Zanco Journal of Human Sciences (ZJHS)

TYPE: Original article
DOI https://doi.org/10.21271/zjhs.30.SpB.36

Application of Krashen's Input Hypothesis in a Listening Comprehension Blended Learning Classroom

OPEN ACCESS
*Corresponding author
miran00496176@uoh.edu.iq
Miran Abdulrahman Abdullah

Miran Abdulrahman Abdullah\ MA Candidate in Applied Linguistics, Department of English Language, College of Basic Education, University of Halabja, Halabja 46018, Kurdistan Region, Iraq

Omar Fouad Ghafor\ Assistant Professor in Applied Linguistics Department of English Language, College of Basic Education, University of Halabja, Halabja 46018, Kurdistan Region, Iraq

RECEIVED: 10/12/2025
ACCEPTED: 10/02/2026
PUBLISHED: 15/04/2026

Keywords:
Krashen's Input Hypothesis, comprehensible input, blended learning, listening comprehension, Kurdish EFL adolescents.

Abstract

Listening remains central in EFL classrooms but is still under-studied, particularly for adolescents in low-resource settings. This study aimed to examine whether meaning-focused Comprehensible Input (CI) delivered in a blended format outperforms conventional accuracy-focused drills. Two intact eighth-grade classes (N = 44) at a K–8 school in Halabja, Kurdistan Region, Iraq, completed eight weeks of instruction. Listening was measured with parallel forms of the Cambridge A2 Key (KET) test, and baseline comparability was checked before analysis. An ANCOVA with pre-test as covariate showed a clear group effect favoring CI, $F(1, 41) = 13.83$, $p < .001$, partial $\eta^2 = .252$; adjusted means were 16.75 (SE = 0.30) for the experimental (CI) group and 15.16 (SE = 0.30) for the control group. A gain-score check pointed the same way ($d = 1.05$). Taken together, the findings suggest that abundant, level-matched input—supported by simple smartphone-based tools—can improve EFL listening for secondary learners. The study used class-level assignment (two clusters) and online engagement was not equated across groups; therefore, the results should be read with caution. To strengthen the evidence and aid classroom adoption, future work should use more classes with cluster randomization, include fidelity checks, and add follow-up testing to see whether gains last.

About the Journal
Zanco Journal of Humanity Sciences (ZJHS) is an international, multi-disciplinary, peer-reviewed, double-blind and open-access journal that enhances research in all fields of basic and applied sciences through the publication of high-quality articles that describe significant and novel works; and advance knowledge in a diversity of scientific fields. https://zancojournal.su.edu.krd/index.php/JAHS/about

1. Introduction

Listening dominates everyday language use (Rost, 2013). Yet, for all its importance to communicative competence, it has received far less scholarly attention than the other skills (Vandergrift & Goh, 2012). In many EFL classrooms, instruction still centres on accuracy-driven drills that isolate discrete items rather than fostering meaning construction (Field, 2009); as a result, learners rarely encounter the volume and variety of input required for steady gains in comprehension (Nation & Newton, 2008).

Krashen's Input Hypothesis argues that language acquisition occurs when students are exposed to comprehensible input, meaning messages they can understand, while that input is pitched slightly beyond their current proficiency, the well-known "i + 1" principle (Krashen, 1982, 1985). The concept remains foundational to contemporary models of implicit learning and is consistently linked to improvements in receptive abilities such as listening and reading (Ellis, 2015). Applied to listening, it suggests that abundant, meaning-focused exposure, supported by visuals, redundancy, and learner control, should promote comprehension more effectively than tasks that interrupt input processing to focus on form (Long, 2015).

Blended learning refers to the planned combination of face-to-face teaching with web-based platforms (Garrison and Kanuka, 2004). In this study, the blended environment was smartphone-mediated and used Telegram, Google Sites, and Google Forms alongside regular classroom lessons, enabling instructors to curate audiovisual listening, adjust delivery, and embed scaffolds such as visuals, transcripts, and playback control without displacing classroom time (Mayer, 2009).

This study examined whether a comprehensible input-based listening programme, delivered within a smartphone-mediated blended environment, would improve listening comprehension for eighth grade Kurdish EFL learners more than conventional accuracy focused practice. Two intact classes at Shahid Azad Basic School for Girls in Halabja participated in an eight-week cluster-randomized pilot comparing Krashen aligned materials with standard textbook activities. The research question was: Does exposure to comprehensible input in a blended learning framework significantly improve the listening comprehension of Iraqi Kurdish eighth grade EFL students compared with traditional accuracy focused instruction?

Key terms in this study are used as follows. Krashen's Input Hypothesis refers to acquisition through exposure to messages that learners understand while being pushed slightly beyond their current level ("i + 1"). Comprehensible input in this study refers to meaning-focused listening supported by visuals, redundancy, and learner control so that learners can understand the message while processing language. Blended learning refers to the planned combination of face-to-face teaching with web-based platforms; in this study, the blended environment was implemented through smartphone-mediated tools (Telegram, Google Sites, and Google Forms) alongside regular classroom instruction. Traditional accuracy-focused instruction refers to the textbook routine described here, where listening work is centred on discrete-item comprehension questions and gap-fills targeting specific words or structures. Listening comprehension refers to performance on the Cambridge A2 Key Listening test described in Section 3.4. Building on this rationale, the next section reviews empirical work on blended listening to position our study.

2. Literature Review

2.1 Theoretical Background

Krashen's (1982, 1985) proposal centres on comprehensible input pitched at i + 1, enabling learners to attend to meaning while incidentally acquiring form. Two long-standing critiques are relevant for classrooms: (a) the hypothesis offers limited guidance for calibrating "+1" and (b) it under specifies the complementary roles of interaction and output (Long, 2015; Swain, 1985). Listening is a

stringent test case: speech is transient and places heavy demands on processing. Reviews and classroom studies show that input-rich environments and metacognitive support can benefit listening (Rost, 2013; Vandergrift and Goh, 2012; Graham and Macaro, 2008). Blended ecosystems amplify access to graded input and allow scaffolds (e.g., transcripts, visuals, playback control) consistent with multimedia learning principles (Mayer, 2009). In short, if level matched, meaning centered input is a principal driver of acquisition, a blended design should yield measurable listening gains for adolescents.

A growing body of evidence indicates that blended learning boosts learning outcomes from primary through tertiary levels, and this benefit extends to language-teaching environments as well (Halverson et al., 2017; Graham, 2019). However, most studies focus on adults in high-resource settings, leaving evidence gaps for adolescents and for lower-resource contexts.

2.2 Previous Studies

Building on the theoretical case outlined above, empirical findings relevant to adolescent EFL listening are now reviewed. When EFL learners receive a steady stream of level-matched and engaging input through technology, listening comprehension improves, and when online work is blended with classroom support, the gains are often larger than with traditional practice alone.

A fully web-based example comes from Al Rashidy and Alsabbagh (2023), who built a seven-week Moodle program for Egyptian English majors with five units and fifteen lessons that moved through pre-listening, during listening, and post-listening stages. Using a one-group pretest and posttest design with twenty-five items, they found a clear improvement in overall listening scores from a mean of 14.65 to 18.16 with a significant t value of 6.263 and p below .01. Subskills showed a mixed picture. Listening for details, listening for specific information, drawing conclusions, and predicting showed significant gains, while main idea, inference, speaker purpose, and attitude did not change reliably. The researchers argued that constant access to authentic audio, instant feedback, and learner control made the difference, which lines up well with the logic of comprehensible input and a low affective filter.

A broader comparison of delivery modes strengthens the case for blending. Zolfaghari et al. (2023) worked with sixty intermediate learners at two Iranian institutes and divided them into three groups of twenty. One group learned through a fully virtual mode. One group followed a blended schedule that combined online work with in-person sessions. One group used traditional face-to-face instruction. All three covered the same Touchstone 3 content and sat listening and reading tests before and after the course. Both experimental groups outperformed the control, and the blended group showed the largest listening improvement by a wide margin with a mean that rose from 8.20 to 14.40 and a paired t of 6.56 with $p < .001$ and a very large eta squared of .86. The control group did not show a significant change. The authors note that blending kept the social and interactional advantages of a classroom while multiplying exposure to meaningful audio outside class time, which again matches an input first account of listening development.

In a similar Middle Eastern context, Al Mansour (2019) conducted a large-scale dissertation study in Saudi Arabia. The research set a blended learning model (at Saudi Electronic University) against a traditional one (at Imam Muhammad ibn Saud Islamic University). The blended group's post-test listening scores were significantly higher ($M = 7.00$) than the traditional group's ($M = 4.15$) on a 19-item test. Al Mansour also found that the blended students were significantly more satisfied. They reported positive perceptions of the model, calling it more motivating and flexible, and specifically noted the benefit of having online video access to native speakers.

Further evidence points in the same direction. In Indonesia, Sholihah et al. (2018) blended ordinary class meetings with an online library of graded stories and short talks. After two action-research

cycles, listening averages and course pass rates improved, and students said the extra digital practice made them feel less anxious and more prepared.

Across these studies, several themes are consistent. When students receive large amounts of understandable input through video, audio, and interactive tasks, their comprehension scores improve. Blended learning arrangements that combine face-to-face classroom instruction with structured online listening activities outperform both traditional classroom-only instruction and purely virtual formats because they maintain classroom interaction while expanding learners' exposure to listening input outside class time. In this context, traditional classroom-only instruction refers to listening lessons based on textbook recordings followed by discrete-item comprehension questions and gap-filling exercises, with no structured online extension of listening practice. Learners also describe the blended approach as more flexible and less stressful, which likely lowers anxiety and helps intake. At the same time, most of these studies have focused on adults in universities. Very few studies examine younger learners or low-resource contexts, and many rely on teacher-made tests instead of standardized measures. The present study addresses these gaps by working with Kurdish eighth graders, using parallel forms of the Cambridge A2 Key listening test, and analyzing the data with ANCOVA to control for initial differences. By combining story based comprehensible input with simple smartphone tools in a blended framework, it extends earlier findings to adolescent learners and tests whether abundant, meaning focused input can enhance listening comprehension more effectively than accuracy drills. Taken together, this literature motivates our two-class, cluster-randomized pilot to test whether a CI-oriented blend outperforms traditional accuracy-focused practice under typical school constraints.

3. Method

This study addresses a common EFL classroom problem in which instruction still centres on accuracy-driven drills that isolate discrete items rather than fostering meaning construction; as a result, learners rarely encounter the volume and variety of input required for steady gains in comprehension. This study addressed that problem by asking: Does exposure to comprehensible input in a blended learning framework significantly improve the listening comprehension of Iraqi Kurdish eighth grade EFL students compared with traditional accuracy focused instruction? The objective was to examine whether a comprehensible input-based listening programme, delivered within a smartphone-mediated blended environment, would improve listening comprehension for eighth grade Kurdish EFL learners more than conventional accuracy focused practice. Because most studies focus on adults in high-resource settings, evidence gaps remain for adolescents and for lower-resource contexts. Methodologically, the study used a two-class, cluster-randomized pilot design with pre- and post-tests.

3.1 Context and Participants

The study was conducted at the public Shahid Azad Basic School for Girls in Halabja, Iraqi Kurdistan Region. Before the study started, the eighth-graders, who are 13 to 14 years old, already had several years of classroom exposure because English is taught there three periods a week starting in Grade 1. The Halabja Directorate of Education and the school principal gave their administrative approval for the study.

There were five complete classes (roughly 110 students) in the eighth grade that year. School midterm records were reviewed to keep the timetable intact while maintaining group comparability. To minimise timetable disruption and maximise comparability, the school's English coordinator selected the two sections that were taught by the same instructor (the researcher) and met on the same daily schedule; the other three sections had different teachers and/or timetables. Each class had 22 students (N = 44 in total). The two classes differed by less than one percentage point in mean

English scores (78.4% vs. 77.7%). The two groups were appropriately similar, as attendance records also showed that none of the students in either class had studied overseas or was enrolled in private language classes.

3.2 Design

The study used a two-class, cluster-randomized pilot design with pre- and post-tests. Two intact eighth-grade classes served as the units of assignment. At the start of the term, the school's English coordinator randomly assigned one class to the comprehensible-input (CI) condition and the other to the traditional, accuracy-focused condition; students stayed in their original classes throughout. Each class contained 22 pupils, for a total sample of 44. The two classes were the only eighth-grade sections taught by the same instructor (the researcher) on the same timetable, which motivated their selection for comparability.

Instruction ran for eight weeks. Each class received five 40-minute lessons per week, and the same instructor taught both sections on the same timetable; in this paper, "the researcher" refers to the instructor who delivered the lessons and implemented the two programmes. The independent variable was the instructional condition. The outcome of interest was listening performance on the Cambridge A2 Key Listening test administered after the intervention, with parallel forms used at pre- and post-test. In the primary analysis, the pre-test score was included as a covariate to adjust for any initial differences between the classes.

Delivery mode was held constant. Both groups followed the same blended arrangement (Telegram, Google Sites, and Google Forms), while the listening materials and follow-up tasks differed by condition as described in Section 3.3.

3.3 Treatments

To test whether a meaning-based approach would outperform a form-focused one, each class followed a distinct listening programme while sharing the same instructor, timetable, and online tools (Telegram, Google Sites, and Google Forms). What differed was the nature of the input and the kinds of follow-up tasks students completed.

The comprehensible-input (CI) group worked with short, level-appropriate stories presented first as video clips that included pictures, gestures, and captions. After each segment the researcher paused for a brief comprehension check so the material stayed just beyond the class's current level, Krashen's i + 1. Learners then read an illustrated version of the story that contained only a few "pop-up" grammar notes. Outside class, they logged into Google Sites to watch the same A2-level videos, consult glossaries and transcripts, and try auto-feedback quizzes. Optional extension stories were available for anyone who wanted more. Students also posted voice-note retellings on Telegram and replied to open-ended prompts, which added an element of light output and peer interaction.

The traditional group followed the textbook's accuracy-oriented routine. In class, they listened to the standard recordings, answered comprehension questions, and filled in gaps that targeted specific words or structures. Online homework mirrored this focus on form: students watched ungraded, standard-speed videos chosen only because the topics matched the unit theme, completed multiple-choice listening tests, and received text-only instructions. Telegram served mainly as a noticeboard for deadlines and reminders; no retellings were required.

The same instructor implemented both conditions, and lesson length and digital platforms were held constant across groups so that the main difference was the type of listening input and follow-up practice. However, the CI group's online retellings likely increased interaction and practice opportunities relative to the control group; this is a primary confound. Future replications will equalize output/interaction across groups (e.g., require the same retellings in both or in neither) so

that any difference reflects input focus rather than extra practice.

3.4 Instruments

The outcome measure was the Cambridge A2 Key (KET) Listening test (25 items) (Cambridge Assessment English, 2020). This instrument was selected because it is aligned with the CEFR A2 descriptors, is widely used in school settings, and offers parallel forms suitable for pre–post designs. In the study, students completed two different forms at pre-test and post-test to reduce practice effects. In this paper, equivalent/parallel forms means that the two tests follow the same format and scoring (25 items; one point per correct item; maximum 25) and are intended to be comparable in difficulty as described in the Cambridge A2 Key handbook (Cambridge Assessment English, 2020). The two forms were treated as equivalent based on the handbook description of parallel forms and scoring. No separate reliability or internal-consistency estimate was calculated for this specific sample. Each form samples brief dialogues and short monologues drawn from everyday contexts; item formats include multiple choice, matching, and short gap-fill, with one point awarded per correct item (maximum 25).

Administration procedures were kept uniform across groups. Tests were given in the regular classrooms under exam conditions using the same audio files, player, and speakers. Before each administration, the researcher ran a volume check and read the standard instructions; clarifications were provided once, in simple English, to avoid coaching. The listening paper took about 30 minutes, including time to transfer answers to the response sheet. Students wrote their school ID codes on the papers; no names were used on the mark sheets. Scoring followed the Cambridge answer keys, with one point awarded per correct item (maximum 25); because scoring is answer-key based, no analytic rubric was used (see Appendix A).

3.5 Analysis plan

Forty-four students completed both the pre-test and the post-test, so no scores were missing.

First, the distribution of each set of scores was inspected. Means, standard deviations, and ranges were calculated, and box plots together with Q–Q plots were examined. No extreme outliers were detected because no standardized residual exceeded ±3.0.

Second, baseline similarity was checked. An independent-samples t-test (Welch's version) was applied to the pre-test scores, and Cohen's d with a 95 % confidence interval was reported. Levene's test was used to confirm that the two groups showed equal variances.

Third, the main effect of instruction was tested with a one-way ANCOVA which compares post-test listening scores between the two groups while statistically adjusting for pre-test score as a covariate. Instructional group (CI versus Traditional) served as the fixed factor, the post-test listening score served as the dependent variable, and the pre-test score was entered as a covariate. Linearity between the covariate and the outcome, the homogeneity of regression slopes, the normality of residuals, and the equality of error variances were all verified. Adjusted means, standard errors, 95 % confidence intervals, the F value, the p value, and partial η^2 were reported.

Fourth, a robustness check was performed. Gain scores (Post−Pre) were compared with an independent-samples t-test, and Cohen's d with a 95 % confidence interval was supplied. The agreement between this test and the ANCOVA supported the stability of the findings.

All statistical tests were two-tailed with $\alpha = .05$. Because treatments were assigned to intact classes but the analyses used individual scores, standard errors may be slightly underestimated; the results are therefore described as preliminary (see Appendix A).

3.6 Validity and Reliability

The study used a standardized listening measure (Cambridge A2 Key Listening), and scoring was objective and answer-key based (one point per correct item; maximum 25). Administration was kept as uniform as possible across groups using the same classrooms, audio files, player, and speakers, with a volume check and standard instructions before each test session, and clarifications provided once to avoid coaching. Two different test forms were used at pre-test and post-test to reduce practice effects; the forms were treated as equivalent based on the Cambridge A2 Key handbook description of parallel forms and scoring (Cambridge Assessment English, 2020). No separate reliability or internal-consistency estimate was calculated for this specific sample, and findings are therefore reported as preliminary.

4. Results and Discussion

This section reports the statistical results from this two-class cluster-randomized pilot and explains what they mean in relation to the study's research question. Unless otherwise noted, all tests are two-tailed with $\alpha = .05$; differences in either direction were considered and results were treated as statistically significant when $p < .05$. For every comparison, group means are shown with standard deviations (M, SD) to indicate score spread, and effect sizes are reported with 95% confidence intervals so that practical importance can be judged alongside significance. Interpreting post-test differences against a baseline and adjusting for that baseline with ANCOVA follows common practice in L2 listening research (e.g., Graham and Macaro, 2008; Vandergrift and Goh, 2012).

4.1 Pre-test Comparison and Interpretation

Table 1a summarizes the baseline listening scores for each class before any instruction; Table 1b reports the between-group test at pre-test and the variance check.

Table 1a. Pre-test descriptive statistics by group

Group	N	Mean	SD	SE
Input	22	13.68	3.358	0.716
Traditional	22	13.95	3.658	0.780

Table 1b. Welch independent-samples *t*-test for the pre-test

t	df	*p*	Cohen's *d*	SE(*d*)	95% CI for *d*	Levene's *F*(1,42)	*p* (Levene)
−0.258	41.70	.798	−0.078	0.302	[−0.668, 0.514]	0.006	.937

The two classes began at broadly similar levels (Input M = 13.68, SD = 3.358; Traditional M = 13.95, SD = 3.658). The Welch *t*-test showed no reliable baseline difference, $t(41.70) = -0.258$, $p = .798$, Cohen's $d = -0.078$ (95% CI [−0.668, 0.514]); variances were equal by Levene's test ($p = .937$). A non-significant result does not prove full equivalence, but it suggests that large initial gaps are unlikely. Accordingly, the pre-test score was included as a covariate in the ANCOVA reported next, which is standard practice for strengthening causal interpretation in quasi-experimental designs (Graham and Macaro, 2008; Vandergrift and Goh, 2012). Given the absence of baseline differences, the instructional effect was then estimated using ANCOVA, adjusting for pre-test listening.

4.2 ANCOVA on Post-Test

To estimate the instructional effect while adjusting for baseline listening, a one-way ANCOVA was conducted with Group (Comprehensible-Input vs. Traditional) as the fixed factor, post-test listening score as the dependent variable, and pre-test score as the covariate. Assumptions were checked and were acceptable: the covariate–outcome relation was approximately linear, the homogeneity of regression slopes held, residuals were approximately normal on the Q–Q plot, and error variances were homogeneous (Levene's test of equality of variances, $F(1, 42) = 2.09$, $p = .156$).

Table 2 reports the ANCOVA statistics for the group effect and the covariate after baseline adjustment.

Table 2
ANCOVA Summary for Post-Test Scores

Source	SS	df	MS	F	p	Partial η²	95% CI for η²
Group	27.49	1	27.49	13.83	<0.001	0.252	[0.057, 0.452]
Pre-test	666.57	1	666.57	335.26	<0.001	0.891	[0.824, 0.926]
Residual	81.52	41	1.99	—	—	—	—

Note. Type III sums of squares; em dashes indicate statistics not defined for the residual term.

The group effect was significant, $F(1, 41) = 13.83$, $p < .001$, partial $\eta^2 = .252$ (95% CI [0.057, 0.452]). The pre-test covariate was strongly related to post-test performance, $F(1, 41) = 335.26$, $p < .001$, partial $\eta^2 = .891$ (95% CI [0.824, 0.926]).

Table 3 presents the adjusted post-test means (±SE) with 95% confidence intervals.

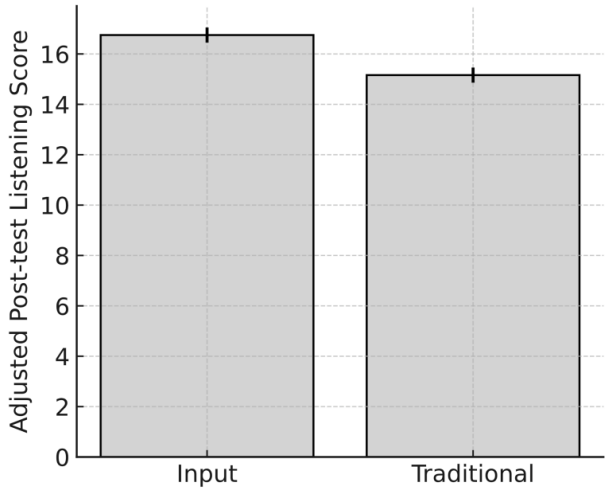
Table 3
Adjusted Marginal Means by Group

Group	Adjusted M	SE	95% CI
Input	16.75	0.30	[16.14, 17.35]
Traditional	15.16	0.30	[14.56, 15.77]

Contrast (Traditional − Input): –1.58, SE = 0.43, t (41) = –3.72, p < 0.001

Figure 1 visualizes the adjusted means with ±1 SE error bars to aid comparison.

Figure 1. Adjusted post-test listening means by instructional group (error bars = ±1 SE).



After controlling for baseline listening scores, participants in the CI (Input) group outperformed their counterparts in the Traditional class by approximately 1.58 points on the 25-point assessment (around a 6% difference), a margin representing a moderate-to-large educational effect. This result aligns with a well-established body of research demonstrating that input-rich, level-matched listening activities yield greater gains than traditional accuracy-focused drills, especially in blended learning contexts (Graham and Macaro, 2008; Vandergrift and Goh, 2012; Zolfaghari et al., 2023). Taken together, this pattern mirrors blended-listening advantages reported in secondary or tertiary EFL settings (e.g., Graham and Macaro, 2008; Zolfaghari et al., 2023). To assess robustness, we then examined raw gain scores to determine whether the pattern held without covariate adjustment.

4.3 Supplementary Gain-Score Analysis

As a robustness check, raw gain scores (Post − Pre) were compared. The CI group (M = 2.91, SD = 1.69) improved more than the Traditional group (M = 1.36, SD = 1.22). The difference was significant, t (42) = 3.48, p = .001, Cohen's d = 1.05, 95% CI [0.41, 1.68]. This converges with the ANCOVA and indicates that the CI programme produced larger improvements even without statistical adjustment.

Table 4a.
Gain-score descriptive statistics by group (Post − Pre)

Group	N	Mean	SD	SE
CI	22	2.909	1.688	0.360
Traditional	22	1.364	1.217	0.259

Table 4b.
Independent-samples t-test on gain scores

t	df	p	Cohen's d	SE(d)	95% CI for d
3.484	42.00	.001	1.050	0.341	[0.413, 1.677]

5. Limitations and Future Research

Treatment was assigned at the class level (two intact classes), but analyses treat students as independent. This can underestimate standard errors and inflate statistical significance. With only two clusters, robust cluster-level models are not feasible; findings should be viewed as preliminary and associational. The CI group completed online retellings, whereas the control group mainly received announcements; this is a primary confound. Additional interaction/practice may partly explain the CI advantage, so future work will equalize output and interaction opportunities across conditions (e.g., identical retellings or none) to isolate the input effect. Participants were female adolescents in a single school; results may not generalise across genders, regions, or proficiency bands. Only listening was assessed; transfer to vocabulary, grammar, or speaking was not measured.

Regarding assumptions and fidelity, ANCOVA checks were limited to linearity between the covariate and outcome, homogeneity of regression slopes, normality of residuals, and equality of error variances. In addition, no independent observation rubric was used to verify how consistently the two programmes were implemented across the eight-week intervention. These limits mean that small assumption violations or uneven delivery across classes could have influenced the estimates. Future work should pre-register an analysis plan, employ independent fidelity observations, use more classes with cluster randomisation, and include delayed post-tests.

6. Conclusion

This study provides preliminary evidence that a comprehensible input–based listening programme,

delivered through a smartphone-mediated blended environment, can enhance adolescent EFL learners' comprehension more effectively than conventional accuracy-focused drills. Across two intact classes, the CI group outperformed the traditional group on both adjusted post-test means and gain scores, with effects in the medium-to-large range. These findings lend support to Krashen's Input Hypothesis and indicate that abundant, level-matched, meaning-centred input, supplemented by visual scaffolds and learner control, can yield measurable gains in listening even in resource-constrained contexts.

At the same time, the results must be interpreted cautiously. The class-level assignment, possible differences in online engagement, and the restricted sample limit the strength and scope of inference. Because the assignment was at the class level and only two clusters were studied, unmeasured class-level factors (e.g., peer dynamics, motivation) cannot be fully ruled out. Therefore, the results should be read as evidence of an association rather than definitive proof of causation: the CI class scored higher, but the design does not by itself establish that CI caused the entire difference. Nonetheless, the pattern observed suggests that story-based, multimodal listening within blended frameworks offers a practical and promising pathway for developing receptive skills in similar settings. Replication with larger samples, additional clusters, and broader outcome measures will be required to establish the robustness and generalisability of the effect.

7. References

- Al Mansour, M.I. (2019) *Investigating the Effectiveness of Blended Learning Approach and Perceptions of EFL Learners' Listening Performance at the Saudi Electronic University*. Unpublished PhD dissertation. Al-Imam Muhammad Ibn Saud Islamic University.
- Al-Rashidy, A.M. and Alsabbagh, A.A. (2023) 'Enhancing listening comprehension skills of EFL learners through using web-based learning', *International Journal of Advanced Humanities Research*, 3(1), pp. 53–69. https://doi.org/10.21608/ijahr.2023.245579.1036
- Cambridge Assessment English (2020) *A2 Key and A2 Key for Schools: Handbook for Teachers (for exams from 2020)*. Cambridge: Cambridge University Press & Assessment. https://www.cambridgeenglish.org/images/504505-a2-key-handbook-2020.pdf
- Ellis, R. (2015) *Understanding Second Language Acquisition*. 2nd edn. Oxford: Oxford University Press.
- Field, J. (2009) *Listening in the Language Classroom*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511575945
- Garrison, D.R. and Kanuka, H. (2004) 'Blended learning: Uncovering its transformative potential in higher education', *The Internet and Higher Education*, 7(2), pp. 95–105. https://doi.org/10.1016/j.iheduc.2004.02.001
- Graham, C.R. (2019) 'Current research in blended learning', in Spector, J.M., Lockee, B.B. and Childress, M.D. (eds.) *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice and Policy*. Cham: Springer. https://doi.org/10.1007/978-3-319-17727-4_34-1
- Graham, S. and Macaro, E. (2008) 'Strategy instruction in listening for lower-intermediate learners of French', *Language Learning*, 58(4), pp. 747–783. https://doi.org/10.1111/j.1467-9922.2008.00478.x
- Halverson, L.R., Graham, C.R., Spring, K.J., Drysdale, J.S. and Henrie, C.R. (2017) 'Blended learning research in higher education and K–12 settings', in Spector, J.M., Lockee, B.B. and Childress, M.D. (eds.) *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice and Policy*. Cham: Springer. https://doi.org/10.1007/978-3-319-17727-4_31-1
- Krashen, S.D. (1982) *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press.
- Krashen, S.D. (1985) *The Input Hypothesis: Issues and Implications*. London: Longman.
- Long, M.H. (2015) *Second Language Acquisition and Task-Based Language Teaching*. Malden, MA: Wiley-Blackwell.
- Mayer, R.E. (2009) *Multimedia Learning*. 2nd edn. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511811678
- Nation, I.S.P. and Newton, J. (2008) *Teaching ESL/EFL Listening and Speaking*. New York: Routledge. https://doi.org/10.4324/9780203891704
- Rost, M. (2013) *Teaching and Researching Listening*. 2nd edn. London: Routledge. https://doi.org/10.4324/9781315833705
- Sholihah, F.A., Permadi, M.B. and Umamah, A. (2018) 'Blended learning and extensive listening: A promising collaboration to enhance listening skill', *ELT Echo: Journal of English Language Teaching in Foreign Language Context*, 3(2), pp. 88–96. DOI: 10.24235/eltecho.v3i2.3323

- Swain, M. (1985) 'Communicative competence: Some roles of comprehensible input and comprehensible output in its development', in Gass, S.M. and Madden, C.G. (eds.) *Input in Second Language Acquisition*. Rowley, MA: Newbury House, pp. 165–179.
- Vandergrift, L. and Goh, C.C.M. (2012) *Teaching and Learning Second Language Listening: Metacognition in Action*. New York: Routledge. https://doi.org/10.4324/9780203843376.
- Vandergrift, L. and Tafaghodtari, M.H. (2010) 'Teaching L2 learners how to listen does make a difference: An empirical study', *Language Learning*, 60(2), pp. 470–497. https://doi.org/10.1111/j.1467-9922.2009.00559.x.
- Zolfaghari, A., Sepehri, M., Hosseini, E. (2023). 'Effects of Virtual Vs. Blended-Language Instruction on the Development of Listening and Reading Skills by Iranian Intermediate EFL Learners', *Iranian Journal of Applied Language Studies*, 15(2), pp. 201-216. https://doi.org/10.22111/ijals.2023.41980.2270

Appendix A. Cambridge A2 Key Listening administration and scoring procedure used in this study

- Test: Cambridge A2 Key (KET) Listening, 25 items; one point per correct item (maximum 25)
- Forms: two different forms used at pre-test and post-test to reduce practice effects
- Setting: regular classrooms under exam conditions
- Audio: same audio files, player, and speakers; volume check before each administration
- Instructions: standard instructions read; clarifications provided once in simple English to avoid coaching
- Timing: about 30 minutes including transfer of answers to the response sheet
- Identification: school ID codes used; no names on mark sheets
- Scoring: Cambridge answer keys used (no analytic rubric)

Appendix B. Results (JASP Statistical Output and ANCOVA Assumption Checks)

Independent Samples T-Test

Independent Samples T-Test

						95% CI for Cohen's d	
	t	df	p	Cohen's d	SE Cohen's d	Lower	Upper
Pre_test	-0.258	41.70	.798	-0.078	0.302	-0.668	0.514

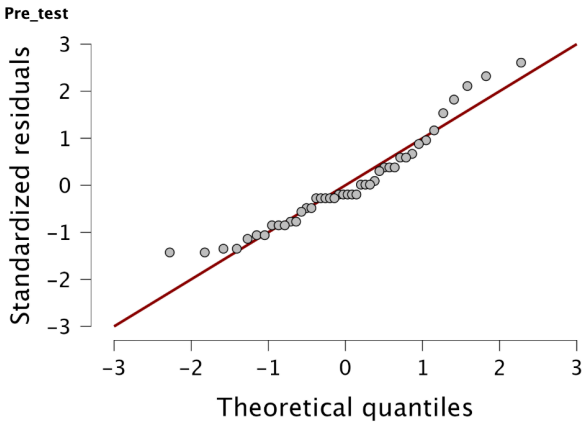
Note. Welch's t-test.

Assumption Checks

Test of Equality of Variances (Levene's)

	F	df_1	df_2	p
Pre_test	0.006	1	42	.937

Q-Q Plots



Pre_test

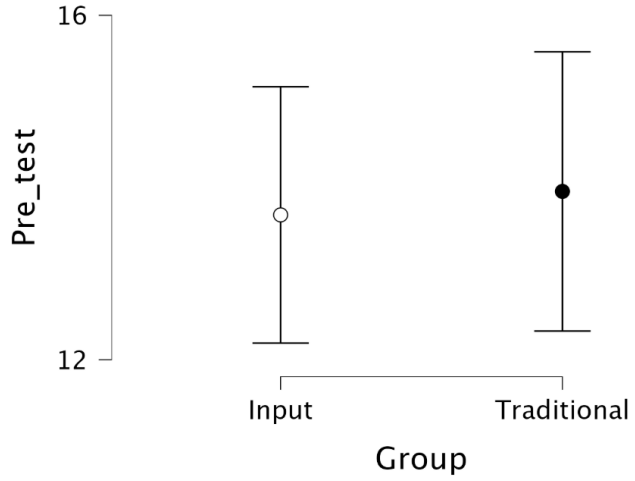
Descriptives

Group Descriptives

	Group	N	Mean	SD	SE	Coefficient of variation
Pre_test	Input	22	13.68	3.358	0.716	0.245
	Traditional	22	13.95	3.658	0.780	0.262

Descriptives Plots

Pre_test



ANCOVA

ANCOVA - Post_test

							95% CI for η^2_p	
Cases	Sum of Squares	df	Mean Square	F	p	η^2_p	Lower	Upper
Group	27.49	1	27.493	13.83	< .001	0.252	0.057	0.452
Pre_test	666.57	1	666.573	335.26	< .001	0.891	0.824	0.926
Residuals	81.52	41	1.988					

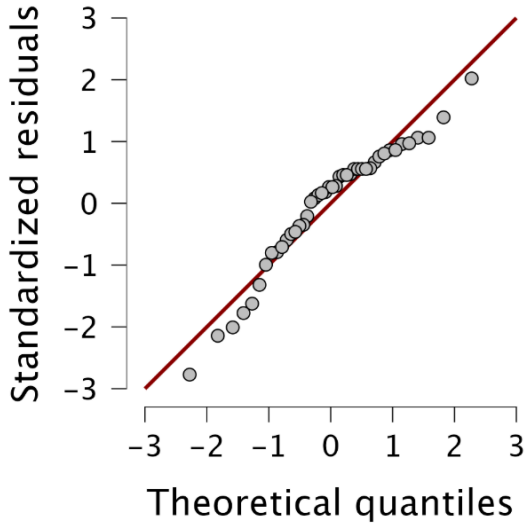
Note. Type III Sum of Squares

Assumption Checks

Test for Equality of Variances (Levene's)

F	df1	df2	p
2.087	1.000	42.00	.156

Q-Q Plot



Contrast Tables

Simple Contrast - Group

		95% CI for Mean Difference					
Comparison	**Estimate**	**Lower**	**Upper**	**SE**	**df**	**t**	**p**
Traditional - Input	-1.582	-2.441	-0.723	0.425	41	-3.719	< .001

Marginal Means

Marginal Means – Group

		95% CI for Mean Difference		
Group	**Marginal Mean**	**Lower**	**Upper**	**SE**
Input	16.75	16.14	17.35	0.301
Traditional	15.16	14.56	15.77	0.301

Independent Samples T-Test

Independent Samples T-Test

							95% CI for Cohen's d	
	Test	**Statistic**	**df**	**p**	**Cohen's d**	**SE Cohen's d**	**Lower**	**Upper**
Difference	Student	3.484	42.00	.001	1.050	0.341	0.413	1.677
	Welch	3.484	38.19	.001	1.050	0.341	0.409	1.680

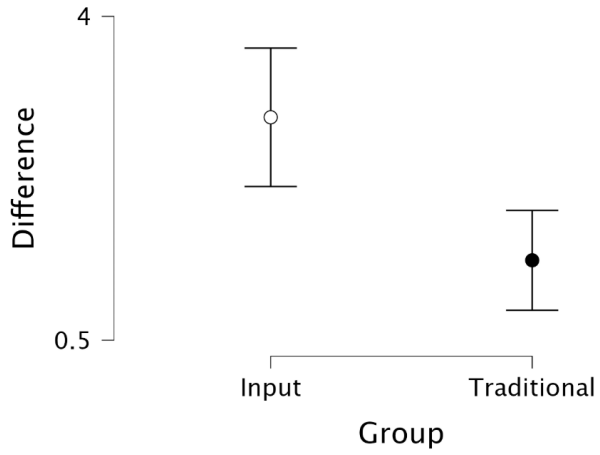
Descriptives

Group Descriptives

	Group	N	Mean	SD	SE	Coefficient of variation
Difference	Input	22	2.909	1.688	0.360	0.580
	Traditional	22	1.364	1.217	0.259	0.892

Descriptives Plots

Difference



جێبەجێکردنی گریمانەی وەرگرتنی کراشن بۆ فێرکردنی تێگەشتن لە گوێگرتن لە پۆلێکی فێربوونی تێکەلاودا

میران عبدالرحمن عبدالله
بەشی زمانی ئینگلیزی، کۆلێژی پەروەردەی بنەڕەت، زانکۆی هەڵەبجە، هەڵەبجە ٤٦٠١٨، هەرێمی کوردستان، عێراق
miran00496176@uoh.edu.iq
عمر فؤاد غفور
بەشی زمانی ئینگلیزی، کۆلێژی پەروەردەی بنەڕەت، زانکۆی هەڵەبجە، هەڵەبجە ٤٦٠١٨، هەرێمی کوردستان، عێراق
omar.ghafor@uoh.edu.iq

پوختە
گوێگرتن بنەمایەکی سەرەکیە لە پۆلەکانی فێرکردنی زمانی ئینگلیزیدا، بەڵام هێشتا لێکۆڵینەوەی کەمی لەسەر کراوە، بەتایبەتی بۆ هەرزەکاران، لەو پۆلانەی کە سەرچاوەی کەمیان تێدایە. ئامانجی ئەم توێژینەوەیە بریتییە لە لێکۆڵینەوە لەوەی کە لەپۆلێکدا خوێندن بەشێوەیەکی تێکەڵاوە (Blended Learning) ئایا شێوازی فێرکردنی وەرگرتنی شایستە بەتێگەیشتن (Comprehensible Input) کاریگەرترە یان شێوازە باوەکان کە جەخت دەکەنە سەر شیکارکردنی ڕاهێنانەکان بە وردی و بەدروستی. دوو هۆبە لە پۆلی هەشتەم، کە کۆی گشتی ٤٤ خوێندکار بوون، لە قوتابخانەیەکی هەڵەبجە لە هەرێمی کوردستانی عێراق، بەشداربوون لەم توێژینەوەیەدا بۆ ماوەی هەشت هەفتە. توانای گوێگرتنیان بە دوو وەشانی هاوشێوە لە تاقیکردنەوەی (KET)ی کامبریج بۆ ئاستی (A2) تاقیکرایەوە. هەروەها دڵنیابووین لەوەی کە هەردوو گروپەکە پێش دەستپێکردن لە ئاستێکی هاوشێوەدا بوون. لەگۆتایدا، گرووپی تاقیکاری (CI) پێشکەوتنێکی بەرچاویان لە توانای گوێگرتندا پیشاندا بەراورد بە گروپەکەی دیکە، هەروەها پشکنینی نمرەی زیادبوو (gain-score) هەمان ئەنجامی پشتڕاستکردەوە. بە گشتی، دەرئەنجامەکان پێشنیاری ئەوە دەکەن کە ڕاهێنانی زۆری گوێگرتن لە ئاستی گونجاودا بە یارمەتی ئامرازی سادەی وەک مۆبایلی زیرەک، دەتوانێت توانای گوێگرتنی قوتابیانی زمانی ئینگلیزی باشتر بکات. بەڵام، توێژینەوەکە دوو پۆلی ئامادەکراوی بەکارهێناوە، و بڕی ڕاهێنانی ئۆنلاین بۆ هەردوو گروپەکە یەکسان نەبووە؛ لەبەر ئەم هۆکارە، پێویستە بە وریاییەوە سەیری ئەنجامەکان بکرێت. بۆ بەهێزترکردنی بەڵگەکان و یارمەتیدانی مامۆستایان بۆ بەکارهێنانی ئەم شێوازە، توێژینەوەی داهاتوو پێویستە پۆلی زیاتر بەکاربهێنێت و بە شێوەی هەڕەمەکی پۆلەکان جیابکاتەوە، وە چاودێری زیاتر لەخۆبگرێت بۆ دڵنیابوونەوە لەوەی کە وانەوتنەوەکە بە دروستی دەوترێنەوە، و تاقیکردنەوەی ماوەدرێژ زیاد بکرێت بۆ ئەوەی دڵنیا ببنەوە کە ئاخۆ پێشکەوتنەکان بەردەوام دەبن.
وشە سەرەکیەکان: گریمانەی وەرگرتنیی کراشن، وەرگرتنی شایستە بەتێگەشتن، فێربوونی تێکەلاوی ئۆنلاین و ناو پۆل، تێگەشتن لە گوێگرتن، هەرزەکارە کوردەکان کە فێری ئینگلیزی دەبن.

استخدام فرضية المدخلات لكراشن لتدريس مهارات الاستماع في فصل دراسي للتعلم المدمج

ميران عبدالرحمن عبدالله
قسم اللغة الإنجليزية، كلية التربية الأساسية،جامعة حلبجة، حلبجة ٤٦٠١٨، إقليم كردستان، العراق
miran00496176@uoh.edu.iq
عمر فؤاد غفور
قسم اللغة الإنجليزية، كلية التربية الأساسية، جامعة حلبجة، حلبجة ٤٦٠١٨، إقليم كردستان، العراق
omar.ghafor@uoh.edu.iq

الملخص
الاستماع مهارة أساسية في فصول تعلم اللغة الإنجليزية، لكنها لا تزال غير مدروسة بما فيه الكفاية، خاصة بالنسبة للمراهقين في المناطق التي تفتقر إلى المصادر و الآلات . تهدف هذه الدراسة إلى معرفة ما إذا كانت طريقة التدريس التي تستخدم "المدخل المفهوم" (Input Comprehensible) في بيئة تعلم مدمجة (Blended Learning) تعمل بشكل أفضل مقارنة بالتدريب التقليدي الذي يركز على الدقة. شاركت فصلان من الصف الثامن (إجمالي 44 طالبًا) من مدرسة في حلبجة، إقليم كردستان، العراق، في هذه الدراسة لمدة ثمانية أسابيع. تم اختبار قدرتهم على الاستماع باستخدام نسختين متكافئتين من اختبار كامبريدج (KET) للمستوى (A2). كما تأكدنا من أن المجموعتين كانتا في نفس المستوى قبل البدء. وفي الختام، أظهرت مجموعة المدخلات المفهومة (CI) مكاسب ملحوظة في مهارات الاستماع في الاختبار البعدي مقارنة بالمجموعة التقليدية، كما أشار تحليل درجات التحسن (gain-score) إلى نفس النتيجة. بشكل عام، تشير النتائج إلى أن تدريبًا مكثفًا على الاستماع بمستوى مناسب، بمساعدة أدوات بسيطة مثل الهواتف الذكية، يمكن أن يحسن استماع اللغة الإنجليزية لطلاب المرحلة الإعدادية. مع ذلك، استخدمت الدراسة مجموعتين جاهزتين (فصلين طبيعيين)، وكان أحد القيود المهمة أن مقدار

التدريب عبر الإنترنت لم يكن متساويًا لكلتا المجموعتين؛ لهذا السبب، يجب التعامل مع النتائج بحذر. لتعزيز الأدلة ومساعدة المعلمين على استخدام هذه الطريقة، يجب أن تستخدم الأبحاث المستقبلية المزيد من الفصول، وتعيينها بشكل عشوائي، وتضمين اختبارات للتأكد من أن التدريس تم بشكل صحيح، وإضافة اختبارات متابعة لمعرفة ما إذا كانت التحسينات مستمرة.

الكلمات المفتاحية: فرضية المدخلات لكراشن، المدخلات المفهومة، التعلم المدمج (عبر الإنترنت والحضوري)، الاستيعاب السمعي، المراهقون الكرد دارسو اللغة الإنجليزية كلغة أجنبية