# KOREAN LANGUAGE MAJORS' PERCEPTIONS OF AND INTENTIONS TO USE A PROPOSED CURRICULUM-ALIGNED AI-INTEGRATED KOREAN GRAMMAR PRACTICE SYSTEM

# NHẬN THỨC VÀ Ý ĐỊNH SỬ DỤNG CỦA SINH VIÊN NGÀNH NGÔN NGỮ HÀN QUỐC ĐỐI VỚI HỆ THỐNG THỰC HÀNH NGỮ PHÁP TIẾNG HÀN TÍCH HỢP AI ĐƯỢC ĐỀ XUẤT THIẾT KẾ THEO CHƯƠNG TRÌNH ĐÀO TẠO

**Le Thi Yen Thu***

*The University of Danang - University of Foreign Language Studies, Vietnam*

*Corresponding author: ltythu@ufl.udn.vn

(Received: June 04, 2026; Revised: August 14, 2026; Accepted: September 14, 2026)

DOI: 10.31130/ud-jst.2026.24(10).386

**Abstract –** This study investigates students' perceptions of and intentions to use a proposed AI-integrated Korean grammar practice system aligned with their program curriculum. The study proposes five independent variables, namely Output Difficulties (OD), Practice Behavior (PB), Feedback Need (FN), Technology Acceptance (TA), and Pedagogical Trust and Alignment (PTA), as well as one dependent variable, Intention to Use (IU). A survey was conducted with 131 students, and the collected data were analyzed using SPSS software. The results indicate that TA, PTA, and OD were significant positive predictors of IU, whereas PB and FN were not. Because the EFA did not establish discriminant validity between TA and IU, the findings are interpreted as exploratory and bounded by the case-study context. Within those limits, they suggest that TA and curriculum alignment are associated with learners' willingness to use such systems.

**Key words -** AI; Korean grammar; intention to use; technology acceptance; curriculum alignment; case study.

**Tóm tắt -** Bài viết khảo sát nhận thức, ý định sử dụng của sinh viên đối với hệ thống thực hành ngữ pháp tiếng Hàn tích hợp AI được đề xuất xây dựng theo chương trình đào tạo tại một cơ sở đào tạo. Nghiên cứu gồm năm biến độc lập: Khó khăn trong sản sinh ngôn ngữ (OD); Hành vi thực hành hiện tại (PB); Nhu cầu phản hồi sửa lỗi (FN); Chấp nhận công nghệ (TA); Niềm tin và sự phù hợp với chương trình đào tạo (PTA); một biến phụ thuộc Ý định sử dụng (IU). Dữ liệu khảo sát từ 131 sinh viên được xử lý bằng phần mềm SPSS. Kết quả hồi quy: TA, PTA và OD là các biến dự báo dương có ý nghĩa thống kê đối với IU; PB và FN không phải là biến dự báo có ý nghĩa. Do EFA chưa xác lập được giá trị phân biệt giữa TA và IU, kết quả được diễn giải ở mức thăm dò. Qua đó gợi ý vai trò của mức độ chấp nhận công nghệ và sự phù hợp với chương trình đào tạo đối với ý định sử dụng hệ thống.

**Từ khóa -** AI; ngữ pháp tiếng Hàn; ý định sử dụng; chấp nhận công nghệ; chương trình đào tạo; nghiên cứu trường hợp.

## 1. Introduction

In the current era of rapid development in digital technology and artificial intelligence (AI), the application of technology in foreign language teaching and learning has become an increasingly common trend. In particular, the emergence of AI-supported learning systems, such as chatbots, grammar correction tools, and automated feedback systems, has opened up new possibilities for personalizing the learning process and enhancing the effectiveness of language practice [1], [2], [3]. The same trend can also be observed in the field of Korean language education, where AI has been shown not only to support learners in retrieving knowledge but also to participate in the processes of practice, feedback, and real-time language correction [4], [5].

However, the actual teaching of Korean in Vietnam shows that many students, despite having systematically studied grammatical structures in class, still encounter difficulties in applying them to real-life communication, particularly in language production activities such as speaking and writing. This observation has a certain empirical basis. A study analyzing 70 essays written by first-year students majoring in Korean Language and Culture at the University of Foreign Languages, Hue University, showed that grammatical errors remained common in writing skills and were identified as an obstacle requiring remedial measures [6]. Learners often understand grammatical structures at a theoretical level but lack the ability to produce language naturally and accurately. This issue is related to Swain's Output Hypothesis [7], [8], [9]. According to the Output Hypothesis, language production plays an important role in the process of foreign language acquisition; learners only truly become aware of their own limitations when they have to directly express ideas through speaking or writing.

Although there are currently many AI tools that support Korean language learning, such as Cake, Duolingo, Talkpal, and HeyKorea, most of these systems are designed in a general manner and are not yet aligned with the specific characteristics of the curriculum of each educational institution. In addition, a study on the application model of ChatGPT for improving writing skills among Korean language majors also reported several limitations in the process of use, including learners' dependence on the tool, their tendency to translate content from Vietnamese into Korean, and the fact that AI-provided learning materials had not been developed based on a synthesis of studies related to the application of AI in Korean language education [10]. These limitations may reduce learners' trust in AI technology in formal academic

environments. This is also the assumption examined in the present study through the variable Pedagogical Trust and Alignment (PTA).

Therefore, the development of an AI-integrated grammar practice system designed based on the sequence and content of the curriculum may contribute to addressing the limitations mentioned above. Such a system would not only enable students to practice language production through sentence-making and communicative practice activities, but also provide corrective feedback with pedagogical orientation, appropriate to learners' proficiency levels and learning content.

Based on this practical context, the present study was conducted to examine students' perceptions, needs, and intentions to use a proposed curriculum-aligned AI-integrated Korean grammar practice system. Since the system is still at the proposal stage, the object evaluated by students was the description of the system's expected functions, presented in Section 2.5, rather than an already operational product.

Specifically, the study focuses on answering the following two research questions:

(1) What factors are associated with Korean language majors' intentions to use the proposed curriculum-aligned AI-integrated Korean grammar practice system?

(2) Among these factors, which have a more prominent predictive role in the context of Korean language education at a university in Vietnam?

Compared with existing studies on the Technology Acceptance Model (TAM) or on the application of AI in foreign language education, this study has three distinctive features. First, it focuses on Korean language majors in Vietnam, a group of learners that has received relatively limited attention in studies on the acceptance of learning technologies, even though Korean and Vietnamese differ considerably in linguistic typology. Second, the study places a group of factors reflecting learning difficulties and learning habits (Output Difficulties - OD, Practice Behavior - PB, and Feedback Need - FN) alongside technology acceptance - TA factors within the same model, in order to compare the predictive roles of the two groups. Third, the study proposes and preliminarily tests the variable "PTA", a factor that reflects the specific characteristics of formal educational settings and has not been directly addressed in common TAMs.

## 2. Theoretical background and literature review

### *2.1. The output hypothesis and the role of language production*

The Output Hypothesis, proposed by Swain [7], [8], [9], emphasizes the important role of language production in foreign language learning. According to Swain, speaking and writing are not merely outcomes of learning, but also mechanisms that promote language development. When learners attempt to express ideas in the target language, they are required to attend to grammatical structures, notice gaps in their knowledge, and self-adjust their language use. Swain also argues that output performs three important functions: noticing linguistic problems, testing linguistic hypotheses, and reflection. Thus, through language production activities, learners have opportunities to experiment with new structures, recognize errors, and adjust their interlanguage competence.

Following this perspective, Izumi [11] argues that language production helps learners shift from semantic processing to deeper syntactic processing. In the process of speaking or writing, learners must perform grammatical encoding and self-monitor their language, thereby becoming more aware of the limits of their linguistic competence. This indicates that language production not only serves communicative purposes but also functions as a mechanism for enhancing linguistic awareness.

In addition, studies in the *Handbook of Research in Second Language Teaching and Learning* [12] also affirm that activities such as speaking practice, sentence-making practice, response writing, and communicative tasks play an important role in developing accuracy and fluency among foreign language learners. From a pedagogical perspective, this suggests that grammar teaching should not stop at recognizing structures or selecting correct answers, but should create conditions for learners to directly produce language through meaningful communicative activities.

In the context of Korean language learning in Vietnam, many students experience difficulties in expressing complete sentences, especially complex sentences or sentences containing complicated grammatical structures [6]. One cause of this problem is the typological difference between Vietnamese and Korean, typically represented by the SVO syntactic order in Vietnamese and the SOV order in Korean. This difference makes learners prone to errors in sentence organization, the use of syntactic components, and idea linkage during language production, thereby affecting the accuracy and naturalness of their spoken and written Korean.

Based on the above arguments, this study considers "OD" and "PB" as factors that may influence students' intentions and perceptions regarding the use of an AI-integrated Korean grammar practice system developed in alignment with the curriculum.

### *2.2. Immediate corrective feedback*

Alongside language production, corrective feedback is also regarded as an important factor in the development of foreign language competence. Many studies show that corrective feedback not only helps learners recognize errors but also supports the adjustment and consolidation of linguistic knowledge during the learning process.

Chaudron [13] laid the foundation for developing a descriptive model of corrective feedback as a process aimed at guiding learners toward more accurate language use. According to Chaudron [13], corrective feedback is not a single action of "correcting what is wrong", but a multi-step interaction between teacher and learner. This sequence of interaction helps learners identify errors, try out new ways of expression, and gradually form more accurate language use. This also serves as an important theoretical basis for

current AI-based correction systems, in which AI not only identifies errors but can also offer suggestions, require learners to revise, and continue providing feedback until an appropriate response is achieved.

In addition, Nagata [14] points out that many traditional computer-assisted foreign language learning systems have not provided detailed and personalized feedback to learners. According to Nagata, effective learning systems need to be capable of analyzing errors and providing more detailed feedback, allowing learners to create complete sentences instead of merely completing traditional grammar exercises, thereby effectively supporting the process of language production practice.

Furthermore, Li et al. [15] emphasize that immediate corrective feedback provided during language production helps learners quickly recognize and adjust errors in subsequent sentences, thereby more effectively supporting the development of language competence.

On the other hand, recent studies show that AI is capable of supporting error analysis and providing rapid feedback to learners both inside and outside the classroom [16], [17]. This is particularly appropriate for language production practice activities such as speaking practice and sentence writing practice.

Based on the above theoretical foundations, this study proposes the variable "FN" to measure students' perceived necessity and desire to receive corrections, explanations of errors, and timely feedback during the process of learning Korean grammar.

### *2.3. Technology Acceptance Model - TAM*

This study also applies the TAM, proposed by Davis [18], to explain students' intentions to use technology. According to TAM, two important factors influencing the intention to use (IU) technology are perceived usefulness and perceived ease of use.

According to Davis [18], perceived usefulness refers to the degree to which users believe that a system can help improve their learning or work performance, while perceived ease of use reflects the degree to which users feel that using the system does not require excessive effort. Later studies by Venkatesh and Davis [19] with TAM2 further extended the model by adding social influence and cognitive instrumental factors, while also reaffirming that usefulness and ease of use have a significant influence on users' IU technology.

Therefore, in this study, the variable "TA" is used to measure students' readiness to use an AI-integrated Korean grammar practice system.

### *2.4. Alignment with pedagogical objectives and the curriculum*

In addition to technological factors, many studies have argued that the effectiveness of educational technology depends greatly on its alignment with pedagogical objectives and learners' proficiency levels. This perspective originates from Biggs's Constructive Alignment model [20], according to which teaching and learning activities need to ensure alignment among learning outcomes, learning activities, and assessment methods in order to guide learners toward the intended learning outcomes. On this basis, studies on educational technology have continued to emphasize that technology is effective only when it is designed in alignment with learning objectives, the educational context, and learner characteristics.

Mentis [21] conducted a study examining the influence of three factors: technology, pedagogy, and educational context. On that basis, the author concluded that technology, pedagogy, and educational context need to be closely integrated in order to realize the potential of e-learning in higher education. Similarly, Debele and Plevyak [22] analyzed 33 successful cases of technology integration in education and indicated that close alignment between technology and pedagogy is one of the important factors determining the effectiveness of technology implementation in teaching and learning. In addition, clearly identifying learning objectives and designing activities aligned with training goals also contributes significantly to the success of technology application projects in education.

From this perspective, AI-based learning systems should be developed on the basis of training objectives, curriculum content, and learners' proficiency levels, rather than relying solely on technological capabilities. Therefore, this study proposes the variable "PTA" to measure students' level of trust in an AI system designed in accordance with the teaching materials used at the university.

### *2.5. Description of the proposed system*

Since the system is the object evaluated by students, this section briefly describes the system expected to be developed. It should be emphasized that, at the time of the survey, the system had not yet been programmed, no trial version was available, and the participants had never used it. To ensure that respondents shared the same understanding of the object being evaluated, the author provided a description of the system's objectives and expected functions before participants answered the questionnaire items. The description included the following five main points:

(1) The practice content bank would be developed according to the progression of the textbook and grammar modules in the curriculum, allowing learners to practice the exact grammar points being studied in class.

(2) Practice activities would focus on language production, including making sentences with given grammar points, completing dialogues, and writing short situational paragraphs, instead of multiple-choice exercises aimed at recognizing structures.

(3) Automated and immediate corrective feedback would be provided by AI: the system would indicate the location of errors, explain the causes of errors, and suggest expressions appropriate to learners' proficiency levels.

(4) The scope of knowledge and expressions generated by AI would be limited to the content of the textbook currently in use, in order to restrict the provision of

structures beyond learners' proficiency levels or different from the content already studied in class.

(5) The instructor's supervisory role would be maintained: instructors would participate in editing the corpus bank, approving sample feedback content, and monitoring learners' practice results.

This description served as the basis for participants to conceptualize the system when responding to the questionnaire items, and also served as the foundation for developing the content of the measurement scales. Evaluation based on a conceptual description rather than actual usage experience is a limitation of the study and is discussed in Section 5.3.

In terms of terminology, from this point onward, the article consistently uses the term "system" to refer to the proposed curriculum-aligned AI-integrated Korean grammar practice system. In the questionnaire sent to students (Appendix A), this system is referred to as a "website" because the expected implementation format is a web-based platform; therefore, the questionnaire items in the appendix are kept in their original wording.

### *2.6. Research model and hypotheses*

Based on the theoretical foundations presented above, the study proposes a model consisting of five independent variables: OD, PB, FN, TA, and PTA; and one dependent variable, IU the system, as presented in Figure 1. The rationale leading to each hypothesis is provided immediately before the statement of that hypothesis.

For H1, according to the Output Hypothesis, learners only recognize gaps in their linguistic competence when they are required to directly produce language [7], [8], [9]. When learners frequently experience difficulties in sentence-making and expression, they are more likely to become aware of the need for a tool that provides additional practice opportunities and feedback. On this basis, the study expects a positive relationship between OD and IU.

*- H1: OD are positively associated with IU the system.*

For H2, PB reflects the extent to which learners actively practice outside class hours. Theoretically, the direction of the relationship between this variable and IU may follow two possibilities: learners who actively practice often have stronger learning motivation and are willing to adopt additional support tools. Conversely, those who have already established stable practice habits may have less need to switch to a new platform. This study adopts the positive direction, based on the argument that motivation to practice is a necessary condition for learners to seek support tools, while also acknowledging the possibility of an opposite direction when interpreting the results.

*- H2: PB is positively associated with IU the system.*

For H3, studies on corrective feedback show that timely feedback with explanations helps learners adjust their language use [13], [14], [15]. The proposed system takes immediate corrective feedback as a central function, as described in Section 2.5. Therefore, the greater the learners' need for feedback, the more likely they are to evaluate the system as useful and form an IU it.

*- H3: FN is positively associated with IU the system.*

For H4, according to TAM, perceived usefulness and perceived ease of use are the two main predictors of IU technology [18], [19]. In the context of this study, students' readiness and comfort in learning through online tools are expected to be positively associated with their IU the system.

*- H4: TA is positively associated with IU the system.*

For H5, the perspective on compatibility among technology, pedagogy, and educational context [20], [21], [22] suggests that technology is effective only when it is aligned with learning objectives and learning content. Therefore, when students believe that the system closely follows the textbook, is accurate, and is supervised by instructors, they are more likely to form a stronger IU it.

*- H5: PTA is positively associated with IU the system.*

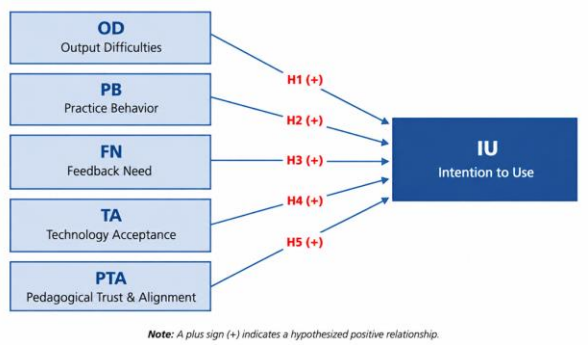



***Figure 1.*** *Proposed Research Model*

*Note: The arrows represent the hypothesized predictive relationships in the model; the plus sign (+) indicates the expected positive direction for all five hypotheses H1–H5.*

To test the research model, the questionnaire was designed using a five-point Likert scale, ranging from 1 (Strongly Disagree) to 5 (Strongly Agree). The questionnaire items were developed based on the theoretical foundations and the characteristics of the Korean language learning context in the university, including reverse-coded items to reduce response bias and enhance the reliability of the measurement scales.

## 3. Research methodology

### *3.1. Research participants*

The research participants were 131 students majoring in Korean Language at N University. The survey sample included first-year students (n = 25, 19.1%), second-year students (n = 61, 46.6%), and third-year students (n = 45, 34.4%), who were selected using convenience sampling.

All students participating in the survey had completed at least one semester in the training program and had taken Korean grammar courses at levels ranging from elementary to intermediate and pre-advanced. The participants' duration of Korean language learning ranged from 6 to 48 months, with a mean of 24.95 months (SD = 9.54). Therefore, they had sufficient learning experience to evaluate difficulties in applying the vocabulary and grammar they had learned to language production practice, as well as their need to use tools that support grammar practice. Regarding Korean

proficiency, the majority of participants had not obtained or had not taken the Test of Proficiency in Korean (TOPIK) (n = 71, 54.2%). Among those who had obtained TOPIK certificates, 10 students had TOPIK Level 1 (7.6%), 17 had TOPIK Level 2 (13.0%), 9 had TOPIK Level 3 (6.9%), 19 had TOPIK Level 4 (14.5%), and 5 had TOPIK Level 5 (3.8%). The sample structure indicates a relatively diverse distribution of Korean proficiency levels, contributing to the representation of different learner groups within the training program.

Because the sample was selected conveniently within a single educational institution, this study is identified as a case study. The results reflect the context of the surveyed institution and are not generalized to all students majoring in Korean Language in Vietnam.

### *3.2. Survey instrument*

The survey instrument used in this study was an online questionnaire designed on the Google Forms platform. The questionnaire was developed based on a synthesis of theoretical foundations related to Swain's Output Hypothesis [7], [8], [9], corrective feedback theory [13], [14], [15], Davis's TAM [18], and perspectives on the alignment between technology, pedagogical objectives, and the curriculum [20], [21], [22]. Before the measurement items, the questionnaire presented a description of the system's expected functions (Section 2.5) to ensure that participants had a shared understanding of the object being evaluated. As stated above, the participants had never actually used the system.

The questionnaire consisted of 35 questions divided into three parts. The first part collected background information about the survey participants, such as academic year, duration of Korean language learning, and TOPIK level. The second part included items measuring students' perceptions and needs regarding the proposed curriculum-aligned AI-integrated Korean grammar practice system. The third part consisted of two open-ended questions designed to collect students' opinions on "difficulties in practicing Korean grammar" and "expectations for an AI-supported learning system at the university".

The variables in the study included OD, PB, FN, TA, PTA, and IU the system. Each variable was measured using five items on a five-point Likert scale, ranging from 1 = "Strongly Disagree" to 5 = "Strongly Agree".

The questionnaire development process was conducted in three steps. First, because no existing set of measurement scales was available that was suitable for the context of a Korean grammar practice system aligned with the curriculum, the questionnaire items were developed by the researcher based on a synthesis of relevant theories. In particular, the two scales of TA and IU the system were developed in the spirit of the TAM model [18], [19], but were reworded to fit the research participants and survey context. Second, the items were reviewed in terms of wording to ensure that the language was clear, appropriate to students' proficiency levels, and did not contain double-barreled content within a single item. Third, the questionnaire was administered in the official survey.

It should be noted that the study did not conduct a pilot test with a small group, nor did it carry out a systematic procedure for consulting experts in Korean language teaching or educational technology before the official survey. This is a methodological limitation, discussed in Section 5.3, and represents an area for improvement in future studies.

### *3.3. Data collection and processing procedures*

Data were collected from April 13, 2026, to May 14, 2026, through an online questionnaire on the Google Forms platform. The survey link was sent to students majoring in Korean Language through class and faculty communication channels. Before participating in the survey, participants were informed of the research purpose, the voluntary nature of participation, and the commitment to protecting their personal information.

After the survey period, 131 valid responses were obtained. The data were checked, coded, and cleaned before analysis to ensure the accuracy and reliability of the research results.

The questionnaire included six reverse-worded items (OD5, PB4, FN5, TA4, PTA5, IU4) to reduce the tendency toward habitual response patterns. All of these items were reverse-coded before calculating scale scores and conducting subsequent analyses, following the principle that the recoded value = 6 − original value for the five-point Likert scale. Thus, all observed variables were aligned in the same direction as the measured construct before reliability testing.

The data were processed using SPSS software. First, descriptive statistics were used to analyze the characteristics of the research sample and the participants' response tendencies. Next, the reliability of the measurement scales was tested using Cronbach's Alpha. Scale items with a Corrected Item–Total Correlation lower than 0.3 were considered for removal, and a scale was accepted when its Cronbach's Alpha coefficient reached 0.6 or higher [23], [24].

After reliability testing, exploratory factor analysis (EFA) was conducted to assess the convergent validity and discriminant validity of the measurement scales. The conditions for conducting EFA included a KMO index greater than 0.5, a statistically significant Bartlett's test ($p < 0.05$), factor loadings greater than 0.5, and a total variance explained of 50% or higher [25], [26], [27].

Finally, multiple linear regression analysis was used to test the research hypotheses regarding the effects of OD, PB, FN, TA, and PTA on IU the website. The level of statistical significance used in the study was $p < 0.05$ [27].

For the data collected from the two open-ended questions, the study used the responses only to support and illustrate the quantitative findings, without conducting a systematic qualitative analysis, such as multi-round coding or inter-coder reliability testing. Specifically, among the 131 participants, 114 answered question (34) on "difficulties in making Korean sentences", and 100 answered question (35) on "expectations for the AI system". All responses were read, synthesized, and

grouped according to similar content, from which five prominent themes were identified for each question. For question (34), the themes included: (1) difficulties related to grammar and vocabulary; (2) difficulties related to structure and word order; (3) difficulties related to context and naturalness; (4) difficulties in creating complex sentences and expressing ideas; and (5) difficulties related to reflexes, speed, and confidence. For question (35), the themes included: (1) alignment with the textbook/curriculum; (2) accuracy and reliability; (3) error correction and explanation; (4) practice, TOPIK, skills, and personalization; and (5) instructor supervision and the university's AI environment. Illustrative quotations were selected based on how representative they were of each theme and were kept in the respondents' original wording. Therefore, the observations drawn from the open-ended question data have only suggestive value, intended to clarify and support the quantitative results, and are not used to infer or confirm the prevalence of opinions across the entire survey sample.

### *3.4. Measurement scales*

The measurement scales in this study were developed based on a synthesis of previous studies on Swain's Output Hypothesis [7], [8], [9], corrective feedback theory [13], [14], [15], Davis's TAM [18], and perspectives on the alignment between technology, pedagogical objectives, and the curriculum [20], [21], [22].

The study used six measurement scales: OD, PB, FN, TA, PTA, and IU the system. Each scale consisted of five observed variables, evaluated on a five-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree).

The operational definitions of the constructs in the study are as follows:

OD refer to the extent to which learners self-assess that they encounter obstacles in forming sentences and expressing ideas in Korean. PB refers to the extent to which learners actively practice sentence-making, speaking, and self-checking errors outside class hours. FN refers to the extent to which learners desire to have errors pointed out, receive explanations of why errors occur, and obtain timely correction suggestions during practice. IU the system refers to the extent to which learners are willing and intend to use the proposed system if it is implemented.

TA is defined as the extent to which learners are ready and comfortable using technology in general and online learning tools in particular for grammar learning, together with their belief that technology helps make learning more effective. This operationalization combines the "perceived usefulness" component and the "comfort and accessibility in use" component of TAM [18] into a unidimensional scale rather than separating them into two distinct scales. The reason is that the object being evaluated does not yet exist in practice; therefore, respondents do not have experiential grounds to distinguish between these two components. It is acknowledged that this operationalization makes some TA items similar in content to IU items. For example, TA1 ("I am willing to use technology in learning grammar") and IU1 ("I am willing to use the website to practice grammar") differ mainly in whether the object referred to is "technology in general" or the "specific website". This is a limitation of the measurement instrument; therefore, results related to TA are interpreted with caution.

PTA is defined as the extent to which learners trust an AI-based learning system implemented by the university and perceive that system as aligned with the curriculum they are following. This construct is operationalized through three aspects: (i) the reliability and accuracy of the content provided by AI (PTA1, PTA2); (ii) the extent to which the system follows the textbook and classroom learning content (PTA3); and (iii) the instructor's academic control over AI-generated content (PTA4). The rationale for combining these three aspects is that, in a formal educational environment, learners' trust in an AI tool does not arise in an abstract manner but is linked to whether the tool falls within the scope of content recognized by the university and is supervised by instructors. In other words, curriculum alignment and instructor supervision are considered foundations for trust. Nevertheless, PTA is a construct proposed by the researcher and has not been validated in previous studies.

## 4. Research results

### *4.1. Reliability testing of the measurement scales*

The reliability of the measurement scales was tested using Cronbach's Alpha before conducting EFA. Generally, observed variables with a Corrected Item–Total Correlation lower than 0.30 are removed, and scales with a Cronbach's Alpha coefficient of 0.70 or higher are considered to meet the acceptable threshold for internal reliability. However, according to Nunnally [23], a Cronbach's Alpha coefficient of 0.60 or higher may be acceptable for exploratory studies, while a level of 0.70 or higher is recommended for confirmatory studies. This view was also synthesized by Straub et al. [28] in their guidelines for scale validation in information systems research. Therefore, this study accepts measurement scales with Cronbach's Alpha coefficients of 0.60 or higher.

After the analysis was conducted, the reliability testing results showed that several observed variables did not meet the reliability requirements and were removed from the scales, including OD5, PB4, FN5, TA4, PTA1, PTA5, and IU4. After adjustment, all measurement scales achieved an acceptable level of reliability, with Cronbach's Alpha coefficients ranging from 0.605 to 0.904.

***Table 1.*** *Cronbach's Alpha Reliability Test Results for the Measurement Scales*

| Construct | Initial Number of Items | Deleted Items | Retained Items | Cronbach's Alpha |
|---|---|---|---|---|
| OD | 5 | OD5 | 4 | 0.831 |
| PB | 5 | PB4 | 4 | 0.605 |
| FN | 5 | FN5 | 4 | 0.904 |
| TA | 5 | TA4 | 4 | 0.816 |
| PTA | 5 | PTA1, PTA5 | 3 | 0.642 |
| IU | 5 | IU4 | 4 | 0.762 |

As shown in the table 1, the FN scale had the highest reliability ($\alpha = 0.904$), followed by OD ($\alpha = 0.831$),

Technology Acceptance (TA) (α = 0.816), and IU the System (α = 0.762). The two scales of PB (α = 0.605) and PTA (α = 0.642) only reached the acceptable threshold for exploratory research and need to be further refined in future studies. With the above reliability levels, the adjusted scales were eligible for subsequent EFA; however, the results related to PB and PTA are interpreted with caution.

One notable point is that all six reverse-worded items (OD5, PB4, FN5, TA4, PTA5, IU4) failed to meet the requirements at the reliability testing stage and were removed from the scales. As presented in Section 3.3, these items had been reverse-coded before analysis; therefore, this result was not due to errors in data coding. The removal of all reverse-worded items may be explained by the fact that negative statements often require an additional step of cognitive processing, making respondents more likely to provide inconsistent answers, especially in an online survey without direct guidance. This is an issue that should be considered when designing questionnaires for future studies with similar participants. A direct consequence is that the number of observed variables in several scales was reduced, with the PTA scale retaining only three variables after PTA1 and PTA5 were removed.

### *4.2. Exploratory Factor Analysis - EFA*

After testing the reliability of the measurement scales, the study conducted EFA on the remaining observed variables to assess the convergent validity and discriminant validity of the scales. Principal Component Analysis combined with Varimax rotation was used in this study. The analysis results showed that the KMO coefficient reached 0.882 (> 0.5), and Bartlett's Test was statistically significant ($p < 0.001$), confirming that the data were suitable for EFA.

***Table 2.*** *Results of the KMO and Bartlett's Tests*

| Statistic | Value |
|---|---|
| KMO | 0.882 |
| Bartlett's Test Chi-square | 1577.174 |
| Sig. | 0.000 |

The EFA results extracted five factors with Eigenvalues greater than 1, with a total variance explained of 65.034%. This level of explained variance exceeds the threshold of 50%, indicating that the extracted factors explained a substantial proportion of the variance in the data. However, the number of extracted factors was five, fewer than the six constructs in the theoretical model, and the factor structure did not fully match the expected scale structure.

Specifically, the rotated matrix (Appendix B) showed that Factor 1 included four variables from the TA scale (TA1, TA2, TA3, TA5), four variables from the IU the System scale (IU1, IU2, IU3, IU5), and one variable from the PTA scale (PTA3). Factor 2 included four variables from the FN scale (FN1–FN4) and PTA4. Factor 3 included four variables from the OD scale (OD1–OD4). Factor 4 included three variables, PB1, PB2, and PB3. Factor 5 included PTA2 and PB5.

Two observations can be drawn from these results. First, the TA and IU scales loaded onto the same factor, meaning that discriminant validity between the two constructs was not established. This is consistent with the observation in Section 3.4 regarding the similarity in wording and content of some items in the two scales. Second, the three remaining variables of the PTA scale were distributed across three different factors (PTA3 in Factor 1, PTA4 in Factor 2, and PTA2 in Factor 5), suggesting that "trust in the accuracy of AI", "alignment with the curriculum", and "the instructor's supervisory role" were not perceived by respondents as manifestations of the same construct. Therefore, the homogeneity of the PTA construct was not established in this study.

In addition, several observed variables had factor loadings that were not entirely clear or showed cross-loadings. The two variables IU5 (0.490) and PTA4 (0.467) did not reach the 0.5 threshold but were retained to preserve the content of the scales. PTA4 also had a loading of 0.431 on Factor 1 and 0.359 on Factor 5; PB3 had a loading of 0.448 on Factor 1 in addition to its primary loading of 0.605; and PB5 had a loading of 0.414 on Factor 4 in addition to its primary loading of 0.629 on Factor 5.

Thus, the data structure was only partially consistent with the proposed measurement model. This represents a limitation of the measurement instrument in this study and does not allow the conclusion that the data structure fully fits the model. Because the scales were developed based on theoretical arguments and to ensure comparability with the proposed hypotheses, the following correlation and regression analyses were still conducted using the six theoretical scales. However, the results are interpreted cautiously and as exploratory, especially for coefficients related to TA and PTA.

***Table 3.*** *Total Variance Explained*

| Factor | Rotated Sums of Squared Loadings | % of Variance | Cumulative % of Variance |
|---|---|---|---|
| 1 | 4.67 | 20.30 | 20.30 |
| 2 | 3.82 | 16.61 | 36.91 |
| 3 | 3.01 | 13.10 | 50.01 |
| 4 | 1.80 | 7.81 | 57.82 |
| 5 | 1.67 | 7.24 | 65.03 |
| Total | 14.96 | 65.03 | - |

*Note: The values reported in the table are the Rotated Sums of Squared Loadings, calculated based on the 23 retained observed variables. The total variance explained reported by the software is 65.034%. The slight discrepancy from the cumulative value shown in the table is due to rounding of the factor loadings in the rotated matrix.*

### *4.3. Pearson correlation analysis*

To assess the relationships among the variables in the research model, Pearson correlation analysis was conducted for the five independent variables, including OD, PB, FN, TA, and PTA, and one dependent variable, IU the System.

The analysis results show that all independent variables had positive correlations with IU the System, and all correlations were statistically significant. Among them,

TA had the strongest correlation with IU (r = 0.713; p < 0.01), followed by PTA (r = 0.590; p < 0.01), FN (r = 0.564; p < 0.01), OD (r = 0.518; p < 0.01), and PB (r = 0.213; p < 0.05).

***Table 4.*** *Pearson Correlation Matrix*

| Variable | OD | PB | FN | TA | PTA | IU |
|---|---|---|---|---|---|---|
| OD | 1 | | | | | |
| PB | 0.124 | 1 | | | | |
| FN | 0.337** | 0.307** | 1 | | | |
| TA | 0.383** | 0.337** | 0.633** | 1 | | |
| PTA | 0.348** | 0.247** | 0.587** | 0.481** | 1 | |
| IU | 0.518** | 0.213** | 0.564** | 0.713** | 0.590** | 1 |

*Note: * p<0.05 ** p<0.01*

In addition, the correlation coefficients among the independent variables were all lower than 0.80, indicating no serious multicollinearity in the research model. Therefore, the data met the conditions for conducting multiple linear regression analysis to test the research hypotheses.

### *4.4. Multiple regression analysis*

#### *4.4.1. Testing the regression model*

To assess the influence of the independent variables on IU the AI-integrated Korean grammar practice website, the study conducted multiple linear regression analysis with five independent variables: OD, PB, FN, TA, and PTA.

***Table 5.*** *Regression Model Fit Statistics*

| Statistic | Value |
|---|---|
| R | 0.798 |
| $R^2$ | 0.636 |
| Adjusted $R^2$ | 0.621 |
| F | 43.691 |
| p | < 0.001 |

The analysis results show that the regression model had an $R^2$ value of 0.636 and an adjusted $R^2$ value of 0.621. This indicates that the five independent variables in the model were able to explain 63.6% of the variance in students' IU the system. The ANOVA test results showed that the regression model was statistically significant, with F = 43.691 and p < 0.001. Thus, the research model fit the survey data and could be used to test the research hypotheses. In addition, the variance inflation factor (VIF) values of the independent variables ranged from 1.151 to 2.045, all below the threshold of 5. This result indicates that multicollinearity did not occur in the model and that the regression estimates had an acceptable level of reliability.

#### *4.4.2. Testing the research hypotheses*

***Table 6.*** *Results of Hypothesis Testing*

| Hypothesis | Path | β | Sig. | Result |
|---|---|---|---|---|
| H1 | OD → IU | 0.232 | 0.000 | Supported |
| H2 | PB → IU | -0.058 | 0.318 | Not Supported |
| H3 | FN → IU | 0.034 | 0.657 | Not Supported |
| H4 | TA → IU | 0.495 | 0.000 | Supported |
| H5 | PTA → IU | 0.265 | 0.000 | Supported |

*Note: β represents the standardized regression coefficient. The VIF values of the independent variables range from 1.151 to 2.045.*

As shown in the table above, the regression results supported only three of the five research hypotheses. Specifically, TA had the strongest positive effect on IU the System (β = 0.495; p < 0.001). This result indicates that when students perceive AI technology as useful, accessible, and aligned with their learning needs, they tend to have a stronger IU the system.

PTA also had a positive effect on IU the website (β = 0.265; p < 0.001). This indicates that students tend to trust and be willing to use AI tools that are designed in alignment with the textbook and guided by the formal educational environment.

Similarly, OD had a positive effect on IU the website (β = 0.232; p < 0.001). This suggests that students who frequently experience difficulties in using Korean grammar have a greater need to seek learning support tools and automated feedback.

In contrast, PB (β = -0.058; p = 0.318) and FN (β = 0.034; p = 0.657) did not have statistically significant effects on IU the System when the other variables were included in the model. Therefore, hypotheses H2 and H3 were not supported by the research data.

In terms of effect size, TA had the strongest effect on IU the System, followed by PTA and OD.

### *4.5. Discussion of the results*

The research results show that three factors, namely TA, PTA, and OD, were statistically significant positive predictors of IU the AI-integrated Korean grammar practice system. Among them, TA had the largest predictive coefficient, followed by PTA and OD. However, because TA and IU did not achieve discriminant validity in the EFA, the magnitude of the coefficients related to TA and the explanatory power of the model may have been overestimated; therefore, the interpretations below are exploratory.

First, the results show that TA was the strongest predictor of IU the System (β = 0.495; p < 0.001). This result is consistent with Davis's TAM [18] and suggests that students tend to be willing to use the system when they believe that AI can effectively support grammar practice and provide rapid feedback. However, as noted in Section 4.2, because TA and IU did not achieve discriminant validity, this coefficient may partly reflect content overlap between the two scales rather than a fully independent predictive relationship.

Second, PTA had a significant positive effect on IU the System (β = 0.265; p < 0.001). This result indicates that students are not only interested in AI technology in general but also pay particular attention to whether that technology is aligned with formal learning content. This finding can be explained through Biggs's perspective [20], according to which learning effectiveness is enhanced when objectives, content, methods, and learning tools are designed consistently with one another. In the context of foreign language learning, an AI system developed based on the textbook and guided by instructors may create a greater sense of trustworthiness and usefulness than general-purpose AI tools.

In addition, some opinions from the open-ended questions, used for illustrative purposes and not subjected to systematic qualitative analysis, are also consistent with this finding. Some students stated that they expected the university's AI system to have high accuracy, closely follow the textbook, and provide more detailed feedback than current popular AI tools. Comments such as "The system is logical and closely follows the textbook", "It has more official and accurate data sources", and "When correcting sentences, it can explain grammatical and vocabulary errors" show that some learners value reliability and alignment with the curriculum.

Third, OD also had a positive effect on IU the System ($\beta = 0.232$; $p < 0.001$). This indicates that students who often encounter difficulties in applying grammar to speaking and writing activities tend to seek learning support tools more actively. This result is consistent with Swain's Output Hypothesis [7], [8], [9], in which language production is viewed as a process that helps learners recognize the gap between their current competence and the target language goals to be achieved. When learners experience difficulties in expression, they have a greater need for support tools capable of providing practice opportunities and timely feedback, such as an AI-integrated grammar practice system.

Furthermore, data from the open-ended question on "difficulties in making Korean sentences" recorded several issues, such as selecting appropriate grammar, lack of vocabulary, and arranging sentence order. Some students reported difficulties in applying grammatical structures to real-life contexts and often felt uncertain about the accuracy of the sentences they produced. Since the open-ended responses were synthesized only for illustrative purposes, these opinions are not used to confirm the prevalence of such issues in the sample. However, their content is consistent with the regression results, in which OD was a positive predictor of IU the System.

Fourth, one notable finding is that although FN had a positive correlation with IU the System ($r = 0.564$; $p < 0.001$), it was no longer a statistically significant predictor in the regression model ($\beta = 0.034$; $p = 0.657$). This result suggests that the need to receive corrective feedback may not be an independent motivation driving IU the system when other factors are controlled. One possible interpretation is that the role of FN was indirectly reflected through other factors such as TA or PTA. In other words, students may not simply want to receive corrective feedback, but may be more concerned with whether such feedback is provided by a technology they trust and whether it is aligned with their learning content. It should be emphasized that this is only a possible interpretation; this study did not conduct mediation analysis to verify it, and this hypothesis needs to be examined in future studies.

Similarly, PB had a positive correlation with IU the System ($r = 0.213$; $p < 0.05$), but it was not significant in the regression model ($\beta = -0.058$; $p = 0.318$). This result suggests that students' frequent grammar practice does not necessarily mean that they need to use a new AI platform. One possible explanation is that students who have already developed their own learning habits may still prefer to continue using familiar methods rather than switch to a new tool. This explanation is consistent with the possible opposite direction of the relationship mentioned when formulating H2 (Section 2.6), although it was not directly tested in this study. Within the scope of the current data, it can only be concluded that the current level of practice was not a statistically significant predictor of IU the System when considered simultaneously with technological factors and alignment with the curriculum.

Overall, within the scope of the data in this study, factors related to perceptions of technology and alignment with the educational context played a more important role than current learning behavior factors in shaping students' IU an AI-integrated Korean grammar practice system.

## 5. Conclusion and practical implications

### *5.1. Conclusion*

This exploratory study was conducted to identify factors influencing perceptions of and intentions to use a proposed curriculum-aligned AI-integrated Korean grammar practice system, with Korean language majors at N University as the research participants. Because the sample was selected conveniently within a single educational institution, this study is a case study, and the conclusions below are tied to that specific context.

The results of data analysis from 131 students show that the research model explained 63.6% of the variance in IU the system. Three factors were statistically significant positive predictors of IU the system, namely TA, PTA, and OD. Among them, TA had the strongest effect. In contrast, PB and FN were not significant predictors when considered simultaneously with the other factors in the model.

Within the research context, the findings show that learners' intentions to use AI-integrated learning tools depend not only on their learning needs but are also associated with their level of TA and the degree to which the technology aligns with the formal curriculum. It should be noted that this study measured IU a proposed system; therefore, it does not allow conclusions to be drawn about actual usage behavior or the learning effectiveness of the system.

### *5.2. Practical implications*

Because this study was limited to self-reported IU a system that has not yet been developed, the following implications should be understood as suggested design priorities for educational institutions, instructors, and developers of AI-integrated learning systems. A clear distinction should be made between IU and learning effectiveness: the results of this study may serve as a reference basis for designing and piloting the system, but they are not sufficient to confirm that the system will improve learners' learning outcomes.

First, educational institutions should consider developing or integrating AI-supported grammar practice platforms into the process of teaching and learning Korean. The research results show that TA was the factor with the

strongest influence on IU. Therefore, ease of use and the capacity to effectively support learning should be prioritized in system design.

Second, the content of the AI system should be developed in close alignment with the curriculum and textbooks currently used at educational institutions. The research results show that students tend to trust and be more willing to use AI tools designed to fit formal learning content. This suggests that developing AI systems specialized for individual training programs may be more effective than using general-purpose AI tools.

Third, the system should focus on supporting language production activities, such as sentence writing, paragraph writing, dialogue completion, or practicing the use of grammar in real-life communicative contexts. These are activities in which students often experience difficulties and are also associated with learner groups that have a higher need to use the system.

### *5.3. Limitations of the study*

Despite certain contributions, this study still has several limitations as follows.

First, the research sample was collected using convenience sampling and focused only on students majoring in Korean Language; therefore, the generalizability of the results to other groups of learners remains limited.

Second, the Korean proficiency levels of the participants were determined mainly through self-reported information and TOPIK results, if available. However, TOPIK was used only as a reference indicator of Korean proficiency, not as a criterion for sample selection. Although many students had not taken the TOPIK, they had still completed grammar courses in the curriculum.

Third, the study was limited to investigating IU and did not evaluate actual usage behavior or the actual learning effectiveness of the AI system.

Fourth, participants evaluated the system based on a functional description in the questionnaire rather than through actual usage experience. Therefore, the results reflect expectations toward a proposed system and may change when learners interact with a real product.

Fifth, the questionnaire was developed by the researcher and had not undergone pilot testing or a systematic expert consultation process before the official survey. Therefore, scale validation remained exploratory.

Sixth, in terms of measurement, all reverse-worded items were removed at the reliability testing stage. The EFA results showed that the TA and IU the System scales did not achieve discriminant validity, while the items of the PTA scale were distributed across multiple factors. Therefore, the measurement scales need to be further refined, particularly by separating the content of TA and IU and clarifying the structure of the PTA construct, followed by retesting using confirmatory factor analysis (CFA).

Seventh, the data from the two open-ended questions were used only for illustrative purposes and were not subjected to systematic qualitative analysis; therefore, they were not used to confirm the prevalence of opinions in the survey sample.

In the future, subsequent studies may expand the survey population, combine qualitative and quantitative methods, and conduct practical trials with AI-supported learning systems to evaluate their effectiveness in real learning environments.

**Acknowledgement:** This article is part of the institutional-level scientific research project coded T2025-05-16.

## REFERENCES

[1] L. Kohnke, B. L. Moorhouse, and D. Zou, "ChatGPT for Language Teaching and Learning," *RELC Journal*, vol. 54, no. 2, pp. 537–550, 2023.

[2] J. Du, "A systematic review of AI-powered chatbots for English as a foreign language speaking," *Computers and Education: Artificial Intelligence*, vol. 6, Art. no. 100230, 2024.

[3] W. Wiboolyasarin, "AI-driven chatbots in second language education: A systematic review of their efficacy and pedagogical implications," *Ampersand*, vol. 14, Art. no. 100224, 2025.

[4] J. Rho and Y.-H. Ko, "A Study on the Effectiveness of AI-Based Korean Conversational Learning," *Journal of Korea Academia-Industrial Cooperation Society*, vol. 26, no. 12, pp. 933–940, 2025.

[5] Lee Ji-eun, "A Study on Korean Language Learning Methods Using Generative AI: - Focusing on Intermediate-level Expression Education," *Master's thesis, Graduate School of Mokwon University,* Daejeon, Republic of Korea, 2025.

[6] P. T. Duyen, "Grammar Errors in Korean Writing Skill of First-Year Students of Korean Language and Culture Department, University of Foreign Languages, Hue University: Current Situation and Solutions," *Journal of Science Hong Bang International University*, Special Iusse 2, p. 165–174, 2024, doi: 10.59294/HIUJS.KHQG.2024.017.

[7] M. Swain, "Communicative competence: Some roles of comprehensible input and comprehensible output in its development," in *Input in Second Language Acquisition*, S. Gass and C. Madden, Eds. Rowley, MA: Newbury House, 1985, pp. 235–253.

[8] M. Swain, "The output hypothesis and beyond: Mediating acquisition through collaborative dialogue," in *Sociocultural Theory and Second Language Learning*, J. P. Lantolf, Ed. Oxford, U.K.: Oxford University Press, 2000, pp. 97–114.

[9] M. Swain, "The output hypothesis: Theory and research," *in Handbook of Research in Second Language Teaching and Learnin*g, E. Hinkel, Ed. Mahwah, NJ, USA: Lawrence Erlbaum Associates, 2005, pp. 471–483.

[10] T. T. Ngoc, N. T. T. Thuy, and N. M. Tam, "A study on the application model of AI ChatGPT to improve writing skills for Korean language-majored students," *The University of Danang - Journal of Science and Technology*, vol. 23, no. 8A, p. 123–128, 2025, doi: 10.31130/ud-jst.2025.23(8A).056.

[11] S. Izumi, "Comprehension and Production Processes in Second Language Learning: In Search of the Psycholinguistic Rationale of the Output Hypothesis," *Applied Linguistics*, vol. 24, no. 2, pp. 168–196, Jun. 2003.

[12] E. Hinkel, Ed., *Handbook of Research in Second Language Teaching and Learning*. Mahwah, NJ, USA: Lawrence Erlbaum Associates, 2005.

[13] C. Chaudron, "A Descriptive Model of Discourse in the Corrective Treatment of Learners' Errors," *Language Learning*, vol. 27, no. 1, pp. 29–46, 1977.

[14] N. Nagata, "Intelligent Computer Feedback for Second Language Instruction," *The Modern Language Journal*, vol. 77, no. 3, pp. 330–339, 1993.

[15] S. Li, Y. Zhu, and R. Ellis, "The Effects of the Timing of Corrective Feedback on the Acquisition of a New Linguistic Structure," *The Modern Language Journal,* vol. 100, no. 1, pp. 276–295, 2016.

[16] B. Elmotri *et al.*, "The Impact of AI-Generated Feedback

Explicitness," *Arab World English Journal*, vol. 16, no. 1, pp. 384–402, 2025.

[17] I. X. Wang *et al.*, "*Neural Automated Writing Evaluation with Corrective Feedback*," arXiv:2402.17613, 2024.

[18] F. D. Davis, "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989.

[19] V. Venkatesh and F. D. Davis, "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies," *Management Science*, vol. 46, no. 2, pp. 186–204, 2000.

[20] J. Biggs, "Enhancing Teaching through Constructive Alignment," *Higher Education,* vol. 32, no. 3, pp. 347–364, 1996.

[21] M. Mentis, "Navigating the e-Learning Terrain: Aligning Technology, Pedagogy and Context," *The Electronic Journal of e-Learning,* vol. 6, no. 3, pp. 217–226, 2008.

[22] M. Debele and L. Plevyak, "Conditions for Successful Use of Technology in Social Studies Classrooms," *Computers in the Schools,* vol. 29, no. 3, pp. 285–299, 2012.

[23] J. C. Nunnally, *Psychometric Theory*. New York, NY, USA: McGraw-Hill, 1967.

[24] P. R. Hinton, C. Brownlow, I. McMurray, and B. Cozens, *SPSS Explained*. London, U.K.: Routledge, 2004.

[25] H. F. Kaiser, "An Index of Factorial Simplicity," *Psychometrika*, vol. 39, no. 1, pp. 31–36, 1974.

[26] M. S. Bartlett, "A Note on the Multiplying Factors for Various Chi Square Approximations," *Journal of the Royal Statistical Society. Series B (Methodological),* vol. 16, no. 2, pp. 296–298, 1954.

[27] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, *Multivariate Data Analysis, 7th ed.* New York, NY, USA: Pearson, 2010.

[28] D. Straub, M.-C. Boudreau, and D. Gefen, "Validation Guidelines for IS Positivist Research," *Communications of the Association for Information Systems*, vol. 13, Art. no. 24, 2004.

## APPENDICES

**Appendix A. Questionnaire**

**Part 1. General information**

1. What academic year are you currently in?

2. What is your current TOPIK level?

3. How long have you been studying Korean?

**Part 2. Items measuring perceptions and intention to use the website**

*(Five-point Likert scale: 1 = Strongly Disagree; 2 = Disagree; 3 = Neutral; 4 = Agree; 5 = Strongly Agree)*

OUTPUT DIFFICULTIES - OD

OD1. I have difficulty expressing ideas in Korean.

OD2. I often do not know how to start a sentence.

OD3. I avoid using complex grammar when speaking/writing.

OD4. I feel unconfident when creating sentences by myself.

OD5. I can easily use grammar when writing sentences. (R)

PRACTICE BEHAVIOR - PB

PB1. I often practice writing Korean sentences outside class hours.

PB2. I actively seek ways to practice speaking using grammar.

PB3. I use online tools to learn grammar.

PB4. I rarely practice making sentences outside class. (R)

PB5. I have the habit of checking my own errors when writing.

FEEDBACK NEED - FN

FN1. I need to have errors corrected immediately when writing/speaking.

FN2. I want to know why a sentence is incorrect.

FN3. I want to receive specific correction suggestions.

FN4. I find it difficult to improve without feedback.

FN5. I do not need error correction during practice. (R)

TECHNOLOGY ACCEPTANCE - TA

TA1. I am willing to use technology in learning grammar.

TA2. I feel comfortable learning through a website.

TA3. I believe that technology helps make learning more effective.

TA4. I have difficulty using technology in learning. (R)

TA5. I like using online learning tools.

PEDAGOGICAL TRUST AND ALIGNMENT - PTA

PTA1. I have greater trust in an AI system designed according to he textbook.

PTA2. I am concerned that external AI tools may be inaccurate.

PTA3. I want grammar practice content to closely follow the classroom curriculum.

PTA4. I feel that AI integrated into the website needs to be supervised by instructors.

PTA5. I do not care whether the content follows the textbook or not. (R)

INTENTION TO USE - IU

IU1. I am willing to use the website to practice grammar.

IU2. I will use the website regularly if it is available.

IU3. I want to use the website instead of traditional practice methods.

IU4. I have no intention of using this system. (R)

IU5. I will recommend the website to my friends.

**Part 3. Open-ended questions**

1. What is your greatest difficulty when making Korean sentences?

2. How do you expect the university's AI system to differ from external tools?

**Note: Items marked with (R) are reverse-worded items and were reverse-coded before analysis, following the formula: recoded value = 6 − original value.*

**Appendix B. Rotated component matrix**

**Rotated Component Matrix[a]**

| | Component | | | | |
|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| TA5 | .753 | .251 | .034 | .168 | -.169 |
| TA2 | .732 | .170 | .274 | .040 | -.041 |
| TA3 | .722 | .376 | .081 | .118 | -.029 |
| IU1 | .689 | .317 | .176 | .031 | .171 |
| IU2 | .668 | .294 | .225 | -.086 | .207 |
| IU3 | .578 | -.115 | .437 | -.171 | .123 |
| TA1 | .545 | .412 | .079 | .126 | .103 |
| PTA3 | .523 | .382 | .080 | -.018 | .323 |
| IU5 | .490 | .230 | .227 | -.009 | .317 |
| FN3 | .391 | .842 | .047 | .069 | .072 |
| FN2 | .354 | .842 | .115 | .045 | .107 |
| FN1 | .167 | .809 | .099 | .120 | .082 |
| FN4 | .276 | .706 | .219 | -.012 | .194 |
| PTA4 | .431 | .467 | .178 | -.213 | .359 |
| OD2 | .073 | .008 | .819 | .052 | -.063 |
| OD3 | .271 | .094 | .775 | .091 | .109 |
| OD4 | .270 | .127 | .763 | -.035 | .025 |
| OD1 | .071 | .300 | .745 | .067 | .117 |
| PB1 | -.127 | .020 | .115 | .755 | -.092 |
| PB2 | .058 | -.024 | -.010 | .712 | .238 |
| PB3 | .448 | .255 | .061 | .605 | -.022 |
| PTA2 | -.021 | .181 | .173 | -.004 | .806 |
| PB5 | .184 | .128 | -.179 | .414 | .629 |

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 7 iterations.