



ANALYSIS OF THE SOURCES OF DIFFICULTIES EXPERIENCED BY UNIVERSITY STUDENTS IN LEARNING A FOREIGN LANGUAGE

Günay Kayhan,

Nilgün Turanlıⁱ,

Neslihan Acar,

Tuba Odabaş,

Yasemin Kütük

School of Foreign Languages,

Ondokuz Mayıs University,

Samsun, Turkey

Abstract:

Background and Objectives: In this study, the difficulties that 114 university students (Male $n = 47$, Female $n = 67$) enrolled in various departments experienced in learning a foreign language were examined using a four-factor scale. **Materials and Methods:** The data were collected through an online survey conducted via the Google Forms platform during the spring semester of the 2024–2025 academic year. All analyses were conducted using Python (scipy, pandas, numpy) and Microsoft Excel. Scale reliability was assessed using Cronbach's alpha. Normality was tested using the Shapiro-Wilk test. Comparisons of gender were conducted using an independent-samples t-test, while comparisons of department and income level were conducted using one-way ANOVA. Effect sizes were reported as Cohen's d for the t-test and as η^2 for ANOVA. The level of statistical significance was set at $\alpha = .05$. **Results:** Descriptive findings showed that participants experienced the greatest difficulty in the social and environmental subscale (Mean = 2.24, SD = 0.66); followed, respectively, by individual learning difficulties (Mean = 2.19, SD = 0.62), exam and material related difficulties (Mean = 1.98, SD = 0.70), and teacher-related difficulties (Mean = 1.88, SD = 0.65). Independent-samples t-test results showed no statistically significant difference with respect to the gender variable (all $p > .05$). One-way ANOVA analysis revealed a statistically significant difference in total difficulty scores in terms of the department variable ($F(5, 108) = 2.754$, $p = .022$, $\eta^2 = .113$); students in departments offering intensive foreign language instruction reported lower difficulty scores. Differences in income level did not yield significant results (all $p > .80$). **Conclusion:** Although the Pearson correlations between subscales were positive and statistically significant (ranging from $r = .45$ to $.76$), the social and environmental subscales showed relatively lower correlations with the other dimensions. The findings

ⁱ Correspondence: email ntgiritli@omu.edu.tr

underscore the importance of holistic and student-centred pedagogical approaches to foreign language instruction.

Keywords: difficulties in learning a foreign language, language learning anxiety, university students, psychometrics

1. Introduction

Foreign language learning is a cognitive and social process that aims to help individuals acquire the ability to communicate in a second language. The challenges students face during this process directly affect both their academic success and their motivation to learn the language. Research shows that difficulties in learning a foreign language have a multidimensional structure and that these difficulties are shaped by factors related to teacher behaviour, teaching methods, insufficient materials, individual cognitive barriers, and social context (Horwitz *et al.*, 1986; Young, 1991).

Studies on foreign language anxiety and learning difficulties among students at Turkish universities indicate that significant problems exist, particularly in the areas of oral production and social interaction. Students report psychological barriers that negatively affect learning, such as avoiding classroom participation, fear of making mistakes, and anxiety about being evaluated by their peers. However, the majority of existing studies have focused on specific language skills, while a systematic and multidimensional examination of these difficulties has not been sufficiently conducted.

The difficulties involved in learning a foreign language encompass a broad conceptual framework that includes linguistic, cognitive, and affective dimensions. In his affective filter hypothesis, Krashen (1982) argued that affective variables such as anxiety, motivation, and self-confidence directly hinder second language acquisition. Horwitz *et al.* (1986) defined foreign language anxiety as a distinct construct and argued that it consists of three core components: communication apprehension, test anxiety, and fear of negative evaluation. These three components show conceptual overlap with the scale dimensions in our study.

Teacher behaviour and teaching methods are also frequently cited as sources of difficulty. Not giving students sufficient opportunities to speak, correcting mistakes in a harsh manner, and failing to provide explanations appropriate to the students' level have been documented as factors that increase anxiety in the classroom (Young, 1991). Exam pressure and a lack of materials stand out as structural barriers that complicate the academic aspect of language learning.

Communicating in a foreign language creates significant social pressures, particularly in contexts where opportunities to speak the target language are limited. Speaking up in large classes, understanding different accents, and finding opportunities for practice to continue language learning are among the environmental difficulties frequently cited by students. In contexts like Turkey, where the foreign language is not widely spoken, these difficulties are expected to be more significant.

How individual variables such as gender, income level, age, and department influence the perception of difficulty in learning a foreign language remains a subject of debate. While some studies report that female students have higher levels of language anxiety, other research suggests that this difference depends on contextual and measurement factors. In terms of the variable of department, the fact that students in programs with intensive foreign language instruction perceive fewer difficulties indicates that level of experience shapes the perception of difficulty.

This study aims to examine the difficulties university students face in learning a foreign language within the framework of a scale that addresses four dimensions: teacher-related difficulties, exam and material-related difficulties, difficulties related to the individual learning process, and social and environmental difficulties. The study seeks to answer the following research questions:

- What is the profile of the difficulties university students experience in learning a foreign language at the subscale level?
- Does the scale demonstrate sufficient reliability for the target sample?
- Do difficulty scores show significant differences in terms of gender, department, and income level?

2. Material and Methods

2.1 Participants

The study sample consisted of 114 university students (Male $n = 47$, 41.2%; Female $n = 67$, 58.8%) enrolled in various departments at a university in Turkey. The distribution of participants by department is as follows: German Language Teaching ($n = 50$, 43.9%), Engineering ($n = 28$, 24.6%), Economics ($n = 14$, 12.3%), Elective Preparatory Program ($n = 9$, 7.9%), English Language Teaching ($n = 7$, 6.1%) and English Translation and Interpretation ($n = 6$, 5.3%). An analysis of the age distribution shows that the vast majority of participants were between the ages of 18 and 20 ($n = 108$, 94.7%). In terms of income level, the middle-income group was the most prevalent ($n = 83$, 72.8%). Informed consent was obtained from all participants prior to data collection.

2.2 Instrument

The scale used in the data collection process was developed by Bayır & Yeşil (2019) and consists of 29 items and four subscales designed to measure the sources of difficulty in foreign language learning. Permission to use the instrument was obtained from the authors prior to data collection. The items, which are rated on a four-point Likert scale (1 = Not at all/Very little, 2 = Somewhat/Partially, 3 = Quite a bit, 4 = Very much), are grouped into the following subscales:

- **Teacher-Related Difficulties (I1–I8, 8 items):** Difficulties caused by the teacher's behaviour, teaching methods, and classroom communication.
- **Test and Material-Related Difficulties (I9–I14, 6 items):** Difficulties caused by the test format, question types, and insufficient course materials.

- **Individual Learning Difficulties (I15–I23, 9 items):** Individual difficulties with translation strategies, vocabulary, grammar, and reading comprehension processes.
- **Social and Environmental Difficulties (I24–I29, 6 items):** Difficulties caused by different speaking styles, crowded environments, a lack of opportunities for practice, and fear of criticism.

2.3 Procedure

Prior to data collection, ethical approval was obtained from the relevant institutional ethics committee. Participants were presented with an informed consent statement confirming that participation was voluntary and assuring them that their responses would remain confidential and be used only for academic research purposes. The questionnaire was administered via an online survey on Google Forms during the spring semester of the 2025-2026 academic year. Participants completed the instrument individually, and completion required approximately 10–15 minutes.

2.4 Data Analysis

All analyses were conducted using Python (scipy, pandas, numpy) and Microsoft Excel. Scale reliability was assessed using Cronbach's alpha. Normality was tested using the Shapiro-Wilk test. The social and environmental subscale, as well as the total score, met the assumption of normality; deviations from normality were identified in the other subscales. Since $N \geq 30$, parametric tests were preferred in accordance with the Central Limit Theorem. Comparisons of gender were conducted using an independent-samples t-test, while comparisons of department and income level were conducted using one-way ANOVA. Effect sizes were reported as Cohen's d for the t-test and as η^2 for ANOVA. The level of statistical significance was set at $\alpha = .05$.

3. Results

3.1 Descriptive Statistics

Table 1 shows descriptive statistics for the scale's four subscales and total score, along with Cronbach's alpha reliability coefficients. The social and environmental subscale has the highest mean difficulty score (Mean = 2.24, SD = 0.66), followed by individual learning (Mean = 2.19, SD = 0.62), exam and material (Mean = 1.98, SD = 0.70), and teacher-related difficulties (Mean = 1.88, SD = 0.65). The total scale mean score is Mean = 2.07 (SD = 0.53). The fact that all subscale mean scores fall below 2.5—the midpoint of the 4-point scale—indicates that the overall perception of difficulty ranges from moderate to low.

Table 1: Descriptive Statistics and Reliability Coefficients for Subscales (N = 114)

Subscale	Number of Items	Mean	SD	Median	Min	Max	α
Teacher-related	8 (I1– I8)	1.88	0.65	1.81	1.00	4.00	.843
Exam and Material	6 (I9– I14)	1.98	0.70	1.83	1.00	4.00	.825
Individual Learning	9 (I15– I23)	2.19	0.62	2.11	1.00	4.00	.871
Social and Environmental	6 (I24– I29)	2.24	0.66	2.17	1.00	4.00	.786
Total Score	29 (I1– I29)	2.07	0.53	2.07	1.00	4.00	.926

SD = standard deviation; α = Cronbach's alpha. All scores are calculated as item averages and range from 1 to 4. Reliability thresholds: $\alpha \geq .90$ = Excellent; $\geq .80$ = Good; $\geq .70$ = Acceptable (George & Mallery, 2003).

Figure 1 visually presents the mean scores for the subscales with their 95% confidence intervals.

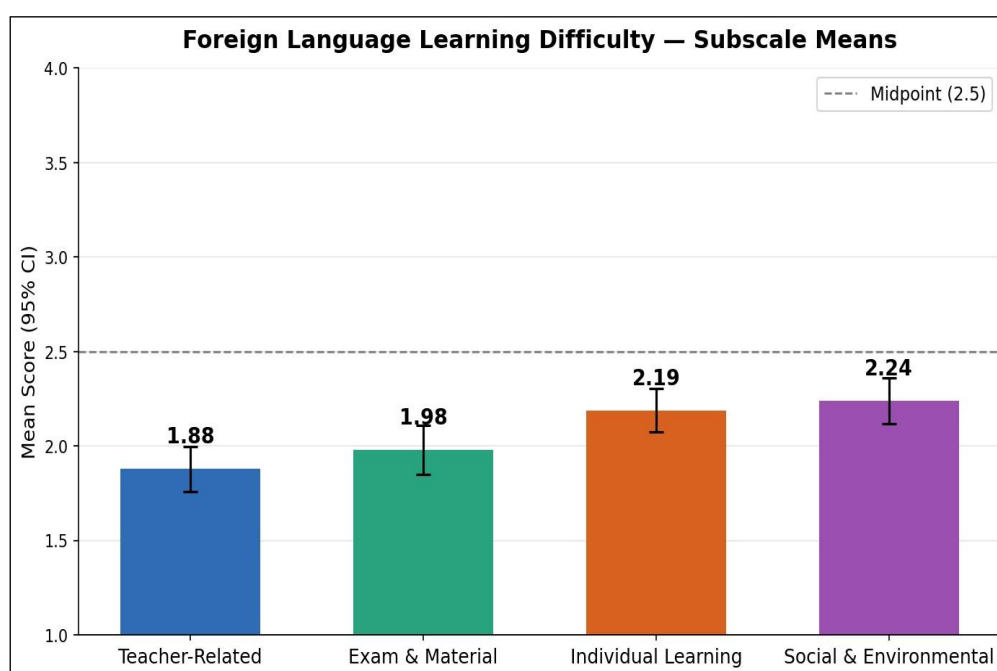


Figure 1: Mean Scores on Foreign Language Difficulty Subscales (with 95% CI, N = 114).

The error bars represent $1.96 \times SE$. The dashed line represents the midpoint of the scale (2.5).

3.2 Item-Level Analysis

When examined at the item level, the items with the greatest average difficulty were, respectively, I24 "Difficulty in understanding when I hear a different speaking style" (Mean = 2.43), I5 "When teachers use boring teaching methods" (Mean = 2.34), and I27 "When I want to speak in crowded classrooms" (Mean = 2.34). The lowest average difficulty scores were for I1, "When teachers do not give us enough opportunity to speak" (Mean = 1.59), I13 "When the same types of questions are asked repeatedly on exams" (Mean = 1.66), and I6 "When teachers do not encourage us to speak" (Mean = 1.72).

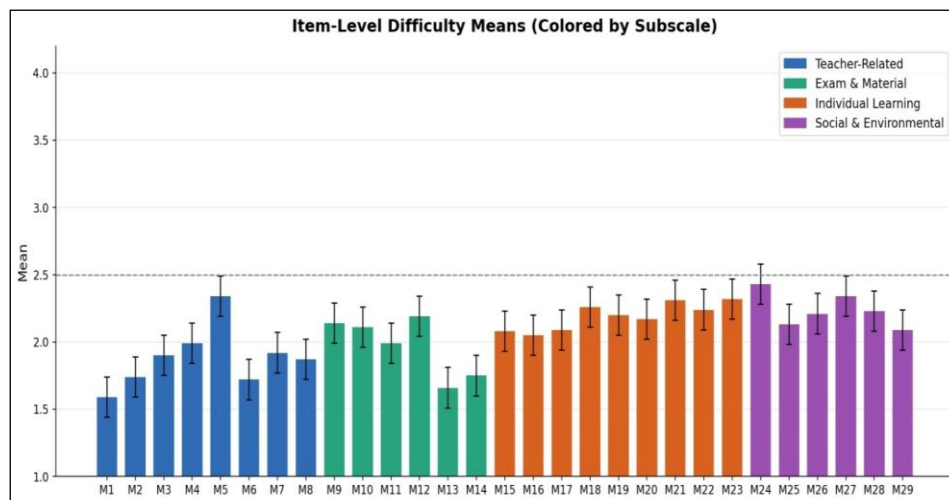


Figure 2: Average Difficulty Scores at the Item Level (Colour-Coded by Subscale)

The error bars indicate 95% confidence intervals.

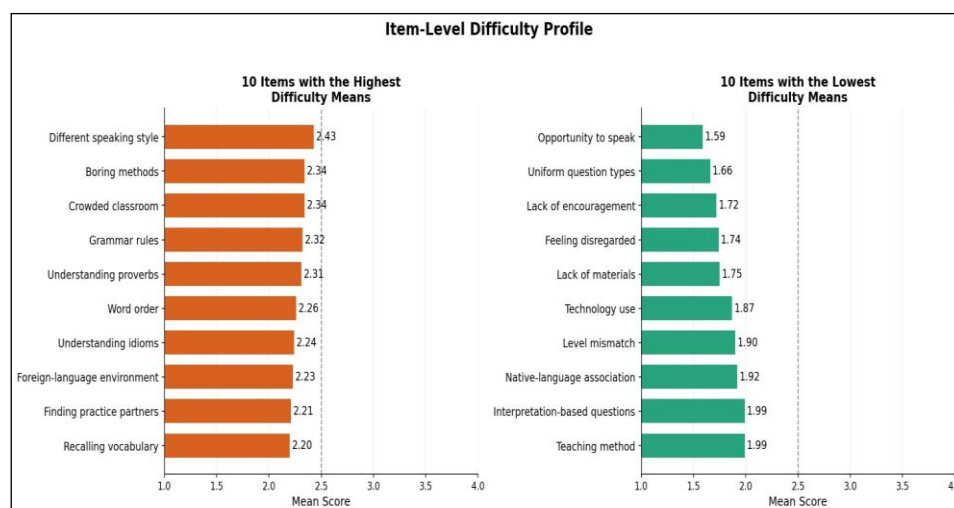


Figure 3: The 10 Items with the Highest and Lowest Average Difficulty Scores

3.3 Reliability Analysis

Cronbach's alpha coefficients are shown in Table 1. The total scale has excellent internal consistency ($\alpha = .926$). At the subscale level, the "individual learning difficulties" subscale demonstrated near-perfect reliability ($\alpha = .871$); while the "teacher-related" ($\alpha = .843$) and "exam and material" ($\alpha = .825$) subscales showed good reliability. The social and environmental subscale has an acceptable level of reliability ($\alpha = .786$). The fact that all subscale alpha values exceed the .70 threshold indicates that the scale meets the psychometric quality requirements for the purposes of this study.

3.4 Correlation Analysis

Table 2 shows the Pearson correlation coefficients between the subscales. All correlations were positive and statistically significant ($p < .001$), ranging from moderate to strong ($r = .451$ to $.758$). The strongest relationship was observed between the teacher-related

subscale and exam and material-related subscales ($r = .758$), reflecting the conceptual proximity of these two dimensions. The social and environmental subscale showed relatively lower correlations with the other dimensions ($r = .451-.635$), indicating a certain degree of conceptual specificity.

Table 2: Pearson Correlation Matrix – Subscales and total score

	Teacher-related	Exam and Material	Individual Learning	Social/ Environmental	Total
Teacher-related	1.000	.758***	.471***	.456***	.832***
Exam and Material	.758***	1.000	.501***	.451***	.826***
Individual Learning	.471***	.501***	1.000	.635***	.821***
Social/ Environmental	.456***	.451***	.635***	1.000	.764***
Total	.832***	.826***	.821***	.764***	1.000

All correlations are significant at the $p < .001$ level. Dark green = $r \geq .70$ (strong); light green = $.50 \leq r < .70$ (moderate); gray = $.30 \leq r < .50$ (weak).

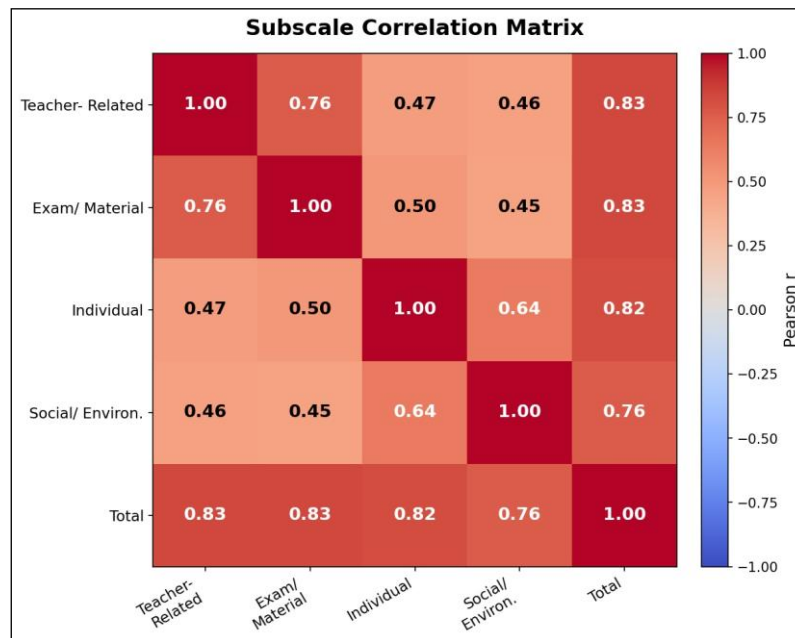


Figure 4: Subscale Pearson Correlation Matrix Heat Map

3.5 Differences in Terms of Gender

Table 3 shows the differences between gender groups based on the results of the independent samples t-test. No statistically significant gender difference was found (all $p > .05$). However, on the social and environmental subscale, female students (Mean = 2.34, SD = 0.64) reported higher difficulty scores than male students (Mean = 2.10, SD =

0.67), and this difference approached the threshold of statistical significance ($t(112) = -1.953$, $p = .053$, $d = 0.37$; small effect size).

Table 3: Differences in terms of gender – Independent-samples t test results

Subscale	Men mean (SD)	Female mean (SD)	t(112)	p	Cohen d	Sig.
Teacher-related	1.95 (0.72)	1.84 (0.60)	0.870	.386	0.17	ns
Exam and material	2.00 (0.76)	1.96 (0.67)	0.270	.787	0.05	ns
Individual learning	2.15 (0.58)	2.22 (0.65)	-0.539	.591	0.10	ns
Social and environmental	2.10 (0.67)	2.34 (0.64)	-1.953	.053	0.37	ns (close)
Total score	2.05 (0.56)	2.08 (0.51)	-0.324	.747	0.06	ns

Note: Men $n = 47$, Women $n = 67$. ns = not significant. Cohen's d : $< .20$ = insignificant; $.20-.49$ = small; $.50-.79$ = moderate; $\geq .80$ = large (Cohen, 1988).

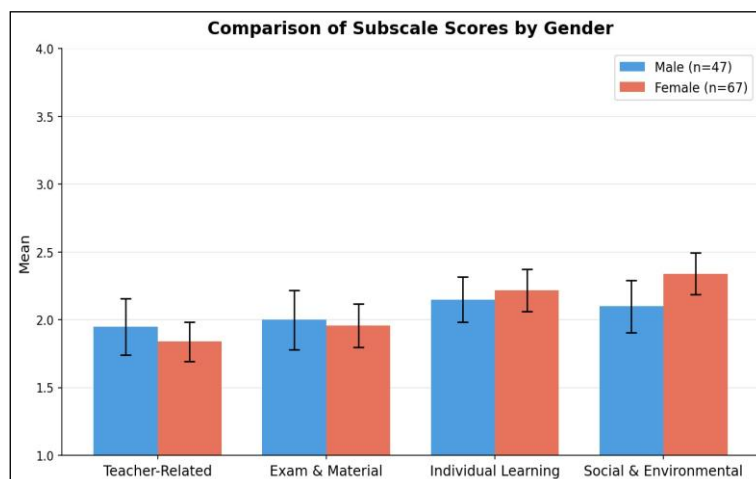


Figure 5: Comparison of Subscale Scores by Gender.
The error bars indicate 95% confidence intervals.

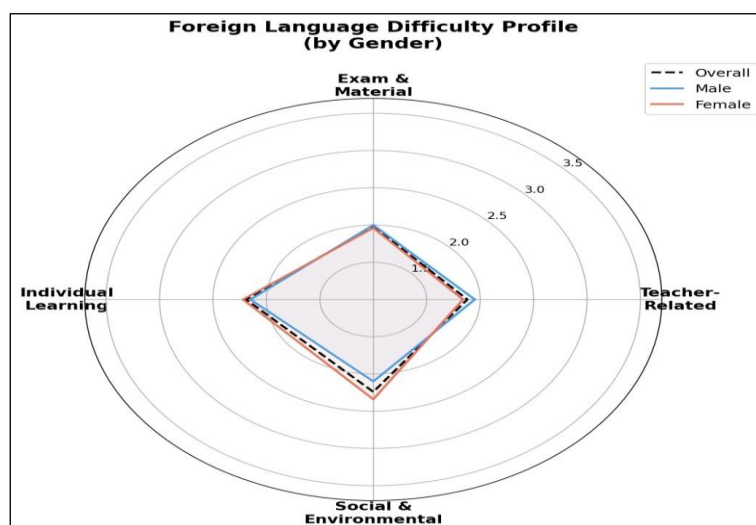


Figure 6: Foreign Language Difficulty Profile Radar Chart (by Gender)

3.6 Differences in Terms of Department

One-way ANOVA showed a statistically significant difference among departments in terms of total difficulty scores ($F(5, 108) = 2.754$, $p = .022$, $\eta^2 = .113$; moderate effect size). The descriptive data presented in Table 4 indicate that the highest difficulty scores were found in the Engineering (Mean = 2.22) and German Language Teaching (Mean = 2.16) departments, while the lowest scores were found in the English Translation and Interpretation (Mean = 1.58) and Elective Preparatory (Mean = 1.78) departments.

Table 4: Total Difficulty Score in terms of Department – One-Way ANOVA Results

Department	n	Mean	SD
German Language Teaching	50	2.16	0.53
Engineering	28	2.22	0.50
Economics	14	1.93	0.45
Elective Preparatory	9	1.78	0.30
English Language Teaching	7	1.95	0.70
English Translation and Interpretation	6	1.58	0.23
F(5, 108)		2.754	
p		.022*	
η^2		.113	

Note: * $p < .05$. η^2 = eta-square. Effect size thresholds: .01 = small; .06 = moderate; .14 = large (Cohen, 1988). The Tukey HSD post hoc test is recommended for identifying significant differences.

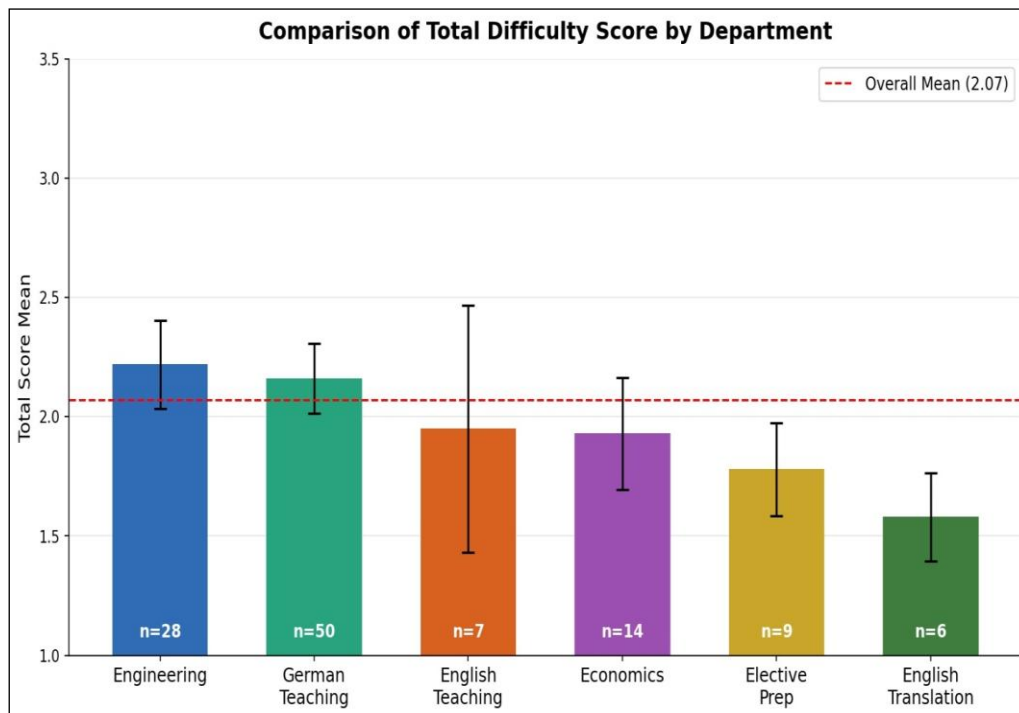


Figure 7: Comparison of Total Difficulty Scores by Department. The error bars represent 95% confidence intervals; the red dashed line indicates the overall mean.

3.7 Differences in Level of Income

One-way ANOVA analysis did not reveal any statistically significant differences in terms of income level for any subscale or total score (all $F < 0.20$, all $p > .80$, all $\eta^2 < .01$). This finding indicates that socioeconomic status does not significantly influence the perception of difficulties in learning a foreign language across this sample.

4. Discussion

The finding that participants reported the highest average level of difficulty in the social and environmental difficulties subscale points to the social burden associated with using a foreign language in contexts where opportunities to speak the target language are limited. The result that the item addressing different speaking styles had the highest average difficulty score (I24, Mean = 2.43) indicates that exposure to different accents and speech patterns constitutes a significant source of listening difficulty for students. Speech anxiety in large classes (I27), the item with the second-highest mean score, also underscores the importance of social anxiety in the classroom setting.

The result that the subscale related to teacher-related difficulties had the lowest average score indicates that participants perceived teacher behaviours as relatively less problematic. This finding may reflect the increasing professional development efforts aimed at foreign language instruction in Turkish higher education in recent years.

The fact that the scale has excellent overall reliability ($\alpha = .926$) confirms that it is a reliable tool for measuring foreign language learning difficulties among university students. The result that the subscale alpha values range from .786 to .871 indicates that researchers can confidently evaluate the subscales as independent constructs.

The finding that no statistically significant difference was found in terms of the gender variable is consistent with similar findings in other Turkish samples. However, the near-significant trend observed in favour of female students on the social and environmental subscale ($p = .053$, $d = 0.37$) suggests that female students may be more sensitive to social evaluation anxiety in a foreign language. This trend needs to be tested with larger and more balanced samples.

The moderate effect size found in terms of the department variable ($\eta^2 = .113$) indicates that the academic program significantly shaped the perception of difficulties in foreign language learning. The finding that English Translation and Interpretation department students reported a significantly lower difficulty score (Mean = 1.58) suggests that intensive and structured foreign language instruction reduces the perception of difficulty. Meanwhile, the finding that engineering students reported the highest difficulty score highlights the need for targeted intervention programs for foreign language instruction in this group.

4.1 Pedagogical Implications

The findings of this study lay the groundwork for various pedagogical implications. First, the fact that social anxiety emerges as the most prominent source of difficulty

underscores the importance of creating a low-anxiety learning environment in the classroom. Small-group activities, structured pair work, and tasks that support confident speaking practice can be designed for this purpose. Second, the fact that different speaking styles and a variety of accents are the most challenging aspects underscores the need to systematically expose students to authentic and diverse audio materials. Third, differences across departments indicate that foreign language education should not be limited to relevant departments but should be integrated into all academic programs. Developing specialized foreign language programs for students in technical fields such as engineering has the potential to reduce this group's perception of difficulty.

4.2 Limitations

Various limitations must be taken into account regarding the generalizability of the findings to broader contexts. The fact that the sample was drawn from a single university limits the generalizability of the findings. The significant imbalance between groups in terms of the age variable ($n = 108$ for ages 18–20 versus $n = 4$ for ages 21–23 and $n = 2$ for ages 24 and older) makes comparisons based on this variable statistically unreliable. It should also be noted that self-report data are susceptible to the social desirability effect. Due to the cross-sectional design, it is not possible to establish causal relationships. It is recommended for future studies to include larger and more balanced samples from different institutions, longitudinal follow-up designs, and objective measures of language performance.

5. Conclusion

This study has identified the difficulties Turkish university students face in learning a foreign language through a comprehensive psychometric profile. The four-subscale structure demonstrated excellent overall reliability ($\alpha = .926$), and internal consistency values above the acceptable level were obtained for all subscales. The prominence of social and environmental difficulties highlights the central importance of social dynamics—such as speaking in large classes, understanding different speaking styles, and anxiety about being evaluated—in foreign language learning. The significant differences observed in terms of department indicate that the foreign language learning experience shapes the perception of difficulty and underscore the need for language teaching interventions tailored to different academic programs. Gender and income level were not found to be significant variables in this sample. These findings collectively underscore the need to strengthen pedagogical approaches that focus on individual differences in foreign language instruction, the social context, and student-centred approaches.

Conflict of Interest Statement

The authors declare no conflicts of interest.

About of Author(s)

Günay Kayhan, School of Foreign Languages, Ondokuz Mayıs University, Samsun, Turkey.

Nilgün Turanlı, School of Foreign Languages, Ondokuz Mayıs University, Samsun, Turkey.

Neslihan Acar, School of Foreign Languages, Ondokuz Mayıs University, Samsun, Turkey.

Tuba Odabaş, School of Foreign Languages, Ondokuz Mayıs University, Samsun, Turkey.

Yasemin Kütük, School of Foreign Languages, Ondokuz Mayıs University, Samsun, Turkey.

References

- Bayır, M. & Yeşil, R. (2019). Yabancı Dil Öğrenme Zorlukları Ölçeğinin Geçerlik ve Güvenirlik Çalışması. *Manas Sosyal Araştırmalar Dergisi*, 8 (2): 1573-1590. <https://doi.org/10.33206/mjss.530574>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. baskı). Lawrence Erlbaum Associates. Retrieved from https://books.google.ro/books/about/Statistical Power Analysis for the Behav.html?id=gA04ngAACAAJ&redir_esc=y
- George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference* (4. baskı). Allyn & Bacon. Retrieved from https://books.google.ro/books/about/SPSS for Windows Step by Step.html?id=AghHAAAAMAAJ&redir_esc=y
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125-132. <https://doi.org/10.2307/327314>
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Pergamon Press. Retrieved from https://books.google.ro/books/about/Principles and Practice in Second Language.html?id=p9toAAAIAAJ&redir_esc=y
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3. baskı). McGraw-Hill. Retrieved from https://books.google.ro/books/about/Psychometric Theory 3E.html?id=6Rf3G58JsC&redir_esc=y
- Young, D. J. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal*, 75(4), 426-437. <https://doi.org/10.2307/329492>