JKSDH

Original Article

The educational effectiveness and learning satisfaction of video lectures utilizing text-to-speech technology

Seung-Hun Lee[iD]

Department of Dental Hygiene, Sahmyook University

Corresponding Author: **Seung-Hun Lee,** Department of Dental Hygiene, Sahmyook Health University, 82 Mangu-ro, Dongdaemun-gu, Seoul-si, 02500, Korea. Tel: +82-2-3407-8621, Fax: +82-2-3407-8639, E-mail: S2022067@shu.ac.kr

ABSTRACT

Objectives: This study investigated the impact of lecture videos that incorporated text-to-speech (TTS) technology on cognitive outcomes, effective learning, and learning satisfaction among dental hygiene students. **Methods:** Cognitive, effective, and satisfaction responses were measured using a 5-point Likert scale and analyzed using t-tests, ANOVA, and correlation analyses. **Results:** The average cognitive outcomes score was approximately 3.0, with several items scoring below this level. Significant grade-level differences were observed in effective learning and learning satisfaction, with first-year students reporting more positive responses than those of higher-year students. Strong positive correlations were found among cognitive outcomes, effective learning, and learning satisfaction (r=0.94–0.95, *p*<0.01). **Conclusions:** Although TTS-based lectures improve accessibility and facilitate repeated revision, further strategies are required to encourage sustained engagement and higher-order cognitive processing, such as incorporating human voices, interactive designs, and contextual reinforcement, are required to encourage sustained engagement and higher-order cognitive processing. The TTS resources are best positioned as supplementary or pre-learning tools in dental hygiene education.

Keywords: Cognition, Dental hygiene, Emotions, Speech synthesis, Student satisfaction

Introduction

The prevailing educational paradigm is undergoing a rapid transformation due to the advent of digital technology, the proliferation of mobile devices, and repercussions of the COVID-19 pandemic [1]. Contemporary university students have been utilizing digital devices since early childhood and have a keen interest in artificial intelligence technology [2].

As higher education continues to evolve in response to digital transformation, diverse instructional formats—such as online video courses and AI-supported learning—have become increasingly prevalent [3]. Among these, text-to-speech (TTS) technology stands out for its unique ability to deliver auditory content, offering new possibilities for multimodal learning. Its applications in language education, particularly in speaking and writing instruction, are well-documented [4].

Concomitant with this evolution in higher education, a plethora of innovative educational models are emerging. These include online video classes, AI-assisted learning, and video lectures that are tailored to different learning levels [3]. In the field of educational technology, the integration of innovative technologies to enhance the appeal and efficacy of teaching and learning processes has emerged as a prominent area of focus. Among these technological solutions, TTS technology has garnered significant attention due to its potential to transform the learning environment. Its utilization in language education, encompassing speaking and writing, is well-documented [4].

Whilst earlier research has indicated a propensity to place excessive reliance on TTS outputs without the need for review or revision, these studies have also emphasized its capacity as an effective learning instrument, thereby compensating for inadequate English proficiency [5,6]. Despite the necessity for enhancement of the naturalness of the audio output, studies have demonstrated the potential for TTS-generated audio to be employed beneficially in the context of language education [7]. The educational effects of listening and reading audio-books have been found to be superior to those of paper-based books in terms of learning motivation, comprehension, and reading ability [8].

Dental hygiene education is a specialized domain that requires the integration of theoretical knowledge and practical skills. Learner comprehension and engagement are especially critical due to the technical and procedural nature of the curriculum. While AI-based TTS technology has been increasingly adopted in general language education, including applications in automatic translation, listening, writing, and reading. There is still a notable lack of empirical research on its effectiveness in professional education contexts, including studies on cognitive and affective learning outcomes or learner satisfaction with TTS-based lecture videos. Research focusing on dental hygiene students is even more limited, with virtually no studies examining the educational impact of TTS-based instruction.

Therefore, the present study aims to address this gap by investigating the educational effectiveness of TTS-based lecture videos in dental hygiene education. The study focuses on learners' satisfaction, perceived comprehension, and engagement with TTS-generated content. To achieve this, a quantitative survey design was employed to analyze learners' responses and provide insights into their experiences and perceptions.

Methods

1. Subjects

The study participants consisted of 313 students from the Department of Dental Hygiene at Sahmyook Health University, including those enrolled in the Advanced Bachelor's Degree Completion Program. A total of 174 students (55.6%) completed the questionnaire via an online form tool (Google Forms, Google LLC, Mountain View, CA, USA) over a 7-day period from October 20 to 26, 2025. In the final data analysis, 27 students were excluded from the study due to either insincere responses or failure to watch the TTS-based lecture videos. As a result, 147 valid responses, constituting 47% of the total, were included in the final analysis.

The minimum required sample size was calculated using the G*power (ver. 3.1.9.7; Heinrich-Heine-University, Düsseldorf, Germany) with a statistical power of 95%, a significance level of 0.05, an effect size of 0.6, and the analysis method set as an independent t-test. The result indicated that a minimum of 122 participants were required for the study. In light of the projected attrition rate of 60%, a minimum of 305 participants was deemed necessary to ensure the adequacy of the final sample. Consequently, the final sample of 313 participants satisfied the stipulated minimum sample size requirement.

The effect size was derived with reference to the cognitive outcomes from a previous study [9]. The study was reviewed and approved by the Institutional Review Board of Sahmyook Health University (IRB No. 2-70094958-A-N-01-250526-HR-003-01), and informed consent was obtained from all participants prior to the study.

2. Methods

All students participated in a three-week lecture series, with each week comprising two 30-minute sessions. The lectures were delivered as TTS video content and made available for one week each via the university's e-Class system (learning management system), allowing students to view the materials flexibly within the designated period.

The TTS audio was generated using Microsoft Azure Neural Text-to-Speech (Microsoft Corp., Redmond, WA, USA). Two distinct voice profiles were applied: a male voice with a neutral and friendly tone was used to represent the instructor, while a female voice with a higher pitch and expressive, curious tone was used to simulate a student persona.

The speech rate and prosody were set to moderate levels to ensure clarity and engagement.

The overall difficulty level of the TTS lectures was slightly lower than that of standard in-person classes, with content simplified for accessibility. Additionally, the lectures were condensed from the typical 50-minute classroom format to 30-minute video segments to enhance focus and accommodate self-paced learning.

The scripts were generated using Copilot (Microsoft Corp., Redmond, WA, USA) to simulate dialogic learning between instructor and student (Socratic method). The creation of visual slides was facilitated by the utilization of PowerPoint 365 (Microsoft presentation software), while the text was converted into speech using Clipchamp (Microsoft Corp., Redmond, WA, USA). The audio and visual components were then edited together to produce the final instructional videos.

The educational content varied by academic year:

First-year students were presented with a series of lectures on the subject of Dental Materials and Practice. The content of these lectures encompassed a variety of topics, including the properties of dental materials, resin and amalgam, and dental alloys and casting.

In their second year, students embarked on an in-depth exploration of Prosthodontics, encompassing the fabrication of fixed prostheses such as crowns and bridges, the design of removable partial dentures, and the construction of complete dentures.

Third-year students engaged with content from Oral Health Statistics, focusing on sampling methods, oral health surveys, and indicators of oral health status (e.g., dental caries, periodontal disease, fluorosis).

In the fourth year, students delved into the domain of digital dentistry, encompassing the intricacies of CBCT (Cone-Beam Computed Tomography), the functionality of intraoral 3D scanners, and the operational mechanisms of dental CAD/CAM systems.

The survey items were revised and supplemented with reference to a previous study [9]. The educational outcomes were assessed using a total of 33 items, comprising 10 items for the cognitive domain, 8 items for the effective domain, 11 items for educational satisfaction, and 4 items for general characteristics of the participants. The measurement of educational effectiveness and satisfaction was conducted using a 5-point Likert scale. The internal consistency reliability (Cronbach's α) was 0.974 for the cognitive, 0.968 for the effective, and 0.975 for satisfaction.

3. Data analysis

The collected data were analyzed using SPSS program (ver. 18.0; IBM Corp., Armonk, NY, USA), with a significance level set at 0.05. The general characteristics of participants were analyzed using descriptive statistics. The differences in educational effectiveness and satisfaction according to general characteristics were examined using independent t-tests or ANOVA, with post-hoc tests performed using Scheffé method. In addition, Pearson correlation analysis was conducted to examine the relationships among cognitive outcomes, effective learning, and learning satisfaction.

Results

The present study found that 98.6% of the participants were female (n=145). The distribution by academic year was as follows: first-year students 26.5% (n=39), second-year students 24.5% (n=36), third-year students 38.1% (n=56), and Advanced Bachelor's Degree Completion Program 10.9% (n=16). Regarding age, 65.3% (n=96) of the subjects were 21 years or younger, while 34.7% (n=51) were 22 years or older.

1. A comparison of cognitive outcomes according to general characteristics

The results of the study indicate that there are differences in cognitive outcomes according to the general characteristics (age and academic grade) of the research participants. Male participants (n=2) were excluded from the study population due to the small sample size, and fourth-year students (n=16) were excluded because they did not meet the criteria for a normal distribution.

As shown by <Table 1>, participants over the age of 22 years demonstrated significantly higher scores in Work performance impact (item 7) ($p<0.05$). In contrast, significant differences were also found in Long-term retention (item 8) and Concentration enhancement (item 9), with participants aged 21 or younger scoring higher ($p<0.01$). The mean scores for all ten cognitive outcome measures remained below 3.0, indicating a moderate level of perceived effectiveness.

This suggests that while students did not report high levels of cognitive benefit, certain areas—particularly those related to foundational understanding—may still hold potential for instructional value.

Statistically significant differences were also identified across academic grade. With regard to Problem-solving support (item 5), first-year students (3.15 points) and third-year students (2.89 points) formed a statistically similar group, both of which scored significantly higher than second-year students (2.11 points) ($p<0.01$).

In a similar vein, the Long-term retention index (item 8) exhibited a highly significant discrepancy across grades ($p<0.001$), with post-hoc analysis indicating that first-year students (3.54 points) scored significantly higher than both second-year (2.44 points) and third-year students (2.75 points), who formed a statistically similar group. This suggests that first-year students formed a distinct higher-performing group in terms of long-term retention, whereas second- and third-year students belonged to a statistically similar lower-performing group.

Furthermore, a significant difference in results of Self-directed learning (item 10) was observed according to grade ($p<0.05$). Post-hoc analysis indicated that first-year (3.15 points) and third-year students (2.93 points) formed a statistically similar group, both scoring significantly higher than second-year students (2.33 points).

Table 1. Differences in cognitive outcomes by the general characteristics   Unit: Mean±SD

Item	Age			Grade				
	≤21(n=96)	>22(n=51)	(t)p^{*}	1(n=39)	2(n=36)	3(n=56)	Total(n=131)	(F)p^{**}
1. Clarity comparison	2.90±1.45	3.31±1.10	(-1.950) 0.053	3.15±1.25	2.89±1.47	2.93±1.50	2.98±1.41	(0.402) 0.670
2. Concept understanding	3.03±1.48	3.39±1.27	(-1.480) 0.141	3.23±1.20	2.89±1.47	2.89±1.49	2.99±1.40	(0.802) 0.451
3. Ease of comprehension	2.95±1.44	3.24±1.37	(-1.173) 0.243	3.08±1.22	2.56±1.36	2.93±1.50	2.87±1.39	(1.414) 0.247
4. Logical flow	2.91±1.37	2.92±1.19	(-0.067) 0.946	3.08±1.16	2.67±1.35	2.79±1.49	2.84±1.36	(0.931) 0.397
5. Problem-solving support	2.88±1.32	2.76±1.11	(-0.508) 0.612	3.15[a]±1.18	2.11[b]±0.89	2.89[a]±1.44	2.76±1.29	(7.322) 0.001
6. Practical application	2.84±1.26	3.04±1.13	(-0.927) 0.355	3.23±1.14	2.78±1.05	2.75±1.39	2.90±1.24	(2.004) 0.139
7. Work performance impact	2.77±1.26	3.22±1.05	(-2.167) 0.032	3.08±1.09	2.89±1.12	2.68±1.40	2.85±1.24	(1.207) 0.303
8. Long-term retention	2.97±1.33	2.33±1.18	(2.289) 0.005	3.54[a]±1.09	2.44[b]±0.97	2.75[b]±1.44	2.90±1.29	(8.164) <0.001
9. Concentration nhancement	2.86±1.51	2.29±1.06	(2.659) 0.009	3.08±1.22	2.33±1.27	2.79±1.58	2.75±1.42	(2.683) 0.072
10. Self-directed learning	2.80±1.31	3.10±1.06	(-1.479) 0.142	3.15[a]±1.25	2.33[b]±1.07	2.93[a]±1.35	2.83±1.28	(4.351) 0.015

*by independent samples t-test; **by one-way ANOVA
[abc]The same characters do not differ significantly by Scheffé's post-hoc test.

2. A comparison of effective learning accoding to general characteristics

<Table 2> presents the differences in effective learning according to age and academic year. No statistically significant differences were found across all effective learning items (p>0.05) based on age.

However, the analysis by academic year revealed significant differences in several items. Specifically, statistically significant differences were found in learning interest (item 1, p<0.05), learning attitude (item 3, p<0.05), and educational values (item 5, p<0.05). Post-hoc analysis indicated that first-year and third-year students formed a statistically similar group, both scoring significantly higher than second-year students in learning interest. The mean scores were 3.23 for first-year students, 2.44 for second-year students, and 3.11 for third-year students.

In learning attitude, first-year students (3.15 points) scored significantly higher than second-year students (2.32 points), while third-year students (2.82 points) did not differ significantly from either group. Thus, first- and third-year students can be interpreted as forming a higher-performing group compared to second-year students.

For educational values, first-year (3.08 points) and third-year students (2.93 points) again formed a statistically similar group, both scoring significantly higher than second-year students (2.31 points). These results suggest that second-year students consistently belonged to a lower-performing group across these effective domains, while first- and third-year students shared similar levels of positive effective responses.

Table 2. Differences in effective learning by the general characteristics   Unit: Mean±SD

Item	Age			Grade				
	≤21(n=96)	>22(n=51)	(t)p^*	1(n=39)	2(n=36)	3(n=56)	Total(n=131)	(F)p^{**}
1. Learning interest	3.03±1.32	3.00±0.98	(0.162) 0.871	3.23^a±0.90	2.44^b±1.26	3.11^a±1.38	2.96±1.26	(4.519) 0.013
2. Learning motivation	2.71±1.49	2.59±0.83	(0.627) 0.532	2.92±1.09	2.33±1.35	2.75±1.52	2.69±1.37	(1.877) 0.157
3. Learning attitude	2.80±1.36	2.82±1.11	(-0.104) 0.922	3.15^a±1.04	2.32b±1.17	2.82^{ab}±1.43	2.79±1.28	(4.046) 0.020
4. Learning willingness	2.78±1.45	2.51±0.86	(1.427) 0.156	3.00±1.05	2.34±1.35	2.75±1.47	2.71±1.34	(2.417) 0.093
5. Educational values	2.81±1.12	2.86±0.87	(-0.291) 0.772	3.08^a±1.01	2.33b±0.83	2.93^{ab}±1.35	2.81±1.16	(4.626) 0.011
6.7Reusability	2.82±1.43	3.12±1.19	(-1.258) 0.211	3.23±1.06	2.67±1.59	2.86±1.45	2.92±1.40	(1.639) 0.198
7. Learning immersion	2.67±1.43	2.69±1.07	(-0.094) 0.925	2.85±1.11	2.22±1.15	2.75±1.54	2.63±1.34	(2.463) 0.089
8. Emotional engagement	2.32±1.43	2.41±1.02	(-0.434) 0.665	2.46±1.41	2.11±1.30	2.39±1.38	2.34±1.37	(0.696) 0.501

*by independent samples t-test; **by one-way ANOVA
^{abc}The same characters do not differ significantly by Scheffé's post-hoc test.

3. A comparison of learning satisfaction to general characteristics

The investigation into age and academic grade to regard on learning satisfaction yielded the following findings <Table 3>. When comparing the two age groups, no significant differences were observed across the majority of items. However, a statistically significant difference was identified in Content clarity (item 1, p<0.05). Participants aged 22 years and older reported higher levels of satisfaction (3.63 points) than those aged 21 years or younger (3.19 points). For the remaining items, no statistically significant age-related differences were identified (p>0.05).

A comparison of results by academic year revealed statistically significant variations across multiple dimensions of learning satisfaction. A number of statistically significant differences were identified in Content clarity (item 1, $p<0.01$), Voice appropriateness (item 2, $p<0.05$), Interest (item 4, $p<0.01$), Learning convenience (item 5, $p<0.01$), Time/Place flexibility (item 6, $p<0.001$), Re-watch intention (item 7, $p<0.01$), Usability (item 8, $p<0.01$), Positive impact on learning (item 9, $p<0.001$), and Overall satisfaction (item 10, $p<0.01$). Post-hoc analysis revealed that in most of these items, first-year and third-year students formed a statistically similar group, both of which scored significantly higher than second-year students.

For example, in Content clarity (item 1), first-year (3.77 points) and third-year students (3.21 points) did not differ significantly, but both scored significantly higher than second-year students (2.78 points). A similar pattern was observed in Voice appropriateness, Interest, Re-watch intention, and Usability.

In Learning convenience (item 5), first-year students (3.38 points) scored significantly higher than both second-year (2.43 points) and third-year students (2.64 points), who formed a statistically similar lower-performing group.

In Time/Place flexibility (item 6), all three groups differed significantly from one another, with first-year students (4.00 points) scoring highest, third-year students (3.18 points) in the middle, and second-year students (2.11 points) lowest.

In Positive impact on learning and Overall satisfaction, first-year and third-year students again formed a higher-performing group, while second-year students reported significantly lower satisfaction.

These findings suggest that second-year students consistently belonged to a lower-performing group across multiple satisfaction dimensions, while first- and third-year students demonstrated similar and significantly higher levels of satisfaction. In some cases, first-year students stood out as the most satisfied group, particularly in areas such as Time/place flexibility and learning convenience.

Table 3. Differences in learning satisfaction by the general characteristics Unit: Mean±SD

Item	Age			Grade				
	≤21(n=96)	>22(n=51)	(t)p^*	1(n=39)	2(n=36)	3(n=56)	Total(n=131)	(F)p^{**}
1. Content clarity	3.19±1.34	3.63±1.18	(-1.963) 0.052	3.77[a]±0.71	2.78[b]±1.25	3.21[ab]±1.56	3.26±1.32	(5.763) 0.004
2. Voice appropriateness	2.90±1.37	3.27±1.15	(-1.688) 0.094	3.11[a]±1.08	2.44[b]±1.18	2.93[ab]±1.50	2.91±1.33	(4.128) 0.018
3. Immersion	2.61±1.44	2.75±1.21	(-0.552) 0.582	2.85±1.25	2.22±1.15	2.71±1.59	2.62±1.39	(2.141) 0.122
4. Interest	2.77±1.46	3.02±1.38	(-1.002) 0.318	3.23[a]±1.20	2.21[b]±1.14	2.82[ab]±1.57	2.78±1.40	(5.183) 0.007
5. Convenience	2.80±1.39	3.04±1.17	(-.1097) 0.275	3.38[a]±1.09	2.43[b]±1.08	2.64[b]±1.58	2.81±1.37	(5.526) 0.005
6. Time/place flexibility	3.17±1.48	3.18±1.26	(-0.040) 0.968	4.00[a]±0.79	2.11[b]±1.01	3.18[c]±1.69	3.13±1.48	(19.880) <0.001
7. Re-watch intention	2.77±1.38	2.65±1.15	(-0.580) 0.563	3.15[a]±1.11	2.22[b]±1.15	2.82[ab]±1.55	2.76±1.37	(4.731) 0.010
8. Usability	2.71±1.46	2.76±1.19	(-0.237) 0.813	3.31[a]±1.22	2.09[b]±1.00	2.64[ab]±1.58	2.69±1.40	(7.574) 0.001
9. Positive impact on learning	2.93±1.33	2.96±1.41	(-0.143) 0.887	3.39[a]±1.16	2.23[b]±1.05	2.94[a]±1.43	2.87±1.32	(7.694) <0.001
10. Overall satisfaction	2.80±1.34	2.79±1.25	(-0.008) 0.994	3.23[a]±1.14	2.12[b]±1.02	2.89[a]±1.49	2.78±1.33	(8.153) 0.001

[*]by independent samples t-test; [**]by one-way ANOVA
[abc]The same characters do not differ significantly by Scheffé's post-hoc test.

4. The correlation between cognitive outcomes, effective learning outcomes, and learner satisfaction

The results of the Pearson correlation analysis revealed strong positive relationships among cognitive outcomes, effective learning, and learner satisfaction <Table 4>. The cognitive outcomes demonstrated a statistically significant correlation with effective learning (r=0.948, p<0.01) and learner satisfaction (r=0.937, p<0.01). Furthermore, effective learning demonstrated a strong correlation with learner satisfaction (r=0.948, p<0.01). The findings indicate a close association between all three variables, with higher levels of one factor corresponding to higher levels of the others.

However, given that the average scores for cognitive outcomes remained below 3.0 across most items, the results do not provide sufficient evidence to conclusively assert the effectiveness of TTS-based lectures in enhancing cognitive learning.

Therefore, it would be more appropriate to interpret these findings as suggesting the potential of TTS-based video lectures to support foundational concept learning, rather than confirming their effectiveness.

Table 4. Correlation between cognitive outcomes, effective learning outcomes, and learner satisfaction

Variables	Cognitive outcomes	Effective learning	Learner satisfaction
Cognitive outcomes	1.000		
effective learning	0.948[*]	1.000	
Learner satisfaction	0.937[*]	0.948[*]	1.000

[*]p<0.01, by Pearson's correlation coefficient

Discussion

This study examined the educational impact of lecture videos that incorporated TTS technology on dental hygiene students, with a focus on cognitive outcomes, effective learning and satisfaction with the learning experience. Although the 60% attrition rate among respondents might appear to be high, it should be noted that comparable or even higher rates of attrition have been reported in the context of online surveys and learning environments in previous studies. For instance, Rahmani et al. [10] and Shaikh & Asif [11] emphasize that attrition rates in online higher education and surveys, particularly in settings with voluntary participation, commonly range between 30% and 60%. Consequently, the 60% attrition rate employed in this study was a conservative estimate that ensured sufficient statistical power and accounting for potential non-response bias.

The TTS-based group achieved overall mean cognitive scores around the midpoint, with several items, particularly problem-solving, long-term retention, and concentration, falling below 3.0. These results indicate that learners perceived constraints in cognitive depth provided by TTS-based instruction rather than simple neutrality toward the format. The consistently low cognitive ratings appear to stem from structural limitations inherent to synthesized speech. Reduced prosodic variation and limited expressive cues likely weakened attentional regulation and hindered semantic organization, thereby lowering perceived clarity and problem-solving support [12,13]. The absence of dialogic interaction and adaptive pacing further constrained higher-order reasoning [14].

Caution is warranted in interpreting these findings, as the observed effects may reflect limited engagement with higher-order cognitive processes such as problem-solving and applied reasoning. Compared with lecture videos recorded by human instructors, which provide paralinguistic cues such as intonation, stress and emotional tone to guide attention and regulate cognitive load, TTS-based narration lacks expressive modulation. This reduced naturalness may hinder learner' ability to process complex concepts and to distinguish essential content from supplementary information during cognitively demanding segments [12,13].

Emerging research indicates that TTS voice profiles influence learning differently. Voices with richer prosody enhance attention and reduce cognitive load, whereas flatter voices diminish engagement [12]. Although this study employed two contrasting profiles, it did not measure voice-specific effects. Future studies should systematically compare acoustic features to determine which profiles optimize cognitive and effective outcomes.

Despite cognitive limitations, learners reported strengths in clarity and ease of review, consistent with studies indicating that TTS enhances accessibility and content revisitation [7,8]. Moreover, previous studies have shown that AI-based learning content can enhance self-directed learning by enabling learners to control their study pace and engage in repeated practice [9]. This finding aligns with the results of the present study, which indicated that specific groups of learners, particularly first-year students with less academic experience, experienced significantly greater cognitive benefits. Consequently, TTS-based lectures may offer potential benefits during the initial phase of learning, particularly in supporting the acquisition of introductory concepts.

However, the lower cognitive ratings found among second and third year students, many of whom fell below the midpoint, suggest that TTS-based lectures may not have met the cognitive expectations of more advanced learners. There learners typically require richer explanatory depth, adaptive feedback, and sophisticated conceptual scaffolding. This interpretation is consistent with research indicating that learners' expectations for instructional depth and interactivity increase with academic level, making low prosody and low interaction TTS formats less suitable for advanced coursework [15,16]. Thus, the grade level differences in cognitive outcomes appear more closely tied to pedagogical fit and instructional design misalignment than to learner ability itself.

No significant differences related to age were observed in the effective learning outcomes, yet substantial grade-level variations were evident. First-year students achieved higher scores in learning interest, learning attitude, and perceived educational value than sophomores. In multiple areas, including interest, attitude and value, their scores also surpassed those of juniors. This phenomenon aligns with principles of motivational design theory. According to Keller's ARCS model, attention and relevance are proximal drivers of satisfaction. Entry-level learners often demonstrate a heightened sensitivity to novelty and relevance cues embedded in new media [17].

Beyond the realm of motivation theory, extensive research in the field of online learning has emphasized the pivotal role of course and instructor design, alongside dialogic interaction, in shaping learner satisfaction and perceived learning outcomes. These findings lend support to the notion that novice learners, who rely more heavily on structure and guidance, exhibit stronger positive emotional responses when the learning material is meticulously scaffolded [15,16]. The findings of this study are consistent with this account. The higher effective ratings, such as interest and attitude, among first-year students align with this perspective. However, the lower ratings observed in second-year students suggest the presence of an expectation-experience gap, necessitating targeted design improvements, as discussed in this study.

Emerging evidence suggests that synthetic voices have the capacity to influence learners' emotions and engagement in different ways. Several comparative studies have reported divergent perceptions of AI-generated and human voices. Indeed, there are instances in which specific synthesized voice styles have been shown to negatively impact attention and learning performance. This may help explain the muted emotional responses observed when prosodic cues are limited [18,19]. As posited by Charpentier-Jiménez [18] and Jing et al. [19], the accessibility and potential for repeated viewing afforded by TTS can support positive emotions by reducing effort and enabling self-paced learning. This is particularly advantageous for novice users [14].

No significant age differences were observed in learning satisfaction, except for content clarity. However, multiple grade-level disparities emerged in areas such as content clarity, voice appropriateness, interest, convenience, time/place flexibility, intention to re-watch, usability, positive impact on learning, and overall satisfaction. First-year students demonstrated the highest levels of satisfaction, while sophomores consistently exhibited the lowest. Notably, time/place flexibility had the strongest significance in the present study. These findings align with established models of online learner satisfaction [15,16]. Such models suggest that course design and structure, instructor support and dialogic interaction are key predictors of satisfaction and perceived learning. These factors tend to benefit novice learners, who rely more heavily on scaffolding and a predictable structure.

The importance of flexibility and usability in our data is consistent with the idea that course flexibility, perceived ease of use, system quality, and course quality are critical determinants of e-learning satisfaction [20,21]. Sun et al. [20] and Martin and Bolliger [21] posit that the combination of on-demand access and the ability to rewatch TTS video lectures plausibly increases convenience and perceived control. This, in turn, supports satisfaction via the attention–relevance–confidence–satisfaction linkages described

by Keller's ARCS model [17]. As Keller [17] asserts, voice characteristics have been shown to influence satisfaction. However, recent experimental findings have reported instances in which synthesized voices diminish attention and engagement compared to human voices. This phenomenon may reduce effective responses and subsequent satisfaction, particularly when prosody is constrained [19]. Jing et al [19] argue in their study that the lower sophomore ratings observed across several items in the sample can be attributed to a growing expectation-experience gap, whereby students progress and seek richer prosody, dialogue, and instructor presence [15,16].

The present study demonstrated strong positive correlations between cognitive outcomes, effective learning, and learning satisfaction (r=0.94−0.95, p<0.01). This indicates that learners with a deeper understanding of the content tend to be more emotionally engaged and report higher satisfaction with the learning experience. This finding is consistent with the theoretical perspective that learning is influenced not only by cognitive processes, but also by emotional responses such as interest, perceived value and attitude. These emotional responses act as motivational drivers of internal engagement and sustained effort [17]. Furthermore, previous research in online and technology-supported learning environments has shown that emotional involvement enhances cognitive processing, and that positive emotions contribute to satisfaction, which in turn reinforces continued learning behaviors [15,16].

Despite these insights, several limitations should be acknowledged. Firstly, the study was conducted at a single institution with a sample limited to dental hygiene students, which may restrict the generalizability of the findings. Future research should include diverse academic disciplines and multiple institutions samples to enhance external validity. Secondly, reliance on self-reported data introduces the possibility of response bias. Incorporating objective measures such as pre/post-tests or behavioral metric would provide a more comprehensive understanding of learning outcomes. Thirdly, the absence of qualitative analyses limited insight into learner' subjective experiences with TTS, including discomfort with voice quality or the role of repeated playback. A mixed-methods approach could capture nuanced perceptions of prosody, fatigue, and learning strategies [14]. Fourthly, the lack of interactivity and human voice elements in TTS videos may have constrained engagement and cognitive processing. Exploring hybrid models that integrate TTS with human narration or adaptive feedback is recommended. Lastly, longitudinal studies are needed to assess the sustained impact of TTS-based instruction on knowledge retention and skill transfer. Given that novice learners in this study reported higher effective satisfaction, future studies should also consider tailored engagement strategies for intermediate learners to bridge potential expectation−experience gaps [14−16].

Taken together, the results highlight that cognitive understanding, emotional involvement and satisfaction are mutually reinforcing elements of the learning process. Therefore, optimizing TTS-based lectures in dental hygiene education should involve enhancing informational clarity. It should also include the deliberate design of emotionally engaging, learner-centered experiences, such as interactive elements, relatable clinical examples, and adaptive pacing support. These improvements are likely to enhance cognitive, emotional, and satisfaction outcomes, thereby strengthening the instructional value of TTS-based resources in both foundational and advanced learning contexts [15−17]. Future research should continue to refine these approaches to ensure that TTS-based resources can serve not only as supplementary tools but also core components of effective, inclusive, and engaging digital education.

Conclusions

This study examined the educational impact of lecture videos incorporating text-to-speech (TTS) technology on dental hygiene students.

1. While TTS-based lectures supported basic conceptual understanding and aided content acquisition, their limited prosody and vocal expressiveness constrained higher-order cognitive processing, including problem solving and long-term retention.

2. Effective responses differed significantly by academic grade. First-year students demonstrated greater learning interest, a more positive learning attitude and a higher perceived value, suggesting greater responsiveness to novelty and structured support. In contrast, sophomores consistently reported lower effective responses, indicating a potential expectation−experience gap.

3. Learning satisfaction also varied by grade level, with first-year students reporting the highest satisfaction and second-year students the lowest, particularly with regard to convenience, usability and time/place flexibility. These results highlight the importance of instructional interaction and learner support in sustaining satisfaction.

4. Strong positive correlations were observed among cognitive outcomes, effective learning and learning satisfaction. This suggests that meaningful learning occurs when cognitive understanding and emotional engagement reinforce each other.

Taken together, TTS-based lecture videos are most effective when used as supplementary or pre-learning materials, particularly for foundational coursework. To maximize effectiveness, TTS resources should be supplemented with human-voiced explanations, interactive learning components, and contextualized clinical examples to in order to support sustained engagement and higher-order cognitive development.

Notes

Author Contributions

The author fully participated in the work performed and documented truthfully.

Conflicts of Interest

SH Lee is a member of the Editorial Committee of the Journal of the Korean Society of Dental Hygiene, but was not involved in the review process of this manuscript. The author declared no other conflicts of interest.

Funding

None.

Ethical Statement

This study was approved by the Institutional Review Board (IRB) of Sahmyook Health University (IRB No. 2-70094958-A-N-01-250526-HR-003-01).

Data Availability

Data can be obtained from the corresponding author.

Acknowledgements

None.

References

1. Lee WG, Kim JM. Curriculum development for AI convergence education. Korean J Converg Humanit 2020;8(3):29–52. https://doi.org/10.14729/converging.k.2020.8.3.29
2. Bennett S, Maton K. Beyond the 'digital natives' debate: towards a more nuanced understanding of students' technology experiences. J Comput Assist Learn 2010;26(5):321–31. https://doi.org/10.1111/j.1365-2729.2010.00360.x
3. Zou Y, Kuek F, Feng W, Cheng X. Digital learning in the 21st century: trends, challenges, and innovations in technology integration. Front Educ 2025;10:1562391. https://doi.org/10.3389/feduc.2025.1562391

4. Park JC. A case study on the use of speech synthesis technology (TTS) in Korean speaking class for academic purpose: focusing on presentation tasks. J Korean Lang Educ 2021;32(3):141–60. https://doi.org/10.18209/iakle.2021.32.3.141

5. Kim KR. Translator-assisted L2 writing, necessary or not?: beginner university learners' perceptions of its validity. J Digit Converg 2020;18(6):99–108. https://doi.org/10.14400/JDC.2020.18.6.099

6. Kim HK, Han SM. College students' perceptions of AI-based writing learning tools: with a focus on Google Translate, Naver Papago, and Grammarly. Mod Engl Educ 2021;22(4):90–100. https://doi.org/10.18095/meeso.2021.22.4.90

7. Park JC. A study on the production of Korean listening education materials utilization of artificial intelligence (AI): focusing on using speech synthesis program (TTS). Biling Res 2021;82:61–84. https://doi.org/10.17296/korbil.2021..82.61

8. Park YM, Kim KS. Korean reading education using audiobooks: effects of reading based on listening. Asia-Pac J Converg Res Interch 2024;10(2):651–62. https://doi.org/10.47116/apjcri.2024.02.50

9. Kim MK. PBL using AI technology-based learning tools in a college English class. Korean J Gen Educ 2023;17(2):169–83. https://doi.org/10.46392/kjge.2023.17.2.169

10. Rahmani AM, Groot W, Rahmani H. Dropout in online higher education: a systematic literature review. Int J Educ Technol High Educ 2024;21:19. https://doi.org/10.1186/s41239-024-00450-9

11. Shaikh UU, Asif Z. Persistence and dropout in higher online education: review and categorization of factors. Front Psychol 2022;13:902070. https://doi.org/10.3389/fpsyg.2022.902070

12. Atkinson RK, Mayer RE, Merrill MM. Fostering social agency in multimedia learning: examining the impact of an animated agent's voice. Contemp Educ Psychol 2005;30(1):117–39. https://doi.org/10.1016/j.cedpsych.2004.07.001

13. Mayer RE. Multimedia learning. 2nd ed. Cambridge: Cambridge University Press; 2009: 41–65, 243–70. https://doi.org/10.1017/CBO9780511811678

14. Widyana A, Jerusalem MI, Yumechas B. The applications of text-to-speech technology in language learning: a systematic review. Proceedings of the Sixth International Conference on Language, Literature, Culture, and Education (ICOLLITE 2022) 2022: 85–92. https://doi.org/10.2991/978-2-494069-91-6_14

15. Eom SB, Ashill N. The determinants of students' perceived learning outcomes and satisfaction in university online education: an update. Decis Sci J Innov Educ 2016;14(2):185–215. https://doi.org/10.1111/dsji.12097

16. Paechter M, Maier B. Online or face-to-face? Students' experiences and preferences in e-learning. Internet High Educ 2010;13(4):292–7. https://doi.org/10.1016/j.iheduc.2010.09.004

17. Keller JM. Motivational design for learning and performance: the ARCS model approach. New York: Springer; 2010: 19–44, 67–112. https://doi.org/10.1007/978-1-4419-1250-3

18. Charpentier-Jiménez W. A comparison of synthetic and human speech: an evaluation by English as a foreign language students in a public Costa Rican university. Comunicación 2023;32(2):41–58.

19. Jing B, Wu C, Pi Z, Zhou Y, Zhang Y, Liu H. Cute computer-synthesized voice hinders learning performance in instructional videos. J Exp Educ 2024;1–18. https://doi.org/10.1080/00220973.2024.2446169

20. Sun PC, Tsai RJ, Finger G, Chen YY, Yeh D. What drives a successful e-learning? An empirical investigation of the critical factors influencing learner satisfaction. Comput Educ 2008;50(4):1183–202. https://doi.org/10.1016/j.compedu.2006.11.007

21. Martin F, Bolliger DU. Developing an online learner satisfaction framework in higher education through a systematic review of research. Int J Educ Technol High Educ 2022;19:50. https://doi.org/10.1186/s41239-022-00355-5

음성 합성(TTS) 기술을 활용한 동영상의 교육적 효과와 만족도

초록

연구목적: 본 연구는 텍스트 음성 변환(TTS) 기술을 적용한 강의 영상이 치위생과 학생들의 인지적 학습 성과, 정서적 학습, 학습 만족도에 미치는 영향을 조사하였다. **연구방법:** 인지적, 정서적, 만족도 반응을 5점 리커트 척도로 측정하고 t-검정, 분산분석, 상관분석을 통해 분석하였다. **연구결과:** 인지적 학습 성과 평균 점수는 약 3.0이었으며, 일부 항목은 이 수준 미만으로 나타났다. 정서적 학습과 학습 만족도에서는 학년별 유의한 차이가 관찰되었으며, 1학년생이 고학년생보다 더 긍정적인 반응을 보였다. 인지적 학습 성과, 정서적 학습, 학습 만족도 간에는 강한 양의 상관관계가 확인되었다(r=0.94-0.95, $p<0.01$). **결론:** TTS 기반 강의는 접근성을 개선하고 반복 학습을 쉽게 하지만, 지속적인 참여와 고차원적 인지 처리를 촉진하기 위해서는 인간 목소리 통합, 상호작용적 디자인, 맥락적 강화와 같은 추가 전략이 필요하다. TTS 자료는 치위생과 교육에서 보조 학습 도구나 사전 학습 도구로 활용하는 것이 가장 적합하다.

색인: 인지, 치위생, 감정, 음성 합성, 학생 만족도