian EFL speakers, as was found in the study comparing the EFL performance between eighth graders in Croatia and Hungary (Mihaljević Djigunović, Nikolov, & Ottó, 2008). Furthermore, Croatian speakers tend to consider their EFL proficiency as rather high. The same was confirmed by Jelovčić (2010) who found that the majority of Croatian university students in her study (87.2%) thought of their EFL proficiency as being very good or good.

Such a great need for English and interest in teaching/learning English has increasingly stimulated studies in which special attention has been also given to individual differences (Dörnyei, 2005; MacIntyre, Gregersen, & Mercer, 2016; Robinson, 2002; Zafar, & Meenakshi, 2012), in particular affective factors (Henter, 2014; Ni, 2012; Rubio, 2007), since their role has been emphasised on numerous occasions and mostly referred to learner attitudes to teaching/learning (Dörnyei, Csizér, & Németh, 2006; Zhao, 2015). Research has confirmed that the positive attitudes of students to English Language Learning (ELL) are related to the perceived importance of English for their

studies, future career opportunities, and communication (Suwannatho & Thepsiri, 2015). Learner attitudes have also been investigated in Croatia, showing that Croatian secondary school learners (Mihaljević Djigunović, 2007) and university students (Martinović & Poljaković, 2010) have positive attitudes to English in general.

## 2. The Study

### 2.1. Study Aims and Hypotheses

The aim of the current small-scale study was to gain insights into the exposure of a group of Croatian university students (N=151) to EFL, their proficiency in EFL, their attitudes to this language, and their perceived importance of being proficient in EFL. Its further aim was to ascertain possible relations referring to the participants' attitudes and exposure to EFL and other foreign languages (FLs), their proficiency in EFL and other FLs, and their overall academic achievement. Finally, the study focuses on the possible differences between two subgroups - English non-majors (N=83) and English majors (N=68) with respect to the above-mentioned variables.

In order to achieve the aims above, the following hypotheses were formulated:
H1: All the participants will report great exposure to and high proficiency in EFL, positive attitudes to EFL, and perceived importance of being proficient in EFL.
H2: There will be relationship between the participants' attitudes and their exposure to and proficiency in EFL and other FLs, their perceived importance of being proficient in English and their overall academic achievement.
H3: There will be significant differences between the two subgroups of the sample, i.e. English majors and English non-majors, on all of the tested variables.

### 2.2. Sample

The sample included 151 second- and third-year university students of the Faculty of Teacher Education, University of Zagreb ($N_{female}$=144, $N_{male}$=7, age Median=21). They were enrolled in three different study programmes: (1) pre-service primary school teachers; (2) pre-service preschool teachers; and (3) pre-service primary school English teachers. For the current purpose, the participants were treated as one group for H1 and H2, and were subdivided into: (A) English non-majors (N=83); and (B) English majors (N=68) for the purposes of H3. It is also important to stress that the two groups significantly differed in their exposure to English in the course of their studies. Namely, in addition to some courses taught in Croatian (e.g. Mathematics, Sciences, Developmental Psychology, Introduction to Pedagogy, etc.), the total study programme for the English majors group entails 1,290 compulsory hours of direct instruction in English and 255 hours of English language teaching and classroom practice over a period of five years (Cindrić, Andraka, & Bilić-Štefan, 2013), whereas the English non-majors group has only 75 (pre-service primary school teachers) or 90 (pre-service preschool teachers) hours of direct instruction in English.

### 2.3. Instrument and Procedure

The participants completed a questionnaire, designed by Mikulec (2016) and adapted for this study (see Appendix). It consisted of 26 open- and closed-ended

questions with the objective to elicit information about the participants and their foreign language learning history. More specifically, the participants were asked about the length of their English and other foreign language learning, and about their average grades in English and other foreign languages in their previous education. We also collected data about their GPA, self-assessment of EFL proficiency, and daily/ weekly exposure to the languages in question. Finally, the participants were asked about the frequency and form of their contact with native speakers of English, their attitudes to English, and the importance they assign to their proficiency in English. The questionnaire was written in Croatian and administered during regular English language classes. The data analysis was based on descriptive statistics, as well as on correlation and nonparametric statistical procedures (the Mann Whitney test). The Software Package for Social Sciences for Windows (SPSS) 15.0 was used for data processing.

## 3. Results and Discussion

### 3.1. Participants' Exposure to EFL, Their Proficiency in EFL and Attitudes to EFL

The emphasised presence of English in everyday life, especially among younger people, has resulted in significant exposure to the language. The current study aimed to ascertain the extent of this exposure in the case of the Croatian student population, and to determine other closely related relevant factors. The results indicate that the average length of English language learning in the target population was M=11.54 (SD=3.061). Further analysis of their language learning history indicates that 43% of the participants reported having additionally learned English in various language courses (mostly between 1-4 years), and as many as 80% of them reported having had contacts with native speakers of English. Next, we were interested in the participants' weekly exposure to English outside their studies, and the results show that 33.8% of them reported exposure of up to 10 hours, 32.5% of them reported up to 15 hours and more, while 29.1% of them reported up to 5 hours. The participants were also asked about the form and length of their daily exposure to English, and the results (Figure 1) show that, for all of the suggested forms (the internet, music, films, books and journals/magazines), exposure was predominantly up to 3 hours a day.

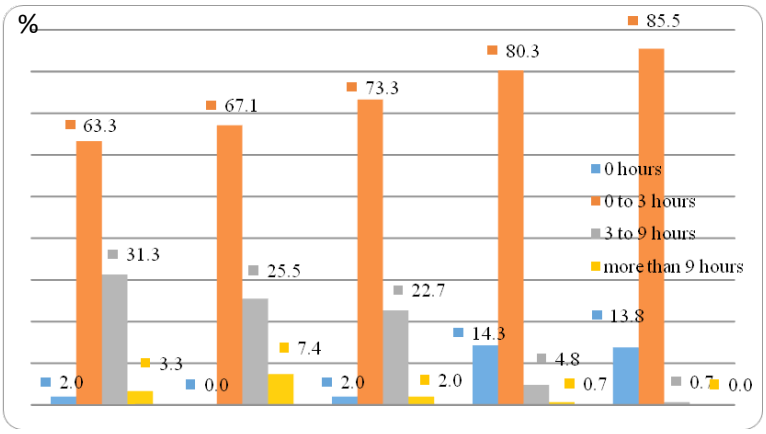


*Figure 1*. Length and form of the participants' daily exposure to English outside their studies, presented in percentages

However, fairly significant exposure (3-9 hours a day) was reported for the internet (31.3%), music (25.5%) and films (22.7%), whereas more than 9-hour exposure was reported for music (7.4%) and the internet (3.3%). These results suggest that the most frequent forms of daily exposure for the target group included the internet and music, while the least frequent forms were journals/magazines and books. These results are in accordance with those in some previous studies which found that internet use was positively correlated with student achievements in English (Josipović Smojver, 2007; Mihaljević Djigunović, 2007). Regarding the participants' exposure to English, it may be concluded that the first part of H1 was confirmed; in other words, the participants reported great exposure.

Language proficiency can be determined on the basis of several indicators (Leclerq & Edmonds, 2014), one of which is self-assessment. According to the Common European Framework of Reference for Languages (2001), self-assessment may prove an efficient addition to learners' assessment, especially when there are no "high stakes" involved, i.e. when their future does not depend on the results of self-assessment, which is the case in this study. Furthermore, Demel (1990) confirmed a high correlation between learners' self-assessment and their real proficiency. On the other hand, since the reliability of data obtained through any form of self-report depends on participants' honesty (Sternberg, 2005), proficiency was based on three variables: (1) primary school average grades in EFL, (2) secondary school average grades in EFL, and (3) self-assessment. The obtained results (Figure 2) indicated that all of the measures were fairly high, especially primary (M=4.81, SD=0.391) and secondary school average grades in EFL (M=4.33, SD=0.800), which grouped towards the higher values. This means that the distribution was negatively skewed, whereas self-assessment (M=3.89, SD=0.799) was somewhat lower than the other two measures. Regardless of these differences, a high correlation was found between the participants' self-assessment on the one side and their primary (r=.352, p=.000) and secondary school grades on the other side (r=.469, p=.000).

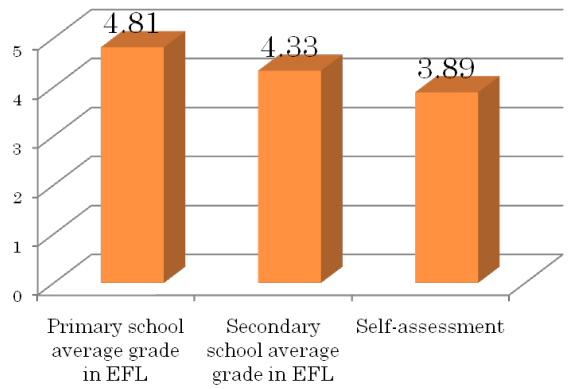


*Figure 2.* Proficiency measures: primary and secondary school average grades in EFL and the participants' self-assessment of EFL proficiency

The current self-assessment results are similar to those obtained by Jelovčić (2010). It may therefore be concluded that the second part of H1 was also confirmed as the participants' self-assessment of EFL proficiency was rather high.

Since attitudes play an important role in foreign language learning (Singleton, 2014), data on the participants' attitudes to English were also collected. The results

showed that a great majority reported positive attitudes (89.3%). Such a high percentage may have resulted from the fact that they were predominantly female students (95.4%) and that 70.7% of them were grammar school graduates. Namely, Mihaljević Djigunović (2007) found a more positive affective profile in a Croatian group of primary and secondary school female learners, and she also determined that grammar school graduates had more positive attitudes to English when compared to graduates from other secondary schools in Croatia. The results of the present study additionally show that as many as 92.7% of the participants reported having learned another foreign language during their education, mostly for four (38.1%) or nine years (23.8%), with a fairly high (M=4.59, SD=0.587) average grade achieved in that foreign language. Although the majority of participants (69.6%) stated that they had no contact with their second foreign language, which may be explained by the fact that it was mostly one of the less present languages (e.g. German or Latin), especially among the young Croatian population, Mihaljević Djigunović (2007) concluded that the learners who had learned more than one foreign language "enhance language awareness, develop learning strategies and decrease language anxiety, which contributes to more positive attitudes to learning English and to a better perception of oneself as a foreign language learner" (p. 119). When the open-ended question related to the participants' attitudes was analysed, it was clear that a majority of the participants expressed positive attitudes to English, which means that they could recognise some present and future benefits of knowing English. Besides, affective reasons, for instance language attractiveness, might have contributed to this. Therefore, it may be concluded that the third part of H1 was also confirmed.

Apart from gaining insights into the participants' attitudes to English, the objective was to ascertain their attitudes to the importance of being proficient in English. Considering the total sample, the percentage was very high again (95.3%). Based on the overall results, the final part of H1 was also confirmed, i.e. most of the participants reported very positive attitudes to English and the importance of being proficient in English, as well.

### 3.2. The Relationship between the Participants' Attitudes and Other Tested Variables

Further data analysis in this paper is related to H2 and focuses on determining possible relationship between the participants' attitudes and exposure to EFL and other foreign languages, their proficiency in English, and their overall academic achievement. The results (Table 1) indicate that the participants' attitudes correlated with the three variables which were used to determine their exposure to EFL: (1) length of EFL learning (r=-0.381, p=.000), (2) learning EFL in language schools (r=0.171, p=.038), and (3) weekly exposure to EFL beyond study programmes (r=-0.386, p=.000). However, only one correlation was positive, i.e. those participants who had learned English in language schools expressed more positive attitudes to English, whereas the length of English learning and weekly exposure to English outside their studies were negatively correlated with the participants' attitudes. All three proficiency measures also correlated with the participants' attitudes as follows: (1) average grade in English in primary school (r=-0.412, p=.000); (2) average grade in English in secondary school (r=-0.329, p=.000); and (3) self-assessment of English proficiency (r=-0.499, p=.000), but again the correlations were negative. That is, the participants with higher primary/ secondary school grades and self-assessment had less positive attitudes to English.

Table 1 Correlations between the participants' attitudes and other tested variables

| Variables | Attitudes |
|---|---|
| Length of ELL | -0.381** |
| Learning English in language schools | 0.171* |
| Weekly exposure to EFL outside studies | -0.386** |
| Learning other foreign languages | 0.085 |
| Exposure to other foreign languages | -0.058 |
| Average grade in English in primary school | -0.412** |
| Average grade in English in secondary school | -0.329** |
| Self-assessment of proficiency in English | -0.499** |
| Grade point average (GPA) | -0.038 |
| Participants' perceived importance of being proficient in English | 0.233** |

Note: * Correlation is significant at the 0.05 level; ** Correlation is significant at the 0.01 level

A positive correlation was found only for the participants' perceived importance of being proficient in English (r=0.233, p=.004). Moreover, no correlation was found between the participants' attitudes on the one hand, and their overall grade point average at the faculty (r=-0.038, p=.324), knowledge of other foreign languages (r=0.085, p=.150) and exposure to them (r=-0.058, p=.241) on the other hand. These results are somewhat unexpected, in particular because previous research revealed a relationship between learning another foreign language and more positive attitudes (Mihaljević Djigunović, 2007). This may be because other foreign languages in our population were obligatory school courses, and, as such, they did not have a positive influence on the participants' attitudes to English. So, it can be concluded that H2 was partially confirmed since a positive correlation with the attitudes was found just for learning English in a language school and for the perceived importance of being proficient in English, while the participants' length of ELL and weekly exposure to English outside studies were negatively correlated. Besides, no correlation was determined for the participants' overall academic achievement and knowledge of other foreign languages.

### 3.3. The Comparison of English Majors and English Non-Majors Subgroups

As stated above, the participants were classified into English non-majors and English majors for the purpose of H3, i.e. for the comparison of the obtained results across the two subgroups. The analysis included the Mann-Whitney test (Table 2) to compare variables related to the participants' language exposure. The first variable for which statistically significant differences were established was average length of English language learning. As expected, the exposure was in favour of the English majors group (Mdn=13), in comparison to the English non-majors[1] group (Mdn=10), U=1109, p=.000. Similar results were obtained when the two groups were compared according

[1] All the results obtained for the English non-majors group were collected in the process of PhD thesis writing by Mikulec (2016).

to: (a) the length of learning another foreign language ($Mdn_{Enon\text{-}majors}$=4; $Mdn_{Emajors}$=8), U=1667, p=.002; (b) weekly exposure to English outside their studies ($Mdn_{Enon\text{-}majors}$=3; $Mdn_{Emajors}$=3), U=1994, p=.001; and (c) contacts with native speakers ($Mdn_{Enon\text{-}majors}$=1; $Mdn_{Emajors}$=1), U=1109, p=.000. In other words, statistically significant differences in favour of the English majors were found for each variable above.

Table 2 Mann-Whitney test results showing significant differences between the English non-majors and the English majors

| Variables | U | p |
|---|---|---|
| Length of ELL in years | 1109 | 0.000 |
| Length of learning another FL in years | 1667 | 0.002 |
| Weekly exposure in hours to English outside studies | 1994 | 0.001 |
| Contacts with native speakers | 1109 | 0.000 |
| Length of daily exposure in hours – the internet | 2072 | 0.002 |
| Length of daily exposure in hours – books | 1934.5 | 0.000 |
| Length of daily exposure in hours – journals/magazines | 2056 | 0.001 |
| Self-assessment of proficiency in English | 1740 | 0.000 |
| Participants' attitudes to English | 2244 | 0.000 |
| Perceived importance of being proficient in English | 2546 | 0.015 |

The length and form of daily exposure to English were also compared between the two groups and the results indicated statistically significant differences, i.e. longer daily exposure for the English majors on three out of the five tested variables: (a) the internet ($Mdn_{Enon\text{-}majors}$=2; $Mdn_{Emajors}$=2), U=2072, p=.002; (b) books ($Mdn_{Enon\text{-}majors}$=2; $Mdn_{Emajors}$=2), U=1934.5, p=.000; and (c) journals/magazines ($Mdn_{Enon\text{-}majors}$=2; $Mdn_{Emajors}$=2), U=2056, p=.001. These results were in accordance with expectations since, as mentioned earlier, the English majors were significantly more exposed to English during their studies, and presumably outside them, due to the fact that teaching English is the core of their future profession. However, statistically significant differences were not determined for the participants' exposure to music and films in English.

Further comparison of the two subgroups was related to their self-perceived proficiency in EFL, and the difference was found to be statistically significant ($Mdn_{Enon\text{-}majors}$=4; $Mdn_{Emajors}$=4), U=1740, p=.000. It is interesting to note that – despite significantly lower exposure to English – the self-assessment within the English non-majors group was rather high (M=3.64, SD=0.820) when compared to that within the English majors group (M=4.19, SD=0.657). It appears that the former group members were less critical of their proficiency in English, which is in accordance with previous research reporting that less proficient learners tend to overestimate their language abilities, while more proficient learners are more likely to underestimate them (Heilenman, 1990; Oskarsson, 1984). That is to say, "the more experience that learners have in a domain ... the more likely they are to be aware of the limitations of their skills and knowledge" (Heilenman, 1990, p. 190), which was obviously the case here.

The participants' attitudes were also compared across the two sub-groups. Naturally, it was found that the percentage of English non-majors who expressed positive attitudes towards English was somewhat lower at 80.5%, although still relatively high in terms of the participants' language learning profile. On the other hand, the entire English majors group reported positive attitudes. The Mann-Whitney test confirmed statistically significant differences between English majors (Mdn=1) and English non-majors (Mdn=1), U=2244, p=.000. The results further indicate that English majors put much more emphasis on affective aspects and the characteristics of the language itself, i.e. they considered English to be easy to learn, claimed that for them everything sounded better in English, and found it easier to express themselves in English. Negative attitudes were expressed only by English non-majors, who mainly listed their personal and affective characteristics, along with the teaching situation and teacher(s), as the main reasons for this. In other words, they pointed out their own failure in the English language learning context, lack of language aptitude, negative experiences with language teachers in the past, and the omnipresence of English ("I don't like English because everyone can speak it, so it is boring") as the main causes of their negative attitudes to English.

The final comparison of the two subgroups was related to their perceived importance of being proficient in English. As expected, the entire English majors group considered this proficiency very important, in comparison to 91.6% of the English non-majors group. The difference was again statistically significant ($Mdn_{Enon-majors}$=1.08; $Mdn_{Emajors}$=1), U=2546, p=.015. The analysis of the participants' responses to the second open-ended question, i.e. the explanation why they found proficiency in English so important, also revealed slight differences between the two groups since the English non-majors group members mostly mentioned the benefits of knowing the language. For example, they commented: "English is a *lingua franca"*, "knowing English facilitates communication with foreigners", "it enables one to read literature, websites, etc., in English", "it is helpful in professional development", and "it facilitates travelling abroad (whether for pleasure, study or work)". As opposed to English non-majors, the English majors group expressed stronger attitudes, i.e. their replies included quantifiers like *very* and *extremely important*. Apart from the reasons above, English majors emphasised proficiency as a requirement for their future profession, expressed a strong wish to become EFL teachers, and stressed personal satisfaction in knowing EFL. Although classified as English non-majors, a very high percentage of the other group also saw their English language proficiency as important.

Based on the presented results, it may be said that H3 was mainly confirmed since statistically significant differences were found for all tested variables, except the participants' exposure to music and films in English.

## 4. Conclusion

Considering the entire sample, this study confirmed great exposure to English as a foreign language, predominantly via the internet and music, as well as rather high self-assessed EFL proficiency.

It also has to be mentioned that the results indicate very positive attitudes in the target population both to English and the importance of being proficient in English, which is in accordance with the current status of English and in light of today's globalisation processes in the world.

Furthermore, the attitudes to English were positively correlated only with learning English in language schools and the perceived importance of being proficient in English, whereas the other variables were either negatively correlated (length of ELL, weekly exposure to English outside studies), or no correlation was determined (the participants' overall academic achievement and their knowledge of other FLs and exposure to them). Positive correlations may indicate that the students who had learnt English in language schools had more positive attitudes to English, although it is hard to discover whether these attitudes reflected their additional exposure to the language or whether it was the other way round, i.e. those who had more positive attitudes were also more likely to learn English in language schools. The positive correlations regarding the perceived importance of being proficient in English are perhaps related to instrumental motivation. Namely, it seems that the positive attitudes of the participants are associated with the possibilities offered to them when they are proficient in English. However, negative and no correlations indicate that additional factors may have affected the participants' attitudes.

Finally, when the results were analysed across the two subgroups, significant differences in favour of the English majors group were found for a number of variables, such as: the length of exposure to English and the length of learning another foreign language; weekly exposure to English outside studies; contacts with native speakers; and the length of daily exposure via the internet, books and journals/magazines. The differences had been expected due to this group's significantly higher level of exposure to English in the course of studies; the differences in their motivation had also been expected since this particular group will depend on English as the language of instruction in their future profession. It may thus be safe to presume that the students who had been learning English longer would choose to study English, and they would then seek and/or have more opportunities to be engaged in communication with native speakers. Further, either because of their own choice or because of their study purposes, it was expected that they would be more exposed to English outside their studies, and would read more in English, whether books, journals, or materials on the internet. However, differences were not found for the participants' daily exposure via music and films. This indicates that, irrespective of their future profession, a majority of the participants are exposed to English in their leisure time activities through watching films and listening to music.

Further studies are recommended that could employ random sampling and additional measures for the following crucial variables: attitudes, FL proficiency and academic achievement.

## 5. References


Cindrić, I., Andraka, M., & Bilić-Štefan, M. (2014). The Student-Teachers' Perspectives on the Significance and Acquisition of Teaching Competences. *Croatian Journal of Education*, 16(1), 11-42

Cindrić, I., & Narančić Kovač, S. (2005). English and Other Foreign Languages in Croatia – A Needs Analysis. *Strani jezici*, 34(3), 189-204.

*Common European Framework of Reference for Languages: Learning, teaching, assessment*. (2001). Cambridge & Strasbourg: Cambridge University Press and Council of Europe.

Crystal, D. (2012). A global language. In P. Seargeant, & J. Swann (Eds.), *English in the world: history, diversity, change* (pp. 151-177). London: Routledge.

Demel, M. (1990). Mogući kriteriji izbora tekstova za rad sa studentima matematike. *Strani jezici*, 19 (4), 227-230.
Dőrneyi, Z. (2005). *The Psychology of the Language Learner*. *Individual Differences in Second Language Acquisition*. New Jersey & London: Lawrence Erlbaum Associates.
Dőrneyi, Z., Csizér, K., & Németh, N. (2006). *Motivation, langauge attitudes and globalisation. A Hungarian perspective*. Clevedon, England: Multilingual Matters.
Heilenman, L. K. (1990). Self-assessment of second language ability: The role of response effects. *Language Testing*, 7, 174-201.
Henter, R. (2014). Affective Factors Involved in Learning a Foreign Language. *Procedia – Social and Behavioral Sciences*, 127, 373-378.. Retrieved from http://www.sciencedirect.com/science/article/pii/S1877042814023659.
Hyland, K. (2009). *Academic Discourse*. London & New York: Continuum International Publishing Group.
Jelovčić, I. (2010). English for specific purposes – student attitude analysis. *Metodika, 20*(1), 124-136.
Josipović Smojver, V. (2007). Listening Comprehension and Croatian Learners of English as a Foreign Language. *Metodika*, 8(14), 137-147
Leclercq, P., & Edmonds, A. (2014). How to Assess L2 Proficiency? An Overview of Proficiency Assessment Research. In P. Leclercq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 Proficiency. Perspectives from SLA* (pp. 3-23). Bristol, Buffalo & Toronto: Multilingual Matters.
MacIntyre, P. D., Gregersen, T., & Mercer, S. (Eds.). (2016). *Positive Psychology in SLA*. Bristol, Buffalo & Toronto: Multilingual Matters.
Martinović, A., & Poljaković, I. (2010). Attitudes toward ESP among university students. *FLUMINENSIA*, *22*(2), 145-161.
Mihaljević Djigunović, J. (2002). *Strah od stranoga jezika*. Zagreb: Naklada Ljevak.
Mihaljević Djigunović, J. (2007). Afektivni profil, aspiracije i zadovoljstvo nastavom engleskoga jezika kod hrvatskih učenika. *Metodika*, 8(1), 104-114.
Mihaljević Djigunović, J., Nikolov, M., & Ottó, I. (2008). A comparative study of Croatian and Hungarian EFL students. *Language Teaching Research*, 12, 3, 433-452.
Mikulec, A. (2016). *Strategije čitanja akademskih tekstova na hrvatskome i engleskome jeziku*. (Unpublished doctoral dissertation). Faculty of Humanities and Social Sciences, University of Zagreb, Zagreb.
Ni, H. (2012). The Effects of Affective Factors in SLA and Pedagogical Implications. *Theory and Practice in Language Studies*, 2, 7, 1508-1513.
Oscarsson, M. (1984). *Self-Assessment of Foreign Language Skills: A Survey of Research and Development Work*. Strasbourg: Council of Europe.
Oxford, R. (Ed.) (1999). *Language Learning Motivation: Pathways to the New Century*. Honolulu: Second Language Teaching and Curriculum Center (2nd edition).
Robinson, P. (Ed.). (2002). *Individual Differences and Instructed Language Learning*. Amsterdam: John Benjamins Publishing Company.
Rubio, F. (Ed.). (2007). *Self-Esteem and Foreign Language Learning*. Newcastle: Cambridge Scholars Publishing.
Singleton, D. (2014). How Do Attitude and Motivation Help in Learning a Second Language. In V. Cook, & D. Singleton (Eds.), *Key Topics in Second Language Acquisition* (pp. 90-107). Bristol, Buffalo & Toronto: Multilingual Matters.
Sternberg, R. J. (2005). *Kognitivna psihologija*. Jastrebarsko: Naklada Slap.
Suwannatho, N., & Thepsiri, K. (2015, January). The Correlation between Low Proficiency Undergraduate Students' Attitudes and Motivation. In L. Barrat

(Ed.), *Proceedings of the 35th Thailand TESOL International Conference "English Language Education in Asia: Reflections and Directions"* (pp. 81-102) . Retrieved from http://www.academia.edu/15222950/Teachers_behaviour_in_integrating_Digital_Storytelling_DST_Software_into_their_teaching_Its_impact_and_implications

Zafar, Sh., & Meenakshi, K. (2012). Individual Differences and Second Language Acquisition: A Review. *Journal of Language Teaching and Research*, 3, 4, 639-646.

Zhao, L. (2015). The Influence of Learner Motivation and Attitudes on Second Language Teaching. *Theory and Practice in Language Studies*, 5, 11, 2333-2339.

## 6. Appendix: Questionnaire

**Podatci o studentici/studentu**

1. Dob_____ Spol M Ž
2. Napišite godinu i smjer studija _____
3. Koliko dugo učite engleski jezik?_____
4. Jeste li učili engleski i izvan škole/fakulteta? Koliko dugo i kada?_____
   _____
   _____
5. Koju ste srednju školu završili? _____
   _____
6. Koju ste prosječnu ocjenu imali iz engleskog jezika u **osnovnoj** školi? _____
7. Koju ste prosječnu ocjenu imali iz engleskog jezika u **srednjoj** školi? _____
8. Volite li engleski jezik? Zašto? _____
   _____
   _____
9. Jeste li učili još neki strani jezik osim engleskog? Koji/Koje? _____
   _____
   _____
10. Koliko dugo? 1. jezik_____ 2. jezik_____
11. Koju ste prosječnu ocjenu imali iz drugog (i trećeg, četvrtog) stranog jezika u **osnovnoj** školi? (navedite jezik i ocjenu) _____
    _____
    _____
12. Koju ste prosječnu ocjenu imali iz drugog (i trećeg, četvrtog) stranog jezika u **srednjoj** školi? (navedite jezik i ocjenu) _____
    _____
    _____
13. Koliko ste **tjedno** izloženi **drugomu** stranom jeziku?
    a. nemam kontakt s tim jezikom
    b. do 5 sati tjedno
    c. do 10 sati tjedno
    d. 15 i više sati tjedno

14. Koliko ste **tjedno** izloženi **trećemu** stranom jeziku?
    a. nemam kontakt s tim jezikom
    b. do 5 sati tjedno
    c. do 10 sati tjedno
    d. 15 i više sati tjedno
15. Kojom biste ocjenom ocijenili svoje znanje **engleskoga** jezika? _____
16. Koji Vam je ukupan prosjek ocjena na studiju?_____
17. Koliko ste sati TJEDNO izloženi **engleskom** jeziku izvan studija?
    a. nisam uopće
    b. do 5 sati tjedno
    c. do 10 sati tjedno
    d. više od 15 sati tjedno. Koliko?_____
18. Jeste li kada imali ili sada imate kontakt s govornicima engleskog jezika? _____
    a. U kojem obliku (razgovor, dopisivanje…)? _____
    _____
    _____
    b. Zašto? Koliko često?_____
    _____
    _____
19. Oblik DNEVNE izloženosti engleskom jeziku izvan studija
    a. Internet
        i. do 3 sata dnevno j. 3-9 sati dnevno k. više od 9 sati dnevno
    b. glazba
        i. do 3 sata dnevno j. 3-9 sati dnevno k. više od 9 sati dnevno
    c. filmovi
        i. do 3 sata dnevno j. 3-9 sati dnevno k. više od 9 sati dnevno
    d. knjige
        i. do 3 sata dnevno j. 3-9 sati dnevno k. više od 9 sati dnevno
    e. časopisi
        i. do 3 sata dnevno j. 3-9 sati dnevno k. više od 9 sati dnevno
20. Koliko Vam je važno da budete uspješni u engleskom jeziku? Zašto? _____
    _____
    _____
    _____

# Frequency and Type of Gesture in the Blind and the Sighted in L1 and L2

**Iva Gugo**
*Faculty of Humanities and Social Sciences, Zagreb, Croatia*
*ivagugo@hotmail.com*

**Renata Geld**
*Faculty of Humanities and Social Sciences, Zagreb, Croatia*
*geldrenata@gmail.com*

## 1. Introduction

In interpersonal communication, speech is often accompanied by gestures, which help convey the intended message in visual terms. The sighted can see gestures which they produce, but the blind cannot. Such a discrepancy raises the question whether gesture plays the same role for the blind as they do for the sighted. Furthermore, we may ask ourselves whether the blind and the sighted gesticulate differently in different languages. In the case of L1 vs. L2, the issue becomes even more complex because L2 speakers use various communication strategies in order to compensate for possible lack of proficiency in a productive skill such as speaking. They tend to paraphrase, describe, use superordinate or subordinate concepts, borrow from other languages, create new words, and so on. In addition to these language-based strategies, they often resort to all sorts of nonverbal strategies that are likely to help them convey the intended message. It is not unusual for L2 speakers to point to themselves or objects and people around them, make facial expressions, change their body posture, and use gestures.

The aim of the study presented in this paper was to investigate frequency and types of gestures in blind and sighted speakers of Croatian as L1 and English as L2. We begin with a theoretical overview of gestures: we explain what gestures are and how they are classified. The outline is followed by a summary of previous research on gestures which served as a basis for our hypotheses. Next, we outline our aims and describe the methodology of the study. Finally, we present and discuss the results.

## 2. Gestures: definition and classification

McNeill (2006) defines the term 'gesture' as "a multiplicity of communicative movements, primarily but not always of the hands and arms" (p. 58). Speakers use their hands and arms to convey a certain message, e.g. where somebody went, or to illustrate what something looks like. It is not just concrete actions and items that are illustrated with gestures. Speakers even convey abstract concepts in gestures. For example, they show that they love somebody by shaping a heart with their hands. As Pavelin-Lešić (2010) notes, speakers materialize abstract concepts and handle them as concrete objects by using gestures. People are not aware of the fact that they gesticulate when they talk. Because speakers focus primarily on speech and the articulation of

their thoughts, gestures emerge subconsciously. As a consequence, gestures may not only transmit information that the speaker is articulating; they may also transmit information not conveyed in speech and may even give the listener insight into the speaker's mental state (Goldin-Meadow, 2000).

While people most often gesticulate in the presence of other individuals, it has been found that gestures emerge even when the speaker is alone, e.g. when talking on the telephone. This finding proves that a person need not see their listener in order to gesticulate. In other words, people do not gesticulate for the listener's sake, but rather their own sake. This largely explains why blind people gesticulate, even though they cannot see their listener and have never seen others produce a gesture (Iverson & Goldin-Meadow, 2001).

Although gestures co-occur with speech, they differ from speech in one important feature: speech is codified, meaning that words are combined into larger units (clauses and sentences) following rules from mental grammar and lexicon. If people wish to be understood, they must adhere to this set of rules. Unlike speech, however, gesture "is idiosyncratic and constructed at the moment of speaking – it does not belong to a conventional code" (Goldin-Meadow, 2000, p. 237). There are only a few conventionalized gestures, and no established rules for gesturing exist. Nevertheless, gestures which co-occur with speech can be classified according to their function in relation to speech. McNeill (1992) defines four primary types of co-speech gestures: iconic gestures, metaphoric gestures, deictic gestures and beats. An iconic gesture is a gesture which "bears a close formal relationship to the semantic content of speech" (McNeill, 1992, p. 78). For example, making a throwing movement with the hand to indicate that somebody throws something would classify as an iconic gesture. While iconic gestures are used to represent concrete entities, metaphoric gestures "present an image of an abstract concept" (McNeill, 1992, p. 80). For example, a speaker says "Next, we went…" and makes a swiping motion to the side while uttering the word 'next'. Of the four types of gestures, metaphoric gestures are the ones that are dependent on context the most (Pavelin-Lešić, 2010). What this suggests is that their meaning cannot be understood in isolation, but only in connection with speech and the overall context that it accompanies. Deictic gestures are gestures of pointing. While talking, speakers use these gestures to point at people or things. The referents of deictic gestures may, but need not be present for the speaker to point at them. Sometimes, a speaker may refer with a deictic gesture to something that was previously present in the room. In such a case, a speaker relies on the fact that the listener will remember who or what was in the place they are pointing at (Yule, 2006).The last primary co-speech gesture type is beats, "movements that do not present a discernible meaning, and they can be recognized positively in terms of their prototypical movement characteristics" (McNeill, 1992, p. 80). Beats serve to highlight words or phrases which are accentuated in speech. Speakers realize them as "small, low energy, rapid flicks of the finger or hand" and produce them repetitively when speaking (McNeill, 1992, p. 80).

In addition to these four primary types of co-speech gestures, there are also several secondary gesture types: speech-framed gestures, emblems, pantomime and signs. We have included only speech-framed gestures into our analysis. These are gestures which, unlike the previously mentioned gesture types, do not occur simultaneously with speech, although they are preceded and followed by it. Instead, they occupy a slot in the sentence that the speaker does not verbalize. For example, a speaker says, "Sylvester went…", but instead of saying how or where Sylvester went, the speaker shows it in gestures (McNeill, 2006).

## 3. Previous research

The aim of this study was to investigate gesture in blind and sighted L2 learners. Unfortunately, few research studies on gestures in L2 speakers have been conducted so far. Nevertheless, there are some studies on gesture in foreign-language environment in general which help shed light on gestures in L2 speakers. For example, Guilberg (2010) reports the important role of gesture in L2 learners and teachers. Learners tend to learn more if their teacher gesticulates. Beats in particular are very useful; since they follow the rhythm of the speech, they help learners internalize prosodic and phonological properties of the target language. Gregersen, Olivares-Cuhat and Storm (2009) also emphasize the importance of gesture in the learning of an L2. In their study conducted on L2 learners, they found that gestures help overcome lexical gaps if they are employed in addition to or instead of speech in cases when a learner lacks the appropriate vocabulary for articulating their thoughts. Their study also showed that speakers gesticulate more often in their L1than in their L2. Further, the authors found the frequency of gestures in L2 to be reciprocal to competence: the most proficient learners gesticulated the most, whereas less proficient learners gesticulated less.

The two studies mentioned above involved sighted native speakers of English. There are only a few studies on gestures of blind native speakers of English and, to the authors' knowledge, none on gestures of blind L2 learners of English. Nevertheless, the existing studies and their findings will serve as important points of reference in this work.

One of the biggest myths about the blind is that they do not gesticulate. Iverson, Tencer, Lany and Goldin-Meadow (2000) debunked this myth when they found that blind infants produced gestures during the language acquisition process, albeit less frequently than sighted infants. The distribution of gesture types, however, was similar in both groups in their study. In another study, Iverson and Goldin-Meadow (1997) investigated gesture in older blind and sighted children in three tasks, one of which was a narrative task. All participants produced very few gestures, but the blind produced the fewest gestures. What is more, they produced exclusively iconic gestures. Although the blind produced only one gesture type, the authors were able to prove that the blind do gesticulate and that vision is not a prerequisite for producing gestures. In a later study, Iverson and Goldin-Meadow (2001) found that blind children gesticulated even when the addressee was blind. These findings suggest that the blind gesticulate because producing gestures is an intrinsic part of the speaking process, and it does not happen for the sake of others. Moreover, the authors observed that the blind frequently gesticulated while thinking out loud. They thus concluded that gestures also function as a channel through which the blind express thoughts which are difficult to articulate.

## 4. Study

### 4.1. Aims and research questions

The main aim of this study was to compare blind and sighted L2 learners' frequency of gesturing and the gesture types they produce. More specifically, we wished to determine the following:

a) Whether blind learners gesticulate more or less than sighted learners in their L1, i.e. Croatian. As a study by Iverson and Goldin-Meadow (1997) showed, blind native speakers of English tend to produce fewer gestures when narrating a story in their mother tongue than sighted speakers. Analogously, it was expected that Croatian speakers who are blind would produce fewer gestures when narrating a story in their first language than speakers who are sighted;
b) Whether blind learners gesticulate more or less than sighted learners in their L2, i.e. English. It was assumed that the findings pertaining to gesticulation in English would mirror those pertaining to gesticulation in Croatian, i.e. that blind learners would gesticulate less than sighted learners in L2, as well;
c) Whether there was a relationship between proficiency in L2 and the frequency of gesturing. Following Gregersen et al.'s (2009) findings, we expected that more proficient learners would gesticulate more in L2 than the less proficient learners;
d) Whether Croatian blind L2 learners gesticulate more in L1 or in L2. We assumed that Gregersen et al.'s (2009) finding about sighted speakers gesticulating less in a foreign language could also be applied to blind learners, meaning that the blind would gesticulate less in L2.

Finally, this study aimed to determine the types of gestures blind speakers produce in L1 and L2 as well as the purpose of their production. Although the blind produced only one gesture type during the narrative task in Iverson and Goldin-Meadow's (1997) study, the fact that the blind, much like the sighted, were found to produce a variety of gesture types during the speech acquisition process (Iverson & Goldin-Meadow, 2000) led us to presume that the gesture types employed by the blind in this study might be similar to those produced by the sighted.

### 4.2. Participants and methods

Fifteen L2 learners participated in the study. Five of them were blind learners of English, aged 16-18, attending a vocational program in a specialized school for the blind and visually impaired. According to their school achievement, their teacher's evaluation and their self-assessment of their competence, they were classified as intermediate learners. The rest of the learners were sighted and they formed two subsamples. Subsample #1 consisted of sighted grammar-school L2learners classified as advanced learners, whereas subsample #2 consisted of intermediate sighted vocational-school L2 learners. The students in the subsamples were 16 and 17 years old.

Prior to the study, all the participants gave their consent to participate in the study. They agreed to be video-recorded and allowed the use of the visual material in the publication of the results. The school principals also permitted the researchers to conduct the study on the premises of their schools. The study was conducted in two phases. In both phases, the participants received a story and were asked to retell it in front of a camera. The story was read to the blind participants, while the sighted participants read the story on their own. In the first phase, they were given a story in Croatian (their L1), while in the second phase, they received the same story in English (their L2).The story used was *The Killer in the Backseat*, chosen because it was expected that its dynamic and exciting character would elicit gestures from the participants. The retellings in L1 in all groups lasted between 90 and 120 seconds.

In L2, both subsamples needed 2 minutes on average to retell the story, whereas the average length of the retold story in L2 in the blind group was 3 minutes. The blind group's longer narratives in L2 were caused by long pauses in their retellings, during which they silently constructed their utterances.

After all the data had been collected, the participants' gestures were counted. In the analysis, Iverson and Goldin-Meadow's (1997) definition of a gesture served as a basis for recognition. Therefore, a hand movement counted as a gesture if it preceded, followed or co-occurred with speech and if it had a clear beginning and ending. The gestures were classified into types, following McNeill's (1992, 2006) classification of gestures. For each group, we also calculated the mean numbers of gestures as well as the mean number of each produced gesture type.

## 5. Results and Discussion

### 5.1. Frequency of gesturing in the blind and the sighted in L1

A comparison of the mean numbers of gestures (Figure 1) shows that the blind produced fewer gestures than both subsamples when narrating the story in Croatian. Therefore, our assumption that the blind would gesticulate less than the sighted in L1 was confirmed.

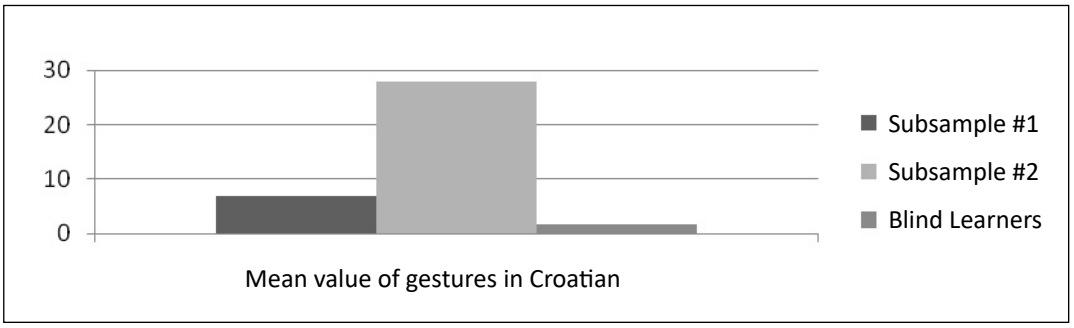


Figure 1.Mean numbers of gestures in both subsamples and the group of blind learners in L1

However, what is also noticeable from the table above is that subsample #2 produced a higher number of gestures than the other two groups. We believe that the discrepancy between the two subsamples may be explained by specific educational circumstances that produce an interplay between affective and cognitive factors. Grammar-school learners (subsample #1) are more skillful speakers; they often have discussions on different topics in their L1, and expressing their ideas and opinions has become a rather effortless task.Vocational-school learners (subsample #2) are not as eloquent as their grammar-school peers. Their discussion classes are less frequent. For that reason, they are often nervous when speaking in class and need to gesticulate more in order to convey a message. On the other hand, the fact that the blind produced so few gestures cannot be attributed to their level of proficiency. As already mentioned, our blind participants were at the lower-level proficiency; they were language learners with relatively weak narrative skills. The only reasonable conclusion is that the blind tend to rely mostly on speech when communicating in their mother tongue.

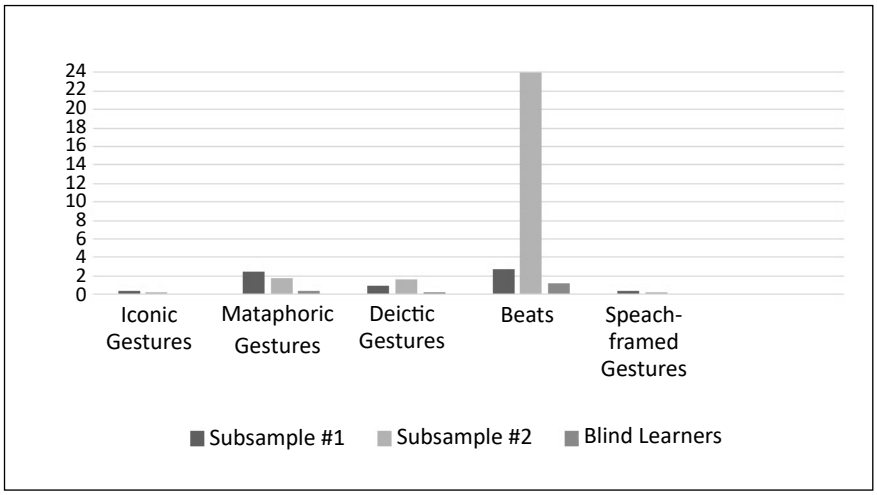


Figure 2. Mean numbers of occurrences of gesture types in both subsamples and the group of blind learners in L1

A further comparison of the mean numbers of occurrences of gestures types cross all groups (Figure 2) shows that all gesture types except beats were produced infrequently in L1. Beats were produced by subsample #2 much more often than any other gesture type. The forms of beats differed from one participant to another; they were formed by raising one hand and bringing it back to the initial position, moving one hand to the side and bringing it back to the resting position, opening and closing palms, etc. (Figure 3).

Figure 3. "ustrašila se" (*she got scared*) –Participant opens and closes his palms, an example of a beat gesture produced in subsample #2.

As Gregersen et al. (2009) suggest, beats help learners organize their thoughts into verbal language and articulate them. This is likely to be the case with subsample #2. In personal communication, several participants from this subsample admitted that they were not sure if they were up to the task as they did not think that they had good speaking skills in either L1 or L2. As already stressed, their school is vocational and its curriculum does not include debates or longer discussions. We may conclude that the high occurrence of beats is probably due to their role in helping our participants narrate the story and complete the task.

### 5.2. Frequency of gesturing in the blind and the sighted in L2

Figure 4 shows that blind L2 learners produced more gestures in L2 than subsample#1, but fewer than subsample#2. The mean number of gestures produced by the blind in English is numerically halfway between the mean values of both subsamples.Whether the blind gesticulate more or less in L2 than their sighted peers is yet to be determined through future research.

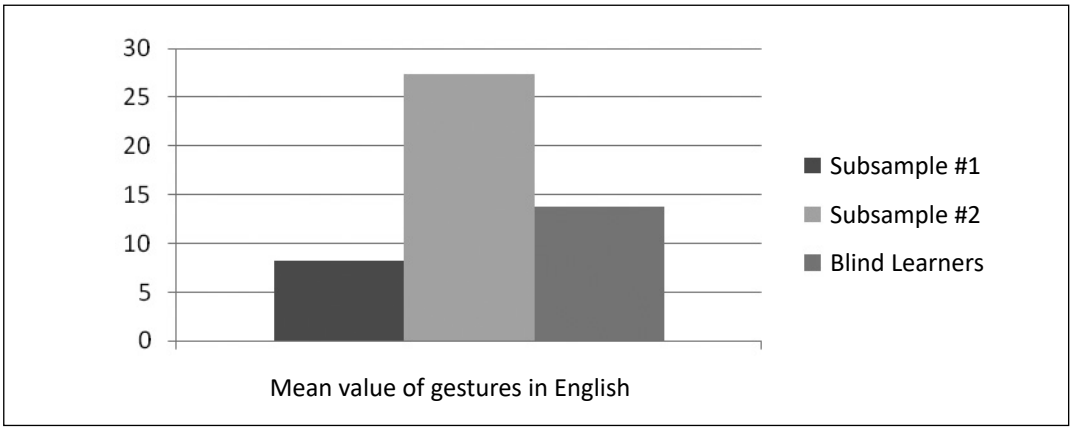


Figure 4.Mean numbers of gestures in both subsamples and the group of blind learners in L2

Following Gregersen et al.'s (2009) findings, it was expected that all groups would gesticulate less in L2. However, the mean number of gestures in L2 in the group of blind participants was higher than their mean number of gestures in L1, whereas the mean numbers of gestures in L1 and L2 for both subsamples of sighted learners were almost equal.

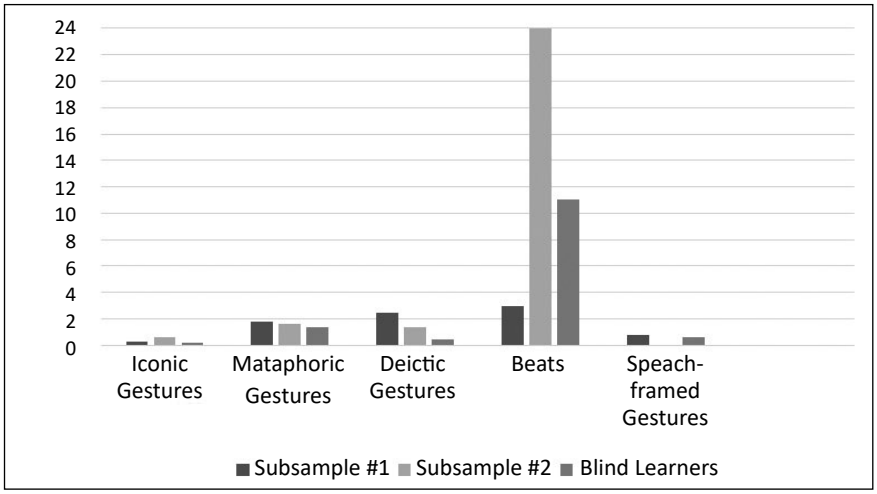


Figure 5. Mean numbers of occurrences of gesture types in both subsamples and the group of blind learners in L2

Subsample #1 produced the fewest gestures of all three groups. It is reasonable to attribute this result to their higher proficiency in English. As grammar-school learners, they have more L2 classes per week than the other two groups, and, as already mentioned, a lot of their classes, irrespective of the school subject, include a lot

of discussion. They are more advanced in the use of vocabulary and grammar and they are less likely to miss "the right words". Therefore, they do not need to compensate by using gestures.

The comparison of the mean numbers of occurrences of gestures types in all groups (Figure 5) shows that the previously established high number of gestures in subsample #2 is mostly due to a high occurrence of beats. As was the case in L1, this finding in L2 may be again explained by the educational circumstances - fewer Croatian and English language classes and, consequently, less advanced language proficiency and fewer chances to practice speaking skills.The use of gestures compensates for their poorer speaking skills and helps them overcome the anxiety they are likely to experience when speaking in L2.

When compared to both subsamples, the group of blind L2 learners demonstrated more similarities to subsample #2. In addition to the fact that both groups consisted of intermediate-level L2 learners, a very significant commonality between them is the tendency towards beats. The possible explanation for why blind learners employed beats so frequently, and why they produced much more gestures in L2 than in L1, is offered in later sections.

### 5.3. Relationship between proficiency in L2 and the frequency of gesturing

Although it was expected that the more proficient participants in this study (subsample#1) would gesticulate more frequently in L2, our findings show that less proficient participants (subsample#2) gesticulated three times more often in English than their more proficient peers. A tentative conclusion which can be drawn from this data (see Figure 4) is that proficiency and the number of gestures produced in L2 seem to be inversely proportional to one another: the more proficient learners are in L2, the fewer gestures they produce.

This result contradicts the finding from Gregersen et al.'s (2009) study. However, when the participants' educational context and mental disposition are taken into account, it can be explained why more proficient L2 learners would employ gestures less frequently than their less proficient peers. As already suggested, they are simply more used to situations in which they are asked to talk for a length of time, and they either learn how to control their hand movements or they simply do not feel the need to gesticulate. In addition, as proficient speakers, they rarely experience communication breakdowns, so that they need not employ gestures in order to compensate for them.

Less proficient sighted L2 learners in our study (subsample#2) go to a vocational school with a lot of specialist subjects where languages and social sciences are not priority. They are not frequently encouraged to debate, defend their views or speak extensively in any language class. So, when faced with a task to orally produce a story in L2, they experience anxiety and communication breakdowns. Just like in L1, they resort to gestures to compensate for their lack of knowledge.

Our blind participants, who were also classified as intermediate-level L2 learners, were found to gesticulate more frequently in L2. What reasons may lie behind this result is explained in the sections that follow.

### 5.4. Frequency of gesturing in the blind in L1 and L2

As seen in Figure 6, the mean number of gestures produced in English (L2) by the blind was much higher than in Croatian (L1). Therefore, the hypothesis that the blind would gesticulate less in L2 was not confirmed.

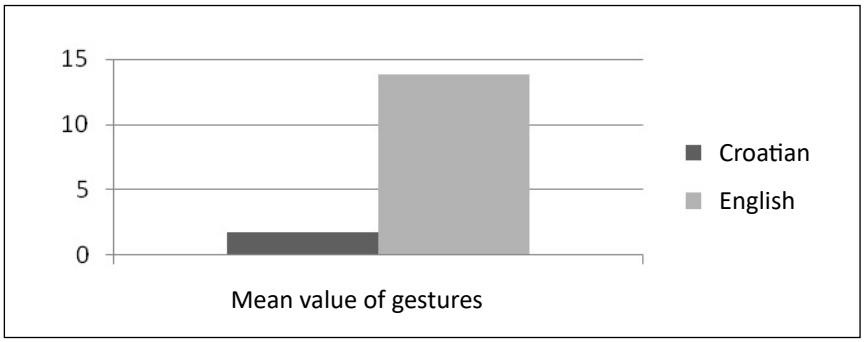


Figure 6. Mean numbers of gestures in blind L2 learners in English and Croatian

As a group, the blind participants in this study were classified to be of intermediate proficiency. While narrating in English, they made a number of grammatical mistakes. They were aware of the fact that their speaking skills were not good and they expressed their concern about it openly. Gestures helped them organize their thoughts, articulate them and explain them in visual terms. Once the blind successfully articulated their thoughts with the assistance of gestures, they were more motivated to continue and complete the task at hand. The conclusion is that gestures seemed to compensate for our blind L2 learners' lacking knowledge, motivate them when speaking and clarify the information they wish to convey to the listener.

### 5.5. Gesture types in blind participants' narratives

Mean numbers of occurrences for each gesture type produced by blind L2 learners in Croatian (L1) are distinctly low (Figure 7), which is in line with the previous finding that the blind do not gesticulate much in L1. Some gesture types (iconic gestures and speech-framed gestures) were not produced at all by the blind participants in Croatian.

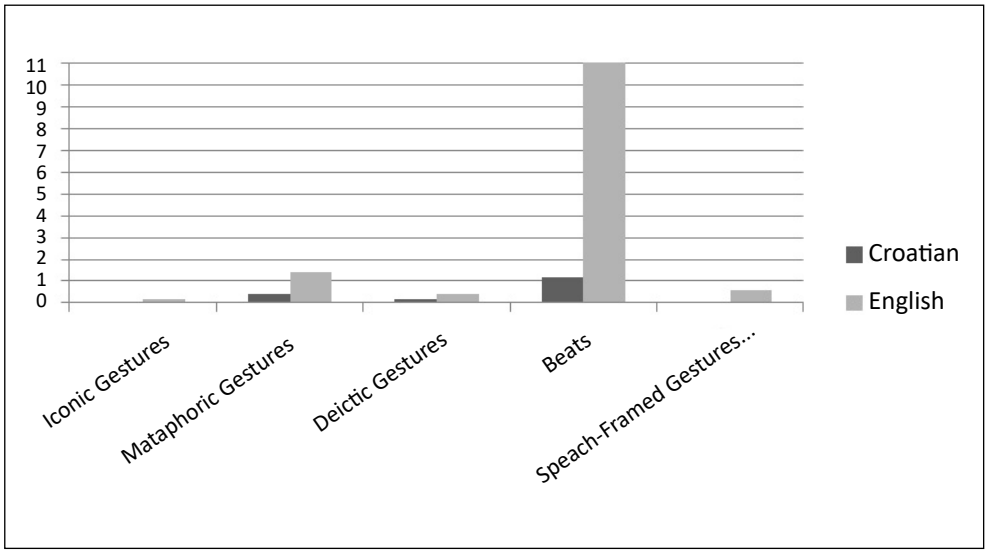


Figure 7.Mean numbers of occurrences of gesture types produced by blind L2 learners in English and Croatian

In the second phase of the study, while narrating in English (L2), blind L2 learners produced all gestures types. Metaphoric and deictic gestures were produced

almost equally in both languages, while the number of beats was significantly more frequent in L2. An example of a blind student producing a beat can be seen in Figure 8.

Figure 8. "in *the* *him *office*" - Learner repeatedly taps his right knee as he speaks the italicized words

For the blind, beats were the most frequently employed gesture type in both L1 and especially L2. When observing the circumstances in which beats occurred, it becomes evident that they had an important role in the blind learners' communication. In many cases, beats tended to occur after long pauses, during which the learners seemed to think about how they would articulate their thoughts. Once they thought of a way to express them, they accompanied their spoken words with a beat gesture. In some participants, this pattern was highly repetitive. The conclusion which can be drawn from this observation is that beats assist the blind in putting their thoughts into utterances, i.e. in organizing them into syntactic units (phrases, clauses and sentences) and articulating them. As a matter of fact, they are likely to be the main facilitator in the process of completing the task, in spite of possible problems in oral communication.

An important finding in this study is that the blind produce metaphoric and deictic gestures, even though the opposite was reported in Iverson and Goldin-Meadow's (1997) study of blind children narrating a story. Figure 9 shows a blind L2 learner producing a metaphoric gesture by sweeping with his hand through the air. Through this gesture, he signalized the transition to the next stage of the story.

Figure 9. "she *go next" - Learner motions with the hand from one side to the other to indicate the transition to the next stage in the narrative

In spite of being visually impaired, the blind also successfully produced deictic gestures. They proved to be capable of organizing objects and people from the story into space and pointing at them as if they were actually in the room. Figure 10 shows an interesting example of a deictic gesture made by a blind L2 learner. Whenever he referenced the main character in the story, he pointed at himself, showing that he identified with her. Even though his oral narrative was produced from a third-person point of view, his gestures indicated a first-person point of view.

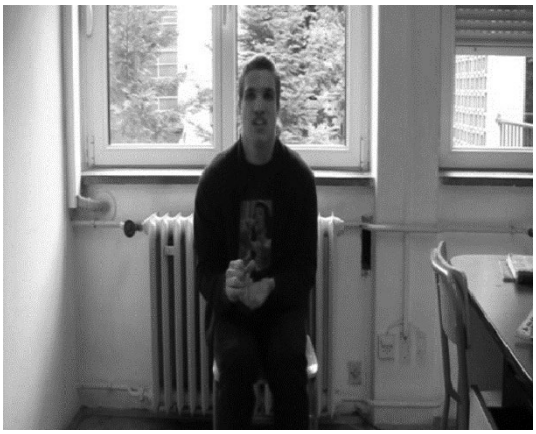

Figure 10 "to her" - Learner points to himself when referencing the main character of the narrative (deictic gesture)

## 6. Conclusion

The aim of this study was to examine blind and sighted L2 learners' gestures. The findings suggest that blind L2 learners gesticulate less in L1 (Croatian) than their sighted peers.While inconclusive findings make it difficult to state whether the blind gesticulate more or less than the sighted in their L2 (English), it is significant that, as a group, they employ gestures more frequently in L2. Also, our findings show that the blind produce all gesture types.

A further conclusion from this studyis that proficiency in L2 and the frequency of gesturing appear to be inversely proportional. When narrating a story in English, less proficient L2 learners in this study gesticulated more frequently than more proficient learners.Both blind and sighted less proficient L2 learners seemed to employ gestures as assistance in speech production; with the help of gestures, they successfully overcame potential communication breakdowns that might have occurred due to their lack of proficiency. Because of this, their motivation to continue with the task increased. Of all gesture types, beats in particular were greatly employed by the less proficient learners when they were trying to communicate a message across. They assisted these learners in organizing their thoughts into speech and articulating them.

This study was qualitative in nature and it was conducted on a small sample of blind L2 learners whose proficiency was at an intermediate level. In the future, it would be desirable to conduct a similar study with a larger sample of blind learners of varied proficiency levels so as to confirm or refute the tendencies observed in this small-scale study, especially those results pertaining to the relationship between proficiency and the frequency of gesturing.

## 7. References


Goldin-Meadow, S. (2000). Beyond Words: The Importance of Gestures to Researchers and Learners. *Child Development*, *71*(1), 231-239.

Gregersen, T., Olivares-Cuhat, G., & Storm, J. (2009). An Examination of L1 and L2 Gesture Use: What Role Does Proficiency Play? .*The Modern Language Journal*, *93*(2), 195-208.

Guilberg, M. (2010).Methodological reflections on gesture analysis in second language acquisition and bilingualism research. *Second Language Research*, *26*(1), 75-102.

Iverson, J. M., & Goldin-Meadow, S. (2001). The resilience of gesture in talk: gesture in blind speakers and listeners. *Developmental Science*, *4*(4), 416-422.

Iverson, J. M., & Goldin-Meadow, S. (1997). What's Communication Got to Do With It? Gesture in Children Blind from Birth. *Developmental Psychology*, *33*(3), 453-467.

Iverson, J. M., Tencer, H. L., Lany, J., & Goldin-Meadow, S. (2000).The Relation between Gesture and Speech in Congenitally Blind and Sighted Language-Leaners. *Journal of Nonverbal Behavior*, *24*(2), 105-130.

McNeill D. (1992).*Hand and Mind: What Gestures Reveal about Thought.* Chicago: University of Chicago Press.

McNeill, D. (2006). Gesture: A Psycholinguistic Approach. In K. Brown (Ed.), *The Encyclopedia of Language and Linguistics* (pp. 58-66). Boston: Elsevier.

Pavelin Lešić, B. (2010).The Metaphorization of Practical Action and Everyday Life Experience in the Words, Emblems, and Coverbal Gestures of Spoken Language. *SRAZ*, *LV*, 223-233.

Yule, G. (2006). *The Study of Language* (3rded.). Cambridge: Cambridge University Press.

# 'Dyslexia-friendly' Approaches in the Teaching Practice of Croatian English as a Foreign Language (EFL) Teachers

**Agnieszka Kałdonek-Crnjaković**
*Independent researcher, London, UK*
*agnieszka.kaldonek@gmail.com*

**Zrinka Fišer**
*University of Zagreb*
*zrinkafiser@gmail.com*

## 1. Introduction

Developmental dyslexia is a specific learning difficulty (SpLD) that mainly affects learning to read and write (Croatian Dyslexia Association, 2016). Consequently, this difficulty can affect learning a foreign language (FL) in many ways (Crombie, 2000). It has been established that traditional FL teaching methods are not effective for dyslexic learners and that they learn best when the teaching is based on a cumulative, explicit, and structured approach with multisensory techniques and raising metalinguistic awareness (Ganschow, Sparks & Javorsky, 1998; Schneider & Crombie, 2003). Yet, the recent research has shown that Croatian English as a foreign language (EFL) teachers' knowledge about effective methods in teaching dyslexic learners is limited (Fišer & Dumančić, 2014; Kałdonek-Crnjaković & Fišer, 2016). One of the reasons of such limited knowledge is that the core EFL teacher training curriculum in Croatia does not entail a module which discusses dyslexia and other SpLDs in the context of FL learning and teaching.

## 2. Dyslexia and foreign language learning and teaching

The language aptitude in dyslexic learners is believed to be weaker and therefore they usually find the phonology, orthography, syntax and structural aspects of a FL difficult to acquire (Crombie, 2000). Consequently, dyslexic learners make very slow progress in learning a FL and it is rather unlikely that they can achieve the highest level of proficiency (Nijakowska, 2008). It was also established that methods based on the 'natural' or 'communicative' approach, as proposed by Krashen (1987), are ineffective in teaching a FL to dyslexic learners and alternative methods based on a multisensory, explicit, structured, and cumulative approach alongside with raising metalinguistic awareness should be adopted (Ganschow, Sparks & Javorsky, 1998; Schneider & Crombie, 2003). A number of studies have suggested a beneficial effect of such approaches in teaching English (e.g., Kahn-Horwitz et al., 2005, 2006; Kałdonek-Crnjaković, 2013; Nijakowska, 2008), German (e.g., Schneider, 1999), as well as French and Spanish (e.g., Sparks & Ganschow, 1993; Sparks, Ganschow, Artzer, & Patton, 1997; Sparks et al., 1991, 1992) as a foreign language. Each of the above-mentioned approaches will be further discussed.

## 2.1. Multisensory teaching

The principle of multisensory teaching is the simultaneous employment of visual, auditory, kinaesthetic and tactile modes. An example of multisensory teaching in a FL classroom is when the teacher uses flash cards to teach new vocabulary. The teacher shows the card with the word, says the word aloud, and practises the spelling by tracing the letters of the word on the card. The students then repeat the word aloud and practise its spelling by, for example, writing the word with their finger on the table, in the sand tray or in the air.

The kinaesthetic and tactile modes are very important for dyslexic learners, especially for younger learners (Kałdonek-Crnjaković, 2013; Schneider & Kulmhofer, 2016), as they cannot fully process and retain what is being taught when the material is presented just orally or visually. Examples of teaching and learning using the kinaesthetic and tactile modes include, for example, touching lips when producing specific sounds, using body motion in teaching vocabulary or interactive games in which touch and movement are present (Schneider & Crombie, 2003).

## 2.2. Explicit teaching

On the assumption that dyslexic learners are unable to acquire information by mere exposure to it, Sparks, Ganschow and Patton (2008) suggest explicit teaching of different aspects of a foreign language, including letter-sound relation, grammatical structures, semantic and socio-pragmatic contexts. The explicit instruction involves a contrastive analysis of language items of the target language and the native language of the learner by consideration of similarities and differences, as well as a synthetic approach, for example in enriching vocabulary activities. All these skills are crucial for information storage and retrieval capacity of long- and short-term memory (Schneider, 1999). Lightbown and Spada (2006) stress that advanced learners may find explicit explanation of great value, especially in terms of independent error-correction, whereas younger learners may rather benefit from immediate explicit forms of feedback on errors.

## 2.3. Structured teaching

The principle of structured teaching is introducing the material gradually where a more complex topic is built on an easier one. Schneider and Crombie (2003) advise that the presented material should refer to the previously learned information. Given the chunkable nature of linguistic patterns of a word and its relation to other words, as well as the existence of a vast set of rules, the structured approach is particularly important in teaching vocabulary and spelling, especially for younger learners or less proficient ones. The specific teaching and learning techniques include making categorisation by a specific letter pattern, the rule of pronunciation or by using keywords that will act as a reminder of the correct pronunciation or spelling. The structure is also important in improving writing skills, especially in the context of producing logical sequences (Schneider & Crombie, 2003; Schneider & Kulmhofer, 2016). It is also important that structure-based approaches aim at integration of the rules so that a learner is not exposed to isolated patterns and forced to practise one structure at a time. The learner should be provided with opportunities to discover how different language features compare and contrast in everyday language use (Lightbown & Spada, 2006).

## 2.4. Cumulative teaching

Cumulative teaching entails overlearning and systematic summarising of the learning outcomes. Overlearning, which allows for a regular revision of the taught material, gives the learner an opportunity to practise through a variety of activities. The amount of overlearning will depend on the learner's severity of learning difficulty, their underlying ability and memory capacity to automatize the taught material. It is also frequently related to the difficulty of a task or of the language aspect. It can be assumed that a learner with a severe learning difficulty and lower cognitive skills will need more time to successfully retain the material and use it in an independent context. Since some learners may take longer to acquire the taught material, Schneider and Crombie (2003) suggest that overlearning should be provided by using a range of activities that include multisensory techniques to avoid boring and rote learning tasks, which may consequently have a detrimental effect on the learner's motivation for learning.

## 2.5. Raising metalinguistic awareness

Metalinguistic awareness means conscious thinking about the patterns and rules of the language. Learners should be encouraged to discover independently how language works and how they can self-correct and monitor their own learning process. Building metalinguistic awareness can be attained by thought-provoking questions (e.g., 'What do you mean?', 'How would you rephrase that?') and non-verbal gestures. Also, making reference to the learner's native language may contribute to building their metalinguistic awareness because of positive language transfer (Schneider & Crombie, 2003; Schneider & Kulmhofer, 2016).

## 2.6. Other considerations in teaching a FL to learners with dyslexia

Apart from multisensory, explicit, structured, cumulative teaching and raising metalinguistic awareness, Schneider and Crombie (2003) suggest that the teacher should consider such accommodations as: presenting the material at a slower pace, providing additional time for task completion, colour-coding, pictograms and graphics, reading materials and worksheets on tainted paper with an enlarged font, and avoiding 'busy' pages. The teachers should also use mnemonic devices that would help the learner remember difficult language concepts and meet the learners' individual needs by teaching according to their learning style, giving an individual pace of work, anticipating possible difficulties, and providing differentiated resources and constructive feedback on student's learning progress.

Furthermore, Miller and Bussman-Gillis (2000) suggest that FL teachers who work with dyslexic learners must understand the speech-sound system and have a good command of spoken and written language. They also need to understand the principles of co-articulation and the physical properties of sounds and be able to recognise how the understanding of language patterns forms the foundation for the study of morphology, syntax and semantics. It is thus important that FL teachers, who work with dyslexic students, have extensive knowledge about both foreign language acquisition and special education (Kormos & Kontra, 2008). Since experience alone is often insufficient to develop an effective teaching practice, teachers should also adopt an on-going reflective approach to their teaching, which allows them to regularly

review their practice and search for the most effective strategies (Richards & Lockhart, 1996; Schneider & Kulmhofer, 2016).

## 3. Aim and research questions

The present study aimed at investigating Croatian EFL teachers' classroom practice in regard to the recommended teaching approaches and accommodations for dyslexic learners ('dyslexia-friendly' approaches).

We asked the following research questions:

1. Which recommended teaching approaches and accommodations do the teachers use?
2. Which language skills do they teach using the recommended approaches?
3. Does the practice of the primary school teachers differ from the secondary school teachers?

### 3.1. Participants

The demographic data about the participants are presented in Table 1. Sixteen schoolteachers, eight primary and eight secondary teachers, participated in the study[2]. The participants came from different parts of Croatia: Karlovac, Rijeka, Slavonski Brod, and Zagreb. Their work experience ranged between ten months and 29 years. Five teachers had up to five years of work experience (for the purpose of this paper the term 'novice' teachers will be used) and eleven had six or more years of experience (for the purpose of this paper the term 'experienced' teachers will be used). Twelve participants had experience in teaching dyslexic learners and seven of them had dyslexic students in the observed lesson.

Table 1 Data about the participants

| | | |
|---|---|---|
| Work place | Primary school | 8 |
| | Secondary school | 8 |
| Work experience | Novice (up to 5 years) | 5 |
| | Experienced (6 or more years) | 11 |
| Experience in teaching dyslexic learners | Yes | 12 |
| | No | 4 |
| Dyslexic students present in the observed lesson | Yes | 7 |

[2] In Croatia, primary teachers teach students aged 7-14, whereas secondary teachers teach students aged 15-18.

### 3.2. Methodology

The data were collected through lesson observation using an observation grid with pre-determined categorisation. The observation grid contained the following eight categories:

1. multisensory teaching;
2. explicit teaching;
3. structured teaching;
4. cumulative teaching;
5. raising metalinguistic awareness;
6. use of mnemonic devices;
7. meeting individual needs; and
8. accommodations.

We observed one lesson of 45 minutes taught by each participant. We noted the approach the participant used and the skill that they taught using a given approach. Before conducting the observations, we obtained oral permission from the schools' head teachers as well as from the participants themselves.

We analysed the data using a qualitative and quantitative method and used simple statistical methods for the number and percentage of the participants who had used a given approach or accommodation. We examined comparatively two groups of participants, primary and secondary school teachers, and used descriptive analysis for the examples of skills that were taught using a given approach.

### 3.3. Results

#### 3.3.1. Multisensory teaching

Table 2 Multisensory teaching

| | Primary School | Secondary School |
|---|---|---|
| Visual | 8(100%) | 8(100%) |
| Auditory | 7(87.5%) | 5(62.5%) |
| Kinaesthetic/tactile | 7(87.5%) | 3(37.5%) |

All the teachers used the visual mode. They used graphic organizers, posters, the board, and colour markers for teaching grammar, spelling and vocabulary. In addition, primary school teachers used flash cards, object presentation, supporting a text with pictures, cartoons, videos, and art work prepared by students.

The auditory mode was used by 12 teachers (75%). More primary (87.5%) than secondary school teachers (62.5%) used this mode (Table 2). The teachers used the auditory mode to teach grammar, spelling, pronunciation, and vocabulary. The examples included stressing the sound or the pitch within a word (e.g., when teaching

the spelling of the word *receipt* and *can't*), explaining the language concept (e.g., when teaching a question structure and the suffix *-ly* for forming adverbs), and peer assisted reading. In addition, primary school teachers used books on tapes, paired reading and computerized text readers, language games, songs, and rhymes. Secondary school teachers also stressed pronunciation of new vocabulary.

The kinaesthetic/tactile mode was used by ten teachers (62.5%). It was used more by primary school teachers (87.5%) than by those in the secondary school (37.5%) (see Table 2). Secondary school teachers used the kinaesthetic/tactile mode to teach grammar and spelling, but only to emphasize words patterns (e.g., when teaching the suffix *-ly* for forming adverbs). Primary school teachers used a wider range of forms to teach grammar, spelling, and vocabulary. They primarily used games such as looking for an object, guessing the words they were miming, touching pictures, the 'Simon says' game, raising coloured pens after hearing the correct word, or throwing a piece of chalk at the right word written on the board. Besides games, primary school teachers also asked their students to follow the text with their finger during listening activities, demonstrated a specific sound by using the vocal apparatus (e.g. *-ed* for past tense and *-er* for comparatives), and used the Total Physical Response method (Asher, 1969).

### 3.3.2. Explicit teaching

All the teachers taught grammar, pronunciation, spelling, and vocabulary explicitly using direct instruction and contrastive analysis by making reference to the students' native language as well as providing detailed explanation and explicit feedback. The examples included the singular and plural form of the noun, present and past tenses, comparatives, prefixes and suffixes, modals and conditionals, homophones, and the pronunciation of the sounds (e.g., /w/ and /v/, the silent *p* in the word *receipt*).

### 3.3.3. Use of mnemonic devices

Table 3. Use of mnemonic devices

| Primary school | Secondary school |
|---|---|
| 6(75%) | 4(50%) |

Ten teachers used mnemonic devices (62.5%). They were used more by primary school teachers (75%) than by those in the secondary school (50%) (Table 3). The teachers used mnemonic devices to teach grammar, pronunciation, spelling, and vocabulary. They used personalisation, visual and auditory association, gestures and motion clues, as well as keywords and key sentences. The examples included comparing words with similar spelling (e.g., *than* and *then*) and playing 'sticky fingers' when explaining the connection between the root word and the suffix. In addition, primary school teachers used picture and sound clues, songs, face expressions, scanning the text for similar words, and making funny stories with the words that have similar pronunciation.

### 3.3.4. Raising metalinguistic awareness

All the teachers raised metalinguistic awareness when teaching grammar, pronunciation, spelling, and vocabulary. They asked thought-provoking questions (e.g., when revising the singular and plural form of nouns, object pronouns, and prepositions), made reference to the students' native language (e.g., translating the sentences written by the student in English into Croatian to provoke self-correction or explaining the structure of conditional clauses and modal verbs), used non-verbal and para-verbal gestures (e.g., the teacher facilitated the reading of the word or prompted the retrieval of the correct preposition by sounding out the first letter), changed the tone of their voice (e.g., when emphasizing pronunciation), and referred to the situations from the students' and the teacher's lives (e.g., when discussing the text in the course book). Primary school teachers also used rephrasing and encouragement.

### 3.3.5. Structured teaching

Table 4. Structured teaching

| Primary school | Secondary school |
|---|---|
| 7(87.5%) | 6(75%) |

Thirteen teachers (81%) used the structured approach (Table 4). The teachers taught grammar, spelling, and vocabulary in a structured way. The introduced material was presented in chunks with a gradual increase of challenge to ensure mastery of learning. The examples included teaching the singular and plural form of the noun, the use of prepositions and new vocabulary in sentences, as well as the structure of present and past modals, conditional clauses, and adverbs. The taught material was stored for later recall in different forms which included flash cards with new vocabulary and students' self-made dictionaries and charts.

### 3.3.6. Cumulative teaching

Table 5. Cumulative teaching

| Primary school | Secondary school |
|---|---|
| 5(62.5%) | 6(75%) |

Eleven teachers (69%) used the cumulative approach. The approach was used more by secondary school teachers (75%) than by those in the primary school (62.5%) (Table 5). The teachers taught grammar and vocabulary in the form of overlearning and made a summary of the learning outcomes. The examples included teaching the

prepositions of place and the singular and plural form of the noun, as well as revising the key words discussed in the lesson.

### 3.3.7. Meeting individual needs

Table 6. Meeting individual needs

| Primary school | Secondary school |
|---|---|
| 4(50%) | 7(87.5%) |

Eleven teachers (69%) met individual needs of the students in the lesson and they were met more by secondary school teachers (87.5%) than by those in the primary school (50%) (Table 6). The teachers allowed an individual pace of work (e.g., when students were writing sentences, retelling a story, or doing a grammar task) and gave constructive feedback when correcting a grammar or writing task. In addition, primary school teachers taught according to the students' learning style and recognised students' learning strengths when playing class games or doing pair work. Secondary school teachers differentiated and modified questions to adjust them to the student's knowledge or when the student did not understand the question at first. Secondary school teachers also gave additional explanation for grammar tasks.

### 3.3.8. Accommodations

Table 7. Accommodations

| Primary school | Secondary school |
|---|---|
| 6(75%) | 6(75%) |

Twelve teachers (75%) provided accommodations. They were provided by an equal percentage of primary and secondary school teachers (Table 7). The accommodations included worksheets with no 'busy' pages and pictures, and giving additional time to complete reading and writing tasks. Primary school teachers also used a seating plan according to students' abilities and provided different opportunities for stress relief such as the use of a stress-ball, eating during lesson, short walks at the back of the classroom, or time-off outside the classroom. They also sat dyslexic students closer to the teacher and the board.

## 4. Discussion

The aim of the present study was to identify 'dyslexia-friendly' teaching approaches in the Croatian EFL teachers' classroom practice based on the data obtained from lesson

observations. To answer the first research question ('Which recommended teaching approaches and accommodations do the teachers use?'), data was collected using an observation grid with pre-determined eight categories of recommended approaches. Apart from noting the approach that was used in the lesson, the observer also noted the language skill that was taught using a given approach, which provided data to answer the second research question ('Which language skills do they teach using the recommended approaches?'). To answer the third research question ('Does the practice of the primary school teachers differ from the secondary school teachers?'), we examined the data comparatively for primary and secondary school teachers. We calculated the percentage of the teachers who used a given approach and the skills that were taught while using it.

The findings suggest that the participants used all the recommended approaches and provided accommodations. All the participants taught explicitly, in a multisensory way, and raised metalinguistic awareness. An equal percentage of primary and secondary school teachers provided accommodations.

However, there were noticeable differences between the primary and secondary school teachers in regard to other approaches. Regarding multisensory teaching, primary school teachers used more activities that relied on the auditory and kinaesthetic/tactile mode and used a wider range of auditory and visual forms. We believe that this difference in the teaching approach is related to the age of the learner. As suggested by Cameron (2001), teaching EFL to young learners should be based on multisensory experience, where the child is provided with opportunities to see, hear, manipulate, touch, and feel.

Moreover, more primary school teachers taught in a structured way and used mnemonic devices, whereas more secondary school teachers used the cumulative approach and met students' individual needs. It is also worth noting that primary school teachers used a wider range of mnemonic devices and both groups of participants met students' individual needs in a different way. We believe that this difference again stems from the methodology of teaching younger learners as the primary school teachers used mnemonic devices that are related to multisensory experience.

Among the participants there were teachers with and without experience in teaching dyslexic learners. We assumed that more of the former would use the approaches recommended in teaching dyslexic learners. However, the findings indicate that more teachers with no experience in teaching dyslexic learners used the kinaesthetic/tactile mode and mnemonic devices, and taught in a structured way. Moreover, more teachers who did not have dyslexic students in their observed lesson used auditory, kinaesthetic and tactile modes of multisensory teaching. We therefore concluded that multisensory and structured teaching and the use of mnemonic devices are an integral part of the contemporary EFL teaching methodology and are not unique for the methods that are recommended in teaching dyslexic learners. Yet it is worth noting that more teachers with experience in teaching dyslexic learners, those with longer work experience, and those who had a dyslexic student present in their observed lesson met students' individual needs. More teachers with a dyslexic student present in their lesson also taught in a cumulative and structured way and provided accommodations. These findings thus suggest that the teachers with experience in teaching dyslexic learners and those who have a dyslexic student present in their lessons adopt a more learner-centered approach.

## 5. Conclusion

The findings of this study indicate that primary and secondary school teachers in Croatia used all the approaches recommended in teaching dyslexic learners as well as provided accommodations and met individual needs of the students. Yet, the practice of primary school teachers differed from those in the secondary school, which is related to the age of the students they teach. The teachers in both phases of education taught a wide range of skills, including grammar, pronunciation, spelling, and vocabulary, using different recommended approaches. We therefore concluded that Croatian EFL teachers' classroom practice is 'dyslexia-friendly', and in this way the teachers create a positive learning atmosphere in a class of students with mixed abilities.

The present study had some limitations. There were only 16 participants and only one lesson was observed. Therefore, future research should involve a larger number of participants from both levels of education and a teacher should be observed over the course of a series of lessons. Future research may also consider a stricter interpretation of multisensory teaching, where the visual, auditory, kinaesthetic and tactical modes are used simultaneously (Kelly & Phillips, 2011). In this study, the modes that constitute multisensory teaching were examined separately.

Moreover, the practice of Croatian EFL primary and secondary school teachers should be further investigated since the findings of this study are not in accordance with the findings of previous research (Fišer & Dumančić, 2014; Kałdonek-Crnjaković & Fišer, 2016), which indicated that Croatian EFL teachers had limited knowledge of dyslexia and recommended methods in teaching dyslexic learners. Therefore, the future research, apart from lesson observation, should examine participants' knowledge about dyslexia in the context of FL learning and teaching and about the use of recommended methods in relation to different language skills and specific needs of the students.

## 6. References


Asher, J. J. (1969). The Total Physical Response approach in second language learning. *The Modern Language Journal, 53(1),* 3-17.

Cameron, L. (2001). *Teaching languages to young learners*. Cambridge: CUP.

Crombie, M. A. (2000). Dyslexia and the learning of a foreign language in school: Where are we going? *Dyslexia, 6(2)*, 112-123.

Croatian Dyslexia Association (2016). Definition of dyslexia. Retrieved on 20 June 2016 from http://hud.hr/definicije-disleksije/.

Fišer, Z., & Dumančić, D. (2014). How do I 'unjumbel' this? Study of EFL teachers' competences and preferences in teaching students with dyslexia. In S. Letica Krevelj & J. Mihaljević Djigunović (Eds.), *UZRT 2014: Empirical Studies in Applied Linguistics* (pp. 20-30). Zagreb: FF Press.

Ganschow, L., Sparks, R. L., & Javorsky, J. (1998). Foreign language learning difficulties: An historical perspective. *Journal of Learning Disabilities, 31(3),* 248-258.

Kahn-Horwitz, J., Shimron, J., & Sparks, R. L. (2005). Predicting foreign language reading achievement in elementary school students. *Journal of Reading and Writing, 18(6),* 527-558.

Kahn-Horwitz, J., Shimron, J., & Sparks, R. L. (2006). Weak and strong novice readers of English foreign language: Effects of first language and socioeconomic status. *Annals of Dyslexia, 56(1),* 167-185.

Kałdonek-Crnjaković, A. (2013). *Multisensory structured metacognitive instruction in the light of the age factor in teaching EFL to dyslexic learners* (Unpublished doctoral dissertation). University of Zagreb, Faculty of Philosophy and Social Sciences, Zagreb, Croatia.

Kałdonek-Crnjaković, A., & Fišer, Z. (2016). *How much do I know, I need to know and I want to know about dyslexia? Moving forward with vocational training for EFL teachers in Croatia*. Poster presented at the 10th International Conference of the British Dyslexia Association, 10-12 March 2016, Oxford, UK.

Kelly, K., & Phillips, S. (2011). *Teaching literacy to learners with dyslexia. A multisensory approach*. London: Sage.

Kormos, J., & Kontra, E. H. (2008). Hungarian teachers' perceptions of dyslexic language learners. In J. Kormos & E. H. Kontra (Eds.), *Language learners with special needs* (pp. 189-213). Bristol: Multilingual Matters.

Krashen, S. (1987). Principles and practice in second language acquisition. Oxford: Oxford University Press.

Lightbown, P. M., & Spada, N. (2006). *How languages are learned*. Oxford: Oxford University Press.

Miller, S., & Bussma-Gillis, M. (2000). The language puzzle: Connecting the study of linguistics with a multisensory language instructional programme in foreign language learning. In L. Peer & G. Reid (Eds.), *Multilingualism, literacy and dyslexia. A challenge for educators* (pp. 218-228). London: David Fulton Publishers.

Nijakowska, J. (2008). An experiment with direct multisensory instruction in teaching word reading and spelling to Polish dyslexic learners of English. In J. Kormos & E. H. Kontra (Eds.), *Language Learners with Special Needs* (pp. 130-157). Bristol: Multilingual Matters.

Richards, J. C., & Lockhart, Ch. (1996). *Reflective teaching in second language classrooms.* Cambridge: Cambridge University Press.

Schneider, E. (1999). *Multisensory structured metacognitive instruction*. Frankfurt am Main: Peter Lang.

Schneider, E., & Crombie, M. (2003). *Dyslexia and foreign language learning*. New York: David Fulton Publishers.

Schneider, E., & Kulmhofer, A. (2016). Helping struggling learners of English as an additional language succeed with interactive multisensory structured strategies. *BELT- Brazilian English Language Teaching Journal, 7(1),* 3-25.

Sparks, R. L., & Ganschow, L. (1993). The effect of multisensory structured language instruction on native language and foreign language aptitude skills of at-risk high school foreign language learners. A replication and follow-up study. *Annals of Dyslexia, 43(1),* 194-216.

Sparks, R. L., Ganschow, L., Artzer, M., & Patton, J. (1997). The foreign language proficiency of at-risk learners over two years of foreign language instruction. A follow-up study. *Journal of Learning Disabilities, 30(1),* 92-98.

Sparks, R. L., Ganschow, L., Kennewig, S., & Miller, K. (1991). Using Orton-Gillingham methodologies to teach language to learning disabled dyslexic students. Explicit teaching of phonology in a second language. *Annals of Dyslexia, 41(1),* 96-118.

Sparks, R. L., Ganschow, L., & Patton, J. (2008). L1 and L2 literacy, aptitude and affective variables as discriminators among high- and low-achieving L2 learners with special needs. In J. Kormos & E. H. Kontra (Eds.), *Language Learners with Special Needs* (pp. 11-35). Bristol: Multilingual Matters.

Sparks, R. L., Ganschow, L., Pohlman, J., Artzer, M., & Skinner, S. (1992). The effects of a multisensory structured language approach on the native and foreign language aptitude skills of high risk foreign language learners. *Annals of Dyslexia, 42(1),* 25-53.