THE EFFECT OF TASK COMPLEXITY AND TASK CONDITION ON SPOKEN LANGUAGE PRODUCTION BY INDONESIAN EFL LEARNERS

A Thesis

By

WORO ZULI ASTUTI

NPM 2423042021

MASTER OF ENGLISH EDUCATION STUDY PROGRAM
LANGUAGE AND ARTS EDUCATION DEPARTMENT
TEACHER TRAININGAND EDUCATION FACULTY
UNIVERSITY OF LAMPUNG
2025

ABSTRACT

THE EFFECT OF TASK COMPLEXITY AND TASK CONDITION ON SPOKEN LANGUAGE PRODUCTION BY INDONESIAN EFL LEARNERS

By

WORO ZULI ASTUTI

This study investigates the effect of task complexity and task condition (gender) on the CAF (Complexity, Accuracy, and Fluency) of spoken language performance by Indonesian EFL learners, framed within Robinson's (2007) Triadic Componential Framework. Adopting an exploratory mixed-methods, the research was conducted in two phases. In the first phase, a Topic Preference Questionnaire identified gender-based topic preferences among twelfth-grade students at MAN 1 East Lampung. Based on these preferences, a series of monologic speaking tasks of varying cognitive complexity were designed and administered in the second phase. The learners' oral performances were analyzed using CAF measures. Statistical analysis using Mann-Whitney U tests revealed significant gender differences in complexity and fluency, with female learners producing more complex and fluent speech, while accuracy showed no significant difference between groups. These findings partially support Robinson's Cognition Hypothesis, indicating that increased task complexity can promote syntactic elaboration and align with Skehan's Trade-Off Hypothesis, reflecting limited attentional resources distributed among CAF dimensions. The study concludes that gender functions as a significant influencing factor in managing cognitive load during task performance. Pedagogically, the findings highlight the need for gender-sensitive and cognitively principled task design to enhance spoken fluency and linguistic sophistication in EFL contexts.

Keywords: *task complexity, task condition, gender, speaking performance, CAF, Triadic Componential Framework.*

THE EFFECT OF TASK COMPLEXITY AND TASK CONDITION ON SPOKEN LANGUAGE PRODUCTION BY INDONESIAN EFL LEARNERS

By

WORO ZULI ASTUTI

A Thesis

Submitted in a partial fulfillment of
The requirements for S-2 Degree

MASTER OF ENGLISH EDUCATION STUDY PROGRAM
LANGUAGE AND ARTS EDUCATION DEPARTMENT
TEACHER TRAININGAND EDUCATION FACULTY
UNIVERSITY OF LAMPUNG
2025

Research Title	: **THE EFFECT OF TASK COMPLEXITY AND TASK CONDITION ON SPOKEN LANGUAGE PRODUCTION BY INDONESIAN EFL LEARNERS**
Student's Name	: Woro Zuli Astuti
Student's Number	: 2423042021
Study Program	: Magister in English Language Teaching
Department	: Language and Arts Education
Faculty	: Teacher Training and Education

APPROVED BY

Advisory Committee

Advisor

Mahpul, M.A., P.hD.
NIP. 196507061994031002

Co-Advisor

Dr. Budi Kadaryanto, M.A.
NIP. 198103262005011002

Chairperson of Department
Language and Arts Education

Dr. Sumarti, S.Pd., M.Hum.
NIP 197003181994032002

Chairperson of Master
In English Language Teaching

Dr. Budi Kadaryanto, M.A.
NIP. 198103262005011002

ADMITTED BY

1. Examination Committee

   Chairperson : **Mahpul, M.A., P.hD.** .....

   Secretary : **Dr. Budi Kadaryanto, M.A.** .....

   Examiners : 1. **Prof. Dr. Muhammad Sukirlan, M.A.** .....

   2. **Dr. Ari Nurweni, M. A.** .....

2. Dean of Teacher Training and Education Faculty

   Dr. Albet Maydiantoro, M.Pd.
   NIP 198705042014041001

3. Director of Postgraduate Program

   Prof. Dr. Ir. Murhadi, M.Si.
   NIP 196403261989021001

4. Graduated on: **18th December 2025**

LEMBAR PERNYATAAN

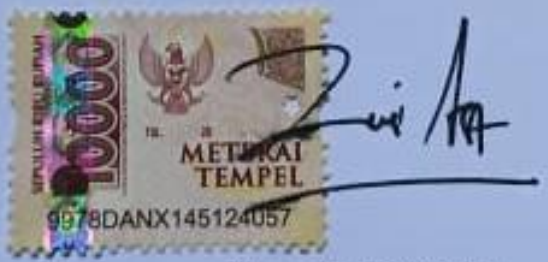
Dengan ini saya menyatakan dengan sebenarnya bahwa:

1. Thesis dengan judul "The Effect of Task Complexity and Task Condition on Spoken Language Production by Indonesian EFL Learners" adalah hasil karya sendiri dan saya tidak melakukan penjiplakan atau pengutipan atas karya penulis lain dengan cara tidak sesuai tata etika ilmiah yang berlaku dalam masyarakat akademik atau yang disebut plagiarisme.
2. Hak intelektual atas karya ilmiah ini diserahkan sepenuhnya kepada Universitas Lampung.

Atas pernyataan ini, apabila dikemudian hari ternyata ditemukan adanya ketidakbenaran, saya bersedia menanggung akibat dan sanksi yang diberikan kepada saya, saya bersedia dan sanggup dituntut sesuai hukum yang berlaku.

Bandar Lampung, 18th December 2025

Yang Membuat Pernyataan



Woro Zuli Astuti

NPM. 2423042021

CURRICULUM VITAE

The writer's full name is Woro Zuli Astuti. She was born in Wonosobo on July 11th, 1981, and is the only daughter of Asiyah and Fatkhurrohman. She began her education at SD Negeri 2 Tanjung Harapan and graduated in 1992. She completed her junior high school education at Mts Muhammadiyah Sekampung in 1995 and her senior high school education at SMA Negeri 1 Punggur Kota Gajah in 1998. She earned her Bachelor's degree from the English Education Department of Purworejo Muhammadiyah University in 2003. She has been serving as an English teacher at MAN 1 East Lampung since 2005. In February 2025, she officially enrolled as a Master's student in the English Education Program at Lampung University.

DEDICATION

I would proudly dedicate this thesis to:

My beloved father and mother

My supporting system, my husband

My lovely daughters

My sisters and brothers

My beloved best friends in MAN 1 East Lampung

My friends in Master of English Education

My alma mater University of Lampung

MOTTO

فَلَهَا نْ أَحْسَنْتُمْ أَحْسَنْتُمْ لِأَنْفُسِكُمْ وَإِنْ أَسَأْتُمْ

(If you do good, you do it for your own good. And if you do evil, you do it against your own self.)

(Q.S. Al Isra: 7)

ACKNOWLEDGEMENT

Alhamdulillah, first and foremost, the writer would like to praise and thank Allah SWT, the Almighty, for His countless blessings and for granting her the opportunity to complete this thesis. This work, entitled "The Effect of Task Complexity and Task Condition on Spoken Language Production by Indonesian EFL Learners," fulfills one of the requirements for her S2 degree in the Language and Arts Department, Faculty of Teacher Training and Education, at the University of Lampung. Greetings and invocations are presented to the Prophet Muhammad SAW, who has guided mankind to the right path blessed by the Lord. The writer has many people to thank and is deeply grateful to all who have supported her in completing this thesis. Although there are too many names to mention individually, several deserve her special and sincere appreciation. Her deepest gratitude goes to Mahpul, M.A., Ph.D., her principal advisor, for his invaluable guidance, scientific insight, unlimited patience, and continuous support throughout the preparation and completion of this thesis. It was an honor to have been supervised by an expert in Task-Based Language Teaching. The writer also extends her sincere thanks to Dr. Budi Kadaryanto, M.A., her co-advisor, for his constructive feedback, encouragement, and commitment, which greatly strengthened this research. She is immensely grateful to Prof. Dr. Muhammad Sukirlan, M.A., the first examiner, for his helpful suggestions and contributions that improved the quality of this thesis. Special appreciation is also due to Dr. Ari Nurweni, M.A., the co-examiner, for generously providing support, input, assistance, and corrections that helped refine this work. The writer thanks all lecturers and administrative staff of the Master's Program in English Education for their practical contributions and outstanding assistance throughout her studies and during the completion of this thesis. To all fellow students of MPBI 24, thank you for your sincere and enduring friendship. Finally, the writer sincerely welcomes any suggestions or constructive feedback that may help enhance the quality of this work.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I
INTRODUCTION

This chapter explains about the background of the research, the research question, objective of the research, benefits of the research, scope of the research, and definition of terms.

1.1 Background

Mastery of English language holds significant importance in today's globalized world. For Indonesian EFL (English as a Foreign Language) learners, mastering English is not merely about acquiring a new language but also about unlocking broader opportunities on the global stage (Zein et al., 2020). This context underscores the urgency of improving English proficiency through effective instructional approaches, particularly within formal education.

However, English proficiency among Indonesian junior and senior high school students often remains limited, especially in productive skills like speaking and writing (Kumayas & Lengkoan, 2023) . Speaking, in particular, is frequently perceived as the most challenging skill. Students cite reasons such as insufficient English vocabulary, difficulties with memorization and pronunciation which differs significantly from Indonesian, fear of making mistakes, anxiety about being ridiculed by peers, and a lack of grammatical knowledge (Megawati, 2016). Consequently, while many students can comprehend written and spoken English to a certain extent, their ability to use English expressively often remains underdeveloped (Ibrahim, 2006). This limitation is frequently attributed to a lack of meaningful English use outside the classroom and the dominance of traditional teaching methods (Shvidko et al., 2015).

In this educational landscape, the role of schools becomes crucial in equipping

students with the necessary skills to communicate effectively in English. Although the Indonesian National Curricula emphasize communicative competence, they sometimes lack specificity in linguistic competence and pragmatic functions, which can pose challenges (Panggua et al., 2022). Furthermore, Indonesian students often face significant hurdles in speaking English, attributed to factors such as the varying competence levels among EFL teachers and inadequate speaking skills training materials (Panggua et al., 2022). Many educators continue to adopt grammar- oriented and teacher-centered methods that focus on linguistic form over communicative function (Abrar, 2018). Therefore, an alternative pedagogical approach that emphasizes communication and real-world language use is urgently needed to bridge the gap between curriculum goals and classroom reality.

One alternative is Task-Based Language Teaching (TBLT), which prioritizes meaning-focused tasks to promote language learning through interaction and purposeful communication (Nunan, 2004). TBLT provides learners with opportunities to use the target language in authentic contexts that require both fluency and accuracy (Albino, 2017). This approach has demonstrated potential in improving various language skills, particularly speaking. For instance, research by Masuram & Sripada (2020) demonstrates that implementing TBLT activities significantly enhances EFL learners' oral fluency and interaction skills. This is supported by context-specific findings from Fitriani & Wirza (2018) who report that speaking tasks such as picture narration and problem-solving improve fluency and spoken structure among Indonesian learners. Moreover, TBLT not only enhances linguistic performance but also increases learners' motivation and engagement through collaboration, critical thinking, and contextualized problem-solving (Córdoba Zúñiga, 2016).

The success of TBLT, however, largely depends on the design and complexity of the tasks used. Effective task design requires careful consideration of various factors, including the cognitive demands of the task and the specific needs of the learners (Robinson, 2011). Well-designed tasks must align with learners' cognitive capacity, linguistic resources, and background knowledge to promote

optimal learning outcomes (Erlam, 2016). As noted by Vivian (2017), task complexity, especially when tailored to students' prior knowledge, has a significant effect on oral performance. To understand and manipulate task complexity more systematically, Robinson's Cognition Hypothesis, particularly his Triadic Componential Framework (TCF), provides an essential theoretical lens.

This TCF framework breaks down task design into three core elements: task complexity (cognitive factors), task difficulty (learner factors), and task condition (interactional factors). Robinson also distinguishes between resource-directing dimensions, factors that push learners to process specific linguistic forms and resource dispersing dimensions, factors that challenge learners' attention management or memory (Robinson, 2001). According to this theory, tasks that are complex in resource-directing dimensions but simple in resource-dispersing ones are most effective in promoting language development (Robinson, 2001). Therefore, understanding and manipulating these dimensions are key to designing effective speaking tasks.

Previous studies have shown that manipulating task complexity typically by altering reasoning demands, planning time, or information structure can influence the fluency, accuracy, and complexity (CAF) of language production (Awwad, 2017; Piri et al., 2012; Rahimi, 2019). However, many previous studies, such as those Ghaderi et al., (2022), Sánchez & Kalamakis (2023) and Xu et al., (2022), have tended to focus on single-dimension manipulations or written tasks. This emphasis means they often neglect the broader interactional or contextual elements of the task, particularly those related to task condition. As a result, the effect of contextual variables such as group composition, gender, and learner preferences on spoken task performance remains largely underexplored.

One under-explored aspect within task conditions is how gender and topic preferences influence learners' engagement and spoken performance. Research shows that spoken task performance is shaped not only by task complexity but

also by individual characteristics, particularly gender. Gender in language learning refers to socially constructed expectations that influence how men and women use and respond to language (Cameron, 2005). Classic work by Lakoff (1975) and Tannen (1990) shows that males and females differ in communication styles, conversational goals, and preferred topics. Haas (1979) similarly found that males tend to choose competitive, factual, or status-oriented topics, while females prefer relational, expressive, and interpersonal themes. Later studies by Bischoping (1993) and Coates (2017) confirm these patterns, noting that males gravitate toward action-based or information-rich subjects, whereas females favor socially and emotionally oriented content.

Empirical studies also show clear gender-based topic preferences: male learners commonly choose themes such as technology, science, sports, and business, while female learners prefer art, creativity, fashion, beauty, and interpersonal relationships (Al-Shibel, 2021; A. Chen, 2012; Svirina & Ashrapova, 2020). Despite these well-documented tendencies, very few studies have examined how gender-linked topic preferences interact with task complexity to influence spoken task performance. Much existing research discusses task complexity without considering social variables and focuses largely on written language (Ayu, 2020; Kamel et al., 2019; Xu et al., 2022).

Therefore, this study aims to fill this identified gap by selectively manipulating task complexity, specifically by making the resource-directing aspects complex while keeping the resource-dispersing aspects simple, and by integrating gender-based topic preferences as task condition variables. Grounded in Robinson's Triadic Componential Framework (TCF), this design helps focus learners' attention on form without overwhelming their cognitive capacity. The study investigates how cognitive and social-contextual factors interact to influence Indonesian EFL learners' oral performance in terms of complexity, accuracy, and fluency (CAF), offering insights for more inclusive and effective TBLT task design.

1.2 Research Questions

1. What topic of task are male and female students mostly interested in?
2. Is there a statistically significant difference in spoken language production generated by female and male students performing tasks developed on the basis of task complexity and gender-specific topic preferences?

1.3 Research Objectives

1. To explore the topic of tasks male and female students are most interested in.
2. To investigate whether or not there is a statistically significant difference in spoken language production generated by female and male students performing the tasks developed on the basis of task complexity and gender-specific topic preferences.

1.4 Research Benefits

1. Theoretical Benefits
    a. Expanding linguistics and educational psychology knowledge regarding gender differences in topic preferences and spoken language production.
    b. Contributing to further research on individual differences in TBLT, specifically by modeling how gender, as a task condition variable, interacts with cognitive task complexity.

2. Practical Benefits
    a. For educators: Providing insights into how gender differences influence students' communication styles and topic choices in academic discussions.
    b. For educational institutions: It is a foundation for designing a more inclusive and effective curriculum or teaching methods based on students' preferences and spoken language abilities.

1.5 Scope of the Research

This research is limited to investigating the influence of task complexity and task condition on the spoken language production of Indonesian EFL learners.

The study focuses on three key dimensions of speaking performance-complexity, accuracy, and fluency (CAF). The participants are twelfth-grade students at MAN 1 East Lampung, and the tasks are designed based on the principles of Task-Based Language Teaching (TBLT). The research includes speaking tasks designed with deliberate manipulations in cognitive demand and incorporates gender-based topic preferences as a key condition variable.

1.6 Definition of Terms

Task-based Language Teaching (TBLT): An approach in language teaching which uses the language in real situation setting not as an object for study (Ellis, 1993).

Task complexity: The result of the attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner (Robinson, 2001).

Resource-directing: Task complexity features that increase cognitive and conceptual demands by directing learners' attention to specific linguistic forms, such as through reasoning demand, causal relationships, or the number of elements to be processed (Robinson, 2001).

Resource-dispersing: Task complexity features that increase procedural demands by dispersing learners' cognitive resources across different aspects of the task, such as planning time, single versus dual tasks, or prior knowledge, without necessarily focusing attention on specific language forms (Robinson, 2001).

CAF (Complexity, Accuracy, Fluency): A triad of performance descriptors for an oral and written assessment of language and have also been used for measuring progress in language learning (Housen, 2009).

Gender in language learning: Refers to the ways in which language use and acquisition are influenced by the social and cultural constructs of gender (Cameron, 2005).

Gender Differences (in communication): Refers to the systematic variations in linguistic behavior, interactional style, and communicative preference that are socially constructed and performed, rather than biologically determined (Cameron, 2005).

CHAPTER II
LITERATURE REVIEW

This chapter explains about Concept of Speaking, Task Based Language Teaching, Definition of Tasks, Differences between Tasks and Exercises, Types of Tasks, Triadic Componential Framework, Measurement of language production, Theoretical assumptions, and Hypothesis.

2.1 Concept of Speaking

One of the most important goals of teachers is to enable learners to use English for speaking. Speaking is one of the main productive language skills and is considered an essential component of communicative competence, enabling learners to actively generate oral output for meaningful interaction in EFL settings. According to Bygate (1997), speaking is a skill that involves the ability to produce and organize ideas in a logical, coherent, and meaningful way, extending beyond mere sound or word production to appropriate language use in social contexts. Similarly, Brown defines speaking as an interactive process of constructing meaning that involves producing, receiving, and processing information, where both linguistic and sociolinguistic competence play important roles in achieving effective communication (Brown, 2014).

Speaking is also viewed as a complex skill that requires control of several elements such as pronunciation, grammar, vocabulary, fluency, and comprehension, as learners must deploy linguistic forms accurately and fluidly in real communication (Harmer, 2007). This aligns with Skehan's (1998) framework, which emphasizes that speaking performance can be measured through Complexity, Accuracy, and Fluency (CAF), with complexity reflecting the sophistication of language use, accuracy showing the correctness of

language forms, and fluency indicating the smoothness and speed of speech delivery.

Furthermore, speaking involves both knowledge and skill components, as Bygate (1987) points out that knowledge refers to knowing what to say and how to say it, while skill refers to using that knowledge automatically in real time. Therefore, teaching and assessing speaking must consider not only the linguistic aspects but also the communicative, cognitive, and contextual factors that influence performance, including recent emphases on self-regulation and digital tools to overcome anxiety in EFL contexts. This view supports the design of task-based speaking activities, such as those in the present study, which aim to elicit authentic spoken language and assess learners' ability through the CAF framework while addressing gender-based topic preferences.

2.2 Task-Based Language Teaching (TBLT)

Task-Based Language Teaching (TBLT) emerged in the early 1980s as a response to the need for more learner-centered and communicative approaches in language teaching. Prabhu (1988) developed a project known as the Communicational Teaching Project in Bangalore, India. This project was informed by Second Language Acquisition (SLA) theory and based on the premise that language teaching practice which focuses primarily on language forms is not effective for developing learner's competence in the target language.

TBLT is an approach to language teaching that emphasizes using tasks as the primary learning unit. Ellis (2003) states that TBLT emphasizes authentic communication in language learning to improve communicative competence. This approach provides task-based learning experiences that resemble real-life situations, allowing learners to develop their language skills naturally. The TBLT approach has a strong theoretical foundation in second language acquisition theory. Long (2015) argues that interaction in authentic tasks provides learners with opportunities to receive rich and meaningful language input. Through interactions that involve meaning negotiation, learners can better

understand the contextual use of language. Consequently, as globalization amplifies the demand for such communicative proficiency, TBLT's effectiveness in mirroring real-world language use and achieving specific outcomes has cemented its status as an increasingly relevant and widely recognized approach in language education (Xu et al., 2022). In its implementation, TBLT consists of three main stages: the pre-task stage, the task cycle stage, and the post-task stage (Ellis, 1993; Erlam, 2016). In the pre-task stage, teachers provide context and clear instructions. The task cycle stage involves communication activities that encourage active language use. Meanwhile, the post- task stage focuses on reflection and evaluation of the task outcomes.

One of the main advantages of Task-Based Language Teaching (TBLT) is its ability to enhance student engagement in the learning process. Research by Richards & Rodgers (2001) indicated that TBLT can increase student motivation by involving them in activities that are meaningful and relevant to real-life contexts. This sense of relevance also fosters greater learner confidence in using the target language. The implementation of TBLT has demonstrated its effectiveness in improving students' language skills. Chen, (2019) showed that technology-mediated collaborative video tasks, implemented within a TBLT framework, significantly improved students' oral communication, increased engagement, and encouraged greater autonomy in learning. Moreover, this approach has been successfully employed in the teaching of other foreign languages. According to Willis & Willis (2007) students who learn through this approach significantly improve fluency and accuracy in speaking. This occurs because they have opportunities to practice in an environment that supports honest communication. Moreover in the context of English language learning, Sabil (2020) found that TBLT could improve students' fluency and accuracy in speaking. They noted that integrating TBLT principles into existing curricula can enhance students' oral communication performance. Additionally, TBLT allows for integrating other language skills, such as reading, writing, and listening. A study by Nunan, (2004) found that this approach helps students understand texts more deeply because they practice using the language in

various contexts. Furthermore, a systematic review by Mudinillah et al., (2024) indicated that TBLT can enhance students' speaking and listening skills while promoting learner autonomy. In the context of foreign language learning, TBLT has been implemented at various educational levels, from elementary schools to universities (Metty, 2025; Munira & Ferdousi, 2020; Natsuko et al., 2012).

However, there are some challenges in implementing TBLT, such as the need for careful preparation and the active role of teachers as facilitators. Ji & Pham (2020) found that teachers must be able to design tasks that match students' proficiency levels and provide appropriate guidance to ensure an effective learning process. In Indonesian context, Saputro et al., (2021) identified that teachers in Indonesia face obstacles such as limited resources and lack of training in applying TBLT. Nonetheless, they also acknowledged its benefits in increasing student engagement and motivation. These studies suggest that TBLT has great potential in improving students' language skills. However, its successful implementation requires careful planning, adequate teacher training, and support from educational institutions. Ellis (2009) outlines a number of principles, which he suggests will facilitate the successful implementation of TBLT in a given educational context. One of these is that teachers need to have a clear understanding of what a language task is. Below is a further explanation of the Task in relation to TBLT

2.2.1 Definition of Tasks

The concept of 'task' is central to Task-Based Language Teaching (TBLT), yet it has been defined in various ways by different linguists and researchers. Here are several representative definitions of task. One of the earliest and broadest definitions was offered by Long (2015), who described a task as a piece of work undertaken for oneself or for others, freely or for some reward. Task is meant the hundred and one things people do in everyday life, at work, at play, and in between.

In the context of Task-Based Language Teaching (TBLT), a task is defined as an activity that involves language processing to achieve a specific communicative goal. Ellis (2003) describes a task as an activity

that requires the use of language in a meaningful and relevant situation for learners. Within TBLT, tasks emphasize interaction and meaning negotiation, distinguishing it from mechanical exercises that are often less contextual.

Nunan (2004) defines a task as a piece of classroom work which involves learners in comprehending, manipulating, producing or interacting in the target language while their attention is principally focused on meaning rather than form. The task should also have a sense of completeness, being able to stand alone as a communicative act in its own right with a beginning, middle and an end. This definition was given from the communicative perspective, focusing on meaning rather than on form.

Adding further nuance, Skehan (1996; 1998) emphasizes that a task is an activity where attention to meaning is paramount, there is a connection to the real world, a communicative problem to be solved, and the assessment of the task is primarily in terms of its outcome rather than linguistic accuracy alone. Furthermore, Bygate (2001) defines a pedagogical task as an activity where learners are encouraged to use language with an emphasis on meaning to achieve a learning objective. Skehan (1998) also made important distinctions between a task and a more traditional exercise, highlighting the communicative and meaning-focused nature of the former. Willis (1996) also contributes to the understanding of tasks by defining them as goal-oriented activities where learners use the target language for a communicative purpose to achieve a specific outcome. This practical, outcome driven conceptualization underscores the functionality of tasks in language learning.

Synthesizing these perspectives, a pedagogical task within TBLT is generally understood as a goal-oriented activity that requires learners to use the target language communicatively, with a primary focus on meaning rather than linguistic form, to achieve a defined outcome. Tasks

aim to mirror real-world language use, engage learners in authentic interaction, and often involve a clear structure or work plan. They are distinct from traditional exercises that tend to focus on the manipulation of language forms in isolation.

2.2.2 Differences between Tasks and Exercises

Ellis (2003) distinguishes between tasks and exercises based on cognitive engagement and communicative goals. Exercises are more structural, focusing on language form, while tasks emphasize meaning and the use of language in real contexts. Tasks allow learners to engage in more authentic language production compared to exercises.

The following is a table of Differences Between Tasks and Exercises based on the explanation by Ellis (2003):

Table 1. Differences Between Tasks and Exercises

Aspect	Tasks	Exercises
Focus	Meaning-focused	Form-focused
Purpose	Communicative goal in real-life contexts	Practice of specific language forms (grammar, vocabulary, etc.)
Cognitive Engagement	High cognitive engagement (problem-solving, decision-making)	Lower cognitive engagement (repetition, drilling)
Authenticity	Reflect real-world language use	Often artificial or decontextualized language use
Learner Involvement	Learner-centered, encourages autonomy and initiative	Teacher-centered, controlled practice
Language Production	Promotes spontaneous and meaningful language production	Encourages accurate reproduction of language forms
Interaction	Encourages communication, negotiation of meaning, peer interaction	Limited interaction, individual focus
Learning Outcome	Development of communicative competence	Mastery of specific linguistic structures

Aspect	Tasks	Exercises
Motivation	Increases learner motivation and self-efficacy	Motivation may rely on external factors
Adaptability	Easily adapted to various contexts and learner needs	Less adaptable, more standardized

Tasks primarily focus on meaning, aiming to engage learners in authentic communication. However, they also provide opportunities to draw learners' attention to linguistic forms as they arise naturally during interaction. This approach is known as Focus on Form (FonF), where attention to grammar and vocabulary occurs incidentally within meaningful language use. In contrast, exercises are typically associated with Focus on Forms (FonFs), where instruction centers on the explicit teaching and practice of discrete linguistic items, such as grammar rules or vocabulary lists, often in isolation from real communicative contexts. Thus, while tasks integrate form and meaning in a communicative framework, exercises prioritize the mastery of forms over communicative function. However, successful implementation of TBLT requires careful planning and consideration of task design. Tasks should be appropriately challenging, culturally relevant, and aligned with learners' language proficiency levels to maximize effectiveness. Teachers must have the skills to design and facilitate tasks promoting meaningful communication and language use.

In conclusion, the distinction between tasks and exercises outlined by Ellis (2003) underscores the importance of meaningful language use in learning. The adoption of TBLT in language education offers numerous benefits, including enhanced communicative competence, increased motivation, authentic language practice, collaborative learning opportunities, learner autonomy, and adaptability to various contexts. As research continues to support the efficacy of TBLT, educators are encouraged to incorporate task-based approaches to enrich language learning experiences.

2.2.3 Types of Tasks

Tasks in learning can be categorized based on their nature and purpose. According to Ellis (2003), several functions are designed to develop students' skills and knowledge through different approaches. Types of tasks commonly used in TBLT include the following:

1. Target vs. pedagogic tasks

   Long (1985) distinguishes between target tasks, which are real-world activities people perform in daily life, and pedagogic tasks, which are simplified classroom versions designed to help learners prepare for future real-life communication. Pedagogic tasks often act as smaller components of a larger target task. For example, completing sections of a job application such as educational background or work experience, may serve as preparatory tasks for writing an actual application. In practice, the line between target and pedagogic tasks is not always clear, since real-life tasks can be carried out for instructional purposes (e.g., asking for directions in an L2 environment or emailing an international partner in a foreign language setting).

2. One-way vs. two-way tasks

   In one-way tasks, only one participant holds the information needed for task completion, making them responsible for most of the communication, while the partner mainly indicates comprehension. In two-way tasks, both participants possess information that must be exchanged for the task to be successfully completed.

3. Open vs. closed tasks

   Open tasks do not require a single, predetermined outcome, allowing multiple possible solutions. Closed tasks, however, have a fixed answer or solution. For example, a spot-the-difference or picture-sequencing activity with one correct order is a closed task. If learners instead create their own story from unrelated pictures, the task would be considered open because no specific solution is required.

4. Convergent vs. divergent tasks

   In convergent tasks, participants must arrive at a shared solution or agreement. In divergent tasks, agreement is not necessary, and participants may produce different outcomes.

5. Focused vs. unfocused tasks

   Ellis (2020) introduced the distinction between focused and unfocused tasks. Unfocused tasks promote general communication without targeting specific language forms. Focused tasks also encourage communication but are designed to elicit a particular linguistic feature. A syllabus can consist entirely of unfocused tasks (e.g., Prabhu's 1987 Communicational Teaching Project) or include focused tasks guided by explicit structural objectives.

6. Input-based vs. output-based tasks

   Ellis (2020) defines input-based tasks as tasks that require learners to process language through listening or reading, without obligating them to produce output although production is not discouraged. Based on Swain's Output Hypothesis (in Duong, 2020), output-based tasks provide opportunities for learners to use the language and promote learning through noticing gaps, testing hypotheses, and reflecting on their language use. Overall, task-based learning offers a powerful and flexible approach that supports both language development and skill training. By engaging learners in meaningful tasks, the approach reduces pressure, increases motivation, and encourages active use of the target language.

2.2.4 Task-Based Language Teaching Methodology

The methodology of Task-Based Language Teaching (TBLT), as outlined by Ellis (2003), consists of three main stages: the pre-task, the task, and the post-task. Each stage plays a crucial role in facilitating effective and communicative language learning. The pre-task stage is a preparatory phase where the teacher provides clear instructions and context. At this stage, students are introduced to the topic, objectives,

and necessary steps to complete the task. Teachers may also provide examples or models to aid students' comprehension. Proper preparation at this stage is essential to ensure students understand the task and are ready to engage.

During the task stage, students actively complete the task designed to stimulate authentic language use. These tasks usually reflect real-life situations, allowing students to practice language skills in a relevant context. Active student participation in this stage can enhance their motivation and communication skills.

The post-task stage is a reflection phase where the teacher and students analyze and evaluate language use during the task execution. In this stage, constructive feedback is given to correct mistakes and reinforce students' understanding. This reflection is essential to help students recognize areas that need improvement and plan future learning strategies.

The following is a table based on the previous text description of the Methodology of TBLT

Tabel 2. Tabel Methodology of TBLT

The Methodology of TBLT	**Activity**	**Description**
1. Pre-Task	Planning Time	Teacher introduces topic, objectives, and task instructions. Students are prepared with context and language support, including examples/models.
	Doing similar task as the During Task	Students may be given a model or sample task to understand what they are expected to do.
2. During Task	Performing task similar to that done in the pre-task	Students complete the main task using authentic language. The focus is on meaning, communication, and active participation.

The Methodology of TBLT	Activity	Description
3. Post-Task	Learner Reports	Students share the outcomes of their task, often through presentations or group discussions.
	Consciousness Raising	Teacher provides feedback, discusses language use, corrects errors, and highlights target forms.
	Repeat Task	Students may repeat the task with improvements based on feedback, reinforcing learning and boosting confidence.

2.3 Triadic Componential Framework

Peter Robinson's Triadic Componential (TFC) Framework offers a comprehensive understanding of the factors influencing task complexity in second language acquisition (SLA). This framework categorizes these factors into three primary domains: Task Complexity (Cognitive Factors), Task Difficulty (Learner Factors), and Task Condition (Interactive Factors). The components of the framework could be seen as followed:

Tabel 3. The Triadic Componential Framework for Task Classification-Categories, Criteria, Analytic Procedures, and Design Characteristics from Robinson (2007)

Task complexity (cognitive factors)	Task condition (interactive factors)	Task Difficulty (learner factors)
(classification criteria: Cognitive demands) (classification procedure: Information-theoretic analyses)	(classification criteria: Interactional demands) (classification procedure: behavior-descriptive analyses)	(classification criteria: ability requirements) (classification procedure: ability assessment analyses)
(a)Resource-directing variables making cognitive/conceptual Demands	(a) Participation variables making interactional demands	(a) Ability variables and task relevant resource differentials
+/- here and now +/ few elements +/- spatial reasoning +/-causal reasoning -/+ intentional reasoning -/+ perspective-taking	+/- open solution +/- one-way flow +/-convergent solution +/- few participations +/- few contributions needed +/- negotiation not needed	h/l working memory h/l reasoning h/l task-switching h/l aptitude h/l field independence h/l mind/intention-reading
(b) Resource-dispersing	(b) Participant variables	(b) Affective variables

Variables making performative/procedural Demands	making interactant demands	and statetrait differentials
+/ planning time +/- single task +/- few steps +/- independency of steps +/- prior knowledge	+/- same proficiency +/- same gender +/- familiar +/- shared content knowledge +/- equal status and order	h/l openness to experience h/l control of emotion h/l task motivation h/l processing anxiety h/l willingness to

Each component of the Triadic Framework is described in more detail in the following sections. The factors of task complexity (cognitive factors) both the resource-directing and the resource-dispersing dimensions, are described in detail below.

2.3.1 *Task Complexity (Cognitive Factors)*

Robinson (2001) in his Cognition Hypothesis, states that task complexity can be categorized into two main dimensions: resource-directing (which directs attention to specific linguistic aspects such as grammar and sentence structure) and resource- dispersing (which involves factors that deplete cognitive resources, such as time pressure and the number of elements in a task). Mahpul (2014) found that dialogic tasks with higher complexity can enhance accuracy, but under unplanned conditions, complex tasks lead to greater fluency compared to more straightforward tasks. This indicates that the relationship between task complexity and language production is not always linear but depends on task conditions.

The Cognition Hypothesis, as proposed by Robinson (2001, 2007, 2011) and Skehan (1995), posits that the complexity of a task directly influences the cognitive demands placed on individuals undertaking it. In second language acquisition, this hypothesis suggests that more complex tasks require more significant cognitive resources, affecting language production in complexity, accuracy, and fluency. Understanding this relationship is crucial for designing effective language learning curricula that optimize learner outcomes.

Task complexity is the result of the attentional, memory, reasoning, and other information-processing demands imposed by the task structure on the language learner (Robinson, 2001). Robinson's Triadic Componential Framework differentiates between two dimensions of task complexity: resource-directing and resource-dispersing. Resource-directing variables, such as the number of elements involved or the reasoning demands, direct learners' attention to specific linguistic features necessary to meet task demands. Conversely, resource-dispersing variables, like planning time and prior knowledge, influence the distribution of cognitive resources without directing attention to specific linguistic features. This framework has been instrumental in guiding research on how varying task complexity impacts language production.

Empirical studies have yielded mixed results regarding the effects of task complexity on language production. For instance, a study by C. Li et al., (2024) examined the impact of cognitive and affective factors on young learners' writing performance. The findings indicated that increased task complexity and positive emotions enhanced the complexity and accuracy of learners' written output, supporting Robinson's Cognition Hypothesis. However, other studies have reported different outcomes. For example, research by Cho (2018) found that increased task complexity did not influence accuracy and syntactic complexity but positively affected fluency, aligning more with Skehan's Limited Attentional Capacity Model, which suggests a trade-off between different aspects of language production under increased cognitive load.

The mode of communication also plays a significant role in how task complexity affects language production. In a study investigating synchronous computer mediated communication (SCMC), L. Li, (2023) found that increased task complexity led to higher lexical complexity but lower accuracy in text-based SCMC. This suggests that the medium through which tasks are performed can mediate the effects of task

complexity, highlighting the need to consider technological contexts in task design.

Planning is another critical factor influencing the relationship between task complexity and language production. Research by Ong (2014) demonstrated that providing extended pre-task planning time led to greater fluency, lexical complexity, and overall writing quality. This finding implies that allowing learners time to plan before engaging in complex tasks can mitigate some of the cognitive demands, thereby enhancing performance. However, these results also suggest that the benefits of planning may vary depending on the specific aspects of language production being measured.

The interplay between cognitive and affective factors is also crucial in understanding the effects of task complexity. C. Li et al., (2024) integrated task-mediated cognitive-affective model of L2 writing emphasizes that positive and negative emotions interact with cognitive demands to influence writing processes and outcomes. This model suggests that learners' emotional states can either facilitate or hinder their ability to manage the mental demands of complex tasks, thereby affecting language production. Moreover, individual differences, such as working memory capacity and language aptitude, have been shown to moderate the effects of task complexity. Robinson (2011) argued that individuals with higher working memory capacity are better equipped to handle increased task complexity, improving both the complexity and accuracy of language production. This underscores the importance of considering learner-specific factors when designing tasks to ensure they are appropriately challenging yet manageable.

In addition to cognitive factors, task characteristics, such as prior knowledge and task structure, significantly influence language production. Rahimpour & Hazar (2007) found that learners with previous knowledge of a task topic produced language with higher lexical complexity but lower accuracy, suggesting that familiarity with content

allows learners to focus on more sophisticated language use, albeit at the expense of accuracy. Similarly, well-structured tasks have been associated with improved accuracy and fluency, highlighting the need for careful task design. Sequence and grading tasks are vital considerations in curriculum design based on their complexity. Saeedi et al., (2012) demonstrated that manipulating task complexity along resource-directing dimensions led to simultaneous increases in complexity and accuracy of language production. This finding supports task complexity as a basis for sequencing tasks in a language syllabus to progressively build learners' linguistic capabilities.

However, not all studies align with the predictions of the Cognition Hypothesis. For example, research by Ismail & Samad (2017) found that increased task complexity did not consistently lead to improvements in language production, suggesting that other factors, such as task familiarity and learner motivation, may play significant roles. These discrepancies highlight the complexity of the relationship between task complexity and language production and the need for further research to unravel these dynamics.

In conclusion, the relationship between task complexity and language production is multifaceted, influenced by cognitive demands, individual learner differences, task characteristics, and contextual factors. While the Cognition Hypothesis provides a valuable framework for understanding these dynamics, empirical findings suggest that its predictions may not universally apply across all contexts. Therefore, language educators and curriculum designers should consider these factors when developing tasks to optimize language learning outcomes.

2.3.2 *Task Condition (Interactive Factors)*

Within Robinson's Triadic Componential Framework (2001, 2007), Task Condition refers to the interactional and participatory demands that shape how learners engage with tasks. Unlike Task Complexity, which concerns the cognitive load built into the design of the task, Task

Condition focuses on the external interactional requirements that learners must follow during task performance. These requirements typically remain constant while complexity is manipulated. Robinson (2011) explains that Task Condition is classified through behavior-descriptive analyses, which evaluate how learners participate in interaction, for example, examining turn-taking behavior, negotiation of meaning, repair sequences, and feedback episodes observable in transcripts of task performance. Task Condition comprises two major components: participation variables and participant variables, each contributing uniquely to the interactional environment of a task.

1. Participation Variables

   Participation variables determine the structural configuration of the interaction required by the task. They specify the extent to which learners must collaborate, exchange information, and coordinate meaning-making. One of the most widely studied distinctions is between one-way and two-way tasks. In one-way tasks, information flows from a single speaker to a listener, such as in picture descriptions or story-telling tasks. These tasks tend to produce longer monologic turns but involve fewer opportunities for negotiation of meaning (Pica, Kanagy, & Falodun, 1993). In contrast, two-way tasks require each participant to hold unique information essential to task completion, such as in spot-the-difference or jigsaw tasks. Empirical work consistently shows that two-way tasks elicit more interaction, clarification requests, confirmation checks, and modified output (Gass et al., 2005) thus promoting richer conditions for L2 development.

   Another important participation variable concerns whether a task has a convergent or divergent solution. Convergent tasks, which require a single agreed-upon outcome, promote collaborative negotiation, precision in language use, and a strong focus on accuracy as learners work toward consensus (Ellis, 2003). Divergent

tasks, such as debates or brainstorming sessions, allow for multiple valid outcomes and therefore encourage more elaborate, personally driven responses and the expression of individual perspectives. The number of required participations also shapes interactional demands. Tasks designed to require few contributions from each participant may produce unequal participation or limit overall language production. Conversely, tasks that necessitate many contributions from all learners can generate extended discourse and push learners toward more sustained and balanced interaction (Robinson, 2011). Furthermore, the necessity for negotiation is a critical variable. Tasks that inherently involve information gaps, opinion gaps, or reasoning gaps tend to stimulate negotiation of meaning. Research shows that this negotiation facilitates noticing, pushed output, and modified interaction, core mechanisms of L2 acquisition (Long, 2015).

2. Participant Variables

   Participant variables refer to the social, interpersonal, and individual characteristics of the learners who perform the task. These factors influence how comfortable learners feel, how actively they participate, and how the interaction unfolds. One such variable is familiarity. Research indicates that familiar partners tend to show lower anxiety, greater willingness to communicate, and more naturalistic interaction, which can lead to increased risk-taking and voluble language production (Gass et al., 2005). Unfamiliar partners may increase communicative caution but can also heighten the need for explicit negotiation.

   Another key variable is proficiency level alignment. When participants share a similar proficiency level, interaction can become more balanced and reciprocal. However, mixed-proficiency pairings may lead to asymmetrical participation, yet they can also create opportunities for scaffolding, where higher-proficiency learners provide support that enables lower-proficiency partners to perform

beyond their independent ability (Storch, 2002)

Shared content knowledge is another factor shaping interaction. When learners possess similar background knowledge relevant to the task, they can interact more fluently. Limited shared knowledge may introduce comprehension challenges but could also stimulate more negotiation and clarification as learners work to establish common ground (Gass & Mackey, 2007).

Gender composition can also influence interaction styles. Research in discourse studies suggests that gender can influence participation patterns, conversational dominance, and the use of collaborative or competitive discourse strategies (Tannen, 1990), which can shape the quantity and quality of language production during tasks.

The final dimension involves status and role equality. Tasks performed between learners of unequal perceived status may result in unequal participation, reduced risk-taking, or dominant–submissive interaction patterns. Tasks that assign equal roles and symmetrical status relationships tend to promote more balanced turn-taking and collaborative dialogue (Storch, 2002). Overall, Task Condition shapes the interactional context in which cognitive demands and learner abilities operate. Robinson argues that manipulating participation and participant variables, for example, by designing a two-way, convergent task between familiar partners of equal.

2.3.3 *Task Difficulty (Learner Factors)*

Within Robinson's Triadic Componential Framework, Task Difficulty represents the learner-centered dimension that explains why individuals approach and experience the same pedagogical task in different ways. While Task Complexity refers to the cognitive demands designed into a task, and Task Condition refers to its fixed interactional requirements, Task Difficulty arises from how these task features interact with each

learner's cognitive, affective, and motivational characteristics. Because this interaction differs across individuals, perceived difficulty is inherently subjective. Robinson (2011) highlights that, unlike complexity, which teachers can sequence in advance, difficulty is reactive and depends on how well the demands of a task align with each learner's internal resources.

Task Difficulty is typically examined through ability requirement analyses that use psychometric measures and self-report instruments to identify individual differences (Robinson, 2001). One major group of variables involves relatively stable cognitive abilities and aptitudes. Working memory capacity, for example, affects learners' ability to hold and process linguistic information during task performance and may restrict performance even on tasks that are objectively simple (Ellis, 2003). Language learning aptitude, which includes phonetic coding ability, grammatical sensitivity, and associative memory, influences how efficiently learners can notice and internalize new forms. Other cognitive traits, such as reasoning ability, field independence, and intention-reading skills, shape learners' performance on tasks that require logical thinking, pattern identification, or understanding an interlocutor's intended meaning.

A second important group of variables relates to affective and motivational factors that shape learners' perceptions of task difficulty. Willingness to Communicate (WTC), defined as the readiness to initiate communication when one is free to remain silent, strongly affects participation in communicative tasks and varies across situations, people, and contexts (MacIntyre, 1998). Motivation, whether intrinsic or extrinsic, also plays a central role: more motivated learners tend to show greater persistence and deeper engagement with demanding tasks (Anjomshoa & Sadighi, 2015). In contrast, processing anxiety may reduce attentional resources and increase perceived difficulty, while strong emotional regulation skills help learners stay focused and manage challenges effectively (Khajavi, 2021).

The combined influence of these cognitive, affective, and motivational factors produces variation in how learners experience the same task. As Robinson (2001) clarifies, complexity is a feature of tasks, but difficulty is a feature of the relationship between the task and the learner. This distinction is evident in practice: a simple here-and-now descriptive task may overwhelm a learner with low working memory and high anxiety, while a cognitively demanding reasoning task may be manageable for a learner with strong aptitude and high interest in the topic. Task Difficulty also interacts with Task Condition. For instance, a two-way task may feel more difficult for a learner with low willingness to communicate, yet familiarity with a partner may reduce anxiety and facilitate smoother interaction (Plough & Gass, 1999).

Recent research continues to highlight the pedagogical importance of person–task interactions. Motivation plays a central role in how learners interpret and respond to task demands. Learners with higher motivation generally show greater persistence, attention, and willingness to engage with tasks, even when they are challenging. Khajavi (2021) demonstrates that learners' motivational states significantly shape their task engagement and willingness to communicate. In addition, studies in second language motivation show that motivated learners tend to evaluate difficult tasks more positively and are more likely to treat them as learning opportunities rather than obstacles (Lamb, 2017). These findings indicate that perceived task difficulty is not determined solely by the task itself but emerges from the dynamic relationship between learner characteristics and task demands. Understanding task difficulty has important pedagogical implications. Tasks perceived as too demanding may cause frustration and reduce motivation, while tasks that are too easy may lead to boredom and limited learning (Cho, 2018). Effective pedagogy therefore requires careful sequencing and appropriate scaffolding, particularly for tasks involving higher-order cognitive demands.

In sum, Task Difficulty provides a valuable framework for understanding why learners respond differently to the same task under identical conditions. It highlights the importance of considering learner variables, alongside cognitive task design, when planning instruction in task-based language teaching. For researchers, it underscores the need to account for individual differences when interpreting task performance and acquisition outcomes, ensuring that conclusions about task effects consider the mediating role of learner-specific characteristics

2.3.4 *The Role of Gender*

Linguistically, gender and sex are two arguable concepts. According to the Oxford English Dictionary the words gender and sex both have the same concept, the state of being male or female, but they are used in different ways: sex usually refers to biological differences, while gender tends to refer to cultural or social ones (Oxford English Dictionary:2008). Research into gender differences in the field of linguistics began with Robin Lakoff in 1975. In her influential work, Lakoff introduced the concept of female language by identifying several of its distinctive features, which subsequently attracted considerable interest among linguists. She observed that women tend to use more specific and nuanced color terms, while men are more likely to employ stronger language and expletives. For instance, women might say "go to hell," whereas men prefer expressions like "shit" or "damn it" (Lakoff, 1975). Additionally, women often use adjectives to express emotions and favor tag questions to seek confirmation, even when they are confident in their statements. Their speech frequently exhibits a rising intonation, suggesting uncertainty, and they generally demonstrate greater politeness through the use of indirect language. Women are also more likely to adhere to formal grammatical rules and are perceived as less successful in the use of humor, as reflected in the dominance of male comedians in popular culture.

Building on Lakoff's deficit model of women's language, genderlect theory, as proposed by Deborah Tannen (1990), offers a more interactional perspective by framing gender differences as variations in conversational styles or "genderlects" shaped by social and cultural expectations. Tannen argues that men and women often speak different dialects of English, not due to inherent deficits but because of differing goals in communication, men adopt a report talk style focused on establishing status, independence, and hierarchy through direct, competitive exchanges, while women use a rapport talk style emphasizing connection, rapport-building, and empathy through indirect, collaborative speech. For example, in mixed-gender conversations, women may interpret men's interruptions as assertions of dominance, whereas men view them as efficient information-sharing, leading to cross-cultural misunderstandings in language use. This theory shifts the focus from linguistic features alone to the contextual dynamics of interaction, highlighting how genderlects influence topic selection and participation in spoken discourse, particularly in educational settings like EFL classrooms where task-based activities can either exacerbate or bridge these divides.

Gender also plays a significant role in spoken language production, as previous research has shown notable differences in how men and women communicate. Studies indicate that men and women not only speak differently but also engage in conversations about distinct topics. Men are more likely to focus on subjects such as business, politics, sports, and money, often using direct and assertive language aligned with Tannen's report style. In contrast, women tend to emphasize topics related to family, home, and emotions, displaying a more supportive and empathetic communication style consistent with rapport talk (Haas, 1979; Tannen, 1990). Recognizing and understanding these gender-based differences is essential for educators seeking to enhance language learning outcomes, as it enables the development of more inclusive and effective pedagogical strategies that account for genderlect variations.

Topic selection emerges as a critical dimension in gender differences, both in terms of broad topic preferences (Svirina & Ashrapova, 2020) and in the level dynamics of how conversational partners initiate and sustain discussions (Bischoping, 2004). In her study of dyadic interactions, Bischoping observed that women are more likely to introduce relational themes that foster intimacy and emotional sharing, leading to longer and more collaborative exchanges, whereas men favor instrumental topics that emphasize information exchange and problem-solving, often resulting in shorter, more directive turns. For instance, in discussing a shared experience like a movie, women might elaborate on emotional impacts and relationships depicted, while men focus on plot mechanics or factual details. This contrast underscores how gender shapes the initiation and elaboration of topics, aligning with Tannen's (1990) rapport vs. report dichotomy and extending it to practical implications for spoken language production in EFL contexts, where topic alignment can enhance fluency and reduce disfluencies observed in mismatched tasks.

In conclusion, gender plays a significant role in how language is produced, particularly in spoken interaction during task-based learning activities. Understanding these dynamics, including genderlect theory's emphasis on stylistic differences, is crucial for educators in designing inclusive tasks that accommodate the diverse linguistic needs and tendencies of all learners, fostering more effective learning outcomes in speaking-focused classrooms. Educators should consider these gender-based tendencies when designing and implementing language tasks. By recognizing and accommodating different language production strategies such as report versus rapport styles, teachers can create more inclusive and effective learning environments (Fauziati, 2016; Tannen, 1990). This approach ensures that male and female students can leverage their strengths during language tasks.

Future research should continue to explore the impact of gender on language production within TBLT, incorporating genderlect theory to examine how stylistic mismatches affect interactional outcomes. Further studies can provide deeper insights into how these differences influence long-term language development and guide the creation of tailored teaching strategies (Strobach & Woszidlo, 2015). Such research can help educators better meet both male and female learners' unique needs. Ultimately, acknowledging the role of gender in language production, as enriched by genderlect theory, enriches the TBLT framework. It encourages educators to adopt flexible teaching methods that cater to the diverse linguistic strategies of their students. This adaptability is key to fostering effective language acquisition for all learners.

2.4 Measurement of Language Production

In research on language production, measuring complexity, accuracy, and fluency (CAF) is crucial in assessing the development of speaking skills in a second language (L2). Foster and Tonkyn (2003) emphasize that complexity refers to the structural and lexical sophistication used in language production, which can be measured by the number of subordinate clauses, the average length of utterances, and the variation of vocabulary used. Accuracy refers to the grammatical and lexical correctness in language production, which can be measured by the number of grammatical errors per 100 words, the rate of linguistic errors, and adherence to syntactic rules in L2. Meanwhile, fluency relates to the smoothness of speech production and can be measured by the number of pauses, repetitions, and speech rate (words or syllables per minute). The Foster and Tonkyn approach provides a more systematic quantitative method for evaluating speaking skills in the context of language learning, considering contextual factors such as task conditions, the level of interactivity in conversations, and the communication strategies used by language learners. In this study, the Foster and Tonkyn model will serve as the basis for measuring spoken language production by English as a Foreign Language (EFL) learners in Indonesia, allowing an analysis of how task complexity and task conditions affect the quality of language production in terms of complexity, accuracy, and

fluency.

In the analysis of language production, three main dimensions are used to measure learner performance: complexity, accuracy, and fluency (Housen & Kuiken, 2008). These three interconnected aspects provide a comprehensive picture of a person's language proficiency.

2.4.1 Complexity

Language complexity refers to using more varied grammatical structures and higher syntactic complexity in language production (Scontras et al., 2015). Learners with high complexity can effectively manipulate language to express more sophisticated ideas. Research indicates that complexity is influenced by task design and learners' exposure to different linguistic structures (Skehan, 2021). In this study, syntactic complexity was measured by dividing the number of clauses by the number of AS-units (Foster et al., 2000). This ratio reflects the degree of subordination and sentence elaboration in learners' spoken production. Lexical complexity was excluded because the data were oral in nature, and syntactic measures are more reliable indicators of structural development in speech (Housen & Kuiken, 2009). Lexical complexity was not measured in this study because the focus was on structural variation rather than lexical range. According to Ellis and Barkhuizen (2005), syntactic measures are more sensitive to task-induced cognitive load, whereas lexical complexity often reflects individual vocabulary knowledge rather than task effects. In addition, syntactic complexity provides a clearer reflection of learners' ability to manage clausal embedding and subordination under different task conditions, which aligns with the aim of this study to examine how task complexity influences spoken performance. It is important to note that the desirable level of syntactic complexity is context-dependent, influenced by both task type and learner proficiency.

2.4.2 Accuracy

Accuracy measures learners' adherence to grammatical rules, including syntax, vocabulary, and spelling (Foster and Tonkyn,2003). High accuracy indicates a deep understanding of language rules and the ability to apply them consistently. Studies show that accuracy improves with explicit instruction and corrective feedback. However, excessive focus on accuracy can sometimes hinder fluency, as learners may become overly concerned with making mistakes(Suzuki & Kormos, 2020). Each AS-unit was examined for the presence or absence of grammatical errors. Only AS-units completely free from morphological, syntactic, or word-order errors were classified as error-free. The same operational definitions and examples were provided during rater training to ensure inter-rater consistency and reliability.

2.4.3 Fluency

Fluency communicates smoothly, with minimal hesitation or self-correction (Tavakoli & Wright, 2021; Foster and Tonkyn, 2003). Fluent learners tend to speak or write more confidently and effectively. Research suggests that fluency is closely linked to practice and exposure to authentic language use (Lambert & Kormos, 2020).

In task-based performance research, fluency is often measured through temporal variables, particularly speech rate. Two widely used measures derived from Lennon's (1990) fluency framework are Speech Rate A and Speech Rate B, both of which capture how efficiently learners produce spoken language. Speech Rate A calculates the total number of syllables produced in unpruned speech, including repetitions, self-repairs, false starts, and other dysfluency markers, divided by the total task time and multiplied by 60. This measure reflects learners' overall production processes, including planning and monitoring behaviors. In contrast, Speech Rate B (pruned speech) excludes repetitions, self-repairs, false starts, and L1 asides, thereby providing a more precise index of articulatory speed (Yuan & Ellis, 2003; Gilabert, 2005). Because it removes dysfluency markers, Speech Rate

B captures learners' actual speed of fluent delivery rather than the time spent planning or resolving breakdowns.

These two measures are widely regarded as comprehensive fluency indicators because they incorporate both temporal aspects and the presence or absence of dysfluency features (Ellis, 2005), and have been consistently used in task-based studies such as Yuan and Ellis (2003), Gilabert (2005) and Ahmadian and Tavakoli (2010). Given its focus on pruned, fluent speech, the present study adopts Speech Rate B as the primary fluency measure, as it more accurately captures learners' actual fluency rather than their planning behavior.

To operationalize CAF within the present framework, the following measures, adapted from Skehan (2009), Foster and Tonkyn (2003), and Housen et al. (2022), are employed and summarized in the table below.

Tabel 4. CAF Measures

CAF Measures		
Complexity	**Accuracy**	**Fluency**
Syntactic Complexity	% of Error-Free AS-Units	Speech Rate B (calculated as pruned syllables per minute, excluding repetitions, repairs, false start and L1 asides)

In conclusion, complexity, accuracy, and fluency are three fundamental pillars in measuring language production. A deep understanding of these aspects not only aids in evaluating language proficiency but also in designing effective learning programs and enriching linguistic research (Ellis & Barkhuizen, 2020).

2.5 Theoretical Assumptions

This research is grounded in several key theoretical assumption, it is primarily assumed that the cognitive complexity of a task directly impacts learners oral language production. Tasks intentionally designed with varying levels of complexity, through the manipulation of resource-directing variables such as the number of elements or reasoning demands, are expected to stimulate learners to utilize more elaborate and varied linguistic structures. This aligns with established frameworks such as Robinson's Cognition Hypothesis, which

posits that increasing certain types of task complexity can guide learners towards more complex and accurate language use. Furthermore, it is assumed that task conditions mediate performance outcomes. In the context of this study, participant gender is considered a significant aspect of task condition as a participant variable. It is posited that gender may interact with task complexity or topic preference, thereby affecting various facets of spoken language production

A fundamental premise is that learners spoken language can be reliably and validly assessed via Complexity, Accuracy, and Fluency (CAF) measures, which are accepted as providing quantifiable and meaningful indicators of language proficiency and development. These combined assumptions thus form the theoretical framework for investigating the impact of task complexity and gender on the oral production of Indonesian EFL learners.

2.6 Hypothesis

Based on the theoretical assumption above, the researcher has her hypothesis as followed:

Research Question 2

H1: The manipulated tasks with different types of task complexity with different topics of interest generate statistically significant differences in spoken language productions.

H0: None of the manipulated tasks with different types of task complexity with different topics of interest generate statistically significant differences in spoken language productions.

CHAPTER III
RESEARCH METHOD

This chapter discusses the methodology adopted in the present study. It then describes the research design, participants, instruments, research procedures, data analysis and hypothesis testing.

3.1 Research Design

This study adopted an exploratory mixed-methods design to investigate how gender and topic preference shape Indonesian EFL learners' spoken language production under an identical task-complexity condition. Following Plano Clark' (2017) sequential model, the first phase collected qualitative data through a questionnaire to identify male and female learners' preferred speaking topics. These findings informed the development of two monologue tasks in the quantitative phase, both designed using Robinson's framework with complex resource-directing demands and simple resource-dispersing demands. In the second phase, quantitative data were collected to examine how gender, as the task-condition variable, influenced learners' spoken performance, which was evaluated in terms of Complexity, Accuracy, and Fluency (CAF). By integrating both phases, the study offers a comprehensive account of how gender interacts with topic preference and task design to shape learners' spoken language outcomes.

3.2 Data (Variables)

This research investigated the influence of two independent variables on one dependent variable, guided by Robinson's Triadic Componential Framework.

These variables were defined as follows:

1. Independent Variable

   There were two independent variables in this research namely Task Complexity and Task Condition.

a. Task Complexity (Cognitive Factors)

This independent variable referred to the inherent cognitive demands associated with the speaking tasks assigned to the learners. Task complexity was manipulated by varying resource-directing dimensions, such as the number of elements involved (- few elements) and the reasoning demands required (+ reasoning demands). The manipulation of these cognitive dimensions of task complexity was expected to influence learners spoken language production in terms of CAF.

b. Task Condition (Interactive Factors - focusing on Participant Variables)

This independent variable referred to specific aspects of the interactive setting in which the tasks were performed, focusing particularly on participant variables, specifically gender. This study aimed to investigate whether the gender of the learners had an effect on their spoken language production when performing tasks of varying complexity, and whether gender interacted with task complexity. This approach aligned with Robinson's Triadic Componential Framework, which categorizes gender under participant variables within Task Conditions

2. Dependent Variable

The dependent variable of this research was Spoken Language Production in term of CAF. This variable represented the main outcome of interest in the study, encompassing the quality and effectiveness of the learners' spoken performance during the tasks. Reflecting the learners' communicative competence under different task complexity levels, the students' spoken task performances were analyzed to examine the effects of task complexity and task condition on their CAF measures.

3.3 Source of Data

The subjects of this study were 60 Indonesian EFL learners enrolled in the 12th grade at MAN 1 East Lampung. The sample consisted of 30 male and 30 female students, providing balanced gender representation for examining topic preference and spoken language performance. This grade level was selected because students at this stage are expected to have developed sufficient English

proficiency to participate effectively in task-based language production activities.

Participants were purposively selected based on their English achievement, using their report card records to ensure that they possessed adequate linguistic readiness for the speaking tasks. All students had studied English for more than two years and were familiar with various speaking activities integrated into the curriculum. Selecting participants from the same school and grade level also ensured a relatively homogeneous group in terms of instructional background and exposure to the standardized national curriculum.

3.4 Research Instrument

This research had 2 instruments. These instruments were as follows:

a. Topic Preference Questionnaire.

The first instrument was a Topic Preference Questionnaire administered directly to male and female students. This instrument addressed the first research question concerning the relationship between gender and students' topic choices for English-speaking tasks. The questionnaire collected basic demographic data and asked respondents to select one preferred topic for discussion from a list reflecting common adolescent interests or to propose an alternative if desired. Each participant wrote a brief explanation of their choice. The questionnaire was designed to identify students' preferred speaking topics and to inform the assignment of speaking topics for the main task, ensuring alignment with gender-based interests. It was developed based on recent research on gender and topic preferences in English language learning. Topic options were adapted from Svirina and Ashrapova (2020) and refined after review by an experienced English language expert to ensure clarity, relevance, and appropriateness for senior high school students. This adaptation tailored the instrument to the specific educational context, supporting content validity and practical effectiveness. The topics represented broad areas such as sports, technology, cloth or fashion, entertainment, food, family and friends.

The Topic Preference Questionnaire was analyzed to determine the most preferred topic within each gender group. The selections made by male and female participants were quantified to obtain their raw frequency distributions. These distributions were subsequently compared to identify the single highest-ranked topic for male students and for female students. This analytical procedure directly addressed the first research question concerning learners' topic preferences.

b. Speaking Tasks.

Spoken language production of the students was acquired through the completion of a task with different topics assigned to each gender. The tasks consisted of two types, designed to be complex by manipulating both resource-directing and resource-dispersing dimensions. The model of these tasks was described as follows:

Table 5. The Model of Tasks

Task	**Task Complexity Variables**		**Gender Topic Preferences**
	Complex Resource Directing	**Simple Resource Dispersing**	
Task 1	(-) Few elements (+) Reasoning Demands	(+) Planning Time (+) Single Task (+) Prior Knowledge	Female Preferred Topic
Task 2	(-) Few elements (+) Reasoning Demands	(+) Planning Time (+) Single Task (+) Prior Knowledge	Male Preferred Topic

In accordance with the model above, both tasks applied the same variables for task complexity. They were complex in resource-directing dimensions because they required reasoning and involved many elements, and they were simple in resource-dispersing dimensions since students were given enough planning time, needed to perform only one task at a time, and could use prior knowledge. The only difference between the two tasks was the topic, which matched either female or male student preferences. The researcher intentionally avoided prompts involving distant or past events, as these could place an excessive cognitive load on learners. Recalling and

describing non-immediate experiences demands more memory retrieval and discourse reconstruction, which may reduce learners' capacity to attend to linguistic form and fluency during real-time performance. performance.

3.5 Validity and Reliability

Several procedures were conducted systematically to ensure the validity and reliability of the two research instruments. Therefore, the validity and reliability of each instrument are discussed separately below.

3.5.1 Questionnaire Validity and Reliability

a. Questionnaire validity

- Content validity

  The content validity of the questionnaire was established through expert judgment from two experienced language teaching professionals. They reviewed the instrument to determine whether the questionnaire items adequately represented learners' topic preferences for English-speaking tasks and ensured that all relevant topics and instructions were included and appropriate for the target population.

- Construct Validity

  Construct validity was ensured by grounding the questionnaire in the work of Svirina and Ashrapova (2020), whose study specifically examined gender-based topic preferences in English language learning. Their findings provided a clear empirical framework for identifying common topic categories preferred by male and female learners. This alignment strengthened the validity of the questionnaire as an appropriate tool for identifying gender-related topic choices in English-speaking tasks.

- Face Validity

  Face validity was ensured by piloting the questionnaire with a small group of non-participant students who shared similar

characteristics to the main sample. The pilot study helped identify ambiguous items and confirmed that the questionnaire was clear and understandable for the intended participants.

b. Questionnaire reliability

The reliability of the questionnaire was ensured through standardized administration procedures and pilot testing. All participants received identical instructions and response options to maintain consistency during administration. Feedback obtained from the pilot testing informed several minor revisions to improve clarity and item interpretation. These procedures collectively supported the internal consistency and reliability of the questionnaire as a data collection instrument.

3.5.2 Speaking Tasks Validity and Reliability

a. Speaking Tasks Validity

- Content validity

Content validity was ensured by aligning the speaking tasks with the English Curriculum for Grade XII (Merdeka Curriculum, phase F), particularly the descriptive and analytical exposition text types practiced at this level. The researcher designed the tasks to require comparing elements, expressing preferences, and justifying opinions, consistent with the curriculum's learning outcomes. Two expert validators reviewed the tasks to confirm that the content, language functions, and expected outcomes were consistent with the English Learning Outcomes for Phase F, ensuring the tasks accurately represented curriculum-based communicative goals. Their feedback, assessed via a validation checklist, led to minor revisions in task instructions, ensuring the tasks accurately measured the intended communicative skills.

- Construct validity

  Construct validity was established by aligning task design with Robinson's (2007) Triadic Componential Framework. Task Complexity was operationalized through resource-directing variables: (+) causal reasoning demands and (-) few elements, creating systematic cognitive load variations. Task Condition was represented through the participant variable of gender, treated as a fixed factor influencing performance.

- Face Validity

  Face validity was ensured by making the speaking tasks look clear and meaningful to the students. The topics were adapted from well-known sources and adjusted to match students' interests and experiences in daily life. The tasks used topics based on students' most preferred themes to make them feel more motivated and confident when speaking. The prompts were reviewed by experts to make sure the language was suitable for Grade XII students and that the tasks reflected real communication situations. This made the tasks appear valid and practical to both teachers and students.

b. Speaking Task Reliability

Reliability was ensured by using standardized procedures in administering, transcribing, and scoring the speaking tasks. Since one of the focuses of this research was to examine the effect of task complexity on students' spoken performance, an area involving subjective assessment, the researcher employed an inter-rater scoring procedure to obtain more reliable results. Two raters were involved in the scoring process, the researcher herself and an English teacher at MAN 1 East Lampung. After both raters completed their assessments, the inter-rater reliability of the scores was statistically tested using SPSS to verify consistency across raters. To establish statistical reliability, the Intraclass Correlation Coefficient (ICC) was calculated using a two-way random effects

model with an absolute agreement definition.

The interpretation of ICC values follows the guideline proposed by Koo and Li (2016)

ICC Range	Reliability Interpretation
< 0.50	Poor
0.50 – 0.75	Moderate
0.75 – 0.90	Good
> 0.90	Excellent

The results demonstrated very high reliability across all three measures: Complexity with an ICC value of 0.998 and a significance level below 0.001, Accuracy with an ICC value of 0.999 and a significance level below 0.001, and Fluency with an ICC value of 1.000 and a significance level below 0.001. All coefficients exceeded the minimum acceptable threshold of 0.75, confirming good to excellent inter-rater reliability. According to the criteria proposed by Koo and Li, ICC values above 0.90 signify excellent consistency, indicating that both raters evaluated the students' spoken performance in a highly consistent manner.

3.5.3 Normality of Speaking Test

The normality distribution test is a test to measure whether our data has a normal distribution or not. The data gained in this research was statistically analyzed by using SPSS.

The result for normality test for CAF measurement of two type of tasks is as follows:

Table 6. Normality Test for CAF Measurement of Tasks by Gender

Variable	Gender	Shapiro–Wilk Statistic	df	Sig. (p-value)
Complexity	Male	.990	30	.010
	Female	.972	30	.028
Accuracy	Male	.984	30	.235
	Female	.954	30	.016
Fluency	Male	.974	30	.659
	Female	.968	30	.494

The table above presents the normality test results for each CAF measurement. In the Shapiro–Wilk test, data are considered normally distributed when the significance value (Sig.) is greater than 0.05 and not normally distributed when Sig. is below 0.05. For the Complexity measure, the Shapiro–Wilk significance values were 0.010 for males and 0.028 for females. Since both values are below 0.05, the Complexity scores did not meet the normality assumption. For the Accuracy measure, males obtained a significance value of 0.235, which is above 0.05 and therefore normally distributed. However, females had a significance value of 0.016, which is below 0.05, indicating non-normal distribution. Thus, normality for Accuracy was only partially fulfilled. For the Fluency measure, both males 0.659 and females 0.494 had significance values above 0.05, demonstrating that the Fluency data for both groups were normally distributed.

Despite Fluency meeting the normality criteria, the Mann–Whitney U test was applied uniformly to all CAF variables. This approach ensured methodological consistency across analyses and avoided mixing parametric and non-parametric tests. Because normality was violated in two of the three CAF dimensions (Complexity and Accuracy), using a single non-parametric procedure minimized potential interpretive bias and provided a more coherent analytical framework. This decision aligns with Azadi & Gholami (2013) who recommend using non-parametric tests when data show partial or inconsistent normality.

3.6 Data Collecting Techniques

Two types of data collection techniques were employed in this study, a questionnaire and speaking tasks. The questionnaire, adapted from Svirina and Ashrapova (2020), was used to identify students' topic preferences, addressing the first research question concerning the relationship between gender and topic interest. It yielded primarily qualitative data through students' selection of their preferred topics and their brief written explanations for these choices. Each student was asked to select one topic from the provided list by giving a tick (✓)

to indicate their choice and to write a short justification explaining the reason for their preference.

The speaking tasks were implemented to elicit students' spoken performance for analysis. These monologic tasks were designed based on students' topic preferences identified from the questionnaire and varied in complexity through the manipulation of resource-directing and resource-dispersing factors, following the principles of task-based design. Each student's spoken performance was audio-recorded, transcribed, and analyzed for Complexity, Accuracy, and Fluency (CAF) measures.

3.7 Research Procedure

This study was conducted in two main stages, focusing on student active participation based on their topic preferences and spoken English production (EFL context). The detailed procedures were carried out as follows:

3.7.1 Qualitative Procedure

Stage 1 aimed to identify students topic preferences and their relationship to gender.

1. Providing a List of Topics

   A list of topics was presented to all participating students (30 male and 30 female) by a questionnaire.

2. Topic Selection by Students

   Each student individually selected their most preferred topic from the provided list by giving a tick (✓) next to their chosen topic and wrote a brief explanation for their choice.

3. Preference Analysis

   Responses were analyzed qualitatively to identify trends differences in topic preferences between male and female students. The results were than used to assign gender-tailored topics for subsequent tasks.

3.7.2 Quantitative Procedure

Stage 2 examined students' spoken language production under varying task complexity and gender-based topic conditions. Each of the 60

participating students performed one monologic speaking task only.

Task 1 was a monologue on a topic representative of those generally preferred by male students (as identified in Stage 1) and task 2 was a monologue on a topic representative of those generally preferred by female students (as identified in Stage 1).

a. Planning Stage

After receiving the assigned topic, students were given time to plan what they would talk about. The planning process involved generating ideas, organizing the structure of their talk, and selecting appropriate vocabulary. The planning time was limited to approximately 10 minutes and standardized for all participants. A 10 minute pre-task planning period has been widely supported by (Ellis, 2009) as an effective duration for enhancing learners' oral performance in task-based language learning, as it allows learners to conceptualize their ideas, organize their discourse, and select appropriate linguistic forms before performing the task.

b. Transcription and Coding

The audio-recorded monologues were transcribed manually, capturing all spoken words along with pauses, fillers, and hesitation markers that reflected real-time processing. Following pruned-speech criteria, repetitions, self-repairs, false starts, and L1 asides were excluded from the transcription to provide a clearer indication of learners' actual articulatory speed. Audacity was used to identify the real-time duration of each speech sample. The transcripts were then reviewed and checked for accuracy to ensure that the data were reliable and complete. The transcribed data were subsequently analyzed using the Complexity, Accuracy, and Fluency (CAF) framework, as operationalized by Foster and Tonkyn (2003).

c. Data Analysis

The analysis was conducted to:

1) Measure the quality of spoken language production based on task complexity and task conditions.
2) Examine differences in spoken performance between male and female students based on their chosen topics.

To achieve these objectives, the collected data were analyzed by measuring three main aspects of spoken language performance: complexity, accuracy, and fluency (CAF). The scoring procedures for all CAF measures, including the complete AS-unit and clause identification rules adapted from Foster et al. (2000).

The detailed procedures were as follows:

1. Complexity

Syntactic complexity was calculated by counting the number of clauses in each AS-unit (Analysis of Speech unit). Each AS-unit was separated using vertical lines (‖), and each clause was marked with the symbol "C".

Example:

I, um, want to tell you about my activities last weekend. (C) ‖ On Saturday, I, uh, went to the park. (C) ‖ I played, um, soccer with my friends. (C) ‖ Then, we, uh, had a picnic. (C) ‖ On Sunday, I, um, stayed home and watched movies. (C) ‖

Total AS-units: 5

Total clauses: 5

Formula:

Syntactic Complexity = $\frac{Total\ Clauses}{Total\ AS\_Unit}$

Syntactic Complexity = $\frac{5}{5} = 1.00$

The researcher analyzes every sentence in the transcription above, and the complexity value is 1.00.

2. Accuracy

Accuracy was measured by calculating the percentage of error-

free AS-units out of the total AS-units in a transcription. By calculating the ratio of the number of error-free AS unit to the total number of AS unit, accuracy is calculated (Mahpul , 2014)

Example:

I, um, want to tell you about my activities last weekend. On Saturday, I, uh, **go** to the park. I played, um, soccer with my friends. Then, we, uh, had a picnic On Sunday, I, um, stayed home and watched movies.

Total AS-units: 5

Error-free AS-units: 4 (The second AS-unit has an error: go should be went.)

Formula:

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{Number\ of\ Error - Free\ AS - Unit}}{\boldsymbol{Number\ of\ AS - units}}\ \boldsymbol{X\ 100}$$

$$\boldsymbol{Accuracy} = \frac{\boldsymbol{4}}{\boldsymbol{5}} \times \boldsymbol{100} = \boldsymbol{80.00}$$

The researcher analyzes every sentence in the transcription above, and the accuracy value is 80.00.

3. Fluency

Fluency was measured using Speech Rate B, calculated as the number of syllables spoken per minute. Pauses and hesitation fillers such as “emmm” and “eee” were included in the total duration, as they reflect planning and processing time, but they were not counted in the syllable total. In line with the definition of Speech Rate B, syllables produced in repetitions, self-corrections, false starts, and any Indonesian or local-language words were also excluded from the syllable count. This measure was selected because it more accurately represents learners’ actual fluency by focusing on pruned, meaningful speech rather than planning behaviors.

The example and the specific rules used for identifying and excluding syllables are provided below.

1. *Ing* forms such as, doing, saying, etc., counted as two syllables.
2. The constructions such as, isn't, doesn't, didn't, were calculated as two syllables.
3. Syllables in Indonesian words were not counted
4. Epenthesis (insertion of sounds in the middle of words) does not count as a syllable, e.g., speak /səpi:k/, instead of /spi:k/.
5. Past /ed/ form was not regarded as a syllable (e.g., looked). But past /ed/ was calculated as a syllable for the verbs ending with *t* or *d* (e.g., "wanted", "landed"), each counted as two syllables. (Mahpul, 2014)

Example:

I (1) um …want (1) to (1) tell (1) you (1) about (2) my (1) activities (4) last (1) weekend (2) umm ... On (1) Saturday (3), I (1) uhh…went (1) to (1) the (1) park (1)...I (1) played (1) umm...soccer (2) with (1) my (1) friends (1)...Then (1), we (1) uh… had (1) a (1) picnic (2)..On (1) Sunday (3), I (1) um… stayed (1) home (1) and (1) watched (1) movies (2).

Total syllables: 48

Total time (including pauses): 30 seconds

Formula

$$Fluency = \frac{Number\ of\ Syllables}{Total\ Seconds} \times 60$$

$$Fluency = \frac{48}{30} \times 60 = 96.00$$

The researcher analyzes every sentence in the transcription above, and the fluency value is 96.00

Based on Housen and Kuiken (2009), CAF measures are intended for relative comparison, such as comparing tasks, comparing groups, or measuring improvement, rather than absolute scoring. In line with this perspective, once the CAF scores were

calculated, the researcher constructed a temporary scoring range to facilitate clearer interpretation of students' performance. Although not formally standardized, this range provides a practical classification of learners' abilities. The categories used in this study are presented in the tables below.

Table 7. Classification of Complexity Performances

Level	Range	Interpretation and Code
Low	≤ 1.89	More Simple Clauses (MSC)
High	> 1.89	More Complex Clauses (MCC)

Note. The threshold of 1.89 is based on the median score (N = 60).

More Simple Clauses (MSC) refers to spoken production characterized by a low clause-to-AS-unit ratio, where learners typically produce AS-units with only one clause and limited structural development. Conversely, More Complex Clauses (MCC) represents a higher clause-to-AS-unit ratio, showing that learners generated AS-units with multiple clauses and displayed more advanced syntactic elaboration.

Table 8. Classification of Accuracy Performances

Level	Range	Interpretation and Code
Low	≤ 20	More Error AS-Unit (MEA)
High	> 20	More Error Free AS-Unit (MEF)

Note. The threshold of 20 is based on the median score (N = 60).

More Error AS-Unit (MEA) represents a lower proportion of error-free AS-units, indicating that learners produced fewer grammatically accurate units during the task. Conversely, More Error Free AS-Unit (MEF) reflects a higher proportion of error-free AS-units, suggesting greater grammatical accuracy and more consistent control of linguistic forms.

Table 9. Classification of Fluency Performances

Level	Range	Description and Code
Low	≤ 127.32	Less Fluent (LF)
High	> 127.32	More Fluent (MF)

Note. The threshold of 127.32 is based on the median score (N = 60).

Learners categorized as Less Fluent (LF) produce fewer syllables per minute, indicating a slower speech rate, more frequent pauses, and reduced temporal flow. By contrast, More Fluent (MF) learners produce more syllables per minute, reflecting faster and smoother speech with fewer pauses and greater ease in maintaining continuous articulation.

3.8 Hyphothesis Testing

Hypothesis testing was conducted to determine the possible outcomes of the research. This study aimed to examine the effects of task complexity, integrating the resource-directing and resource-dispersing dimensions, on learners' speaking performance and topic preferences in English speaking tasks.

The following hypotheses are formulated and tested:

Research Question 2

H1: The manipulated tasks with different types of task complexity with different topics of interest will generate statistically significant differences in spoken language productions.

H0: None of the manipulated tasks with different types of task complexity with different topics of interest will generate statistically significant differences in spoken language productions.

To test these hypotheses, the Mann–Whitney U test was used to determine the statistical significance of differences in CAF scores between male and female students.

CHAPTER V
CONCLUSIONS AND SUGGESTIONS

This chapter describes the conclusions of the research and also the suggestions for Teaching English Foreign Language (TEFL) and for further research.

5.1 Conclusions

This study investigated the effects of task complexity and task condition (gender) on Indonesian EFL learners' spoken performance. The study also explored topic preferences, where male learners predominantly chose sports, while female learners preferred movies. The analysis of spoken performance was conducted based on tasks that were uniform in complexity but revealed these underlying preference patterns.

The second phase of the study examined the influence of task complexity and gender on learners' oral performance, operationalized through the dimensions of Complexity, Accuracy, and Fluency (CAF). A clear pattern emerged female learners produced more syntactically complex utterances and demonstrated higher fluency than their male peers, while both groups exhibited comparable levels of grammatical accuracy.

These findings indicate that male and female learners may adopt different strategies when confronted with cognitively demanding tasks. The greater linguistic complexity displayed by female learners aligns with Robinson's (2003) Cognition Hypothesis, which argues that increases in task complexity can elicit more sophisticated linguistic production. Notably, their fluency did not decline alongside the rise in complexity, contrary to common expectations, suggesting a particularly effective management of cognitive resources.

Conversely, the absence of gender differences in accuracy lends support to Skehan's (1998) Trade-Off Hypothesis, which posits that learners have limited

attentional capacity and may struggle to maintain grammatical precision when simultaneously attending to fluency and complexity under high cognitive load.

A plausible explanation for the female learners' comparatively stronger performance may lie in evidence from the first phase of the study, which showed that they were more strongly engaged with the topics used in the speaking tasks. Taken together, these findings underscore the importance of considering not only cognitive constraints but also individual factors such as gender and topic preference when interpreting learner performance in task-based speaking activities.

In sum, the findings demonstrate that spoken performance is shaped not only by the specific cognitive demands of a task, such as the need for causal reasoning and the integration of multiple elements, but also by social factors, particularly gender and its associated topic preferences. For female learners, familiarity and interest in expressive topics likely helped them handle the cognitive load of these demands, making it easier to produce more fluent and complex speech. These results underscore that the relationship between the design of a task and a learner's output is dynamically influenced by individual and social factors.

5.2 Suggestions

5.2.1. For Teaching English Foreign Language (TEFL)

1. The results indicate that cognitively demanding tasks can enhance learners' fluency and syntactic complexity. TEFL programs should therefore incorporate task-based principles into curriculum design, ensuring that speaking tasks include reasoning, comparison, and decision-making elements.
2. Gender-related patterns found in the study demonstrate the importance of aligning tasks with learners' cognitive tendencies and communicative orientations. TEFL frameworks should account for individual differences, such as gender, learning preferences, and affective factors, when designing task types and sequencing.

3. Since topic preferences significantly influenced students' engagement and performance, TEFL material developers are encouraged to select topics that reflect learners' interests. Incorporating enjoyable, relatable, and meaningful themes in speaking tasks can improve motivation and willingness to communicate.
4. The findings highlight the importance of evaluating spoken language through multiple dimensions: Complexity, Accuracy, and Fluency (CAF). TEFL assessment practices should therefore move beyond accuracy-only evaluation and adopt multidimensional performance indicators to ensure a more holistic understanding of learners' communicative abilities.
5. Sequencing tasks from simple to complex, as supported by the Cognition Hypothesis, can help learners develop gradually. TEFL curriculum planners should integrate clear progression patterns to support learners' cognitive and linguistic development.

5.2.2. For Further Research

1. Future studies should involve a larger and more diverse sample from multiple schools or regions to enhance the generalizability of the findings beyond the context of MAN 1 East Lampung.
2. Further research should examine dialogic or interactive speaking tasks to explore how gender and task complexity influence interactional features such as negotiation of meaning, turn-taking patterns, and collaborative speech production.
3. Subsequent studies should complement CAF analysis with measures of conceptual content, such as idea density, topical relevance, argument coherence, or the use of supporting details. This would provide a more holistic assessment of oral performance that accounts for both linguistic form and conceptual substance.
4. Future research should explore whether aligning tasks with gender-based topic preferences contribute to sustained improvements in fluency, complexity, and overall communicative competence over longer periods, rather than only immediate task performance.

5.3 Limitations

Several limitations should be acknowledged in interpreting the findings of this study.

1. The participant sample was limited to 60 students from a single senior high school, which may restrict the generalizability of the results. Including a larger and more diverse group of learners from different schools or regions would strengthen the external validity of future studies.
2. The study also examined only monologic speaking tasks, which means the findings may not fully reflect how gender and task complexity influence performance during interactive or dialogic communication. Future research could incorporate pair or group speaking tasks to explore interactional features such as turn-taking and negotiation of meaning.
3. The analysis relied exclusively on CAF (Complexity, Accuracy, Fluency) measures to assess oral performance. While CAF provides a robust profile of linguistic production, it does not account for the quality, coherence, or richness of the ideas expressed. The CAF framework is designed to quantify linguistic form and delivery, not to evaluate the conceptual depth, topical relevance, or logical coherence of the message itself.

REFERENCES

Abrar, M. (2018). If our English isn't a language , what is it ? Indonesian EFL Student Teachers ' Challenges Speaking English. *The Qualitative Report*, *23*, 129–145.

Agustina, V. (2017). *The use of task complexity in spoken performance* [Master's thesis, Universitas Lampung]. Digilib UNILA. **http://digilib.unila.ac.id/2927**

Al-Shibel, A. G. (2021). Gender differences in classroom interactions and preferences. *Journal of Language and Linguistic Studies*, *17*(1), 534–552.

Albino, G. (2017). Improving speaking fluency in a task-based language teaching approach :The Case of EFL Learners at PUNIV-Cazenga. *SAGE Open*, *7*(2). https://doi.org/10.1177/2158244017691077

Anjomshoa, L., & Sadighi, F. (2015). The Importance of Motivation in Second Language Acquisition. *International Journal on Studies in English Language and Literature*, *3*(2), 126–137.

Awwad, A. A. (2017). *The effects of task complexity manipulated by intentional reasoning demands on second language learners ' speech performance : Interaction with language proficiency and working memory (*Doctoral dissertation, University of Barcelona).

Ayu, M. et all. (2020). The Use of Task Complexity in Written Data, *Journal of English Language Teaching and Linguictics*. 274–282.

Azadi, G., & Gholami, R. (2013). An Analysis of Errors in English Writing Made by Chinese and Korean University Students. *Theory and Practice in Language Studies*, *3*(8), 1299–1495.

Bischoping, K. (1993). Gender Differences in Conversation Topics *, 1922-19901. 28.Social Forces*.

Brown, G. (2014). Understanding Spoken Language (Vol. 12, Issue 3).*TESOL Quarteely,*221-238

Bygate, M. (1997). Speaking. *In Languagge Teaching*: *A scheme for teacher education.* Oxford University Press 14–20.

Cameron, D. (2005). Language, gender, and sexuality: Current issues and new directions. *Applied Linguistics, 26*(4), 482–502.https://doi.org/1 0.1093/applin/ami027

Chen, A. (2012). Application of Gender Difference and Topic Preference to Promote Students' Motivation for Online EFL Learning. *The EuroCALL Review, 20*(1), 47. https://doi.org/10.4995/eurocall.2012.16040

Chen, K. T. (2019). The Effects of Technology-Mediated TBLT on Enhancing the Speaking Abilities of University Students in a Collaborative EFL Learning Environment. *Jurnal of Asia TEFL*, 1–22.

Cho, M. (2018). Task complexity , modality , and working memory in L2 task performance. *System*, *72*, 85–98. https://doi.org/10.1016/j.system.2017.10.010

Coates. (2017). *Women , Men and Language: A sociolinguistic account of gender differences in language.* Routledge.

Córdoba Zúñiga, E. (2016). Implementing Task-Based Language Teaching to Integrate Language Skills in an EFL Program at a Colombian University. *PROFILE Issues in Teachers' Professional Development*, *18*(2), 13. https://doi.org/10.15446/profile.v18n2.49754

Ellis. (2009). The Differential Effects of Three Types of Task Planning on the Fluency , Complexity , and Accuracy in L2 Oral Production. *Applied Linguistics*, 474–509. https://doi.org/10.1093/applin/amp042

Ellis, R. (1993). *The study of second language acquisition.* Oxford University Press.

Ellis, R. (2003). *Task-based Language Learning and Teaching*. Oxford University Press.

Erlam, R. (2016). I'm still not sure what a task is*:* Teachers designing language tasks. *Language Teaching Research*, *20*(3), 279–299. https://doi.org/10.1177/1362168814566087

Fitriani, I., & Wirza, Y. (2018). Task-Based Language Teaching in Promoting Students ' Speaking Fluency. *Lingua, Jurnal Bahasa & Sastra*, *19*(1), 19–26.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring Spoken Language : A Unit for All Reasons. *Applied Linguistics*, 354–375.

Gass, S., Mackey, A., & Ross-feldman, L. (2005). Task-Based Interactions in Classroom and Laboratory Settings. *Language Learning,* 575–611.

Ghaderi, E., Rouhi, A., Reza, A., & Tabrizi, N. (2022). Writing Task Complexity , Task Condition , and the Efficacy of Feedback. *Journal of Language and Education,8*(4), 73–87.

Haas, A. (1979). Male and female spoken language differences: Stereotypes and evidence. *Psychological Bulletin*, *86*(3), 616–626. https://doi.org/10.1037/0033-2909.86.3.616

Harmer, J. (2007). *The Practice of English Languag Teaching* (4th ed.). Pearson Longman

Housen, K., & Kuiken, F. (2008). Complexity, Accuracy and Fluency in Second Language Acquisition. *Applied Linguistics*, 1147–1151.

Huang, Z. (2023). *Exploring Gender Differences in Language Learning*. I*n Proceedings of the 2nd International Conference on Education, Language*

and Art (ICELA 2022). (pp. 543–554). Atlantis Press. https://doi.org/10.2991/978-2-38476-004-6

Ibrahim, J. (2006). The Implementation of EMI (English Medium Instruction) in Indonesian Universities: Its Opportunities, its Threats, its Problems, and its Possible Solutions. *K@ Ta*, *3*(2), 121–137.

Ismail, L., & Samad, A. (2017). Effects of Task Complexity Variables on Complexity , Accuracy and Fluency of Second Language Production : A Critical Review. *International Journal of Leraning*, *9*(4). https://doi.org/10.5296/ijl.v9i4.11609

Ji, Y., & Pham, T. (2020). Implementing task-based language teaching (TBLT) to teach grammar in English classes in China: using design-based research to explore challenges and strategies. *Innovation in Language Learning and Teaching*, *14*(2), 164–177. https://doi.org/10.1080/17501229.2018.1545021

Kamel, M., Mohammad, M., & Alwi, N. A. (2019). The Impact of Task Complexity on EFL Learners ' Writing Production across Gender. *International Journal of Engineering and Advanced Technology, 8958*(5), 1218–1220. https://doi.org/10.35940/ijeat.E1172.0585C19

Khajavi, G. (2021). A Closer Look at Grit and Language Mindset as Predictors of Foreign Language Achievement. *Journal of Multilingual and Multicultural development, 43*, 379–402.

Kobayashi, Y. (2010). The Role of Gender in Foreign Language Learning Attitudes : Japanese female students ' attitudes towards English learning. *Gender and Education,* https://doi.org/10.1080/09540250220133021

Kumayas, T. A., & Lengkoan, F. (2023). The Challenges of Teaching Grammar at the University Level : Learning From the Experience of English Lecturer. *Journal on Education*,*11*(1), 98–105.

Lakoff, R. (1975). *Language and woman's place.* Harper & Row. *2*(1), 45–80.

Lamb, M. (2017). The Motivational Dimension of Language Teaching. Language Teaching, 301–346. https://doi.org/10.1017/S0261444817000088

Li, C., Wei, L., & Lu, X. (2024). Task complexity and L2 writing performance of young learners : Contributions of cognitive and affective factors. *The Modern Language Journal,* 741–770. https://doi.org/10.1111/modl.12954

Li, L. (2023). An Overview on Task-based Language Teaching. *Frontiers in Educational Research, 3*(10), 24–29.

Long, M. (2015). *Second Language Acquisition and Task-Based Language Teaching*. Wiley-Blackwell.

MacIntyre, P. D. (1998). Conceptualizing Willingness to Communicate in A L2: A Situational Model of L2 Confidence and Affiliation. *The Modern Language Journal, 82*, 545–562.

Mahpul. (2014). *Task Difficulty in Dialogic Oral Production by Indonesian EFL Learners*. (Doctoral dissertation, University of Melbourne).

Masuram, J., & Sripada, P. N. (2020). Developing spoken fluency through task-

based teaching. *Procedia Computer Science*, *623-630.* https://doi.org/10.1016/j.procs.2020.05.080

Megawati, F. (2022). Penguatan Kemampuan Literasi Bahasa Inggris Melalui Media Pembelajaran untuk Meningkatkan Motivasi Belajar bagi Siswa Sekolah Dasar. *Jurnal Basicedu*

Metty, A. (2025). Revisiting Task-based Language Teaching to TEYL : Evidence from Indonesian Young Learners. *Journal on English as a Foreign Language*, 298–309.

Mudinillah, A., Rahmi, S. N., & Taro, N. (2024). Task-Based Language Teaching : A System atic Review of Research and Applications. *EduLite: Journal of English Education, Literature and Culture, 3*(2), 102–115.

Munira, S., & Ferdousi, S. A. (2020). A Study of TBLT Approach : An Experiment of Sample Lesson and Preparing Lesson Plan for EFL Classroom. *International Journal of English Language and Literature Studies,*1-10*.*

Natsuko, S., Teaching, T. L., & Difference, S.-. (2012). Task-Based Language Teaching at Elementary Schools in Japan : Issues and Possibilities*.* *JALT Journal,* 207–212.

Nunan, D. (2004). *Task-based Language Teaching.* Cambridge University Press.

Ong, J. (2014). How do Planning Time and Task Conditions Affect Metacognitive Processes of L2 Writers ? *Journal of Second Language Writing*, *23*, 17–30. https://doi.org/10.1016/j.jslw.2013.10.002

Panggua, S., Kristen, U., & Toraja, I. (2022). Need Analysis : Developing Speaking Skills Training Materials for High School EFL Teacher in Indonesia. J*ournal of Languages and Language Teaching*, 441-451. https://doi.org/10.33394/jollt.v10i3.5435

Piri, F., Barati, H., & Ketabi, S. (2012). The effects of pre-task, on-line, and both pre-task and on-line planning on fluency, complexity, and accuracy - The case of Iranian EFL learners' written production. *English Language Teaching*, *5*(6), 158–167. https://doi.org/10.5539/elt.v5n6p158

Plano Clark, V. L. (2017). *Mixed methods research (2nd ed.).* SAGE Publications.

Prabhu, N. S. (1988). *Second Language Pedagogy*. Oxford University Press.

Rahimi, M. (2019). Effects of increasing the degree of reasoning and the number of elements on L2 argumentative writing. *Language Teaching Research*, *23*(5), 633–654. https://doi.org/10.1177/1362168818761465

Rahimpour, M., & Hazar, F. (2007). Topic Familiarity Effect on Accuracy, Complexity, and Fluency of L2 Oral Output. *Journal of Asia TEFL*,*4*(4), 191–211.

Richards, J. C., & Rodgers, T. S. (2001). *Approaches and Methods in Language Teaching (2nd ed.). Cambridge University Press.*

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*,

22(1), 27–57. https://doi.org/10.1093/applin/22.1.27

Robinson, P. (2007). *Designing and Conducting Mixed Methods Research. Australian and New Zealand Journal of Public Health*, *31*(4), 388. https://doi.org/10.1111/j.1753-6405.2007.00096.x

Robinson, P. (2011). Task-Based Language Learning*: A Review of Issues*. *Language Learning*, *61*(s1), 1–36. https://doi.org/10.1111/j.1467-9922.2011.00641.x

Sabil, M. (2020). The Use of Task-Based Language Teaching in Improving Students ' Speaking Proficiency. *Journal of Languages and Language Teaching*, 31–36. https://doi.org/10.33394/jollt.v8i1.2238*4*(1), 31–36.

Saeedi, M., Ketabi, S., & Kazerooni, S. R. (2012). The Impacts Of Manipulating Task Complexity On EFL Learners ' Performance. *Theory and Practice in Language Studies*, 1057–1075.

Sample, E., & Michel, M. (2014). An Exploratory Study into Trade-off Effects of Complexity , Accuracy , and Fluency on Young Learners ' Oral Task Repetition. *TESL Canada Journal, 31*(8), 23–46. https://doi.org/10.18806/tesl.v31i0.1185

Sánchez, L., & Kalamakis, S. Z. (2023). The interplay of task complexity (± Here-and-Now) and proficiency in shaping written narratives in Spanish as a foreign language. *System*, 1–12. https://doi.org/10.1016/j.system.2022.102949

Saputro, T. H., Hima, A. N., & Farah, R. R. (2021). Benefits and challenges of doing task- based language teaching in Indonesia : Teachers ' perception. *Kata Kita: Journal of Language, Literature, and Teaching7*(1), 131–142.

Scontras, G., Badecker, W., Shank, L., & Lim, E. (2015). Syntactic Complexity Effects in Sentence Production. *Cognitive Science,* 559–583. https://doi.org/10.1111/cogs.12168

Shvidko, E., Evans, N. W., & Hartshorn, K. J. (2015). Factors affecting language use outside the ESL classroom: Student perspectives*. System*, *51*, 11–27. https://doi.org/10.1016/j.system.2015.03.006

Skehan, P. (n.d.). *Processing Perspective on Task Performance*. John Benjamins.

Skehan, P. (1996). A Framework for the Implementation of Task- based Instruction. *Applied Linguistics*, https://doi.org/10.1093/applin/17.1.38

Storch, N. (2002). *Patterns of Interaction in ESL Pair Work*. *Language Learning,* 119–158.

Suzuki, S., & Kormos, J. (2020). Perceived Fluency : An Investigation of Complexity, Accuracy, and Fluency features of L2 speech*. Studies in Second Language Acquisition,. 42*. https://doi.org/10.1017/S0272263119000627

Svirina, L., & Ashrapova, A. (2020). Gender differences in selections among topics for producing meaningful utterances. *Utopia y Praxis Latinoamericana*, *25*(Extra12), 287–293.

https://doi.org/10.5281/zenodo.4280149

Tannen, D. (1990). *You Just Don ' t ~ Understand*: W*omen and men in conversation.* William Morrow.

Willis, D., & Willis, J. (2007). *Doing Task-based Teaching*. Oxford University Press.

Xing, J. (2023). How do EFL learners process oral tasks with different complexity : An exploratory study. *Frontiers in Psychology*, 14, 1241964. https://doi.org/10.3389/fpsyg.2023.1241964

Xu, T. S., Zhang, L. J., & Gaffney, J. S. (2022). Examining the Relative Effects of Task Complexity and Cognitive Demands on Students' Writing in a Second Language. *Studies in Second Language Acquisition*, *44*(2), 483–506. https://doi.org/10.1017/S0272263121000310

Zein, S., Sukyadi, D., Hamied, F. A., & Lengkanawati, N. S. (2020). English language education in Indonesia: A review of research (2011-2019). *Language Teaching*, *53*(4), 491–523. https://doi.org/10.1017/S0261444820000208