CHAPTER II

REVIEW OF RELATED LITERATURE

This chapter provides a comprehensive literature review, covering previous studies on the effectiveness of YouTube in vocabulary teaching, the theoretical foundations of the characteristics of young learners learning, and the research framework. Specifically, this chapter highlights the gap in previous studies regarding the technical details of classroom implementation and offers a solution through the integration of visual media from the KidsTV123 channel to enhance mastery of pronunciation, spelling and word meaning.

2.1 Previous Studies

The use of YouTube videos to teach vocabulary has demonstrated many advantages, making it a valuable tool in modern language education. As proposed by Sorohiti et al. (2020), incorporating YouTube into vocabulary teaching significantly improves student learning outcomes by providing dynamic multimedia content that suits a variety of learning styles. Their quantitative study, which involved pre- and post-vocabulary mastery tests, revealed a noticeable improvement in students' ability to remember and apply words, with participants showing a 20–25% increase in scores. This led to recommendations for teachers to integrate YouTube videos while supporting them with self-reflection activities, such as keeping a journal of words learned and structured sessions that include group discussions to reinforce retention.

This aligns with findings from Sabgini (2023), which showed that YouTube-based vocabulary teaching can be effectively applied to young learners, resulting in increased mastery and positive student responses, including higher motivation and reduced anxiety in language classes. This study used qualitative methods, including classroom interaction observations and open-ended interviews with students aged 7–10, highlighting how videos with animations and songs make abstract concepts easier to understand.

However, previous studies often lack detailed guidance on specific classroom methods needed to successfully improve vocabulary skills, such as step-by-step lesson plans or integration with other teaching strategies. This gap is reflected in the conclusions of Roman and Mujiyono (2021), who emphasize that the use of YouTube for educational purposes requires a methodical approach to organizing content and designing lessons to ensure real effectiveness, including selecting videos that meet curriculum standards and adapting them appropriately.

Despite its many benefits, the use of YouTube in the classroom presents several challenges. As noted by Arsyad (2011), the effective implementation of YouTube channels in teaching depends on a stable internet connection, which can be a constraint in areas with limited infrastructure. This is reinforced by Putri et al. (2024), who found that students often experience buffering issues and incomplete video loading due to poor connectivity, negatively impacting comprehension and participation. Their study reported that up to 40% of students experienced disruptions. Therefore, teachers must prepare backup activities, such as printed transcripts or offline audio recordings, to ensure learning continuity.

Vocabulary itself serves as an essential foundation for mastering the four core English skills: speaking, reading, writing, and listening. Without sufficient vocabulary, students struggle to express ideas clearly and understand messages effectively. Vocabulary does not simply refer to a list of words but involves a complex system of lexical knowledge. According to Schmitt (2000), vocabulary knowledge includes word form (pronunciation and spelling), meaning, and use in context. Similarly, Nation (2001) explains that vocabulary learning should focus on three main aspects: form, meaning, and use. Based on these theoretical perspectives, this study narrows its focus to three measurable elements of vocabulary mastery: pronunciation, spelling, and meaning, which are considered fundamental for elementary students.

Several previous studies have also examined pronunciation improvement through YouTube. Nurdianti (2022) found that YouTube provides engaging and enjoyable learning experiences that improve students' pronunciation accuracy. Sulistiyaningsih (2017) similarly reported significant improvement in pronunciation test scores after using YouTube videos. However, these studies mainly concentrated on pronunciation without integrating spelling and meaning as part of comprehensive vocabulary mastery.

Although many previous studies reported positive results, several limitations can be identified. First, most research primarily measured vocabulary improvement through numerical test scores without clearly specifying which components of vocabulary were developed—whether pronunciation accuracy, spelling recognition, or depth of meaning comprehension. Second, detailed pedagogical procedures were often not described clearly, leaving limited practical guidance for teachers regarding step-by-step classroom implementation. Third, while studies such as Madani et al. (2023) and Arifah and Auliya (2021) highlighted increased motivation and positive perceptions toward YouTube, they focused more on engagement rather than structured vocabulary components.

Therefore, this study differs from previous research in several important aspects. It clearly defines vocabulary based on established theoretical frameworks and focuses specifically on pronunciation, spelling, and meaning as measurable components. In addition, it integrates YouTube videos with structured learning tasks such as guided viewing sessions, pause-and-repeat exercises, vocabulary worksheets, contextual sentence construction, interactive games, and reflective activities. Furthermore, this study anticipates technical limitations by preparing alternative offline strategies to ensure that learning remains effective even in conditions of unstable internet access.

2.2 Theoretical Review

2.2.1 The Concept of Vocabulary

Vocabulary is one of the most crucial components in the process of learning a foreign language, especially for young learners. According to Schmitt (2000), vocabulary is not just a list of protected words, but a complex system of lexical knowledge that includes single words, phrases, idioms, and multi-word units. These units are known as lexical items or lexemes, which function as single entities of meaning even though they consist of several words. For example, phrases such as kick the bucket or break a leg to wish someone good luck are considered a single unit of meaning, not just a combination

of words.

Schimitt (2000) emphasizes that comprehension plays a central role in language learning because it is the foundation of communication. Without adequate knowledge, learners find it difficult to understand or produce language effectively, even if they have mastered the rules of grammar. Vocabulary enables learners to express ideas, describe objects, and interact socially, so deficiencies in this area can hinder overall progress.

In the context of elementary school students, language learning becomes increasingly important because they are in the early stages of language development. These children have cognitive capacities that are still developing, so learning comprehension must be tailored to basic and concrete aspects. The main focus is on words that can be visualized or associated with everyday experiences, such as the names of objects for example apple, cat, simple actions such as sleep or eat, and basic emotions such as happy or sad. This approach helps children build a strong foundation before moving on to more abstract knowledge.

To support comprehension learning at the elementary school level, teachers can implement various engaging and interactive strategies. One effective method is the use of pictures, games, and songs, which capitalize on children's tendency to learn through sensory experiences. For example, the word wall technique in the classroom can display pictures and new words, allowing students to repeat them visually. In addition, activities such as role-playing or interactive stories can help children understand the context of spreadsheet use, such as conveying simple dialogues about daily routines.

Additional research, such as that conducted by Nation (2001), suggests that learning should involve active discovery and use. Children can be encouraged to play games such as matching games, where they match words with pictures. This not only improves retention, but also makes the learning process enjoyable. Common challenges

in vocabulary learning for children include difficulty distinguishing between synonyms or homonyms, as well as low motivation if the material is too abstract. Therefore, instruction should be tailored to their developmental level, avoiding excessive cognitive load.

2.2.2 Teaching Vocabulary to Young Learners

Teaching vocabulary to young learners requires engaging, meaningful, and interactive learning activities that match their developmental characteristics. According to Piaget's theory of cognitive development, children aged 7-11 years old in the concrete operational stage symbolically, meaning they learn by connecting words to pictures or objects (Piaget, 1972). That is why, this study focuses on involving audio and visual support through YouTube videos as learning media to teach vocabulary to 4th grade students. By using this media, students are expected to be more engaged and master the vocabulary aspect better. Recent studies emphasize that young learners learn vocabulary more effectively through multisensory experiences that combine visual, auditory, and kinesthetic elements.

According to Abidin & Razak (2020), young learners improve vocabulary retention when learning activities involve images, videos, gestures, and hands-on activities because these stimulate attention and curiosity. Similarly, Fitriani and Pratama (2021) assert that enjoyable learning materials such as songs, animations, and games can increase motivation and make vocabulary learning more memorable. Suryani et al. (2023) add that teachers should provide simple and meaningful contexts so children can relate new words to real-life experiences.

Researchers suggest several steps for teaching vocabulary to young learners, including:

1. Introducing new words with visuals
2. Modelling correct pronunciation.
3. Explaining meaning in simple contexts.
4. Practicing the word through fun activities
5. Reviewing and repetition
6. Encouraging use in real communication.

2.2.3 The Use of YouTube in English Language Learning

YouTube has grown into a very popular educational platform, providing authentic and engaging learning materials for students at various levels. With millions of videos available for free, this platform allows easy access to content that is relevant to real life, making the learning process more dynamic and enjoyable. According to Kabooha & Elyas (2018), YouTube videos can improve students listening and pronunciation skills because they are exposed to natural language use. For example, students can watch news videos, vlogs, or tutorials that use every day English, which helps them understand the intonation, accent, and rhythm of authentic conversations, rather than just rigid textbook texts.

In addition, YouTube videos allow students to visualize vocabulary through images, animations, and real-life contexts, which significantly improves memory retention. Watkins & Wilkins (2011) explain that this visual approach utilizes the principle of multisensory learning, where the combination of audio and visual aids the brain in storing information longer. For example, when students watch videos about travel or cooking recipes, they not only hear words such as “adventure” or “ingredients,” but also see related images, making the vocabulary easier to remember and apply in practical contexts.

The interactive and multimedia nature of YouTube makes it an effective complementary tool in English language teaching. The platform not only provides passive content, but also allows interaction through comments, playlists, and sharing features, which encourage students to participate actively. Teachers can integrate these videos into the curriculum, such as using short clips for listening exercises or language analysis tasks. Furthermore, YouTube supports differentiated learning, where students with different learning styles—for example, visual or kinaesthetic can find content that suits them. However, it is important to select high-quality videos and guide students to stay focused on the learning objectives, avoiding distractions such as advertisements or irrelevant content.

2.2.4 KidsTV123 as a Learning Resource

KidsTV123 is a YouTube channel specifically designed to provide educational videos that focus on basic English language learning, such as the alphabet, numbers, colours, and everyday vocabulary. With a fun and accessible approach, this channel targets early childhood to elementary school children, utilizing video technology to make the learning process more interesting and less boring. The videos on KidsTV123 use elements such as songs, repetition, and colourful animations, which have proven to be very effective for young children as they are appropriate for their stage of development. For example, simple songs with catchy melodies help children remember vocabulary through rhythm and music, while repetition of words or phrases reinforces understanding without making them bored.

Previous studies have shown the positive impact of using KidsTV123 videos in

English lessons. Kurniawati (2020) found that students exposed to these videos experienced a significant increase in vocabulary mastery, especially in pronunciation and spelling. This is due to the repetitive and musical nature of the videos, which allow children to practice pronouncing words such as “apple” or “red” repeatedly with a pleasant rhythm, thereby reducing spelling errors and increasing confidence. Sari (2022) reported similar results, where students who used KidsTV123 as supplementary teaching material showed faster progress in understanding basic vocabulary compared to traditional methods, thanks to the audio-visual combination that stimulated multiple senses.

In addition, colourful animations on KidsTV123 help children visualize abstract concepts, such as colours or numbers, through cute and interactive cartoon images. This not only makes learning more fun, but also improves long-term memory retention. Teachers or parents can integrate these videos into daily routines, such as watching short clips before bedtime or as part of educational games. However, it is important to choose age-appropriate videos and monitor children's interactions to keep them focused on learning, avoiding overly long content or external distractions.

2.3 Theoritical Framework

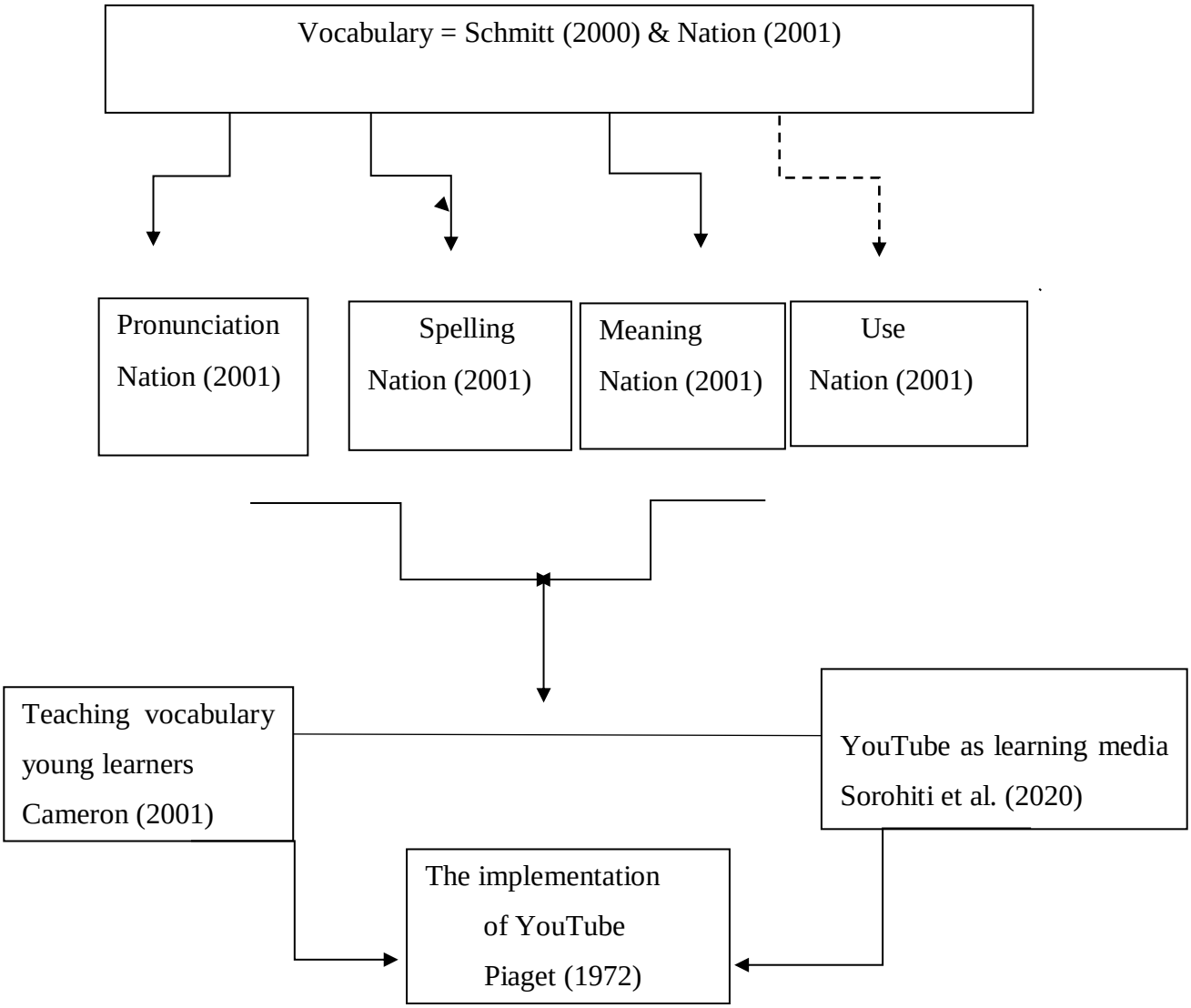


Figure 2.1 Theoretical Framework

1**.** Vocabulary

The concept of vocabulary in this study refers to the theories proposed by Schmitt (2000) and Nation (2001).

According to Schmitt (2000), vocabulary is not merely a list of words but includes knowledge of word forms, meanings, and usage. Similarly, Nation (2001) explains that vocabulary knowledge consists of three main aspects: pronunciation spelling, meaning, and use.

Relationship**:**

The “Vocabulary” box at the top of the diagram represents the main concept, which is then divided into four components:

- Pronunciation
- Spelling
- Meaning
- Use

These components are derived from Nation’s (2001) theory of vocabulary knowledge.

2. Pronunciation

Pronunciation is based on Nation’s (2001) explanation that “form” includes the spoken form of a word.

Harmer (2007) also states that correct pronunciation is essential because it helps learners understand others and be understood in communication.

Relationship:

Pronunciation is part of the “form” aspect in Nation’s theory. In this study, YouTube supports pronunciation learning by providing authentic audio models that students can listen to and imitate.

3. Spelling

According to Nation (2001), spelling belongs to the written form of vocabulary knowledge.

Young learners benefit from seeing and hearing words simultaneously, which strengthens their understanding of the connection between sounds and written forms.

Relationship:

Spelling is closely related to pronunciation since both belong to the “form” aspect. YouTube videos often display written text along with audio, reinforcing students’ recognition of word forms.

4. Meaning

Meaning, in Nation's (2001) framework, refers to the relationship between the word form and the concept it represents.

Based on Piaget's (1972) theory of cognitive development, children aged 7–11 years are in the concrete operational stage. At this stage, they understand concepts more effectively through concrete objects, pictures, and visual representations.

Relationship**:**

YouTube provides images, animations, and contextual visuals that help students understand word meanings in a concrete and engaging way.

5. Use

Use, according to Nation (2001), refers to how a word is used in context, including grammatical patterns and appropriate situations.

Schmitt (2000) emphasizes that knowing a word means knowing how to use it appropriately in communication.

Relationship:

YouTube videos frequently present vocabulary within dialogues or situational contexts, helping students understand how words are used naturally.

6. Teaching Vocabulary to Young Learners

Cameron (2001) states that teaching vocabulary to young learners should be meaningful, engaging, and supported by visual media.

Furthermore, Abidin and Razak (2020) argue that learning activities involving images, videos, gestures, and hands-on activities improve vocabulary retention among young learners.

Relationship:

Because young learners require visual and interactive learning experiences, YouTube is considered an appropriate medium to support vocabulary instruction in this study.

7. YouTube as a Learning Medium

Sorohiti et al. (2020) found that integrating YouTube into classroom instruction improves students' motivation and comprehension due to its dynamic multimedia content.

This is also supported by Mayer's (2009) Multimedia Learning Theory, which states that students learn more effectively from a combination of words and pictures than from words alone.

Relationship**:**

YouTube supports:

- Pronunciation → through audio input
- Spelling → through on-screen text
- Meaning → through visual representations
- Use → through contextual dialogues

8. The Implementation of YouTube

The implementation of YouTube in this study represents the integration of all the theories discussed above.

The overall relationship within the theoretical framework:

Vocabulary theory (Nation & Schmitt) explains the components of vocabulary knowledge.

1. Young learners' theory (Piaget & Cameron) explains the characteristics of elementary school students.
2. Multimedia learning theory (Mayer) and research on YouTube (Sorohiti et al.) explain the effectiveness of audio-visual media.
3. These theories are integrated in the implementation of YouTube to improve students' pronunciation, spelling, meaning, and use.

This diagram illustrates the strategic framework for vocabulary instruction for young learners utilizing digital media**.** The following are the core components of the research:

1. Vocabulary Dimensions

Vocabulary acquisition is categorized into four fundamental pillars. However, this study assigns different levels of emphasis to these aspects:

a. Pronunciation, spelling, and meaning (indicated by solid lines). These constitute the primary focus of the study, as they are directly instructed and assessed within the instructional setting. These aspects represent the foundational components of vocabulary knowledge, particularly for young learners who are still developing their basic language skills. Emphasizing these three dimensions ensures that students first acquire accurate word forms and clear conceptual understanding before progressing to more complex communicative use.

b. Use (indicated by dashed lines). This signifies that the functional application of vocabulary in active communication was not fully implemented. The limited emphasis on "use" is due to several considerations. First, based on Piaget's theory (1972), young learners in the concrete operational stage tend to learn more effectively through concrete representations rather than abstract linguistic production. The productive use of vocabulary in varied communicative contexts requires higher cognitive and linguistic competence, which may not yet be fully developed at this stage.

Second, Nation (2001) explains that vocabulary "use" involves grammatical patterns, collocations, and appropriate contextual application, which demand a deeper level of mastery. Given the scope and time limitations of the study, the instructional focus was directed toward building receptive knowledge (form and meaning) as a foundational stage before encouraging productive use.

Therefore, while "use" remains an important dimension of vocabulary knowledge, it was not fully implemented in this study because the instructional priority was placed

on establishing accurate pronunciation, spelling, and meaning as the core components of early vocabulary development.

2. Integration of Pedagogy and Media

The instructional process is designed by synthesizing the pedagogical characteristics of young learners who require significant visual stimulation with YouTube as an instructional medium, which offer rich and engaging audio-visual content.

3. Research Output

The synthesis of these elements including the integration of KidsTV123 videos into classroom activities, targeted vocabulary aspects such as pronunciation, spelling, and meaning, and pedagogical strategies applied in practical teaching implementation. This phase represents the core application of the research framework, in which educators actively utilize YouTube videos as a dynamic tool to improve vocabulary mastery among students. By aligning video content with learning objectives, teachers create an immersive learning experience that combines audiovisual elements, thereby increasing engagement and retention.

In this implementation phase, the focus shifts to evaluating how the practical use of YouTube videos facilitates students in effectively mastering the pronunciation, spelling, and meaning of vocabulary. Spelling is reinforced through on-screen text and word formation, where students can visually connect letters with sounds.