



ISSN: 1817-6798 (Print)

Journal of Tikrit University for Humanities

available online at: [www.jtuh.org/](http://www.jtuh.org/)
**JTUH**  
 مجلة جامعة تكريت للعلوم الإنسانية  
 Journal of Tikrit University for Humanities
**Afnan Hamoud Hussein Hassan**

Kirkuk Education Directorate

**Prof. Istabraq Tariq Jawad Al Azzawi (Ph.D.)**Tikrit University  
College of Education for Humanities  
Department of English Language\* Corresponding author: E-mail :  
[afnanhmood@st.tu.edu.iq](mailto:afnanhmood@st.tu.edu.iq)**Keywords:**EFL students  
conversation performance  
Tikrit University  
Kirkuk University  
gender difference**ARTICLE INFO****Article history:**

Received	4 Jan 2023
Received in revised form	17 Aug 2023
Accepted	17 Aug 2023
Final Proofreading	28 Nov 2023
Available online	30 Nov 2023

E-mail [t-jtuh@tu.edu.iq](mailto:t-jtuh@tu.edu.iq)©THIS IS AN OPEN ACCESS ARTICLE UNDER  
THE CC BY LICENSE<http://creativecommons.org/licenses/by/4.0/>

## Investigating EFL Students' Performance in Conversation at Tikrit and Kirkuk Universities

### A B S T R A C T

This study aimed to investigate the performance of EFL (English as a Foreign Language) students in conversation at Tikrit and Kirkuk Universities in Iraq. The research questions addressed whether there were statistically significant differences in the students' performance in conversation between the two universities and according to gender. A sample of 259 third-year college students was randomly selected from the population of 403 undergraduate students from the English departments at both universities. The students' performance in conversation was assessed using a proficiency test designed based on the Celce-Murcia model of conversational competence. The test consisted of five sections, assessing both recognition and production levels.

© 2023 JTUH, College of Education for Human Sciences, Tikrit University

DOI: <http://doi.org/10.25130/jtuh.30.11.2.2023.22>

استقصاء أداء طلبة اللغة الانكليزية لغة اجنبية في المحادثة في جامعتي تكريت وكركوك

الباحثة: افنان حمود حسين حسن / مديرية تربية كركوك

أ.د استبرق طارق جواد العزاوي/ جامعة تكريت / كلية التربية للعلوم الانسانية

**الخلاصة:**

هدفت هذه الدراسة إلى التعرف على أداء طلبة اللغة الإنجليزية كلغة أجنبية في المحادثة في جامعتي

تكريت وكركوك في العراق. تناولت أسئلة البحث ما إذا كانت هناك فروق ذات دلالة إحصائية في أداء الطلاب في المحادثة بين الجامعتين وبحسب الجنس. تم اختيار عينة من 259 طالبا جامعا في السنة الثالثة بشكل عشوائي من مجموع 403 طلبة جامعيين من أقسام اللغة الإنجليزية في كلا الجامعتين تم تقييم أداء الطلاب في المحادثة باستخدام اختبار الكفاءة المصمم على أساس نموذج سلسي مورشيا لكفاءة المحادثة.

الكلمات المفتاحية: طلبة اللغة الانكليزية، لغة اجنبية، اداء المحادثة، جامعة تكريت، جامعة كركوك، الفرق بين الجنسين.

## Section One Introduction

### 1.1 Statement of the problem:

Effective communication skills, particularly in conversational contexts, are essential for English as a Foreign Language (EFL) students. However, there is a need to investigate the performance of EFL students in conversation at Tikrit and Kirkuk Universities in Iraq. Although previous studies have explored various aspects of language learning and proficiency, there is a gap in the literature regarding the specific performance of EFL students in conversation at these universities.

The problem addressed in this research is the lack of comprehensive understanding of EFL students' performance in conversation and the factors that may influence their performance. It is crucial to assess and compare the conversational competence of EFL students at Tikrit and Kirkuk Universities, as well as to examine potential differences in performance based on gender.

By investigating EFL students' performance in conversation, this study aims to bridge the existing gap in the literature and provide valuable insights into the strengths and weaknesses of EFL education in these institutions. The findings will contribute to the development of targeted instructional approaches and interventions to enhance students' conversational skills.

Addressing this problem will not only contribute to the field of EFL education but also have practical implications for language educators, curriculum designers, and policymakers. Understanding the specific challenges faced by EFL students in conversation can lead to more effective teaching strategies and the implementation of interventions that address the identified gaps.

Therefore, conducting a comprehensive investigation into EFL students' performance in conversation at Tikrit and Kirkuk Universities will provide valuable insights and contribute to improving the quality of English language education at these institutions.

## 1.2 Aim of the Study

**The current study aims to:**

- Identify and compare:
  - 1- EFL university students' performance in conversation.
  - 2- The difference among EFL university students' performance in conversation according to gender.

## 1.3 Questions of the Study

**The current study tries to answer the following questions:**

- 1- Are there any statistically significant differences among EFL university students' performance in conversation at the two universities?
- 2- Are there any statistically significant differences among EFL university students' performance in conversation according to gender at the two universities?

## 1.4 Limits of the Study

The present study is limited to:

- 1- Third-year EFL university students at Tikrit and Kirkuk universities during the first course of the academic year 2022-2023.
- 2- The material selected is taken from the listening and speaking textbook (Real Listening & Speaking, by M. Craven), the units of real listening and speaking that are used in the test are units one, two, three, four, six and seven. The titles of units are, unit one "How are things?", unit two "Can I take your coat?", unit three "I'm looking for a flat?", unit four "I'd like a refund please", and unit seven "I could do with a break".

## 1.5 Definitions of Basic Terms

### 1.5.1 Investigating

The process of trying to find out all the details or facts about something in order to discover who or what caused it or how it happened (Macmillan Dictionary, 2012: 99).

**The operational definition:** investigating is the observation by close examination and systematic inquiry.

### 1.5.2 EFL Students'

A phrase usually used for non-native English speakers learning English in a country where English is neither commonly spoken nor a medium of instruction

(Rezaee,& Shabani,2019).

### 1.5.3 Performance

Performance refers to the degree of the achievement of objectives or the potentially possible accomplishment regarding the important characteristics of an organization for the relevant stakeholders. Performance is therefore principally specified through a multidimensional set of criteria (Ghalem, Okar, Chroqui, & Semma, :2016).

Performance is about deploying and managing well the components of the causal model that leads to the timely attainment of stated objectives within constraints specific to the firm and to the situation(ibid).

As the EFQM (2003) defined Performance is the level of attainment achieved by an individual, team, organization or process.

**The operational definition:** the actual use of language in concrete situations.

### 1.5.4 Conversation

According to Levinson's influential textbook (1983: 284), the conversation is "the predominant kind of talk in which two or more participants freely alternate in speaking, which generally occurs outside specific institutional settings". On the basis of more recent research, one can view types of conversation as forming a continuum with a mundane talk at one end and carefully pre-planned interviewing or some other strictly role and status-dependent forms of institutional interaction at the other end. What all forms of conversation share, however, is the fact that it is through them that we, as human beings, manage our daily affairs and construct and make sense of our life and activities.

The conversation is a joint activity in which two or more participants use linguistic forms and nonverbal signals to communicate interactively (Brennan, 2010:1).

The term conversation can refer to a variety of communicative situations, including room chats, workplace discussions, ordering food at a restaurant, or even text messages sent via smartphone. What most conversations have in common is the fact that they involve interactive, communicative exchanges between two or more people(Horton,2018:3).

**The operational definition:** an informal, usually private, talk in which two or more people exchange thoughts, feelings, or ideas, or in which news or information is given or discussed.

## Section Two Literature Review

### 2.1 performance in Conversation

Performance in conversation refers to an individual's ability to engage in effective and meaningful spoken interactions with others. It involves the use of appropriate language, communication skills, and social cues to convey ideas, exchange information, and maintain a conversation flow. In the context of the current study, performance in conversation specifically relates to the English as a Foreign Language (EFL) university students' proficiency in conversation.

The assessment of performance in conversation can be done through various methods, such as structured interviews, role-plays, group discussions, and oral presentations. In the current study, the performance in conversation was evaluated using a specific test or evaluation tool, which measured the students' abilities in terms of fluency, accuracy, pronunciation, vocabulary usage, and interactional skills.

### 2.1.1 Conversational Competence

Conversational competence refers to the ability to engage in effective and meaningful conversations in a given language. It involves various skills and strategies, including linguistic, sociolinguistic, and pragmatic knowledge, that enable individuals to participate in conversations, understand and produce appropriate language, and interact successfully with others.

According to Canale and Swain (1980), who developed the influential Communicative Competence Model, conversational competence falls under the broader umbrella of communicative competence. It encompasses the ability to use language in a socially appropriate and contextually meaningful manner. Conversational competence goes beyond mere grammatical knowledge and includes understanding the rules of conversation, such as turn-taking, topic maintenance, repair strategies, and appropriate use of nonverbal cues.

Research has highlighted the importance of developing conversational competence in second language learning. Conversations provide learners with opportunities to practice and apply their language skills in real-life communication contexts. By engaging in conversations, learners can develop their vocabulary, grammar, pronunciation, and pragmatic skills, as well as enhance their listening and speaking abilities (Richards & Schmidt, 2010).

Moreover, conversational competence plays a crucial role in promoting learner autonomy and confidence. When learners can effectively communicate in conversations, they feel more motivated and empowered to engage with the target language and culture. Conversations also allow learners to negotiate meaning, clarify misunderstandings, and develop strategies for effective communication (Johnson, 2012).

Conversational competence inherent to the turn-taking system in conversation described by Sachs et al. (1974) but may be extendable to other dialogic genres:

- ☐ how to open and close conversations,
- ☐ how to establish and change topics,
- ☐ how to get, hold, and relinquish the floor,
- ☐ how to interrupt, and
- ☐ how to collaborate and backchannel (Clece-Murcia, 1995:).

Conversation is not only one of the most apparent and important modes of discourse, but also a very good way to gain/probe the social and interactive nature of communication. It is considered a cooperative venture, i.e. what are the rules of conversation? How do you get attentions of others? How do we begin topics? Holding the floor? Yield the floor? Terminate topics? Avoid Topics? and How does a person interrupt, correct, or seek clarification. So it is important to come with an area that invites the reconceptualization of language (Brown,2014:219).

### 2.1.2 Fundamental Components of Conversational Competence

The first component is attention-getting, children get the attention of the hearer, for example by crying or anything to turn a parent's attention to themselves. Then, both verbal and nonverbal attention-getting conventions are assimilated. The second component is topic nomination, which initiates conversation when speakers have secured the hearer's attention to begin an exchange as in, sure is hot today, isn't it? The third component is topic development, i.e. holding the floor when participants use strategies for continuing the conversation by using techniques like hesitation signals e.g. um, uh, well, and, I mean, and like. The fourth one is turn-taking which is a culturally oriented set of rules that require finely tuned perception in order to communicate effectively (Brown, 2014:220).

The fifth component, topic clarification is the ability to ask questions for clarification, which may arise from inaudibility "What did you just say", disagreement "I see your point, but have you considered...." Or lack of understanding "What does 'competence' mean? The sixth one is repair which involves seeking or giving repair of linguistic forms that contain errors especially between second language learners and native speakers. The seventh components are shifting, avoiding and interrupting which mean conversational abilities that may be effected through both verbal and nonverbal signals. For example, changing a topic, dancing around certain topics and interrupting politely are especially difficult for an L2 learner to acquire the rules for which vary widely across cultures and languages (Brown, 2014:220).

### 2.1.3 Elements of Conversational Competence

#### 2.1.3.1 Appropriateness of Response



A speaker's to produce an appropriate utterance in the foreign language is assumed to be an essential part of his conversational skill. Appropriateness of a response refers to the suitability of the language used to a situation. In order to communicate or share effectively in a conversation, participants should know what to say, with whom, where and when (Chapman, 1982: 113).

It is important that the user of a foreign language should choose the appropriate usage to his immediate social situation. Appropriateness of an utterance is not determined by its grammaticality and meaningfulness; a sentence may be grammatical and meaningful, yet inappropriate to the situation, e.g.:

- a. Would you sit down, please.
- b. Sit down. (ibid).

Thus, it is unsuitable for a speaker to use (b) when talking to a stranger or a person of higher position; it is only appropriate between friends or to a person of low position. So, appropriateness of one's utterance means having a knowledge of the different styles of communication. The choice of a suitable style depends on certain non-linguistic criteria such as: age, sex, familiarity and the social roles of participants (Chapman, 1982:114).

However, Jacobson (1976), reviewed in Richards and Sukwivat (1983: 118), mentions five styles of daily interaction. They are:

- a. Peers' style,
- b. Formal style,
- c. Style of children,
- d. Informal style between family members, and
- e. Informal style between friends.

Though the use of these styles is not rigid, the following example offered by Eckard and Kearny (1981: 3) may illustrate the various situations of using each style:

*Two college students meet and discuss a party (style A); a college professor approaches them to cancel his conference (style B); one student's baby brother runs toward his older brother and talks with him briefly (style C); their mother joins the group to speak with her older son at first (style D) and with his friend later on (style E).*

### 2.1.3.2 Fluency and Ease of Expression

Fluency in using a foreign language is a highly complex notion. It refers to the ability of a learner to “approximate the speed and ease with which native speakers of the language typically produce their utterances” (Harris, 1969:82).

Fluency is usually measured according to the speech rate of a learner. However, Lesson (1975: 170) thinks that fluency is considered in relation to “rapid articulation, minimal pauses, structural economy and fast retrieval of lexical items”.

It is maintained that the main factor which affects the degree of a learner’s fluency is the hesitation pauses in speech. Consequently, the speed of speech is measured according to the amount of such pauses. Pauses in speech are of two kinds: ‘filled pauses’ such as the vocalizations: ah, er, uh, mm, ayyy (lengthening a), tooo (lengthening to) and ‘silent pauses’ when there is a complete silence. However, considered from a psychological point of view, these hesitations in speech are linked to periods of creative thinking; they give the speaker time to organize his thoughts and decide how to express them (Lesson ,1975: 170).

Considering the matter of fluency in speech, Corder (1973: 257) maintains that having a good ability in pronunciation does not necessarily mean fluency in speaking the foreign language. However, it could be said that fluency in using the foreign language is mostly affected by the degree of language problems.

### 2.1.3.3 Comprehension

Further to the mentioned elements of conversational competence, comprehension represents the final basic element in the sequence. According to Crystal and Davy (1975: 85), understanding one other speech depends on the “recognizability” of the words and sentence patterns of speech”. Comprehension refers to the ability of all conversationalists to understand the speech of one another.

Thinking of the significance of reciprocal understanding in a conversation, Rivers (1968: 193) confirms that comprehension of the elements of a message is necessary for the listener in order not to make “a non-committal acknowledgment of the fact that he has been addressed”. However, comprehension is not an easy task since it involves various psychological processes for the interpretation of a message. Thomas (1983:91) affirms that we sometimes misunderstand other people not because of the inability to hear them or understand their words, but rather because of the difficulty in understanding their intention.



#### 2.1.3.4 Strategies to Develop Conversational Competence

**Turn-taking** that demands students focusing attentively on what is going on so that the conversation goes smoothly. It is quite necessary to train the students how to (a) pre-plan a turn; (b) take a turn; (c) interrupt a turn; (d) hold a turn; (e) pass a turn; (f) repair a turn; (g) and close a turn. **Exchange** Generally includes two steps: initiation and answer. Conversational structural system model is made up of act, move, exchange, transaction and interaction. An exchange is composed of more than two moves (turns), i.e. initiation and response. **Group Work** entails students getting along with others. It can be translated into cooperative group work in the classroom so as to enhance students' growth in language, building on conversational skills, they already possess and easing adaptation to school routines. **Games** creates experiences with language and ideas, and "experience is the glue that makes learning stick" (Garagri and Mohammed; 2018: 266).

**Songs** are language-based. They are popular with all ages and can be successful throughout the grades. **Dramatization** provides students with a variety of contextualized and scaffold activities that gradually involve more participation and more oral language proficiency. **Word-guessing game**, it is important to start from a simple point, i.e. to speak correctly in sentences before acquiring oral communicative competence. **Description** aims to train students to speak in paragraphs, and to be able to make themselves understood (ibid).

**Topic talking** helps students put together a formal presentation by providing them with an outline or a few key words. **Familiarizing students** with the nonlinguistic features of the language by (a) introducing cultural background knowledge, (b) Introducing pragmatic knowledge about the foreign language by enabling students to know when, where, and how to apply this knowledge to different contexts so as to arrive at natural, casual, and appropriate communication with native speakers. **Information-Gap Activities** enable students to normally communicate in order to get information they do not possess (ibid:267).

**Jigsaw Activities** are based on the information-gap principle. Typically, the class is divided into groups and each group has part of the information needed to complete an activity. The class must fit the pieces together to complete the whole. **Task-Completion Activities** represented by puzzles, games, map-reading, and other kinds of classroom tasks which focus on using one's language resources to complete a task. **Opinion-sharing activities** where students compare values, opinions, or beliefs, such as a ranking task in which students list six qualities in order of importance. **Information-Transfer Activities** require students to take information presented in one form, and represent it in a different form. An example is reading information about a subject and then representing it as a graph. **Reasoning-Gap Activities** involve deriving some new information from given

information through the process of inference and practical reasoning (Garagri and Mohammed; 2018: 268).

### Section Three Methodology

#### 3.1 Sample

According to Ary, et al. (2018:148), a sample is a number of individuals, objects, or events selected for a study from a population, usually in such a way that they represent the large group from which they are selected. A sample is a group of people, objects, or items that are taken from a large population for measurement. So, to get accurate results, sampling is done (Bhardwaj, 2019).

For the current study, the sample has been randomly selected in order to be truly representative of population characteristics without any bias and in order to obtain valid and reliable results.

The total number of the sample is 259 third-year college students has been chosen intentionally and randomly. The intentional choice is represented through choosing a college and random choice is achieved by selecting a representative number of students from the morning studies in the Department of English at the Colleges of Education, in Tikrit and Kirkuk Universities in the academic year 2022-2023. The sample of the current study represents 64% of the population as indicated in the following Table (1):

Table (1)

*The Population and Sample of the Study*

University	College	No. of Population	No. of Pilot study	No. of the Sample		Total	Percentage
Tikrit	The College of Education for Humanities	203	15	M.	F.	127	62%
				44	83		
Kirkuk	The College of Education for Humanities	200	15	M.	F.	132	66%
				28	104		
Total		403	30	259			64%

\* **Note.** The pilot sample is not represented in the percentage. The percentage only covers the final sample

### 3.2 Proficiency Test

Bachman (1990) states that a key issue in testing proficiency is how the constructs of language ability are specified. The tasks that test takers are required to perform must be legitimate samples of English language use in a defined context.

Brown's state (2004:44) a proficiency test is not limited to any one course, curriculum or single skill in the language; rather, it tests overall ability. Proficiency test have traditionally consisted of standardized multiple choice items on grammar, vocabulary, reading comprehension, and oral comprehension. Sometimes a sample of writing is added, and more recent tests also include oral production performance.

A proficiency test is devised to measure how much of a language someone has learned. It is not linked to any particular course of instruction, but measures the learner's general level of language mastery, the aim of a proficiency test is to determine whether this language ability corresponds to specific language requirements (Hung, & Huang, 2019).

In order to achieve the aims of this study, a proficiency test is constructed. The test questions are constructed to find out the EFL university students' performance in conversation. The test is designed by the researcher depend on the Celce-Murcia model of the component of conversational competence on both recognition and production levels and it consists of five questions. The first two questions test students' recognition levels, each question contains ten items. The first question involves a multiple-choice type that tests the student's ability to select the most appropriate choice that goes best with the context, while the second question demands to state whether a sentence is true or false. The last three questions test the students' production level. The third question contains five items. The fourth question includes two branches A and B. The first part contains five items. The second part contains three items. The last question contains five items.

### 3.2 Scoring Scheme of the Proficiency Test

The test sample includes five questions. The test is scored out of 100Ms. Forty marks are specified for questions at the recognition level which are Q1 and Q2. Two marks are specified for each correct item and zero for each wrong answer in the first and second questions. The second Sixty marks are specified to the questions at the production level which are Q3, Q4 A, B Q5, which are scored according to the rubric for scoring students' Performance, (See table 2)

Questions three, four, and five have been scored by a scoring committee that consists of Assist. Lect. Ihan Abd Al- Ghafoor, Lect. Hamza Abd, and the researcher. After the students completed their responses, answers to questions three, four, and five were copied and scored by the scoring committee.

### 3.4 Rubric for scoring students' Performance

A rubric is an assessment tool that has a description of the expected performance for each criterion in order to achieve a grade or certain outcomes. A rubric is a systematic method to collect data regarding knowledge and skills as stated by Churches (2015) in his study. Garfolo (2016) agreed that rubrics can be used to measure certain behavior.

The specific or standard expectations of performance to evaluate learning outcomes (Aiken, 1996; Company et al., 2017; Stevens & Levi, 2013) are a key part of the rubric as it does not only serve as a tool of assessment but also serves as a learning tool as quoted by Andrade and Du (2005). "Rubrics can teach as well as evaluate" Therefore, this obvious rubric application can benefit any discipline (Montgomery, 2002).

Rubrics for scoring students' Performance have been scored by the scoring committee also.

**Table (2 )**  
*Rubric for scoring students' Performance*

<b>Conversational competence components</b>	<b>No. of Question</b>		
<b>1- How to open a conversations</b>	<b>Q3</b>	<b>0</b>	<b>Not answering the question</b>
		<b>1-2</b>	<b>Somehow little grammatical mistake</b>
		<b>3-4</b>	<b>Missing many components including the key vocabulary</b>
<b>2- How to close a conversation</b>	<b>Q4/A</b>	<b>0</b>	<b>Not answering the question</b>
		<b>1-2</b>	<b>Didn't recognize characters, and grammatical /vocabulary mistakes</b>
		<b>3</b>	<b>recognize the characters with good language and idea</b>
<b>3- how to establish and changetopics</b>	<b>Q4/B</b>	<b>1-3</b>	<b>Not written in inappropriate languageand has grammatical and vocabulary mistakes</b>
		<b>4-7</b>	<b>write a conversation without usinginterruption with Some grammar/vocabulary errors,</b>

			but meaning is still clear
		7-10	write a correct idea and, Dialogue meets length requirements and contains all required materials
4- how to get, hold, and relinquish the floor	Q5	0	Not answering the question
		1-5	inappropriate conversation, Dialogue missing more than two requirements.
5- how to interrupt		5-10	did not complete the conversation and did not use all the dialogue genres
6- how to collaborate and backchannel, etc.		10-15	write a good idea with good language and Dialogue is very entertaining and students use props or creative situations in dialogue.

### 3.5 Pilot Administration of the Research Instrument

A pilot study is a small feasibility study designed to test various aspects of the methods planned for a larger, more rigorous, or confirmatory investigation (Arain, Campbell, Cooper, & Lancaster, 2010).

In order to know if there is any difficulty in the items of the test, a random sample consisting of (30) male and female students in the third year/ Colleges of Education/ English Department/University of Tikrit and Kirkuk are involved in the pilot study.

The diagnostic test has been administrated on the 20th and 21st of December and the SJT has been administrated on the 26<sup>th</sup> and 27<sup>th</sup> in the first semester of the academic year (2022-2023).

The time required for the student to answer all of the questions ranges between 45 and 60 minutes. The pilot study is conducted at the same time for the two tests.

**Table( 3)**  
*Sample of the Pilot Study*

University	Males	Females	Total
University of Tikrit	7	7	14
University of Kirkuk	8	8	16
Total	15	15	30

### 3.6 Item Analyses of Conversation Scale

#### 3.6.1 Difficulty Level

The difficulty level is specified as the ratio of the students who replied correctly to each item (Rosas, 2000:3).

Item difficulty refers to the extent to which an item appears to be complicated or facilitated for a given number of tests. It just reflects the percentage of learners who respond correctly to the object. The most suitable test item will have item difficulty varying between 0.15 and 0.85 (Brown, 2010: 70-1). The items of this test are considered acceptable if they range from 0.35 to 0.73 which indicates the suitability of the items.

#### 3.6.2 Discrimination Power

Discrimination power means "calculating the degree to which a particular item's results correspond with the results of the entire test" (Alderson, 1995:80). This means that an object is deemed to have weak power of discrimination if it is correctly scored by high-skilled students as well as low-skilled students. Item discrimination refers to the degree to which an object makes a difference between good and poor testers.

An object has good power of discrimination if it collects the right answers from the good students and the wrong answers from the bad students. It is worth noting that the high power of discrimination will be close to 1.0, and no power of discrimination will be nil at all (Brown, 2010: 71). The test item discrimination power was found to have a range of (0.27) - (0.73). The test items for DP and DL are shown in the following table:

**Table ( 4 )**  
*Difficulty level and Discrimination power*

Question	Item	Higher	Lower	Difficulty	Discrimination
Q1	1	58	28	0.67	0.50
	2	38	22	0.63	0.27
	3	40	20	0.67	0.33
	4	50	22	0.63	0.47
	5	52	18	0.73	0.57
	6	48	22	0.70	0.43
	7	46	21	0.50	0.42



	8	40	19	0.67	0.35
	9	40	19	0.67	0.35
	10	38	20	0.63	0.30
Q2	1	48	22	0.53	0.43
	2	40	19	0.67	0.35
	3	32	10	0.53	0.37
	4	44	16	0.38	0.47
	5	36	18	0.60	0.30
	6	51	23	0.72	0.47
	7	35	15	0.58	0.33
	8	46	17	0.47	0.48
	9	40	21	0.67	0.32
	10	33	12	0.55	0.35
Q3	1	70	30	0.70	0.67
	2	72	37	0.37	0.58
	3	66	45	0.52	0.35
	4	70	37	0.70	0.55
	5	62	18	0.72	0.73
Q4/A	1	60	24	0.55	0.60
	2	55	21	0.43	0.57
	3	61	17	0.38	0.73
	4	46	14	0.45	0.53
	5	53	20	0.35	0.55
Q4/B	1	187	86	0.71	0.67
Q5	1	241	101	0.46	0.62

### 3.7 Face Validity

Face validity refers to the degree to which a test appears to measure what it claims to measure. It is a global answer as a quick assessment of what the test is measuring (Mohajan, 2017: 16).

Taherdoost (2016) points out that face validity is the subjective evaluation of measuring instruments of relevance and presentation by the researcher to determine whether the items in the instrument are clear and reasonable.

In order to ensure the validity of the two instruments of the study, they have been submitted to a jury of teachers and specialists in the field of methods of teaching the English language. All the notes and modifications stated by jurors have been considered.

### 3.8 Construct Validity

Mc Burney and White (2007) defined construct validity as the property of a test that the measurement actually measures the constructs they are designed to measure.

Construct validity is the extent to which a measurement tool is truly assessing what it has been designed to assess (Drew, 2022).

Table (5 )

Percentage of the Conversation Test

Content		Objectives					
		Open a conversation	Close a conversation	Establish and change the topic	get, hold, and relinquish the floor	Interrupt	collaborate, and backchannel
Q1		1,5	3,6	2	7	4,10	9
Q2		3,5,7	1	10	2,6,9	4	8
Q3		5	3		2	4	1
Q4	A			1	2		
	B	1	3			2	
Q5		1	6	4	2	3	5
Percentage of each Item		0.21	0.15	0.10	0.18	0.15	0.10

## Section Four

### Analysis of data

#### 4.1 Results Related to the First Question

##### 4.1.1 Identifying the Difference Between Theoretical Mean and EFL university students' Performance in Conversation at Tikrit and Kirkuk Universities

To verify the second question, which states, "*Are there any statistically significant differences among EFL university students' performance in conversation at the two universities?*" The mean scores of the student's performance in the conversation test are 36.00 and the theoretical mean is 40 with a standard deviation of 10.31, respectively. The T-test formula for one sample is used. The calculated t-value is 6.236, which is found to be higher than the tabulated t-value, which is 1.972 at the 0.05 level of significance when the degree of freedom is 258. The results of Table 12 indicate there are statistically significant differences between the theoretical mean and students' performance in conversation tests at the two universities in favor of the theoretical mean. Thus, according to these results, the students of the two universities showed significant weakness in the conversation test.

**Table (6 )**

*The Mean Scores, Standard Deviations, and One Sample T-Value of the Students' students' performance in conversation Test at the Two Universities*

Group	No. of students	Mean	SD.	Theoretical Mean Score	T-Value		DF	Level of Significance
Conversation	259	36.00	10.31	40	Calculated	Tabulated	258	0.05
					6.236	1.972		

##### 4.1.2 Comparison Between the Mean Scores of the Tikrit and Kirkuk Universities at the Conversation Test

To find out whether "*There are any statistically significant differences among EFL university students' performance in conversation at the two universities*", all mean scores are obtained and compared. Statistics show that the mean score of Tikrit University is 33.96 and that of Kirkuk University is 37.96. By using the t-test formula for two independent samples, the calculated t-value is found to be 3.182, which is higher than the tabulated t-value which is found to be 1.972 at the degree of freedom 257, and the level of significance 0.05. This means indicates that there is a significant difference between the two universities in favor of Kirkuk University in the conversation test, as shown in Table 7.

**Table (7 )**

*The Mean Scores, Standard Deviations, and Two Independent Sample T-Value of the Students' Performance Conversation Test Between the Two Universities*

Groups	No. of students	Mean	SD.	T-Value		DF	Level of Significance
Tikrit	127	33.96	10.12	Calculated	Tabulated	257	0.05
Kirkuk	132	37.96	10.14	3.182	1.972		

#### 4.1.3 Comparison Among EFL University students' performance in conversational model components at the two universities

In order to find out the EFL university students' performance in conversational model components at the two universities", the one-way analysis of variance was used by the researcher, as shown in the following table:

**Table (8 )**

*One-Way Analysis of Variance (ANOVA) Among conversational model components at the two universities*

	Sum of Squares	DF	Mean Square	F-value		Sig.
				Calculated	Tabulated	
Between Groups	2222.530	5	444.506	60.780	2.22	0.05
Within Groups	11321.035	1548	7.313			
Total	13543.565	1553				

Table 8 shows that the calculated F-value value (60.780) is higher than the tabulated F-value 2.22 at the 0.05 level of significance and DF = 5, 1548. This indicates that there are significant differences between EFL university students in their performance in the test of conversational model components.

**Table (9 )***Comparisons of Means Among Two Universities (Scheffe<sup>a</sup>)*

Groups	N	Subset for alpha = 0.05			
		1	2	3	4
<b>Interrupt</b>	259	4.6988			
<b>Establish and change topics</b>	259	4.9266			
<b>Collaborate and backchannel</b>	259	5.2934	5.2934		
<b>Get, hold, and relinquish the floor</b>	259		6.0811		
<b>Open a conversation</b>	259			7.0888	
<b>Close a conversation</b>	259				8.0116
<b>Sig.</b>		0.282	0.052	1.000	1.000
a. Uses Harmonic Mean Sample Size = 259.					

According to Table 9 above, the comparisons of means showed that the mean scores of the Interrupt are 4.6988, Establish and change topics is 4.9266, Collaborate and backchannel is 5.2934, Get, hold, and relinquish the floor is 6.0811, Open a conversation is 7.0888, and Close a conversation is 8.0116 with harmonic mean sample size = 259. This means indicates that there is a significant difference among Iraqi EFL university students' performance in conversational model components at the two universities and in favor of open a conversation component.

#### **4.1.4 Comparison Among EFL University students' performance in conversational model components at Tikrit University**

In order to find out the EFL university students' performance in conversational model components at Tikrit University", the one-way analysis of variance was used, as shown in the following table:

**Table ( 10 )**

*One-Way Analysis of Variance (ANOVA) Among conversational model components at Tikrit University*

	Sum of Squares	DF	Mean Square	F-value		Sig.
				Calculate d	Tabulated	
<b>Between Groups</b>	1029.030	5	205.806	27.954	2.23	0.05

<b>Within Groups</b>	5565.906	756	7.362			
<b>Total</b>	6594.936	761				

Table 10 shows that the calculated F-value value 27.954 is higher than the tabulated F-value 2.23 at the 0.05 level of significance and DF = 5, 756. This indicates that there are significant differences between EFL Tikrit university students in their performance in the test of conversational model components.

**Table (11)**

*Comparisons of Means Among Tikrit University (Scheffe<sup>a</sup>)*

Groups	N	Subset for alpha = 0.05			
		1	2	3	4
<b>Interrupt</b>	127	4.1732			
<b>Establish and change topics</b>	127	4.7244	4.7244		
<b>Collaborate and backchannel</b>	127	5.1102	5.1102		
<b>Get, hold, and relinquish the floor</b>	127		5.7244	5.7244	
<b>Open a conversation</b>	127			6.8425	6.8425
<b>Close a conversation</b>	127				7.4803
<b>Sig.</b>		0.183	0.126	0.057	0.622
a. Uses Harmonic Mean Sample Size = 127.					

According to Table 11 above, the comparisons of means showed that the mean scores of the Interrupt are 4.1732, Establish and change topics is 4.7244, Collaborate and backchannel is 5.1102, Get, hold, and relinquish the floor is 5.7244, Open a conversation is 6.8425, and Close a conversation is 7.4803 with harmonic mean sample size = 127. This means indicates that there is a significant difference among EFL university students' performance in conversational model components at Tikrit University and in favor of close a conversation component. (Manal, 2023)

#### **4.1.5 Comparison Among EFL University students' performance in conversational model components at Kirkuk University**

In order to find out the EFL university students' Performance in conversational model components at Kirkuk University", the one-way analysis of variance was used, as shown in the following table:



**Table (12)**

*One-Way Analysis of Variance (ANOVA) Among conversational model components at Kirkuk University*

	Sum of Squares	DF	Mean Square	F-value		Sig.
				Calculate d	Tabulated	
<b>Between Groups</b>	1224.314	5	244.863	34.674	2.23	0.05
<b>Within Groups</b>	5550.583	786	7.062			
<b>Total</b>	6774.898	791				

Table 12 shows that the calculated F-value value 34.674 is higher than the tabulated F-value 2.23 at the 0.05 level of significance and DF = 5, 786. This indicates that there are significant differences between EFL Kirkuk University students in their performance in the test of conversational model components.

**Table (13)**

*Comparisons of Means Among Kirkuk University (Scheffe<sup>a</sup>)*

Groups	N	Subset for alpha = 0.05			
		1	2	3	4
<b>Establish and change topics</b>	132	5.1212			
<b>Interrupt</b>	132	5.2045			
<b>Collaborate and backchannel</b>	132	5.4697	5.4697		
<b>Get, hold and relinquish the floor</b>	132		6.4242	6.4242	
<b>Open a conversation</b>	132			7.3258	
<b>Close a conversation</b>	132				8.5227
<b>Sig.</b>		0.951	0.131	0.181	1.000

a. Uses Harmonic Mean Sample Size = 132.

According to Table 13 above, the comparisons of means showed that the mean scores of the Establish and change topics are 5.1212, Interrupt is 5.2045, Collaborate and backchannel is 5.4697, Get, hold and relinquish the floor is 6.4242, Open a conversation is 7.3258, and Close a conversation is 8.5227 with

harmonic mean sample size = 132. This means indicates that there is a significant difference among EFL university students' performance in conversational model components at Tikrit University and in favor of close a conversation component.

#### 4.1.2 Results Related to the Second Question

##### 4.1.2.1 Comparison Between the Mean Scores of Conversation According to Gender Variable at the Two Universities

To find out whether "*There are any statistically significant differences among EFL university students' performance in conversation according to gender at the two universities*", all mean scores are obtained and compared. Statistics show that the mean scores of the males are 36.55 and that of the females are 35.78. By using the t-test formula for two independent samples, the calculated t-value is found to be 0.540, while the tabulated t-value is found to be 1.972 at the degree of freedom 257 and level of significance 0.05. This means indicates that there is no significant difference between males and females in the conversation test at the two universities, as shown in Table 14.

**Table (14)**

*The Mean Scores, Standard Deviations, and two independent Samples T-Value of the Students' Conversation According to Gender*

Group s	No. of students	Mean	SD.	T-Value		DF	Level of Significanc e
Male	72	36.55	10.21	Calculated	Tabulated	257	0.05
Femal e	187	35.78	10.38	0.540	1.972		

##### 4.1.2.2 Comparison Between the Mean Scores of Conversation According to Gender Variable at the Two Universities

To find out whether "*Are there any statistically significant differences among Iraqi EFL university students' performance in conversation according to gender at the two universities?*", all mean scores are obtained and compared. Statistics show that the mean scores of the males are 36.55 and that of the females are 35.78. By using the t-test formula for two independent samples, the calculated t-value is found to be 0.540, while the tabulated t-value is found to be 1.972 at the degree of freedom 257 and level of significance 0.05. This means indicates that there is no significant difference between males and females in the conversation test at the two universities, as shown in Table (15).

**Table (15)**

*The Mean Scores, Standard Deviations, and two independent Samples T-Value of the Students' Conversation According to Gender*

Group s	No. of students	Mean	SD.	T-Value		DF	Level of Significance
Male	72	36.55	10.21	Calculated	Tabulated	257	0.05
Female	187	35.78	10.38	0.540	1.972		

#### 4.1.2.3 Comparison Between the Mean Scores of Conversation According to Gender Variable at Tikrit University

To find out whether "Are there any statistically significant differences among Iraqi EFL university students' performance in conversation according to gender at Tikrit University?", all mean scores are obtained and compared. Statistics show that the mean scores of the males are 35.79 and that of the females are 32.96. By using the t-test formula for two independent samples, the calculated t-value is found to be 1.503, while the tabulated t-value is found to be 1.972 at the degree of freedom 125 and level of significance 0.05. This means indicates that there is no significant difference between males and females in the conversation test at Tikrit University, as shown in Table (16).

**Table (16)**

*The Mean Scores, Standard Deviations, and Two Independent Sample T-Value of the Students' Conversation Test at Tikrit University*

Group s	No. of students	Mean	SD.	T-Value		DF	Level of Significance
Male	44	35.79	10.82	Calculated	Tabulated	125	0.05
Female	83	32.96	9.70	1.503	1.980		

#### 4.1.2.4 Comparison Between the Mean Scores of Conversation According to Gender Variable at Kirkuk University

To find out whether "Are there any statistically significant differences among Iraqi EFL university students' performance in conversation according to gender at

*Kirkuk University?*", all mean scores are obtained and compared. Statistics show that the mean scores of the males are 37.75 and that of the females are 38.02. By using the t-test formula for two independent samples, the calculated t-value is found to be 0.129, while the tabulated t-value is found to be 1.972 at the degree of freedom 130 and level of significance 0.05. This means indicates that there is no significant difference between males and females in the conversation test at Kirkuk University, as shown in Table (17).

**Table (17)**

*The Mean Scores, Standard Deviations, and Two Independent Sample T-Value of the Students' Conversation Test at Kirkuk University*

Group s	No. of students	Mean	SD.	T-Value		DF	Level of Significance
Male	28	37.75	9.25	Calculated	Tabulated	130	0.05
Female	104	38.02	10.41	0.129	1.980		

#### 4.1.3.6 The Correlation Between Iraqi EFL university students' TK and their performance in Conversation

In order to investigate the correlation which is between "*EFL university students' TK and their performance in conversation*", a question must be verified stating that "*Is there a correlation between Iraqi EFL university students' TK and their performance in conversation?*". Accordingly, the Pearson correlation coefficient is utilized. The results calculated reveal that the r- value is (0.059) and critical one (0.194) at a level of significance (0.05) and sample size (259). Consequently, this indicates that there is no correlation coefficient between Iraqi EFL university students' TK and their performance in conversation, as shown in Table (18).

**Table (18)**

*The Correlation between TK and Conversation*

Sample Size	R-Value	Critical value	Significance 0.05
259	0.059	0.194	Not Sig.

## Conclusions

This research highlights the performance of EFL university students in conversation at Tikrit and Kirkuk Universities. The findings reveal weaknesses in

students' conversation skills and significant differences in performance between the two universities. However, no statistically significant differences were observed based on gender. These results emphasize the importance of targeted interventions to enhance conversation abilities and suggest the need for further exploration of instructional methods and cultural factors that impact students' conversational competence.

## References

- Aiken, L. R. (1996). **Rating scales and checklists: Evaluating behavior, personality and attitudes**. New York: John Wiley & Sons, Inc.
- Andrade, H., & Du, Y. (2005). **Student perspectives on rubric-referenced assessment**. *Practical Assessment, Research & Evaluation*, 10, 1-11. <https://doi.org/10.1080/02602930801955986>
- Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). **Two minds, one dialog: Coordinating speaking and understanding**. In *Psychology of learning and motivation* (Vol. 53, pp. 301-344). Academic Press.
- Brown, H. D. (1990). **Language assessment: Principles and classroom practices**. London: Longman.
- Brown, H. Douglas. (2014). **Principles of Language Learning and Teaching**. NY: Pearson.
- Canale, M., & Swain, M. (1980). **Theoretical bases of communicative approaches to second language teaching and testing**. *Applied Linguistics*, 1(1), 1-47. <http://dx.doi.org/10.1093/applin/1.1.1>
- Celce-Murcia, M. (1995). **The Elaboration of Sociolinguistic Competence: Implications for teacher education**.
- Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1995). **Communicative competence: A pedagogically motivated model with content specifications**. *Issues in Applied Linguistics*, 6(2), 5-35.
- Chapman, R. F. (1982). **Chemoreception: The significance of receptor numbers**. *Advances in insect physiology*, 16, 247-356.
- Chomsky, N. (1965). **Aspects of the theory of syntax**. Cambridge, MA: MIT Press.
- Churches, A. (2015). **A guide to formative and summative assessment and rubric development**. 21st Century Project.
- Company, P., Contero, M., Otey, J., Camba, J. D., Agost, M. J., & Pérez-López, D. (2017). **Web-Based system for adaptable rubrics case study on CAD assessment**. *Educational Technology and Society*, 20(3), 24-41. <https://doi.org/10.2307/26196117>
- Corder, S. P. (1973). **Introducing applied linguistics**. Middlesex: Hazell Watson and Viney Ltd.
- Crystal, D., & Davy, D. (1975). **Advanced Conversational English**. London: Longman Group Ltd.

- Drew, C. (2022, September 16). **9 Types of Validity in Research**. Helpful Professor. <https://helpfulprofessor.com/types-of-validity/>
- Eckard, R. D., & Kearny, M. A. (1981). **Teaching Conversation Skills in ESL, Language in Education: Theory and Practice**. Washington: Center for Applied Linguistics.
- Garagri, H. A., & Mohammed, H. S. (2018). **Investigating instructors' techniques to develop EFL university students' conversational competence**. Academic Journal of Nawroz University (AJNU), 7(4), 261-271.
- GHALEM, Â., Chafik, O. K. A. R., CHROQUI, R., & EL ALAMI, S. E. M. M. A. (2016). **Performance: A concept to define! La performance: Un concept à définir!**.
- Harris, D. P. (1969). **Testing English as a Second Language**.
- Horton, B. K. (2018). **Sedimentary record of Andean mountain building**. Earth-Science Reviews, 178, 279-309.
- Hymes, D. H. (1972). **On communicative competence**. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected Readings* (pp. 269-293). Harmondsworth: Penguin.
- Jacobson, A. (1976). **Application of the "Wits" appraisal**. American Journal of Orthodontics, 70(2), 179-189.
- Levinson, S. C. (1983). **Conversational structure**. Pragmatics, 284-370.
- Lesson, R. (1975). **Fluency and Language Teaching**. London: Longman Group Ltd.
- Macmillan Dictionary. (2012). **Online Dictionary**. Retrieved from <http://www.macmillandictionary.com/dictionary/british/investigation>.
- Rezaee, A. A., & Shabani, E. A. (2019). **Self-and Peer-Assessments in the Iranian Context: A Systematic Review**. In *Handbook of Research on Curriculum Reform Initiatives in English Education* (pp. 270-282).
- Richards, J. C., & Sukwiwat, M. (1983). **Language transfer and conversational competence**. Applied Linguistics, 4(2), 113-125.
- Rivers, W. M. (1968). **Teaching Foreign Language Skills**. Chicago: The University of Chicago Press.
- Thomas, J. (1983). **Cross-cultural pragmatic failure**. Applied Linguistics, 4(2), 91-113.
- Manal omar mousa ,(2023) measuring thecreative writing skills of Efl university students 3(1)p,24
- Wongrassamee, S., Simmons, J. E., & Gardiner, P. D. (2003). **Performance measurement tools: The Balanced Scorecard and the EFQM Excellence Model**. Measuring Business Excellence, 7(1), 14-29.