

CHAPTER I

INTRODUCTION

1.1. Background of The Research

In the contemporary educational landscape, particularly in the post-pandemic era, classroom interaction has undergone significant transformation as teaching and learning increasingly rely on complex forms of communication. Recent studies indicate that classroom communication is no longer dominated solely by spoken language but involves the integration of multiple communicative modes (Jewitt, 2009; Kress, 2010; Norris, 2004). However, in classroom interaction, particularly in English as a Foreign Language (EFL) settings, communication primarily occurs through verbal resources (speech) and embodied nonverbal resources such as gesture and gaze, which function together to support meaning-making processes during instruction (Kress, 2010; Jewitt, 2009).

Educational research emphasizes that effective teaching is not only determined by the delivery of cognitive content or instructional strategies but also by how teachers communicate through these communicative resources during classroom interaction. Teachers' ways of giving instructions, asking questions, providing feedback, using gestures, and maintaining eye contact function as essential resources that shape students' interpretation of the lesson and classroom atmosphere (Walsh, 2011; Jewitt, 2009). In EFL classrooms, where learners rely heavily on teachers' explanations to understand the target language, the integration of verbal and nonverbal communication becomes particularly important in facilitating comprehension and participation (Walsh, 2011; Sert, 2015).

In the Indonesian context, particularly at the Vocational High School (SMK) level, English teaching emphasizes practical communication skills that prepare students for workplace demands. Although vocational education often involves procedural and technical explanations, in EFL classrooms the main challenge lies in helping students understand linguistic content through classroom interaction. In this context, teachers frequently combine verbal explanation with gesture and gaze

to clarify meaning, emphasize key points, and manage interaction during classroom activities (Jewitt, Bezemer, & O'Halloran, 2016; Kress, 2010).

However, classroom realities often show variations in how teachers integrate these communicative resources. While some teachers rely heavily on verbal explanation, others effectively combine speech with gesture and gaze to guide students' understanding and participation. These differences influence how classroom interaction is organized and how meaning is constructed during the learning process (Walsh, 2011).

Previous studies have examined teacher communication from different perspectives. Some studies focus on verbal communication, such as questioning, explanation, and feedback, while others emphasize nonverbal communication such as gestures and eye contact (Sert, 2015; Kendon, 2004). However, these studies tend to examine verbal and nonverbal communication separately and often rely on perception-based data rather than real-time classroom interaction. As a result, they provide limited understanding of how verbal and nonverbal resources operate simultaneously in authentic classroom settings (Jewitt, 2009).

In response to this gap, the present study adopts a multimodal perspective of communication. From this perspective, communication is understood as a meaning-making process that emerges through the interaction of multiple semiotic modes (Kress & van Leeuwen, 2001; Kress, 2010). Although multimodal communication encompasses a wide range of semiotic resources, this study narrows its focus to verbal (speech) and embodied nonverbal resources, specifically gesture and gaze. This focus is grounded in the consideration that these modes are most salient in real-time classroom interaction and play a central role in facilitating meaning-making in EFL contexts (Jewitt, 2009; Norris, 2004). Furthermore, the interaction between these modes can be explained through the concept of intersemiotic complementarity, which highlights how different communicative resources work together to complement and reinforce meaning during classroom interaction (O'Halloran, 2004).

Despite the growing interest in multimodal classroom communication, empirical studies employing multimodal analysis in Indonesian EFL contexts remain limited, particularly at the Vocational High School level. Many existing studies rely on questionnaire-based approaches, which provide limited access to the dynamic interactional processes occurring in classrooms (Setiawan & Arifin, 2021). Consequently, there is still limited qualitative evidence explaining how English teachers integrate verbal and nonverbal communicative resources during real-time classroom interaction.

Recent research, however, indicates that multimodal classroom communication has become an increasingly important area of investigation in English language education. A systematic review published by Rahmanu and Molnár (2024), which examined 34 studies published between 2013 and 2024, found that multimodal approaches increasingly incorporate nonverbal resources such as gesture, kinesics, spatial positioning, facial expression, and gaze alongside linguistic resources. The review also reported positive relationships between multimodal approaches and the development of English language proficiency and communicative ability. Nevertheless, the review primarily focused on higher education contexts, indicating that multimodal communication in other educational settings remains an area requiring further investigation.

Recent classroom-based studies have further demonstrated that multimodal resources perform specific pedagogical and interactional functions. For example, research on EFL classroom lead-ins found that teachers combined verbal and visual resources to establish cognitive frameworks for new topics, while eye movements, facial expressions, and gestures helped attract students' attention and encourage participation. These findings indicate that multimodal communication should not be understood merely as the simultaneous presence of several modes, but as the purposeful coordination of different resources according to particular pedagogical contexts.

Similarly, Wang and Dai (2025) examined multimodal teacher–student interaction in EFL classrooms and emphasized the importance of examining how

verbal, gestural, and contextual resources interact dynamically during classroom activities. Their study demonstrates that multimodality involves the orchestration of different resources rather than the independent use of individual modes. This perspective is particularly relevant to the present study because speech, gesture, and gaze may occur simultaneously and perform complementary functions during authentic classroom interaction.

Research published in 2025 has also shown that teacher gestures can function as embodied interactional resources for eliciting student participation. Şimşek Tontuş and Kuru Gönen (2025) found that teachers' gestures were used as embodied elicitation strategies that helped facilitate and balance student participation in language classrooms. This finding strengthens the argument that gesture should not be treated merely as an accompanying behaviour but as an interactional resource that contributes to the organization of classroom discourse.

These developments suggest that the study of teacher communication needs to move beyond an isolated examination of verbal or nonverbal behaviour. Instead, contemporary multimodal research increasingly emphasizes how communicative resources are coordinated in real time to construct meaning, direct attention, elicit responses, and organize participation. However, despite these developments, relatively limited research has examined the simultaneous coordination of speech, gesture, and gaze in naturally occurring Indonesian EFL classroom interaction, particularly within vocational high school contexts. This contextual and analytical gap provides a strong rationale for the present study.

Based on this gap, the present study focuses on analyzing English teachers' communication in classroom interaction through a multimodal perspective, with particular emphasis on verbal and nonverbal resources. Conducted in an EFL classroom at a Vocational High School in Bandung Regency, this study adopts a qualitative case study approach to examine how teachers integrate speech, gesture, and gaze during instruction and how meaning is constructed through these interactions. The central focus of this study is not merely whether teacher communication matters, but how verbal and nonverbal communicative resources

are orchestrated and integrated within authentic classroom interaction in vocational EFL settings.

1.2. Research Problems

Although teacher communication has long been recognized as a crucial element of effective classroom interaction, many existing studies tend to examine verbal and nonverbal communication as separate phenomena. Research on teacher talk often focuses on questioning strategies, explanations, and feedback, while studies on nonverbal communication emphasize resources such as gesture and gaze as independent variables. Such fragmentation limits our understanding of how verbal and nonverbal resources operate together in authentic classroom interaction.

Recent multimodal research provides evidence that this fragmentation is increasingly being addressed by examining classroom communication as an interactional configuration of multiple semiotic resources. Studies have demonstrated that gesture, gaze, facial expression, and other embodied actions can work together with speech to manage attention, support comprehension, elicit responses, and organize participation. For example, recent EFL research has shown that teacher gestures can function as embodied elicitation strategies, while gaze and other embodied resources may contribute to the organization of classroom participation.

However, the existing literature does not yet provide sufficient evidence concerning how speech, gesture, and gaze are simultaneously coordinated in authentic Indonesian vocational EFL classrooms. Recent studies have often focused on particular activities, online environments, higher education, or specific multimodal resources. For instance, systematic reviews have identified a strong concentration of multimodal research in higher education and digital learning environments, while recent classroom-based studies have examined particular pedagogical activities or specific embodied resources.

The problem addressed in the present study is not simply whether teachers use verbal and nonverbal communication, but how these resources are coordinated as an integrated multimodal system during naturally occurring classroom

interaction. This distinction is important because the meaning and function of one communicative resource may depend on its relationship with other resources occurring at the same moment. The present study consequently investigates speech, gesture, and gaze as interconnected semiotic resources within authentic EFL vocational classroom interaction.

The central research problem addressed in this study concerns how English teachers employ and integrate verbal and nonverbal communicative resources as multimodal resources in classroom interaction, particularly within the context of Vocational High Schools in Bandung Regency. By examining classroom communication through a multimodal analytical perspective, this study seeks to explore how teachers' communicative practices structure interaction and support meaning-making processes in vocational EFL classrooms.

1.3. Research Questions

Based on the research problem and the research gap discussed above, this study focuses on examining how English teachers employ verbal and nonverbal communicative resources in classroom interaction. In particular, the study examines speech, gesture, and gaze as interconnected multimodal resources and explores how these resources contribute to meaning-making and the organization of classroom interaction in EFL vocational classrooms. Accordingly, the research questions of this study are formulated as follows:

1. How do English teachers employ verbal and nonverbal communicative resources (speech, gesture, and gaze) in classroom interaction?
2. How do these verbal and nonverbal communicative resources function to construct meaning and organize classroom interaction in EFL vocational classrooms?

1.4. Purposes of The Study

In line with the research questions, this study aims to examine the use and functions of verbal and nonverbal communicative resources in EFL vocational classroom interaction. More specifically, the study seeks to identify how English

teachers employ speech, gesture, and gaze during classroom interaction and to analyze how these resources work together to construct meaning and organize classroom interaction. The purposes of this study are as follows:

1. To identify the verbal and nonverbal communicative resources (speech, gesture, and gaze) employed by English teachers during classroom interaction.
2. To analyze how these verbal and nonverbal communicative resources function in constructing meaning and organizing classroom interaction in EFL vocational classrooms.

1.5. Significance of The Research

This study is significant because it provides an in-depth understanding of how English teachers employ verbal and nonverbal communicative resources (speech, gesture, and gaze) during classroom interaction in EFL settings. Rather than viewing classroom communication merely as verbal instruction that produces psychological learning outcomes, this study approaches classroom interaction as a multimodal meaning-making process in which speech, gesture, and gaze work together to construct instructional meaning. In doing so, the study contributes to a growing shift in English Language Teaching (ELT) research from a narrow focus on teacher talk toward a broader perspective that recognizes the role of integrated verbal and nonverbal communication in classroom interaction.

The significance of this study is strengthened by recent developments in multimodal classroom research. Contemporary studies increasingly demonstrate that classroom communication involves the coordinated use of linguistic and embodied resources rather than language operating independently. Recent empirical evidence shows that gesture, gaze, facial expression, and other embodied resources can contribute to attention management, student participation, meaning-making, and pedagogical interaction.

The present study extends this development by providing qualitative evidence from an Indonesian vocational high school context. This context is particularly important because vocational education combines academic learning with practical and occupational orientations. Examining multimodal teacher communication in

such a setting may provide a more context-sensitive understanding of how teachers use speech, gesture, and gaze to explain instructional content and organize classroom interaction.

1.5.1. Theoretical Significance

Theoretically, this study contributes to research on classroom discourse and multimodal communication by applying a social semiotic perspective to the analysis of teacher–student interaction. The study highlights how verbal and nonverbal communicative resources, particularly speech, gesture, and gaze, operate together as integrated semiotic resources during instructional interaction. In particular, the study contributes to the understanding of intersemiotic complementarity, which explains how these communicative modes complement and reinforce one another in constructing meaning within classroom interaction. By examining classroom interaction in an Indonesian EFL vocational school context, this research also expands empirical evidence in multimodal discourse studies, which has largely focused on broader or Western educational settings.

1.5.2. Practical Significance

Practically, the findings of this study provide insights for English teachers and educators regarding how verbal and nonverbal communication (speech, gesture, and gaze) can support more effective classroom interaction. In EFL vocational classrooms, teachers often need to explain linguistic concepts and guide student participation through interaction. Such instructional situations require teachers to strategically combine speech with gesture and gaze to clarify meaning and manage classroom interaction. By highlighting how these communicative resources function during instruction, this study can help teachers develop greater awareness of how verbal and nonverbal communication can be used to support students' comprehension and engagement. Furthermore, teacher education programs and professional development initiatives may use these insights to strengthen teachers' communicative competence in EFL vocational classrooms.

1.6. Conceptual Framework

This study is grounded in a multimodal perspective that views classroom communication as an integrated system of semiotic resources rather than as separate verbal and nonverbal elements. The conceptual framework of this study is built upon four main theoretical components: Classroom Interactional Competence (CIC), Social Semiotic Theory, Intersemiotic Complementarity, and Multimodal Meaning-Making. Classroom Interactional Competence (Walsh, 2011) provides the foundation for understanding how teachers use interaction as a pedagogical tool to facilitate learning. It emphasizes how teachers organize classroom discourse through questioning, feedback, and turn-taking to create learning opportunities.

Social Semiotic Theory (Kress, 2010) explains that communication involves multiple semiotic resources such as speech, gesture, and gaze, which are combined strategically in social interaction. Although this theory acknowledges a wide range of semiotic modes, this study focuses specifically on those that are most salient in real-time classroom interaction. Intersemiotic Complementarity (O'Halloran, 2004) further explains how different communicative modes work together to complement and reinforce meaning. In classroom interaction, verbal explanation is often supported by gesture and gaze to clarify meaning and manage interaction.

Multimodal Meaning-Making (Jewitt, 2009; Norris, 2004) highlights that meaning is constructed through the simultaneous orchestration of multiple modes. In EFL classrooms, teachers integrate verbal and nonverbal resources to organize interaction and support students' understanding, particularly through the use of speech, gesture, and gaze.

Based on these theoretical perspectives, this study conceptualizes teacher communication as multimodal communicative resources that include verbal (instruction, explanation, questioning, and feedback) and nonverbal modes (gesture and gaze). These resources are orchestrated simultaneously to organize classroom interaction and construct meaning in EFL vocational classrooms.

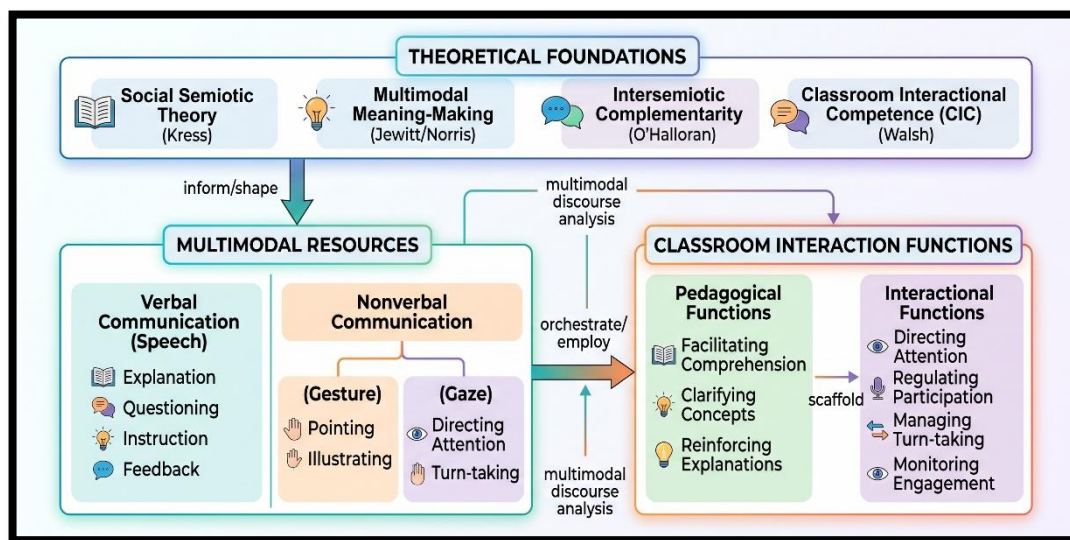


Figure 1. 1 Multimodal Conceptual Framework of English Teachers' Verbal and Nonverbal Communication.

As illustrated in Figure 1.1, the conceptual framework integrates theoretical foundations, multimodal resources (speech, gesture, and gaze), and classroom interaction functions, demonstrating how meaning is co-constructed through intersemiotic complementarity.

1.7. Contributions of The Study

This study contributes to the field of educational communication by addressing conceptual and contextual gaps in previous research on teacher communication in EFL classrooms. By adopting a multimodal analytical framework with a specific focus on verbal and nonverbal communicative resources (speech, gesture, and gaze), the study offers a deeper understanding of how teachers employ these semiotic resources during authentic classroom interaction.

1.7.1. Theoretical Contribution

Theoretically, this study contributes to multimodal discourse and classroom interaction research by demonstrating how meaning is constructed through the integration of verbal and nonverbal communicative resources, particularly speech, gesture, and gaze, in EFL teaching. In particular, the study contributes to the development of a more focused understanding of multimodal communicative resources in classroom interaction by identifying how these

selected modes are employed during instructional activities. By examining these resources in the context of vocational high schools in Indonesia, the study also expands empirical evidence in multimodal discourse studies, which have been largely conducted in broader or Western educational contexts.

1.7.2. Practical Contribution

Practically, this study provides concrete insights for English teachers regarding how verbal and nonverbal communication (speech, gesture, and gaze) can be used strategically in classroom practice. The findings may inform teachers' reflective practices by illustrating how these communicative resources function together to organize interaction, manage classroom discourse, and support student participation.

Furthermore, the study contributes to the development of a set of multimodal indicators that teachers may use to reflect on and evaluate their own teaching practices, particularly when reviewing video recordings of classroom lessons. These indicators can help teachers identify how speech, gesture, and gaze work together to clarify instructional explanations and support students' comprehension and engagement. In addition, these insights can support teacher education and professional development programs in strengthening teachers' communicative competence in EFL vocational classrooms.

1.8. Previous Study

Previous studies on classroom communication have increasingly moved from a predominantly linguistic perspective toward a multimodal understanding of classroom interaction. This development recognizes that teachers and students construct meaning through the coordinated use of speech and embodied, visual, and contextual resources.

Rahmanu and Molnár (2024) conducted a systematic review of 34 empirical studies on multimodal English language learning published between 2013 and 2024. Their review found that multimodal approaches increasingly incorporate linguistic and nonverbal resources, including gesture, kinesics, spatial positioning, facial expression, and gaze. The review also reported that multimodal approaches

were associated with the development of learners' English language proficiency and communicative abilities. However, the reviewed studies were predominantly situated in higher education contexts. This indicates the need for further investigation of multimodal communication in other educational settings, particularly vocational high schools.

Watanabe (2024) examined a teacher's recurrent embodied practice in an EFL classroom and found that a specific hand gesture was systematically used as an interactional resource for allocating turns and managing speaker shifts. This finding demonstrates that teacher gestures may perform specific interactional functions and contribute directly to the organization of classroom participation rather than merely accompanying verbal communication. Nguyen (2024) investigated the embodied nature of students' engagement and participation during a total physical response activity in a beginning second-language classroom. Using multimodal conversation analysis, the study demonstrated that students' engagement and participation were expressed through embodied actions in response to teacher demonstrations, commands, and classroom activities. The study highlights the importance of examining classroom interaction as an embodied and multimodal process.

Şimşek Tontuş and Kuru Gönen (2025) examined teachers' gestures in synchronous online language classrooms and found that gestures functioned as embodied elicitation strategies for facilitating student participation. Their findings demonstrate that gestures may contribute directly to teacher–student interaction by inviting student responses and regulating participation. Wang and Dai (2025) investigated multimodal teacher–student interaction in EFL classrooms through the integration of verbal, gestural, and contextual resources. Their study demonstrated that multimodal interaction involves dynamic patterns of coordination among different communicative resources. This finding is particularly relevant to the present study because speech, gesture, and gaze may occur simultaneously and perform complementary functions during authentic classroom interaction.

Badem (2025) examined multimodal repair initiations in video-mediated EFL classroom interactions, focusing on screen-based and embodied actions. The study demonstrates that teachers employ embodied and digital resources to address

interactional difficulties and create opportunities for student self-repair. Although the study was conducted in an online environment, it reinforces the importance of analyzing multimodal resources through naturally occurring interaction rather than relying exclusively on participants' perceptions. Tseng (2025) investigated multimodal teaching and motivation in a rural EFL classroom and found that multimodal teaching could support students' motivation by engaging different senses, drawing on cultural and personal resources, and creating more equal classroom relationships. Although the educational context and research focus differ from the present study, this research demonstrates that multimodality may influence classroom experiences beyond the transmission of instructional information.

These studies demonstrate that contemporary multimodal research increasingly views classroom communication as an integrated interactional process rather than as separate verbal and nonverbal behaviors. Recent research also shows that gesture, gaze, and other embodied resources may perform specific pedagogical and interactional functions, including directing attention, eliciting responses, regulating participation, and supporting comprehension. More recently, research in 2026 has begun to emphasize the dialogic and scaffolding functions of embodied modes (Rakhmadani et al., 2026; Zhou & Lopez-Ozieblo, 2026), further validating the need to investigate gesture and gaze not as isolated behaviors, but as integral components of classroom interactional competence.

Nevertheless, relatively limited research has examined the simultaneous coordination of speech, gesture, and gaze in naturally occurring Indonesian EFL classroom interaction, particularly in vocational high school contexts. Many recent studies have focused on higher education, online environments, particular pedagogical activities, or specific multimodal resources. The present study extends this recent trajectory into the under-researched context of Indonesian vocational education. Through a qualitative case study and video-based classroom observation, the study investigates how speech, gesture, and gaze are employed and coordinated as interconnected multimodal resources during authentic classroom interaction to construct meaning and organize classroom interaction.