Language Teaching Research Quarterly **EUROKD**

2026 (56), 207–231
https://doi.org/10.32038/ltrq.2026.56.10

Adaptation and Validation of the Portuguese Foreign Language Listening Anxiety Scale (FLLASp) for EFL Contexts

Luciana Cabral P. Bessa
Department of Letters, Arts and Communication, University of Trás-os-Montes e Alto Douro, Portugal
Centro de Investigação Transdisciplinar «Cultura, Espaço e Memória», Universidade do Porto, Portugal
Guilherme Gonçalves*
Department of Engineering, University of Trás-os-Montes e Alto Douro, Portugal
HUMANISE, Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Portugal
Elisabete Mendes Silva
Department of Foreign Languages, Bragança Polytechnic University, Portugal
Transdisciplinary Research Centre in Education and Development (CITeD), Bragança Polytechnic University, Portugal
Miguel Melo
Department of Engineering, University of Trás-os-Montes e Alto Douro, Portugal
HUMANISE, Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Portugal
Maximino Bessa
Department of Engineering, University of Trás-os-Montes e Alto Douro, Portugal
HUMANISE, Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Portugal

Correspondence
Email: guilhermeg@utad.pt

Abstract

Listening in a foreign language is often one of the most anxiety-inducing language skills because it is transient and processed in real time. This study adapted and validated the Portuguese Foreign Language Listening Anxiety Scale (FLLASp) for English as a Foreign Language learners in Portugal. Using back-translation and expert review, the original instrument was culturally and linguistically adapted and administered to 272 undergraduates from two public higher education institutions; after removal of multivariate outliers, analyses were based on 261 cases. Confirmatory factor analyses compared the original two-factor model,

ARTICLE HISTORY
Received: 08 July 2026
Revised: 22 October 2026
Accepted: 29 October 2026

KEYWORDS
Foreign Language Listening Anxiety, Psychometrics, Portuguese Context, Second

©The Author(s), 2026. Published by EUROKD. This article is licensed under a Creative Commons Attribution 4.0 International License, which grants permission to use, share, adapt, distribute and reproduce in any medium or format provided that proper credit is given to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if any changes were made. www.EUROKD.com

Kimura's (2008) three-factor model, and a unidimensional solution. The Portuguese data fit Kimura's (2008) three-factor structure best. Model refinement removed two weak reverse-coded items, yielding a final 24-item scale with improved fit (CFI = 0.937, TLI = 0.930, RMSEA = 0.099, SRMR = 0.056), strong internal consistency, and satisfactory convergent validity. The findings support the FLLASp as a reliable and valid multidimensional measure of foreign language listening anxiety in Portuguese EFL university contexts.

Language Teaching, Factor Analysis

How to cite this article (APA 7th Edition):
Bessa, L. C. P., Gonçalves, G., Silva, E. M., Melo, M., & Bessa, M. (2026). Adaptation and validation of the Portuguese foreign language listening anxiety scale (FLLASp) for EFL contexts. *Language Teaching Research Quarterly*, *56*, 207–231. https://doi.org/10.32038/ltrq.2026.56.10

Introduction

Proficiency in foreign languages is essential for effective communication in an increasingly globalised society and is associated with cognitive, academic, and intercultural benefits (Fox et al., 2019). In contemporary higher education and employment contexts, English as a Foreign Language (EFL) is particularly important because it remains central to global academic, social, and professional interaction (Erdem & Tutkun, 2016; Ethnologue, 2025).

In EFL teaching and learning, identifying the challenges students face is crucial for improving outcomes and refining pedagogical practice. Among the affective variables involved in second language learning, anxiety has been one of the most extensively studied and is widely recognised as a barrier to language learning and academic performance (Botes et al., 2020; Horwitz et al., 1986; Kralova & Petrova, 2017; MacIntyre & Gardner, 1989; Papi & Khajavy, 2023). In this context, foreign language anxiety refers to emotional distress that arises when learners are required to learn, perform, or communicate in another language, and it may hinder performance and learners' engagement with the target language (Horwitz et al., 1986; Papi & Khajavy, 2023).

Listening is a particularly important focus because it is central to communication and often perceived as one of the most demanding language skills (Ji et al., 2022; Lei et al., 2023; Vandergrift, 2007). Unlike reading, listening typically unfolds in real time and offers learners limited control over input, which can intensify the difficulty posed by factors such as speed, accent variation, background noise, and unfamiliar vocabulary (Riddell, 2014; Vandergrift & Goh, 2012). These demands help explain why listening is frequently associated with tension, apprehension, and worry in EFL settings (Ji et al., 2022; Vogely, 1998).

Foreign language listening anxiety, therefore, refers to the tension and apprehension experienced when learners process spoken input under real-time constraints and limited control over the message (Horwitz et al., 1986; Kim, 2000; Vogely, 1998). Because linguistic and cultural context can affect how this construct is measured, simple translation of an instrument is insufficient; careful adaptation and psychometric

validation are required to establish whether a scale retains validity in a new context (Beaton et al., 2000; Ji et al., 2022). This study addresses the lack of research on EFL listening anxiety among Portuguese university students by adapting and validating the FLLASp, with particular attention to its factor structure in the Portuguese context.

State of the Art

According to Horwitz et al. (1986), anxiety is a subjective feeling of tension, apprehension, nervousness, and worry. More recent literature maintains a similar view, describing anxiety as a mental and physical state expressed through emotional, physical, and behavioural symptoms (Kralova & Petrova, 2017). In language learning research, it has largely been treated as a situation-specific affective variable rather than as a general clinical condition.

In second and foreign language learning, anxiety has been conceptualised as a situation-specific affective factor that may interfere with learners' classroom performance and language-learning experience. Horwitz et al. (1986) conceptualised foreign language anxiety as a distinct construct involving learners' self-perceptions, beliefs, feelings, and behaviours related to classroom language learning. Within this broader field, foreign language listening anxiety gradually emerged as a skill-specific construct, particularly from the late 1990s onwards, with studies examining its reported sources and pedagogical implications (Vogely, 1998), its measurement through instruments such as the Foreign Language Listening Anxiety Scale (FLLAS) (Kim, 2000), and its relationship with listening comprehension and performance (Elkhafaifi, 2005).

Foreign language listening anxiety can be understood as a form of situation-specific anxiety, involving tension, apprehension, and worry (Horwitz et al., 1986), experienced when learners process spoken input in a foreign language under real-time conditions and limited control over the message (Kim, 2000; Vogely, 1998). In EFL contexts, listening may be particularly demanding because learners need to process spoken input as it unfolds in real time (Vandergrift & Goh, 2012, p. 37), through complex cognitive and metacognitive processes (Vandergrift & Goh, 2012, p. 17), while affective variables may also influence listening success (Vandergrift & Goh, 2012, p. 97). This difficulty may be intensified by features commonly associated with classroom and authentic listening, such as speech rate, accent variation, unfamiliar vocabulary, reduced forms, and background noise (Riddell, 2014). These concerns led to the development of instruments to measure listening anxiety among foreign language learners, including the FLLAS, originally developed by Kim (2000).

Subsequent studies refined the conceptualisation and measurement of foreign language listening anxiety. Kimura (2008), for example, identified factors such as emotionality, worry, and anticipatory fear, while Yamauchi (2014) proposed a revised structure with

more differentiated anxiety sources. Other studies examined listening anxiety in relation to listening comprehension and performance (Elkhafaifi, 2005; Zhang, 2013).

Anxiety in Foreign Language Learning and Listening

In the context of foreign language teaching and learning, acknowledging anxiety matters because it can inform pedagogical decisions and classroom strategies that improve engagement and achievement. Research on language anxiety has long linked anxious classroom experience to performance outcomes, course grades, and learner self-perception (Horwitz et al., 1986; MacIntyre, 1992; Zhang, 2019). Building on this line of work, other broader classroom measures, such as the Foreign Language Classroom Anxiety Scale (FLCAS) (Lee & Ye, 2023), were developed. Later, more focused research on anxiety associated with particular language skills and contexts (Bárkányi & Brash, 2025; Ghorbanpoor et al., 2026; Özdemir & Seçkin, 2025) was also produced.

More recently, Schimpff and Botes (2026) conducted a broader study that went beyond the traditional classroom setting and the traditional foreign language learners. The authors developed a Context Inclusive Language Skill Anxiety Scale (CILSAS) to measure anxiety in all four skills - reading, writing, listening and speaking in both classroom and non-classroom contexts. The scale was validated with traditional foreign language learners, heritage language speakers and online language learners.

Within this broader field, listening anxiety has been conceptualised as a situation-specific form of anxiety associated with the possibility of misunderstanding spoken language under real-time processing demands (Ji et al., 2022). It may function both as a reaction to difficulty and as a factor shaping subsequent performance, thereby making it theoretically and pedagogically relevant (Ji et al., 2022; Zhang, 2013). Because listening comprehension requires rapid integration of linguistic and non-linguistic cues, limited control over input, and tolerance of ambiguity, learners may experience worry, confusion, and tension during listening tasks (Chang, 2008; Riddell, 2014; Sari & Fithriyana, 2019). Measuring EFL listening anxiety among Portuguese university students is therefore scientifically and pedagogically relevant, particularly when the instrument is theoretically grounded, clearly operationalised, and supported by appropriate validity evidence.

The Foreign Language Listening Anxiety Scale (FLLAS)

FLLAS is a 33-item instrument developed by Kim (2002) to measure the anxiety experienced by learners during foreign language listening tasks. It was originally introduced in the author's PhD thesis (Kim, 2000). The scale was created as a central component of a doctoral study that examined the relationship between foreign language listening and anxiety in a sample of 245 Korean university students studying English as a Foreign Language (EFL) (Kim, 2002). The study's aim was to provide empirical evidence

for the existence of listening-specific anxiety and develop a reliable instrument for its measurement.

Through factor analysis, Kim (2002) identified two primary constructs that constitute the FLLAS:

- *Tension and worry over English listening*, which captures the cognitive and affective stress associated with processing the spoken language in real-time.
- *Lack of self-confidence in listening*, which reflects learners' negative self-perceptions and beliefs regarding their own listening capabilities.

The study established the FLLAS's construct validity by demonstrating that scores were significantly correlated with both general foreign language anxiety and objective listening proficiency, as measured by a listening Test of English as a Foreign Language (TOEFL). The results indicated a weak negative correlation between Foreign Language Listening Anxiety (FLLA) and TOEFL listening scores, suggesting that higher levels of anxiety were associated with lower listening performance. Furthermore, a multiple regression analysis revealed that the Lack of self-confidence in listening factor was the single strongest predictor of listening proficiency, highlighting the critical role of learners' self-perception in performance (Kim, 2002).

Although lack of self-confidence is not equivalent to anxiety in a strict psychological sense, it has been treated in L2 research as a self-perceptual component associated with affective responses to language use. In the broader social-psychological tradition, L2 self-confidence has been linked to perceived communicative competence and anxiety (Clément et al., 2003). In this study, lack of self-confidence is interpreted as a self-perceptual component closely associated with listening anxiety, but not as anxiety per se within the psychological definition of anxiety as tension, apprehension, nervousness, and worry (Horwitz et al., 1986). In conclusion, Kim's (2002) work supported that foreign language listening anxiety is a distinct and measurable construct. The FLLAS provides a tool to identify and assess this anxiety (Kim, 2002).

Kim's (2000) FLLAS Adaptations

Given the influence of Kim's (2002) FLLAS, the instrument has been adapted across various linguistic contexts and empirically revised to investigate its latent structure. While some researchers have attempted to replicate Kim's factor structure, others have modified the item pool or treated the scale as unidimensional.

In a significant structural validation study, Kimura (2008) administered a Japanese translation of the FLLAS (FLLAS-J) to 452 Japanese university freshmen. The original 33-item scale was translated via a double-translation method. Following a Principal-Axis Factoring analysis, Kimura removed seven items that loaded weakly or double-loaded, reducing the scale to 26 items ($\alpha = 0.91$). This analysis yielded a distinct factor structure

compared to the original. Whereas Kim (2002) identified two factors ("Tension and Worry" and "Lack of Confidence"), Kimura (2008) identified a three-factor solution: "Emotionality," "Worry," and "Anticipatory Fear." Notably, the "Emotionality" component emerged as the primary factor in the Japanese context, whereas it was secondary in Kim's original study. The label "Anticipatory Fear" should be interpreted with caution, as fear and anxiety are conceptually distinct constructs in psychology (Bourne, 2010). Although Kimura (2008) used the label "Anticipatory Fear" in the context of foreign language listening anxiety, this factor is more appropriately understood as anticipatory apprehension related to listening situations, consistent with definitions of anxiety as tension, apprehension, nervousness, and worry (Horwitz et al., 1986). Several studies have used the FLLAS with modifications that render them distinct from the original instrument. Wang (2010) administered a Chinese adaptation to 125 English majors. However, the researcher explicitly adapted, deleted, and added items to suit the research context, resulting in a 35-item scale that included questions regarding classroom behaviours (e.g., daydreaming) and exam anxiety. Wang classified these items into three categories: "Tension and Worry," "Lack of Confidence," and "Prior Knowledge". Due to the alteration of the item pool, this structure is not statistically directly comparable to the standard 33-item FLLAS.

Similarly, Yamauchi (2014) conducted a large-scale study (N = 1435) with Japanese university students to develop a revised version of the scale. Arguing that negative wording distorts correlation coefficients, Yamauchi (2014) reworded all reverse-scored items into positive statements, deleted three original items due to conceptual ambiguity, and added five new items specific to the Japanese context (resulting in a 35-item instrument). Using both Exploratory and Confirmatory Factor Analysis, Yamauchi proposed a complex six-factor model grouped into three categories: Context (Real-life vs. Classroom), Material (Lack of Knowledge vs. Difficulty), and Cognitive Processes (Bottom-up vs. Metacognitive). While methodologically rigorous, the fundamental alteration to the items limits comparisons with the standard FLLAS.

In a more recent advancement of the measurement of listening anxiety within the Japanese context, Kimura (2017) revisited the construct using a self-presentational framework. Unlike their previous work, which validated a 26-item structure (Kimura, 2008), this study utilised a large sample (N = 1177) to validate the "Shortened Scale of Foreign Language Listening Anxiety" (SSFLLA). This 27-item instrument was derived by removing seven non-loading items from the original scale, adding one new item regarding social interactions, and crucially, rewriting all negatively worded items into positive statements to avoid methodological artefacts. Using Rasch analysis, Kimura (2017) proposed that listening anxiety is a higher-order construct comprising two distinct dimensions: "Self-Focused Apprehension" (concern about social evaluation) and "Task-Focused Apprehension" (worry about processing aural input). Although this work offers a large sample size, the modification of item wording (positive vs. negative) and the

alteration of the item pool make this work unsuitable for direct statistical comparison with the standard 33-item FLLAS.

Conversely, other researchers have employed the FLLAS as a unidimensional measure without attempting to extract latent factors. Serraj and Noordin (2013) administered the 33-item FLLAS to 210 Iranian EFL students. Although they did not perform factor analysis, their study provided evidence of criterion validity, reporting that FLLA was negatively correlated with listening comprehension ($r = -0.414$) and moderately correlated with general Foreign Language Anxiety ($r = 0.513$). The authors concluded that FLLA imposes a specific and problematic burden on listening performance distinct from general anxiety.

Likewise, Capan and Karaca (2013) investigated 159 Turkish majors using the original English version of the scale. Treating the FLLAS as a unidimensional construct, they analysed the total scores in relation to reading anxiety. Their findings indicated a moderate positive correlation between listening and reading anxiety ($r = 0.327$), supporting the view that while listening anxiety shares commonality with other language anxieties, it functions as a distinct construct.
Table 1 provides a comprehensive summary of the main FLLAS adaptations discussed in the literature, detailing their key characteristics and psychometric properties.

Research Gap and Objectives

The absence of a validated instrument to assess foreign language listening anxiety in European Portuguese highlights the need to adapt and validate a scale suitable for use with university students in Portugal. Moreover, considering that Portuguese is one of the most widely spoken languages in the world, the availability of a validated instrument for this specific population is of significant relevance. Accordingly, given the wide adoption of and strong support for Kim's (2002) FLLAS in the literature (Ji et al., 2022), we selected this instrument for cultural–linguistic adaptation to European Portuguese and validation in a university context.

This study also aims to provide data that can inform didactics and Applied Linguistics research on listening-specific EFL teaching. By identifying the anxiety experiences learners report while listening to English, the study can help support more appropriate classroom strategies and targeted pedagogical intervention.

Table 1

Comparative Analysis of Factor Structures and Methodologies in FLLAS Studies (Adapted from Kim (2002))

Study	Sample	Language	Items	α	Methodology	Factor Structure	CFA Suitability
Kim (2002)	245 Korean	Korean	33 (original)	0.93	PCA (Varimax)	Factors: 1.Tension & Worry 2.Lack of Confidence	Yes
Kimura (2008)	452 Japanese	Japanese (Trans+Back)	26	0.91	Principal Axis (Varimax)	Factors: 1.Emotionality 2.Worry 3. Anticipatory Fear	Yes
Wang (2010)	125 Chinese	Chinese (Adapted)	35 [a]	0.84	Descriptive (No EFA)	Categories: 1.Tension/Worry 2.Confidence 3. Prior Knowledge	No (Incompatible items added/modified)
Serraj and Noordin (2013)	210 Iranian	English (Modified)	33	0.79	Correlation (No EFA)	Unidimensional (Total Scores only)	Yes (Test as a 1-factor model baseline)
Capan and Karaca (2013)	159 Turkish	English (original)	33	None [d]	Correlation (No EFA)	Unidimensional (Total Scores only)	Yes (Test as a 1-factor model baseline)
Yamauchi (2014)	1435 Japanese	Japanese (revised)	35 [b]	0.90 [c]	EFA & CFA	6 Factors: (Grouped into 3 categories: Context, Material, Cognitive)	No (Reverse items reworded to positive; new items added.)
Kimura (2017)	1177 Japanese	Japanese	27 (SSFLLA) [e]	None [f]	Rasch Model	Dimensions: 1.Self-Focused Apprehension 2.Task-Focused Apprehension	No (Reverse items reworded to positive; 1 item added)

Note. Adapted and expanded from Kim (2002).

a. Wang (2010) added items regarding classroom behaviour and testing (e.g., "I often daydream") and performed adaptations to existing items. Items that cross-loaded between factors (Item 13) were also reported.

b. Yamauchi (2014) altered the psychometrics by rewriting negative items into positive statements. Added 5 new items specific to the Japanese context.

c. Alpha reported for the final 18-item model derived from the 35-item pool.

d. Reliability was not reported for the specific sample.

e. Kimura (2017) "Shortened Scale of FLLA." Removed 7 items, added 1 new item, and reworded negative items to be positive.

f. No Cronbach's *α* reported.

Method

This study draws on data collected from Portuguese undergraduate students enrolled in English language course units at the University of Trás-os-Montes and Alto Douro and the Bragança Polytechnic University during the 2023/2024 and 2024/2025 academic years. The primary objective is to adapt and validate the FLLASp within the Portuguese context, using the original items developed by Kim (2002) as a reference. The study examines the scale's psychometric properties, particularly its reliability and statistical validity. To this end, we aim to assess whether the FLLAS retains its original factor structure in the Portuguese version, given findings from previous research reporting structural variations across different translations. The sample consists of university students attending foreign language classes, specifically English, consistent with the participant profiles in earlier FLLAS validation studies.

Sample

The participants were Portuguese students aged 18–58, with a mean age of 20 years old. In general, they had prior experience learning English, primarily acquired during secondary education. The dataset comprises responses from 272 foreign language learners enrolled in English courses at two Portuguese higher education institutions: the University of Trás-os-Montes and Alto Douro and the Bragança Polytechnic University. We collected responses from 4 A1, 9 A2, 58 B1, 108 B2, 58 C1, and 35 C2 level students (self-reported), according to the Common European Framework of Reference for Languages (CEFR). All participants took part in the study voluntarily.

Instrument

The original FLLAS developed by Kim (2002) and used in this study was translated into European Portuguese by members of the research team who are also EFL teachers. Following procedures adopted in previous FLLAS validation studies, back translation was used to verify semantic equivalence, and any discrepancies were solved by consensus. Additionally, some wording adjustments were made to guarantee linguistic clarity and cultural appropriateness to the Portuguese context, while preserving the original meaning, item order, reverse-coded items, and response format. The full translated FLLASp questionnaire can be consulted in Table A1.

Procedures

Students were informed about the purpose of the study, particularly the importance of analysing and identifying potential sources of anxiety and constraints that may affect their overall learning performance and assessment in EFL classes. Students completed an online questionnaire during class using a link shared by their EFL teachers. They were instructed to respond to each questionnaire item carefully, to ensure the accuracy and reliability of the data collected. It was made clear that participation was anonymous and that no individual responses would be linked to specific students. Additionally, students

were assured that all collected data would be used exclusively for research purposes within the scope of this study. This study was conducted in accordance with the ethical principles of the World Medical Association Declaration of Helsinki. Prior to data collection, all participants provided free and informed consent. Participation was voluntary and anonymous, participants retained the right to withdraw at any time without penalty, and measures were taken to ensure confidentiality and protection against harm. The questionnaires were administered between the beginning of the second semester of the 2023/2024 academic year and the second semester of 2024/2025, during EFL classes at the University of Trás-os-Montes and Alto Douro and Bragança Polytechnic University.

The study participants were enrolled in various undergraduate programmes, including BA degrees in:

- Culture and Digital Transformation – English B2.2
- Languages, Literatures, and Cultures – English B2.2 and C1.2
- Languages and Business Relations – English B2.1, C1.1 and C2.1
- Languages for International Relations – English B2.2, C1.2, C2.1
- English translating practices – English B1 to B2
- Foreign Languages: English and Spanish – English B2.2, C1.2 and C2.1
- Tourism – English B2

No questions or concerns were raised by students regarding the interpretation of the questionnaire items, suggesting that the Portuguese translation of the original English version was perceived to be well understood.

Statistical Analysis Procedure

Recent systematic evidence shows that the dimensional structure of foreign language listening anxiety instruments, including the FLLAS, has not been stable across replications and languages, with re-analyses frequently yielding different factor solutions or shifting toward source-of-anxiety and ability-related constructs rather than a single invariant model (Ji et al., 2022).

Given the existence of multiple adaptations of the FLLAS with varying factor structures across different linguistic and cultural contexts, in this study we tested through CFA how our data fits in each of the proposed factor structures from previous adaptations and the original scale.

We then followed up by selecting the factor structure that best fit our data and refined the model by removing items with low contributions to the latent factors.

Results

Statistical analyses were performed using R (RStudio v.2024.04.2), JASP v0.19.3, and SPSS Statistics 27. Four items (6 R, 14 R, 25 R, 31 R) were reverse-coded prior to analysis. Univariate normality for individual items and subscales was examined via skewness and kurtosis and met conventional criteria (|Skewness| < 3 and |Kurtosis| < 10); observed skewness ranged from 1.623 to −0.353 and kurtosis from 2.589 to −1.251, suggesting little deviation from normality (Campos et al., 2020; Kline, 2023).

A Mahalanobis distance test in SPSS was used to identify multivariate outliers using the chi-square cumulative distribution function with 33 degrees of freedom and a $p < 0.001$ to identify cases as potential outliers was used (Tabachnick et al., 2019). This resulted in finding 11 multivariate outliers which were removed, yielding a final analysis sample of 261 participants. Multivariate normality was then evaluated using Mardia's test (Mardia, 1970) and the energy (E-statistic) test in R (Korkmaz et al., 2014); both tests indicated significant departures from multivariate normality ($p < 0.001$). Internal consistency of the 33-item scale was assessed using Cronbach's α, which yielded $\alpha = 0.957$ (95% CI [0.949, 0.963]).

Prior to conducting factor analysis, assumptions were evaluated using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy and Bartlett's test of sphericity (computed in JASP). We considered all items for this test. The overall KMO was 0.934, indicating excellent sampling adequacy (Tabachnick et al., 2019). Bartlett's test of sphericity was significant, $\chi2(528) = 7676.768$, $p < 0.001$, indicating that factor analysis was appropriate.

We then proceeded to conduct a CFA to test the fit of the data from our Portuguese sample to the original two-factor structure proposed by Kim (2002) (model was tested using the items retained in the published two-factor structure), the three-factor structure identified by Kimura (2008), and the unidimensional structure opted for by Serraj and Noordin (2013) and Capan and Karaca (2013).

The Weighted Least Squares Mean and Variance adjusted (WLSMV) estimator with robust standard errors was used, which is robust to non-normality and appropriate for ordinal data (Li, 2016). The CFI, TLI, and RMSEA reported below are the corresponding robust fit statistics for this ordinal-data CFA.

Fit indices used to evaluate model fit included the Chi-square statistic ($\chi2$), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA) with 90% confidence intervals, and Standardised Root Mean Square Residual (SRMR). Conventional cutoffs for acceptable fit were applied: CFI and TLI values above 0.90 were considered to indicate acceptable fit (Bentler, 1990; Bentler & Bonett, 1980), while SRMR values below 0.08 indicated good fit (Cho et al., 2020; Hu & Bentler, 1999).

Additionally, RMSEA values below 0.05 were interpreted as indicating close fit (Browne & Cudeck, 1993). Table 2 summarises the goodness-of-fit indices for each tested model. We then tested the item reliability and convergent validity of each model by calculating Cronbach's alpha (α), McDonald's (ω), and Average Variance Extracted (AVE) for each factor. Based on Hair et al. (2010) the supported thresholds are Cronbach's alpha values above 0.70 (or 0.60 for exploratory research) and AVE values above 0.50. Discriminant validity was assessed using the Heterotrait-Monotrait ratio of correlations (HTMT), with values below 0.85 indicating adequate discriminant validity (Henseler et al., 2015). For this work, we considered loadings above 0.50 as acceptable (Hair et al., 2010).

From the results presented in Table 2 and Table 3, we can observe that, overall, the three-factor model proposed by Kimura (2008) demonstrated the best fit to our Portuguese sample data, with CFI = 0.920, TLI = 0.912, RMSEA = 0.104 (90% CI [0.097, 0.110]), and SRMR = 0.064. It also exhibited satisfactory AVE values (> 0.5 except for Emotionality factor with a value of 0.497), strong internal consistency (α and ω > 0.8 for all factors), and adequate discriminant validity (HTMT < 0.85 for all factor pairs). The two-factor model from Kim (2002) also showed acceptable fit, but with slightly lower indices (CFI = 0.906, TLI = 0.893, RMSEA = 0.132, SRMR = 0.072). However, it had lower AVE values, as low as 0.353 for Lack of Confidence, and weaker item reliability for the same factor. The unidimensional model exhibited similar fit to Kim (2002) model, although with better RMSEA. The AVE of the unidimensional scale was 0.525, with good item reliability (α = 0.957, ω = 0.986).

Given these results, we proceeded with Kimura (2008) factor structure. To refine it, we removed items with loadings below 0.40. As shown in Table 4, this resulted in the removal of 2 reversed-coded items, both on the Emotionality factor (31 R and 6 R) (Table 4), yielding a final 24-item scale with improved fit indices: CFI = 0.937, TLI = 0.930, RMSEA = 0.099 (90% CI [0.092, 0.106]), and SRMR = 0.056 (Table 2). Although the RMSEA remained above the traditional close-fit threshold, the remaining evidence from CFI, TLI, and SRMR support an acceptable fit. All individual item loadings in this 24-item model were statistically significant ($p < 0.001$) with loadings above 0.5 (Table 4). Furthermore, the AVE values improved for the Emotionality factor (0.578), further improving item reliability for Emotionality (α = 0.906, ω = 0.941), while Worry and Anticipatory Fear factors maintained strong reliability and validity metrics. Discriminant validity remained adequate (HTMT < 0.85) except between Emotionality and Anticipatory Fear (HTMT = 0.854) (Table 3), with a value slightly above the threshold.

Consistent with these findings, the standardised factor covariances revealed strong significant correlations between the latent factors ($p < 0.001$): Emotionality was highly correlated with Worry ($r = 0.886$) and Anticipatory Fear ($r = 0.877$), while Worry and Anticipatory Fear showed a correlation of $r = 0.819$.

Table 2

Comparison of Goodness-of-Fit Indices for the Tested Factor Structures

Model	χ2	df	p	CFI	TLI	RMSEA	90% CI	SRMR
Kim (2002) (2 Factors)	738.247	134	<0.001	0.906	0.893	0.132	[0.122,0.141]	0.072
Kimura (2008) (3 Factors)	1122.985	296	<0.001	0.920	0.912	0.104	[0.097,0.110]	0.064
Revised Kimura (2008) (3 Factors)	878.518	249	<0.001	0.937	0.930	0.099	[0.092,0.106]	0.056
Unidimensional (1 Factor) [a]	1866.572	495	<0.001	0.900	0.893	0.103	[0.098, 0.108]	0.073

Note. n = 261. χ2 = chi-square; df = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual.

This model tests the unidimensional structure adopted by Serraj and Noordin (2013) and Capan and Karaca (2013), who used a single total score for the instrument.

Table 3

Reliability, Convergent Validity, and Discriminant Validity (HTMT) of the Models

Model	Factor	α	ω	AVE	HTMT
Kim (2002)	F1. Tension & Worry	0.922	0.940	0.636	F1 – F2: 0.884
	F2. Lack of Confidence	0.721	0.705	0.353	
	Total Scale	0.915	0.948	-	
Kimura (2008)	F1. Emotionality	0.880	0.896	0.497	F1 – F2: 0.847
	F2. Worry	0.927	0.941	0.618	F1 – F3: 0.820
	F3. Anticipatory Fear	0.831	0.840	0.799	F2 – F3: 0.796
	Total Scale	0.947	0.977	-	-
Revised Kimura (2008)	F1. Emotionality	0.906	0.941	0.578	F1 – F2: 0.840
	F2. Worry	0.927	0.941	0.618	F1 – F3: 0.854
	F3. Anticipatory Fear	0.831	0.840	0.799	F2 – F3: 0.796
	Total Scale	0.955	0.981	-	-
Unidimensional	F1. Listening Anxiety (General)	0.957	0.986	0.525	n/a

Note. n = 261; α = Cronbach's alpha; ω = McDonald's omega; AVE = average variance extracted; HTMT = heterotrait-monotrait ratio of correlations. HTMT values below 0.85 indicate established discriminant validity. The Unidimensional model does not have HTMT values as there are no distinct latent factors to compare. The Kim (2002) total-scale reliability is computed only over the 18 items retained in the published two-factor structure.

Table 4

Item Loadings Across Kim, Kimura, Revised Kimura, and Unidimensional CFA Models

Item	Kim (2002)		Kimura (2008)			Kimura (2008) Revised			Uni.
Factor	F1	F2	F1	F2	F3	F1	F2	F3	F1
1		**0.591**		**0.624**			**0.625**		0.609
2				**0.747**			**0.747**		0.724
3									0.697
4				**0.815**			**0.816**		0.795
5	**0.809**			**0.844**			**0.844**		0.813
6_R		0.283	0.282						0.274
7				**0.756**			**0.757**		0.733
8									0.708
9				**0.78**			**0.78**		0.758
10		**0.657**	**0.718**			**0.718**			0.713
11			**0.818**			**0.818**			0.796
12			**0.851**			**0.848**			0.847
13									0.822
14_R		**0.574**	**0.599**			**0.591**			0.583
15			**0.812**			**0.811**			0.786
16	**0.803**			**0.82**			**0.82**		0.788
17	**0.872**		**0.889**			**0.887**			0.864
18	**0.798**			**0.818**			**0.817**		0.79
19		**0.73**	**0.738**			**0.737**			0.717
20		**0.596**	**0.603**			**0.603**			0.59
21		**0.863**		**0.86**			**0.859**		0.826
22									0.848
23	**0.762**			**0.776**			**0.775**		0.755
24	**0.676**								0.664
25_R									0.205
26	**0.797**				**0.841**			**0.841**	0.77
27	**0.877**				**0.944**			**0.944**	0.847
28									0.797
29	**0.795**		**0.816**			**0.816**			0.811
30			**0.777**			**0.777**			0.738
31_R		0.046	0.042						0.046
32			**0.698**			**0.698**			0.685
33	**0.764**			**0.78**			**0.78**		0.747

Note. Uni. = unidimensional model

Loadings > 0.50 are marked in bold.

Following the scoring method described in Appendix A, based on the revised Kimura three-factor structure, we computed the three subscale means and the total score. Descriptive statistics are reported in Table 5, with the scores illustrated in Figure 1. Mean scores were 2.135 (SD = 0.753) for Emotionality, 2.872 (SD = 0.945) for Worry, and 2.165

(SD = 1.006) for Anticipatory Fear. The mean Total FLLASp score was 2.475 (SD = 0.799).

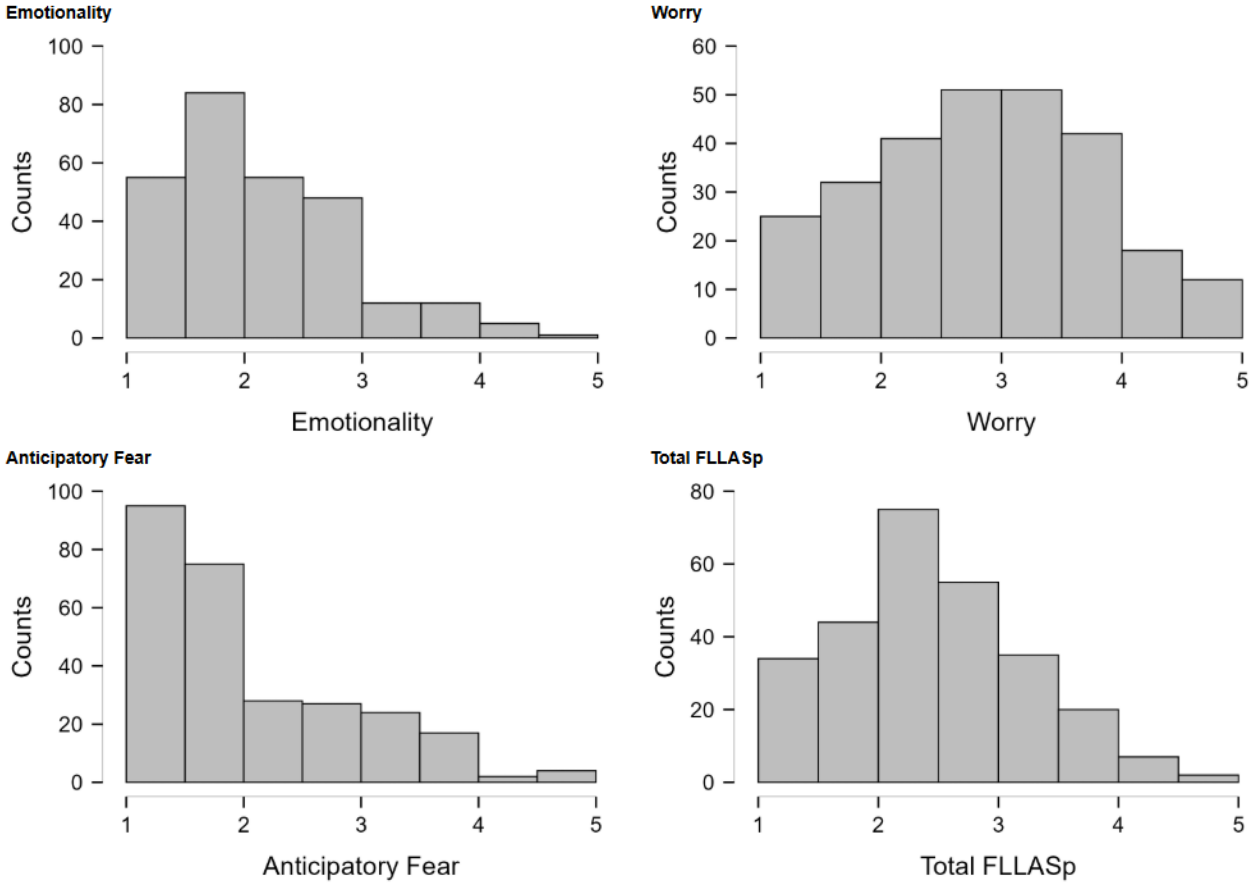
Table 5

Descriptive Statistics for FLLASp Subscale and Total Scores

Statistic	Emotionality	Worry	Anticipatory Fear	Total FLLASp
Median	2.000	2.909	2.000	2.417
Mean	2.135	2.872	2.165	2.475
Std. Deviation	0.753	0.945	1.006	0.799
Skewness	0.818	-0.019	0.642	0.354
Kurtosis	0.763	-0.644	-0.320	-0.188

Figure 1

Distribution of FLLASp Subscale and Total Scores, Calculated According to the Scoring Method Described in Appendix A



Note. Scores range from 1 to 5, with higher values indicating greater listening anxiety.

Discussion

The psychometric analysis provided support for the reliability and validity of the Portuguese version of the FLLAS, although some findings should be interpreted with caution. Of the three models examined (the original two-factor structure, Kimura's (2008) three-factor model, and a unidimensional solution) the three-factor model showed the best fit to the Portuguese data. In this context, foreign language listening

anxiety therefore appears to be better conceptualised as a multidimensional construct encompassing Emotionality, Worry, and Anticipatory Fear.

The Portuguese dataset fit the three-factor structure proposed by Kimura (2008) better than the competing solutions. The retained factors, Emotionality, Worry, and Anticipatory Fear, were supported by acceptable fit and by satisfactory reliability, convergent validity, and discriminant validity after refinement to a 24-item version. In practical terms, this indicates that the instrument functions as a coherent multidimensional measure whose factors are related but sufficiently distinct to be interpreted separately in the Portuguese EFL context.

The three-factor structure as referred has therefore practical implications because it allows teachers to move beyond a possibly ordinary idea that a learner is "anxious" and to identify more appropriately specific patterns of difficulty which may help to benefit both teaching and learning processes.

More specifically, the three-factor structure may help teachers adapt and apply tailored didactic strategies that are thought of to tackle the very problematics resulting from Emotionality; clearer strategies and task preparation for Worry; and progressive familiarisation with demanding listening situations for Anticipatory Fear, although these applications require further empirical validation.

A noteworthy finding in the refinement process was the consistent under-performance of the reverse-coded items 6_R ("It is easy to guess about the parts that I miss...") and 31_R ("English stress and intonation sound familiar..."). These items exhibited low factor loadings across all tested structures: the original Kim (2002) model, the unidimensional model, and the Kimura (2008) model. Furthermore, item 25_R ("I have no fear of listening in English as a member of an audience") was also weak in the unidimensional structure and had already been problematic in prior work (Kim, 2002; Kimura, 2008).

The poor fit of these items is likely attributable to a combination of methodological artefacts and semantic ambiguity. Reverse-coded items can produce a method effect in anxiety scales because they require a shift in response framing that may introduce inconsistency (Woods, 2006; Yamauchi, 2014). However, this alone cannot explain the pattern, since item 14_R loaded adequately across all tested structures.

The content of these problematic items may have been interpreted by Portuguese students as distinct from anxiety. Item 6_R ("It´s easy to guess about the parts that I miss while listening to English") refers to "guessing", which may have been perceived as measuring a learning strategy or linguistic competence rather than an emotional state in Kimura (2008) factor structure, or as a lack of confidence in Kim (2002) structure.

Notably, this item may be interpreted as tapping into a guessing strategy during listening, rather than a purely emotional reaction.

Similarly, Item 31_R may have been interpreted as a measure of familiarity with prosody or prior exposure rather than an affective response. By contrast, item 14_R refers directly to confidence while listening, which likely explains why it remained interpretable as an anxiety-related indicator despite being reverse coded.
Removing these items, therefore, reduced measurement noise and produced a cleaner instrument for assessing Emotionality, Worry, and Anticipatory Fear in the Portuguese sample.

In addition to supporting the reliability and validity of the adapted instrument, these results reinforce the importance of conceptualising listening anxiety as a multidimensional construct that may vary across cultural and linguistic contexts. Recognising this variability reinforces the acknowledgement of diverse learner experiences and contexts. Furthermore, by adjusting the FLLAS scale to the Portuguese context, we have managed to narrow the broad nature of its anxiety constructs. Furthermore, the scale, as presented, showcases the idiosyncrasy of specific elements of the cultural and emotional factors, helping to shed some light on cultural issues that may interfere with the teaching and learning experiences in the Portuguese context. Therefore, teachers will be able to find remedial strategies that are thought of to tackle cultural bias as well as emotional related issues.

The revised scale offers a more useful pedagogical tool for monitoring students' affective responses over time, evaluating didactic strategies, and identifying areas where listening activities may need adjustment, such as vocabulary preparation or context setting. The high correlations among factors are also theoretically relevant, as they support models that treat anxiety as an active component of input processing in second language learning. Overall, the adapted 24-item FLLASp provides a practical instrument for EFL teaching and a clearer empirical basis for future research in Portuguese-speaking contexts. This study has important implications for policy and teaching practice in the context of formal language learning in Portugal. Specifically, the findings emphasise the significance of recognising listening anxiety as a multidimensional factor that can impact learners' engagement and performance. It also highlights the necessity of incorporating emotional support into language curricula and teacher development initiatives. Additionally, the findings are aligned with relevant previous studies, namely the works from Bourne (2010), Horwitz et al. (1986), Kim (2002), and Kimura (2008), which research the impact of anxiety on foreign language learning. From all the tested scales regarding listening anxiety in a foreign language, Kimura´s sample is the one that most closely resembles the Portuguese case. From a pedagogical perspective, understanding the three-factor structure of listening-related anxiety can help teachers identify different patterns of anxiety and select more targeted and strategic forms of support.

High Emotionality scores may suggest a need for gradual exposure to listening tasks, low-stakes practice, and a supportive classroom environment. In contrast, high Worry scores can be addressed through instructional techniques such as pre-listening preparation, strategy instruction, and activities that alleviate uncertainty about task demands. For Anticipatory Fear, it may be beneficial to gradually familiarise learners with social and communicative listening situations before introducing more challenging tasks.

Additionally, the FLLASp can be administered at various points throughout a course to monitor changes in students' emotional responses and assess the effectiveness of teaching strategies. However, it is important to use the scores as pedagogical indicators rather than definite or diagnostic measures.

Even though our main purpose was not to analyse the results obtained but to compare and validate a non-existent FLLAS for the Portuguese context, this work displays data which may alert or confirm teachers and educators' practices and help to assist them with a proper designed and tested tool for acknowledging anxiety and overall emotional factors affecting listening performance in students. Our sample signals a specific reality where it is possible to single out students who may suffer from higher anxiety, which, according to literature (Bourne, 2010; Elkhafaifi, 2005; Kimura, 2017), may affect their performance., even though the results as obtained in this sample did not reveal critical negative results (based on the average values being below the mid point of the scale), For example, if from all dimensions analysed, Anticipatory Fear has not greatly impacted students who were tested in this study, one should not ignore the fact that there may be a small number of students who possibly need adequate didactic approach as shown in Figure 1.

Conclusion

This study aimed to adapt and validate the Foreign Language Listening Anxiety Scale (FLLAS) in the Portuguese EFL university context, to obtain a reliable instrument suitable for identifying effective variables in foreign language learning. According to the findings collected, this study identified strong empirical evidence that the adapted version of FLLAS constitutes both a reliable and valid instrument for assessing listening-related anxiety among Portuguese undergraduate English learners. The psychometric analyses proved solid internal consistency and recognised a multidimensional conceptualisation of listening anxiety. Among all tested models, the three-factor structure proposed by Kimura (2008), comprising Emotionality, Worry, and Anticipatory Fear, yielded the best overall fit to the data. The following refinement of this referred model resulted in a tight 24-item scale with improved fit indices and stronger convergent validity. These results emphasise the importance of precise validation procedures when adapting affective scales across different linguistic and educational contexts. This study addresses FLLA as a complex phenomenon, rather than a single, unitary construct.

These findings also contribute to possible important educational insights in EFL listening practice. Therefore, given the multidimensional nature of the FLLASp as adapted and validated in this study, students' distinct anxiety realities may be addressed through differentiated didactic strategies and approaches. Furthermore, the three-factor structure model provides educators with broad and specific data that may signal the more appropriate activities to be implemented in each case, ideally for each student, contributing to a more inclusive and effective teaching and learning environment.
Despite the study's relevant contributions, certain limitations are also identified, namely that the sample was limited to students from two Portuguese higher-education institutions, restricting its potential broader generalisation. A further limitation is that the final model did not meet the traditional close-fit RMSEA threshold; thus, the fit should be interpreted as acceptable but not ideal, especially given the ordinal WLSMV estimation.

Future research should also explore the potential applications of the adapted FLLASp in a broader range of learning contexts and environments, namely those with learners with different levels of English proficiency, including other Portuguese higher-education institutions, to test the generalisability of the findings. Future work should also test measurement invariance across specific groups within categories such as gender, proficiency level, institution, field and other relevant ones. Further studies would also benefit from additional longitudinal studies to investigate how listening anxiety develops or changes over a semester or academic year and whether specific pedagogical strategies can reduce Emotionality, Worry, or Anticipatory Fear. Additional research topics, including the potential application of the scale in other educational levels beyond higher education and cross-cultural comparisons with learners from other countries, are also interesting areas for future study and are relevant to examining whether the three-factor structure is replicated across linguistic and educational contexts.

ORCID
https://orcid.org/0000-0001-9602-9729
https://orcid.org/0000-0002-3264-587X
https://orcid.org/0000-0002-1782-2567
https://orcid.org/0000-0003-4050-3473
https://orcid.org/0000-0002-3002-704X

Publisher's Note
The claims, arguments, and counter-arguments made in this article are exclusively those of the contributing authors. Hence, they do not necessarily represent the viewpoints of the authors' affiliated institutions, or EUROKD as the publisher, the editors and the reviewers of the article.

Acknowledgements
Not applicable.

Funding
This work is funded by national funds through FCT – Fundação para a Ciência e a Tecnologia, I.P., under the support UID/50014/2025 (https://doi.org/10.54499/UID/50014/2025).

CRediT Authorship Contribution Statement
Luciana Cabral P. Bessa: Conceptualization, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing.
Guilherme Gonçalves: Formal Analysis, Data Curation, Writing – Original Draft, Writing – Review & Editing.
Elisabete Mendes Silva: Conceptualization, Methodology, Investigation, Writing – Original Draft, Writing – Review & Editing.
Miguel Melo: Writing – Review & Editing, Validation.
Maximino Bessa: Conceptualization, Resources, Methodology, Funding Acquisition.

Generative AI Use Disclosure Statement
The authors used ChatGPT (OpenAI) to proofread and improve the clarity and phrasing of the manuscript text. All substantive intellectual content and final decisions remain the responsibility of the authors.

Ethics Declarations
The authors declare that the above-referenced study involving human participants was conducted in full compliance with the ethical standards and institutional guidelines applicable to research at INESC TEC. In accordance with the Institutional Code of Ethics – specifically its provisions on self-assessment and submission to the Ethics Committee, and its appendix listing research projects to be submitted to the Ethics Committee – all research projects are subject to a mandatory self-assessment procedure to determine whether formal ethical review and approval by the Ethics Committee is required. Prior to conducting the study, the official self-assessment checklist provided by the institution was completed. This evaluation determined that: the research posed no significant ethical risks to participants; the procedures involved were non-invasive, voluntary, and followed standard ethical practices; and the study did not meet the criteria requiring formal submission to or approval by the Ethics Committee. As per institutional policy, when no ethical concerns are identified during this self-assessment process, a formal Ethical Clearance Letter is not required, and the study may proceed under the responsibility of the researcher while adhering to ethical best practices. Furthermore, all participants were treated in accordance with widely accepted ethical principles, including informed consent, confidentiality and anonymity of participant data, the right to withdraw, and protection from harm. Link to ethical guidelines:

https://repositorio.inesctec.pt/server/api/core/bitstreams/b0243350-09a9-4009-9b6d-ba50a7a5578b/content

Consent for Publication

The manuscript does not contain any identifiable personal information or individual participant data requiring consent for publication.

Competing Interests

The authors declare no competing interests.

Data Availability

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

References

Bárkányi, Z., & Brash, B. (2025). Foreign language speaking anxiety online: Mitigating strategies and speaking practices. *ReCALL*, *37*(3), 421–440. https://doi.org/10.1017/S0958344025000060

Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, *25*(24), 3186–3191. https://doi.org/10.1097/00007632-200012150-00014

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*(3), 588–606. https://doi.org/10.1037/0033-2909.88.3.588

Botes, E., Dewaele, J.-M., & Greiff, S. (2020). The Foreign Language Classroom Anxiety Scale and academic achievement: An overview of the prevailing literature and a meta-analysis. *Journal for the Psychology of Language Learning*, *2*(1), 26–56. https://doi.org/10.52598/jpll/2/1/3

Bourne, E. J. (2010). *The anxiety and phobia workbook* (5th ed.). New Harbinger Publications.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage Publications.

Campos, L. A., Marôco, J., John, M. T., Santos-Pinto, A., & Campos, J. A. D. B. (2020). Development and psychometric properties of the Portuguese version of the Orofacial Esthetic Scale: OES-Pt. *PeerJ*, *8*, Article e8814. https://doi.org/10.7717/peerj.8814

Capan, S. A., & Karaca, M. (2013). A comparative study of listening anxiety and reading anxiety. *Procedia - Social and Behavioral Sciences*, *70*, 1360–1373. https://doi.org/10.1016/j.sbspro.2013.01.198

Chang, A. C.-S. (2008). Sources of listening anxiety in learning English as a foreign language. *Perceptual and Motor Skills*, *106*(1), 21–34. https://doi.org/10.2466/pms.106.1.21-34

Cho, G., Hwang, H., Sarstedt, M., & Ringle, C. M. (2020). Cutoff criteria for overall model fit indexes in generalized structured component analysis. *Journal of Marketing Analytics*, *8*(4), 189–202. https://doi.org/10.1057/s41270-020-00089-1

Clément, R., Baker, S. C., & MacIntyre, P. D. (2003). Willingness to communicate in a second language: The effects of context, norms, and vitality. *Journal of Language and Social Psychology*, *22*(2), 190–209. https://doi.org/10.1177/0261927X03022002003

Elkhafaifi, H. (2005). Listening comprehension and anxiety in the Arabic language classroom. *The Modern Language Journal*, *89*(2), 206–220. https://doi.org/10.1111/j.1540-4781.2005.00275.x

Erdem, S., & Tutkun, O. F. (2016). Problems in English language teaching according to secondary school students. *The Turkish Online Journal of Educational Technology*, Special Issue for INTE 2016, 268–275. https://files.eric.ed.gov/fulltext/ED629488.pdf

Ethnologue. (2025). *What are the top 200 most spoken languages?* https://www.ethnologue.com/insights/ethnologue200/

Fox, R., Corretjer, O., & Webb, K. (2019). Benefits of foreign language learning and bilingualism: An analysis of published empirical research 2012–2019. *Foreign Language Annals*, *52*(4), 699–726. https://doi.org/10.1111/flan.12424

Ghorbanpoor, M., Mathew, B. P., Al Ofi, A.-H. A. N., & Khattak, Z. I. (2026). The development and validation of the English-as-a-Foreign-Language Listening Test Anxiety Scale. *Language Testing in Asia*, *16*, Article 6. https://doi.org/10.1186/s40468-025-00420-6

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Prentice Hall.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125–132. https://doi.org/10.2307/327317

Hu, L.-T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Ji, S., Qin, X., & Li, K. (2022). A systematic review of foreign language listening anxiety: Focus on the theoretical definitions and measurements. *Frontiers in Psychology*, *13*, Article 859021. https://doi.org/10.3389/fpsyg.2022.859021

Kim, J.-H. (2000). *Foreign language listening anxiety: A study of Korean students learning English* [Doctoral dissertation, The University of Texas at Austin]. ProQuest Dissertations & Theses Global. https://www.proquest.com/openview/c4688d2bbd2299b53984998e8b2d4771/1

Kim, J.-H. (2002). Anxiety and foreign language listening. *English Teaching*, *57*(2), 3–34. https://scholar.kyobobook.co.kr/article/detail/4010023014630

Kimura, H. (2008). Foreign language listening anxiety: Its dimensionality and group differences. *JALT Journal*, *30*(2), 173–196. https://doi.org/10.37546/JALTJJ30.2-2

Kimura, H. (2017). Foreign language listening anxiety: A self-presentational view. *International Journal of Listening*, *31*(3), 142–162. https://doi.org/10.1080/10904018.2016.1222909

Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). Guilford Publications.

Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). MVN: An R package for assessing multivariate normality. *The R Journal*, *6*(2), 151–162. https://doi.org/10.32614/RJ-2014-031

Kralova, Z., & Petrova, G. (2017). Causes and consequences of foreign language anxiety. *XLinguae*, *10*(3), 110–122. https://doi.org/10.18355/XL.2017.10.03.09

Lee, K., & Ye, Y. (2023). Investigating the reliability of the Foreign Language Classroom Anxiety Scale (FLCAS): An application of generalizability theory. *Research Methods in Applied Linguistics*, *2*(1), Article 100036. https://doi.org/10.1016/j.rmal.2022.100036

Lei, L., Deng, Y., & Liu, D. (2023). Research on the learning/teaching of L2 listening: A bibliometric review and its implications. *Studies in Second Language Learning and Teaching*, *13*(4), 781–810. https://doi.org/10.14746/ssllt.40216

Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, *48*(3), 936–949. https://doi.org/10.3758/s13428-015-0619-7

MacIntyre, P. D. (1992). *Anxiety and language learning from a stages of processing perspective* [Doctoral dissertation, University of Western Ontario]. https://hdl.handle.net/20.500.14721/17509

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

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*(3), 519–530. https://doi.org/10.2307/2334770

Özdemir, O., & Seçkin, H. (2025). Exploring foreign language anxiety in higher education: Multifaceted insights into causes, impacts, and coping strategies. *Social Sciences & Humanities Open*, *11*, Article 101364. https://doi.org/10.1016/j.ssaho.2025.101364

Papi, M., & Khajavy, G. H. (2023). Second language anxiety: Construct, effects, and sources. *Annual Review of Applied Linguistics*, *43*, 127–139. https://doi.org/10.1017/S0267190523000028

Riddell, D. (2014). *Teach English as a foreign language: Teach yourself* (New ed.). Hachette UK.

Sari, N., & Fithriyana, R. (2019). Exploring EFL students' problems in listening comprehension. *JEES (Journal of English Educators Society)*, *4*(1), 47–52. https://doi.org/10.21070/jees.v4i1.1722

Schimpff, A. M., & Botes, E. (2026). The development and validation of the Context Inclusive Language Skill Anxiety Scale (CILSAS): Measuring reading, writing, listening and speaking anxiety in any learning context. *Journal of Multilingual and Multicultural Development*. Advance online publication. https://doi.org/10.1080/01434632.2026.2654821

Serraj, S., & Noordin, N. B. (2013). Relationship among Iranian EFL students' foreign language anxiety, foreign language listening anxiety and their listening comprehension. *English Language Teaching*, *6*(5), 1–12. https://doi.org/10.5539/elt.v6n5p1

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2019). *Using multivariate statistics* (7th ed.). Pearson.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, *40*(3), 191–210. https://doi.org/10.1017/S0261444807004338

Vandergrift, L., & Goh, C. C. M. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.

Vogely, A. J. (1998). Listening comprehension anxiety: Students' reported sources and solutions. *Foreign Language Annals*, *31*(1), 67–80. https://doi.org/10.1111/j.1944-9720.1998.tb01333.x

Wang, S. (2010). An empirical study of Chinese English major students' listening anxiety of classroom learning activity at the university level. *Journal of Language Teaching and Research*, *1*(5), 562–568. https://doi.org/10.4304/jltr.1.5.562-568

Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment*, *28*(3), 186–191. https://doi.org/10.1007/s10862-005-9004-7

Yamauchi, Y. (2014). Revised version of the foreign language listening anxiety scale: Precise description of subordinate concepts' influence on learners. *ARELE: Annual Review of English Language Education in Japan*, *25*, 143–158. https://doi.org/10.20581/arele.25.0_143

Zhang, X. (2013). Foreign language listening anxiety and listening performance: Conceptualizations and causal relationships. *System*, *41*(1), 164–177. https://doi.org/10.1016/j.system.2013.01.004

Zhang, X. (2019). Foreign language anxiety and foreign language performance: A meta-analysis. *The Modern Language Journal*, *103*(4), 763–781. https://doi.org/10.1111/modl.12590

Appendix A

Translated FLLAS Items

Table A1

Translated Original FLLAS

Item ID	Item (PT-PT)
1	Quando ouço inglês, geralmente não compreendo uma ou duas palavras.
2	Fico nervoso/a se o excerto que estou a ouvir é lido apenas uma vez durante os testes de compreensão auditiva.
3	Quando alguém pronuncia palavras de forma diferente da minha, tenho dificuldade em entendê-las.
4	Quando alguém fala inglês muito depressa, tenho receio de não conseguir compreender tudo.
5	Fico nervoso/a quando estou a ouvir inglês se não estou familiarizado/a com o tópico.
6_R	Torna-se fácil adivinhar as partes que me faltam quando ouço inglês.
7	Se me distrair, ainda que brevemente, enquanto estou a ouvir inglês, receio perder ideias/partes importantes.
8	Quando estou a ouvir inglês, preocupo-me se não consigo acompanhar a expressão labial e facial da pessoa que está a falar.
9	Durante os testes de compreensão auditiva em inglês, fico nervoso/a e confuso/a quando não percebo todas as palavras.
10	Quando ouço inglês, é difícil distinguir as palavras umas das outras.
11	Sinto-me desconfortável na aula quando ouço inglês sem apoio do texto escrito.
12	Tenho dificuldades em compreender instruções orais em inglês.
13	É difícil concentrar-me no que os falantes de inglês estão a dizer, a não ser que os conheça bem.
14_R	Sinto-me confiante quando estou a ouvir em inglês.
15	Quando estou a ouvir inglês, fico muitas vezes tão confuso/a que não me consigo lembrar do que ouvi.
16	Receio não ter um conhecimento prévio adequado de alguns tópicos quando estou a ouvir inglês.
17	Os meus pensamentos ficam baralhados e confusos quando estou a ouvir informações importantes em inglês.
18	Preocupo-me quando não tenho muito tempo para pensar no que ouvi em inglês.
19	Quando estou a ouvir inglês, frequentemente acabo por traduzir palavra a palavra sem compreender os conteúdos.
20	Preferia não ter de ouvir as pessoas a falar inglês.
21	Preocupo-me quando não sou capaz de ouvir inglês ao meu próprio ritmo.
22	Estou sempre a pensar que toda a gente, exceto eu, percebe muito bem o que um falante de inglês está a dizer.
23	Fico chateado/a quando não tenho a certeza se compreendo aquilo que estou a ouvir em inglês.
24	Se uma pessoa fala inglês muito baixinho, preocupa-me não a conseguir entender/compreender.
25_R	Não receio ouvir em inglês num contexto alargado de pessoas.
26	Fico nervoso/a quando estou a ouvir um falante de inglês ao telefone ou a imaginar uma situação na qual ouço um falante de inglês ao telefone.
27	Fico tenso/a quando estou a ouvir em inglês, enquanto estou num contexto social.
28	É-me difícil ouvir inglês mesmo que haja um pequeno ruído de fundo.
29	Ouvir novas informações em inglês deixa-me inquieto/a.
30	Fico aborrecido/a quando me deparo com palavras que não compreendo quando estou a ouvir inglês.
31_R	A acentuação e entoação em inglês soam-me familiares.
32	Quando estou a ouvir inglês, frequentemente compreendo as palavras mas ainda não consigo perceber o que o falante quer dizer.
33	Assusta-me/Preocupa-me não conseguir identificar uma palavra-chave de um excerto auditivo em inglês.

Scoring the FLLASp

Based on the validated 24-item three-factor structure, we recommend calculating scores by averaging the items within each factor. This procedure involves reverse-coding Item 14_R and then calculating the mean for each subscale, as well as a global index (Total FLLASp) by averaging all 24 items. This approach accounts for the unequal number of items across factors and places all results on the original 1 to 5 Likert metric, where higher scores indicate higher anxiety.

The subscales are composed of the following items:

- Emotionality: items 10, 11, 12, 14, 15, 17, 19, 20, 29, 30, and 32.
- Worry: items 1, 2, 4, 5, 7, 9, 16, 18, 21, 23, and 33.
- Anticipatory Fear: items 26 and 27.

Items 3, 6, 8, 13, 22, 24, 25, 28, and 31 are excluded from the final scale.