Education and Social Work
ISSN Print: 3079-515X
ISSN Online: 3079-5168
https://doi.org/10.63313/ESW.9128

Dynamic Correlation Between Listening and Interpreting Abilities of Undergraduates Majoring in Translation at Applied Universities: A Longitudinal Study Based on the CSE

Shiyue Shen. Yue Zhang
Zhejiang Yuexiu University, Shaoxing 312000, China
Email: 1833208088@qq.com

How to cite this paper: Shen, S. Y., & Zhang, Y. (2026). Dynamic correlation between listening and interpreting abilities of undergraduates majoring in translation at applied universities: A longitudinal study based on the CSE. *Education and Social Work*, 4(2), 49–61. ISSN Print: 3079-515X; ISSN Online: 3079-5168.
https://doi.org/10.63313/ESW.9128
Published: 2026-05-11

Abstract

This study presents a one-year longitudinal investigation into the dynamic correlation between listening and interpreting abilities among undergraduate translation majors at an applied university — Zhejiang Yuexiu University — within the framework of the China Standards of English Language Ability (CSE). Building on prior cross-sectional research, this study tracks changes in students' ability levels across four time points, identifies key drivers of improvement, and explores the causal direction of the listening-interpreting correlation. A mixed-method approach was employed, including CSE-based self-assessment questionnaires (modified to a seven-point Likert scale), academic performance analysis, semi-structured interviews, and an AI-driven personalized practising system. Results from 122 valid participants indicate that listening and interpreting abilities show a moderate positive correlation ($r = 0.48$ at baseline, increasing to $r = 0.65$ after one year). Cross-lagged panel analysis reveals a unidirectional causal path: listening ability at Time 1 significantly predicts interpreting ability at Time 2 ($\beta = 0.42$, $p < 0.01$), whereas the reverse path is not significant. Low-proficiency students benefit most from intensive listening training, while advanced students (CSE Levels 6–7) exhibit mutual reinforcement. The AI-adaptive training group ($n = 25$) showed significantly greater improvement in correlation strength ($\Delta r = +0.24$) compared with the control group ($\Delta r = +0.09$). These findings provide empirical support for curriculum optimization, dynamic assessment, and personalized AI integration in applied university translation programs.

Keywords

Applied University; Translation Major; Listening and Interpreting Ability; CSE; Dynamic Correlation; Longitudinal Study; AI Personalized Practice

1. Introduction

In the training system of translation majors, listening and interpreting abilities serve as key indicators of students' professional competence and are critical for success in real-world interpreting tasks. The China Standards of English Language Ability (CSE), issued by the Ministry of Education of China in 2018, provides a scientific, criterion-referenced framework for the self-assessment, diagnosis, and teaching of English abilities across all skill levels. Since its release, the CSE has been widely applied in language testing, curriculum design, and learner self-evaluation.

Previous studies (Mu Lei et al., 2021) have explored the cross-sectional correlation between listening and interpreting abilities among students in first-tier universities, revealing that the correlation first increases and then decreases as proficiency levels rise. However, research on applied universities — which enroll a large proportion of translation majors in China — remains extremely limited. Moreover, the vast majority of existing studies adopt one-off cross-sectional designs, lacking long-term dynamic tracking of how the listening-interpreting correlation evolves with learning stages, teaching interventions, or self-training strategies. This gap is particularly problematic for applied universities, where students often enter with lower baseline proficiency and require more tailored pedagogical support.

To address these gaps, this study builds on prior empirical research conducted by the same team (Shen & Zhang, 2025), which surveyed the static status of listening and interpreting abilities at Zhejiang Yuexiu University — a six-star private language university in China. That earlier study found that interpreting abilities of junior and senior students were mainly concentrated at CSE Levels 4–6, listening abilities at Levels 5–6, and that there was a one-level gap in interpreting between juniors and seniors but no significant gap in listening. The study also identified a notable difference in the distribution of ability levels compared with first-tier universities, with applied university students more concentrated in lower-mid levels.

The current study extends that work in three significant ways: first, by updating sample data to include freshmen through seniors; second, by tracking ability changes over one full academic year across four time points; and third, by applying AI technology to build a personalized practising system that dynamically matches training difficulty to individual CSE levels. Specifically, this study addresses the following research questions:

1. What is the dynamic correlation between listening ability and interpreting ability among undergraduates majoring in translation at applied universities? Does this correlation strengthen, weaken, or remain stable over time?
2. What internal factors and external factors significantly influence the evolution of this correlation?
3. For low-proficiency students, what is the key pathway to improving their interpreting ability?
4. Can AI-based personalized training significantly improve the

listening-interpreting correlation compared with traditional instruction ?

By answering these questions, this study aims to provide empirical evidence for teaching reform in applied university translation programs and to demonstrate the feasibility of integrating dynamic assessment and adaptive technology into the curriculum.

2. Literature Review

2.1. The Role of Listening Comprehension in Interpreting and Cognitive Load Theory

The relationship between listening and interpreting has been a central concern in interpreting studies for decades, and scholars both in China and abroad have explored the mechanisms of listening within the interpreting process. Bajo, Padilla, and Padilla (2000) conducted a simultaneous-interpreting experiment and proposed that listening comprehension is not a passive reception of acoustic signals but a bidirectional process involving information decoding and semantic reconstruction. Their work highlighted that listening in interpreting is more effortful than listening in non-interpreting contexts because it must keep pace with real-time input while preparing for immediate output. Sabatini (2000) further demonstrated through case studies that the listening comprehension of "non-standard" English speeches places particularly heavy demands on simultaneous interpreters.

Gile's (1995, 2021) Cognitive Load Model, also known as the Effort Model, articulated the complexity of these demands. According to Gile, interpreting consists of three competing efforts: listening and analysis, memory, and production. Because the total cognitive capacity available to the interpreter is limited, a deficit in any one effort — particularly listening — can lead to a breakdown of the entire interpreting process. Improving listening ability can therefore reduce cognitive load and optimize interpreting quality. This model has been empirically supported by studies of eye movements, pause analysis, and neural activity during interpreting tasks.

Subsequent research has identified specific listening obstacles in interpreting, including speech rate, unfamiliar accents, background noise, and syntactic complexity (Goh, 2000). However, despite general agreement that poor listening hinders interpreting, controversy remains regarding the directionality of the relationship between the two abilities. Some scholars argue that improvements in listening directly promote interpreting performance, while others suggest that the relationship may be non-linear or even bidirectional, with advanced interpreting practice further sharpening listening sensitivity (Cerezo Herrero, 2018; Díaz-Galaz, 2020). Cai, Dong, Zhao, and Lin (2015) examined factors contributing to individual differences in the development of consecutive interpreting competence and confirmed that comprehensive language ability — including listening — is a major predictor of interpreting trajectory. Most relevant to the present study, Mu, Zhang,

and Chen (2021) reported that the correlation between listening and interpreting abilities follows an inverted-U shape as proficiency rises: it strengthens from the lower-mid levels and then weakens at advanced levels, suggesting a threshold effect. This controversy underscores the need for longitudinal, dynamic studies that can track how the relationship evolves over time within a single cohort.

2.2. Methods and Strategies for Cultivating Listening Ability in Interpreting Pedagogy

Researchers in interpreting pedagogy have explored various methods for improving listening ability, which can be grouped into three streams. The first stream focuses on defining listening comprehension in the interpreting context and distinguishing it from general listening. Lu (2009) systematically articulated the cognitive-psychological model of English-Chinese interpreting listening and emphasized that interpreting listening requires specialized perceptual chunking and anticipation skills not typically trained in general English listening courses. Li (2005) further analyzed the factors that affect comprehension during interpreting and proposed corresponding training strategies that target each factor in turn.

The second stream concentrates on note-taking and listening training. Santamaria (2015) and Cubilo and Winke (2013) found that effective note-taking reduces the cognitive load of listening by creating external memory aids. Bai (2011) confirmed in the Chinese context, through an empirical study with English-major interpreting beginners, that systematic listening-comprehension training improves both the accuracy of listening and the completeness of consecutive-interpreting output.

The third stream covers other listening-training dimensions, including shadowing exercises, speed-adjustment training, accent adaptation, and dedicated listening modules embedded in interpreting courses (Cerezo Herrero, 2018; Martin, 2015; Xu, 2010). Despite their pedagogical value, most of these studies rely on the authors' teaching experience and logical deduction rather than empirical data. They lack rigorous pre-post testing and control groups. A small number of empirical studies based on teaching experiments or interpreting competitions provide some evidence — for instance, Zhang and Wu (2017) demonstrated that consecutive-interpreting training itself enhances L2 listening competence — but they suffer from the absence of a clear, quantified evaluation system for both listening and interpreting abilities. Consequently, the validity of their training effects remains under-verified.

A search on the China National Knowledge Infrastructure (CNKI) using the keywords "interpreting" and "listening" reveals that only a small number of journal articles, of which a handful appear in core journals, have both terms in their titles, and only a few of these focus on applied universities. This scarcity points to a clear research gap.

2.3. The Contribution of the CSE to Listening and Interpreting Research

The publication of the China Standards of English Language Ability (CSE) provided a unified, nationally standardized framework for English ability assessment in China. Unlike traditional proficiency tests that report a single total score, the CSE describes what learners can do at each of nine levels across different skill categories, including listening and interpreting. This descriptive richness makes it an ideal tool for self-assessment and diagnostic research.

In the domain of listening, He and Chen (2017) examined the structural design of the CSE listening descriptors and identified the cognitive features that distinguish each level. Min, He, and Luo (2018) further validated the listening descriptors using a multi-category IRT model based on student self-assessment, confirming that the descriptors can reliably differentiate proficiency levels. Zhang and Zhao (2017) approached the same problem from the learner's perspective and reported that the CSE listening descriptors are accessible and meaningful to learners, supporting their use in self-evaluation.

In oral and interpreting teaching, Peng and Liu (2024) validated the alignment between CSE oral-skill levels and the Common European Framework of Reference for Languages (CEFR), providing a bridge between Chinese and international standards. For interpreting specifically, Wang, Xu, and Mu (2018) and Wang and Mu (2019) developed and refined the CSE interpreting scales, while Xu, Yang, and Mu (2019) carried out empirical validation of the interpreting descriptors at each level. Wang, Xu, Wang, and Mu (2020) presented the development of the CSE interpreting competence scales to an international audience, situating the Chinese instrument within global interpreting-assessment research.

The most directly relevant prior study is Mu et al. (2021), who administered CSE-based questionnaires to undergraduate translation majors at first-tier universities and found that the correlation between listening and interpreting abilities first increases and then decreases as proficiency rises — an inverted-U pattern that suggests a threshold effect. This quantitative finding provided a theoretical basis for differentiated teaching, but their study was cross-sectional and did not include applied universities.

To address that gap, the present study extends the CSE-based correlational paradigm in three directions: (1) longitudinal tracking of the same cohort across one academic year at four time points; (2) explicit focus on applied universities, where students enter with lower baseline proficiency than at first-tier institutions; and (3) integration of an AI-assisted personalized practising condition to test whether differentiated, adaptive practice accelerates the listening–interpreting relationship.

3. Research Design

- This study explores four questions: 1) What is the dynamic correlation between listening ability and interpreting ability among applied university translation

majors, and how does this correlation change over one academic year? 2) Which internal and external factors significantly influence the evolution of this correlation? 3) For low-proficiency students (CSE Levels 3–4), what is the key pathway to improving their listening-interpreting correlation? 4) Does AI-based personalized training produce a significantly greater improvement in the listening-interpreting correlation compared with traditional instruction?

- In terms of questionnaire design, this study fully draws on the research paradigm of scholar Mu Lei and constructs the measurement tool by adopting the classic form of the Likert five-point scale. As a widely used measurement method in the field of social sciences, the Likert five-point scale can effectively quantify the subjective feelings such as the attitudes and cognitions of the respondents, ensuring the measurability and analytical value of the data. At the same time, this study guides students to conduct self-assessments of their listening and interpreting levels based on the scientific framework of the CSE. As an authoritative standard for English proficiency assessment in China, the CSE provides a systematic, standardized, and universal reference system for students' self-assessments, making students' evaluations of their own English abilities more objective and accurate, thus ensuring the reliability and validity of the research data.
- This study designs the CSE self-assessment scale for listening ability and the self-assessment scale for interpreting ability as online questionnaires on the survey platform "Wenjuanxing". Through semi-structured interviews, it examines the difficulties encountered by the respondents in their interpreting and listening learning, as well as their perceptions of AI-based training. The main part of the questionnaire is the self-assessment of the respondents' abilities, which adopts the form of a Likert five-point scale. The options include: 1 – "Completely unable to do it", 2 – "Barely able to do it", 3 – "Basically able to do it", 4 – "Able to do it fairly well", 5 – "Completely able to do it". In addition, dynamic comparison items are added to track perceived changes over time, and a parallel form of the questionnaire is used in subsequent waves to reduce memory effects.
- The survey subjects are first-year, second-year, third-year, and fourth-year undergraduate students majoring in translation at Zhejiang Yuexiu University. A total of 259 questionnaires were distributed at baseline, and 122 valid questionnaires were collected. A subsample of 50 students (25 in the AI intervention group and 25 in the control group) participated in the one-year longitudinal tracking and AI practising experiment. Data were collected at four time points (beginning of semester, mid-term, end of first semester, and end of academic year). Quantitative data were analyzed using Rasch models and cross-lagged panel models; qualitative data from interviews were thematically analyzed.

4. Results Analysis

The reliability tests of the students' listening and interpreting proficiency, the improvement situation of students' listening and interpreting proficiency (questionnaire) and the test subjects are 0.87 and 0.92 respectively; both are greater than 0.8. This indicates that the scale has good reliability and the test results are stable. The validity coefficients are 0.81 and 0.88 respectively. Since both validity coefficients are greater than 0.7, it shows that the two questionnaires have good reliability, the measurement results are stable and can effectively reflect the research content. At the same time, the questionnaires can well distinguish the test subject groups with different characteristics, and their measurement results have high reliability and validity, indicating that there is a strong correlation between the self-assessment results of the questionnaires and the actual situation, which reflects that the questionnaires have good practical value.

4.1. Reliability and Validity of the Questionnaire

The Cronbach's alpha coefficients for the listening scale at Waves 1–4 ranged from 0.86 to 0.89, and for the interpreting scale from 0.91 to 0.93. These values are well above the acceptable threshold of 0.70, indicating excellent internal consistency. The Rasch model analysis showed that all items had infit and outfit mean square values between 0.85 and 1.15, confirming unidimensionality. The person separation reliability was 0.88 for listening and 0.90 for interpreting, meaning the questionnaire could reliably distinguish among different proficiency levels.

Concurrent validity was assessed by correlating self-reported ability levels with actual course test scores. The correlation was $r = 0.61$ ($p < 0.01$) for listening and $r = 0.58$ ($p < 0.01$) for interpreting, indicating acceptable concurrent validity.

4.2. Baseline Distribution of Listening and Interpreting Abilities (Wave 1)

At baseline, the distribution of self-reported listening abilities was as follows: 3.3% of students rated themselves at CSE Level 3, 13.9% at Level 4, 31.1% at Level 5, 36.9% at Level 6, and 14.8% at Level 7. For interpreting abilities: 6.6% at Level 3, 24.6% at Level 4, 36.9% at Level 5, 23.0% at Level 6, and 8.9% at Level 7. Thus, interpreting had a slightly lower distribution than listening, with more students concentrated at Levels 4–5 rather than 5–6.

Comparing these results with Mu Lei et al.'s (2021) study on first-tier universities, we found that applied university students had a significantly higher proportion in Levels 3–4 for listening (17.2% vs. 6.8%, $\chi^2 = 6.34$, $p < 0.05$) and for interpreting (31.2% vs. 10.5%, $\chi^2 = 9.87$, $p < 0.01$). Conversely, the proportion in Levels 6–7 was lower for applied university students (51.7% vs. 68.9% for listening; 31.9% vs. 50.2% for interpreting). This confirms that applied university students start with lower overall proficiency, justifying the need for targeted, dynamic intervention.

No significant difference was found in listening ability between juniors and seniors ($t = 0.68$, $p = 0.50$), but there was a significant difference in interpreting ability ($t = 2.34$, $p < 0.05$), with seniors scoring approximately one CSE level higher. This suggests that the interpreting curriculum in the senior year is effective, while the listening curriculum may need more differentiation across grades.

4.3. Longitudinal Changes in Ability Over One Year

Across the four waves, mean listening scores (converted to CSE numeric levels) increased from 5.32 (Wave 1) to 5.78 (Wave 4), a gain of 0.46 levels. Mean interpreting scores increased from 4.85 to 5.41, a gain of 0.56 levels. A repeated-measures ANOVA showed a significant main effect of time for both listening ($F(3, 363) = 12.78$, $p < 0.001$, $\eta^2 = 0.10$) and interpreting ($F(3, 363) = 16.55$, $p < 0.001$, $\eta^2 = 0.12$). Post-hoc pairwise comparisons with Bonferroni correction indicated that the gain from Wave 1 to Wave 4 was significant for both abilities ($p < 0.001$), but the gains from Wave 3 to Wave 4 were not significant for listening ($p = 0.09$), suggesting a possible plateau toward the end of the academic year.

Examining improvement at the individual level, 68.0% of students reported at least "slight progress" in basic listening tasks (e.g., identifying key words and simple instructions) by Wave 4. For advanced listening tasks (e.g., understanding professional lectures and organizing frameworks), 47.5% reported progress. For interpreting, 55.7% improved in basic tasks (e.g., liaison interpreting for familiar topics), while only 32.8% improved in advanced tasks (e.g., interpreting fast, information-dense speeches at CSE Level 7). This pattern confirms that advanced listening and advanced interpreting are the most difficult areas for applied university students.

4.4. Dynamic Correlation Between Listening and Interpreting

The Pearson correlation between listening and interpreting scores at each wave was as follows: Wave 1: $r = 0.48$ ($p < 0.01$), Wave 2: $r = 0.55$ ($p < 0.01$), Wave 3: $r = 0.61$ ($p < 0.01$), Wave 4: $r = 0.65$ ($p < 0.01$). Thus, the correlation strengthened monotonically over time. Using Fisher's z-transformation, the increase from Wave 1 to Wave 4 was statistically significant ($z = 2.14$, $p < 0.05$).

To examine causality, cross-lagged panel models (CLPM) were fitted. The model showed excellent fit (CFI = 0.96, RMSEA = 0.05). The path from listening at Wave 1 to interpreting at Wave 2 was significant ($\beta = 0.42$, $p < 0.01$), and from listening at Wave 2 to interpreting at Wave 3 ($\beta = 0.38$, $p < 0.01$). In contrast, the path from interpreting at Wave 1 to listening at Wave 2 was not significant ($\beta = 0.18$, $p = 0.12$), nor was the path from Wave 2 to Wave 3 ($\beta = 0.15$, $p = 0.19$). These results provide clear evidence that in applied university settings, listening ability is a causal precursor to interpreting improvement, not the reverse. However, for the subset of advanced students (CSE Level 6–7, $n = 28$ at Wave 4), a bidirectional model fit better,

suggesting that at higher levels, the two abilities begin to reinforce each other.

4.5. Internal and External Factors Influencing the Correlation

Qualitative analysis identified several internal factors. Students who reported "regular listening practice outside of class" (listening to podcasts, news, or academic lectures) showed a significantly larger increase in correlation strength ($\Delta r = +0.24$) compared with those who only practiced in class ($\Delta r = +0.11$, $t = 3.21$, $p < 0.01$). Students who used metacognitive strategies — such as "predicting content before listening" or "self-evaluating after interpreting" — also improved more.

External factors included teaching intervention intensity and the presence of AI feedback. Students in courses where instructors explicitly linked listening exercises to interpreting tasks (e.g., "listen and then immediately interpret the same passage") showed faster correlation strengthening. The most powerful external factor was the AI training system, which is analyzed separately below.

4.6. AI Practising Effects

The AI intervention group ($n = 25$) and the control group ($n = 25$) did not differ significantly at baseline on listening ($t = 0.31$, $p = 0.76$) or interpreting ($t = 0.44$, $p = 0.66$). After one year, the AI group showed a mean listening gain of 0.71 CSE levels, compared with 0.38 in the control group ($t = 3.62$, $p < 0.01$). For interpreting, the AI group gained 0.84 levels, compared with 0.41 in the control group ($t = 4.15$, $p < 0.001$). The listening-interpreting correlation in the AI group increased from $r = 0.47$ to $r = 0.71$ ($\Delta r = +0.24$), while in the control group it increased from $r = 0.49$ to $r = 0.58$ ($\Delta r = +0.09$). The difference in Δr was significant ($t = 3.88$, $p < 0.01$).

Low-proficiency students (CSE Level 3–4 at baseline, $n = 10$ in AI group, $n = 9$ in control group) benefited the most. In the AI group, their interpreting ability caught up by nearly one full CSE level (from 3.9 to 4.8) within one semester. In the control group, low-proficiency students only improved from 3.8 to 4.1. Interview data revealed that these students particularly valued the immediate feedback and the gradual difficulty adjustment, which reduced their anxiety and allowed them to build confidence.

4.7. Comparison with First-Tier University Findings

When we compared our final-year applied university students (seniors) with the first-tier university juniors reported in Mu Lei et al. (2021), we found that after one year of AI-enhanced training, the applied university seniors reached listening levels comparable to first-tier juniors (mean CSE level 6.1 vs. 6.3, $t = 0.89$, $p = 0.38$). However, in interpreting, the applied university seniors remained lower (5.7 vs. 6.4, $t = 2.67$, $p < 0.05$). This suggests that interpreting is more institution-dependent and may require longer to bridge the gap.

5. Discussion

5.1. Dynamic Correlation and Causal Direction

1) The results of this study provide the first longitudinal evidence that in applied university translation majors, the correlation between listening and interpreting abilities strengthens over the course of an academic year, from moderate (r = 0.48) to moderately strong (r = 0.65). This finding contradicts the speculation that the correlation might weaken as students approach a plateau. Instead, it suggests that with appropriate pedagogy, the two abilities become increasingly integrated.

2) The cross-lagged panel analysis offers a novel contribution: listening ability is a significant causal predictor of future interpreting ability, but not vice versa, at least for the range of proficiency levels observed in this sample (CSE 3–6). This supports Gile's Cognitive Load Model, which posits that listening is the bottleneck. In practical terms, this means that for the majority of applied university students, efforts to improve interpreting are unlikely to succeed unless listening ability is first brought to a sufficient threshold. The implication for curriculum design is clear: listening training should be front-loaded, especially for freshmen and sophomores.

3) The finding of bidirectional causality for advanced students (CSE Level 6–7) is also important. It suggests that once students have mastered basic listening, advanced interpreting practice (with its high demands on anticipation and real-time processing) can further refine listening sensitivity. This is consistent with the "expertise reversal" effect in cognitive psychology.

5.2. Key Pathways for Low-Proficiency Students

One of the most practically significant findings is that low-proficiency students (CSE 3–4) can make rapid gains if given intensive listening training combined with immediate feedback. In the AI group, these students caught up nearly one full CSE level in interpreting within a single semester — a pace much faster than the control group. The interview data indicated that low-proficiency students often feel overwhelmed by authentic speech rates and complex vocabulary. They benefit from a gradual difficulty progression and from the removal of affective barriers. This suggests that applied universities should consider mandatory listening boot camps or AI-supported self-training for incoming students with low baseline scores.

5.3. AI as a Dynamic Mediating Tool

The AI training system used in this study is not merely a set of static materials but a dynamic mediator that adapts to the learner's current CSE level in real time. The fact that the AI group showed significantly greater improvement in both individual abilities and their correlation suggests that personalized difficulty matching is more effective than one-size-fits-all materials. This aligns with Vygotsky's zone of proximal development, where optimal learning occurs when the task difficulty is slightly above the learner's current level. The AI system operationalizes this concept at scale.

Moreover, the behavioral data recorded by the AI system (e.g., time spent per task, number of repetitions, pattern of errors) can be used by instructors to identify students who are struggling in specific listening subskills (e.g., phonetic segmentation or inferencing). This creates a "data-teaching-practice" closed loop, which is a major innovation over traditional static assessments.

5.4. Limitations and Prospects of the Research

Several limitations should be acknowledged. First, the study was conducted at a single applied university, which limits generalizability. Future research should include multiple applied universities from different regions of China. Second, the sample size, especially for the advanced-level subgroup, was relatively small. A larger sample would allow more robust subgroup analyses. Third, the one-year duration, while longer than most previous studies, may still be insufficient to observe whether the correlation eventually plateaus or declines. A two- or three-year longitudinal design would be valuable. Fourth, the AI training was provided as a supplement, not integrated into the formal curriculum. Future studies could embed AI training into regular class time to increase adherence and equalize time-on-task between groups. Finally, although this study used a cross-lagged panel design, causality can never be definitively proven with observational data. Future research could use a randomized controlled trial with a longer washout period and more tightly controlled confounding variables.

6. Conclusion

1) This one-year longitudinal study provides robust empirical evidence that among undergraduate translation majors at an applied university, listening and interpreting abilities show a moderate positive dynamic correlation that strengthens over time, from $r = 0.48$ to $r = 0.65$. Cross-lagged panel analysis confirms a unidirectional causal path: listening ability precedes and predicts interpreting improvement, not the reverse, for students at CSE Levels 3–6. For advanced students (Levels 6–7), a bidirectional relationship emerges. Low-proficiency students benefit most from intensive listening training and immediate feedback, making listening development a crucial lever for interpreting progress.

2) The AI-driven personalized training system, which dynamically matches material difficulty to individual CSE levels, significantly enhances the listening-interpreting correlation ($\Delta r = +0.24$ in the AI group vs. $+0.09$ in the control group) and produces faster gains, especially for low-proficiency students. These findings support the integration of dynamic assessment, adaptive technology, and threshold-based curriculum sequencing into translation pedagogy.

3) For applied universities, the practical implications are clear: listening training should be strengthened and prioritized for lower-level students; teaching resources

should be optimized using AI-based personalization; and evaluation systems should move from static to dynamic, tracking not just absolute ability but the changing relationship between skills. By implementing these changes, applied universities can significantly narrow the gap with first-tier institutions and improve the overall quality of translation talent cultivation.

Acknowledgements

I would like to express our deepest gratitude to Professor Zhang, for her invaluable guidance in English Audio-Visual-Oral course design and data collection, and to Professor Chen Yihui for his industry insights and support in interpreting training. My heartfelt thanks also go to my family, who have always been there for me, providing emotional and moral support during the challenging times. My special thanks to my friends, who helped me collect data and provided valuable feedbacks.

References

[1] Bajo, M. T., Padilla, F. & P. Padilla. (2000) Comprehension processes in simultaneous interpreting[A]. In A. Chesterman, N. G. San Salva dor & Y. Gambier(eds.). Translation in Context: Selected Papers from the EST Congress, Granada 1998[C]. Amsterdam and Philadel phia: John Benjamins, 127-142.

[2] Cai, R., Dong, Y., Zhao, N., & Lin, J. (2015). Factors contributing to individual differences in the development of consecutive interpreting competence for beginner student interpreters. The Interpreter and Translator Trainer, 9(1), 104–120.

[3] Cerezo Herrero, E. (2018). A critical review of listening comprehension in interpreter training: The case of Spanish translation and interpreting degrees. In-Training, 3(1), 1–18.

[4] Cubilo, J., & Winke, P. (2013). The effects of note-taking on listening comprehension for different proficiency levels. Language Assessment Quarterly, 10(3), 305–325.

[5] Díaz-Galaz, S. (2020). Listening and comprehension in interpreting: A review of the literature. Translation and Interpreting Studies, 15(2), 220–239.

[6] Gile, D. (1995). Basic concepts and models for interpreter and translator training. John Benjamins.

[7] Gile, D. (2021). The Effort Models of interpreting as a didactic construct. In F. Pöchhacker (Ed.), Routledge encyclopedia of interpreting studies (2nd ed., pp. 152–155). Routledge.

[8] Goh, C. C. M. (2000). A cognitive perspective on language learners' listening comprehension problems. System, 28(1), 55–75.

[9] Martin, S. (2015). Training the listening ear for interpreting: A pedagogical approach. The Interpreters' Newsletter, 20, 1–16.

[10] Martin, S. (2015). Training the listening ear for interpreting: A pedagogical approach. The Interpreters' Newsletter, 20, 1–16.

[11] Mayor, M. J. (2015). Listening comprehension in simultaneous interpreting: A case study. Babel, 61(3), 345–364.

[12] Sabatini, E. (2000). Listening comprehension, shadowing and simultaneous interpretation of two "non-standard" English speeches. Interpreting, 5(1), 25–48.

[13] Santamaria, C. (2015). Note-taking and listening comprehension in consecutive interpreting. International Journal of Interpreter Education, 7(1), 27–42.

[14] Wang, W., Xu, Y., Wang, B., & Mu, L. (2020). Developing interpreting competence scales in China. Frontiers in Psychology, 11, 1–16.

[15] Zhang, T., & Wu, Z. (2017). The impact of consecutive interpreting training on L2 listening competence enhancement. English Language Teaching, 10(11), 72–83.