Journal of English Language and Education
ISSN 2597- 6850 (Online), 2502-4132 (Print)
Journal Homepage: https://jele.or.id/index.php/jele/index

OPEN ACCESS Article

English Vocabulary Acquisition through YouTube Exposure in a Three-Year-Old Child: A Qualitative Case Study

https://doi.org/10.31004/jele.v11i5.3195

*Keysi Muretya, Titin Rahmiatin Rahim, Ririn Syahriani[abc]
[123]Universitas Muhammadiyah Kendari, Indonesia
Corresponding Author: muretyakeysi@gmail.com

A B S T R A C T

Young children may have limited opportunities to encounter English outside formal instruction, making digital media an increasingly relevant source of language input. This study examined how a three-year-old child acquired English vocabulary through regular exposure to English-language YouTube content in a home-based learning environment. The study focused on vocabulary acquisition, retention and use in daily communication, and the role of family support in the child's English learning. A qualitative case study design was employed. Data were collected through nine naturalistic observations and semi-structured interviews with the child's mother and caregiver and analyzed using thematic analysis. The findings revealed three major themes. First, the child acquired vocabulary through repeated visual and auditory exposure to English-language YouTube content. Second, previously acquired vocabulary was retained and spontaneously used during play, daily routines, and emotional expressions. Third, family support through guided viewing, responsive interaction, and corrective feedback reinforced vocabulary development and meaningful language use. These findings suggest that home-based YouTube exposure can provide accessible English input for young children, particularly when supported by active family involvement and appropriate supervision.

Keywords: *Second Language Acquisition, Early Childhood, English Vocabulary, Youtube Exposure, Qualitative Case Study*

Article History:
Received 19th August 2026
Accepted 25th September 2026
Published 05th October 2026

INTRODUCTION

Early childhood represents an important period for cognitive and language development and is often regarded as a favorable stage for second language learning. During these early years, children are highly responsive to linguistic input and can develop additional language knowledge through meaningful and repeated exposure (Hoff, 2013; Lightbown & Spada, 2021). Their ability to absorb language from their surrounding environment makes early childhood particularly conducive to second language acquisition (Lenneberg, 1967; Lightbown & Spada, 2021). In contrast to adult learners, young children commonly develop vocabulary through natural exposure, interaction, and repeated use rather than through direct and explicit instruction (Krashen, 1985; Vygotsky, 1978). The rapid expansion of digital technology has further increased children's access to foreign-language input, particularly English, which is widely available through various online platforms (Gowenlock et al., 2024; Linebarger & Vaala, 2020).

YouTube is among the digital platforms that increasingly provide young children with opportunities to encounter English outside formal learning environments. Its diverse content, including songs, animated programs, stories, and educational videos, exposes children to English vocabulary within audiovisual and contextualized settings. The combination of verbal information and visual representation can make language input more meaningful and accessible to young learners (Mayer, 2021). Previous research has indicated that audiovisual materials can facilitate vocabulary development by providing visual support, repetition, and contextual clues that help children understand unfamiliar words (Anggraini et al., 2022; Dewi

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Keysi Muretya, et al.

& Anggraeni, 2023). Repeated exposure to vocabulary in different contexts can also contribute to stronger retention and more durable lexical knowledge (Webb & Nation, 2017). Nevertheless, exposure to YouTube alone does not necessarily guarantee effective language learning. Children's engagement with the content and the involvement of family members can influence how digital exposure contributes to language development (Gowenlock et al., 2024). Through guidance, interaction, and feedback, family members can extend children's learning beyond passive viewing and create opportunities for meaningful language use (Lightbown & Spada, 2021).

Young children's vocabulary development through digital media should therefore be viewed as a gradual process rather than simply the memorization of individual words. Repeated encounters with English across different situations can help children establish connections between linguistic forms and their meanings. Through these experiences, children may associate spoken words with visual objects, actions, and events while gradually developing contextual understanding of how vocabulary is used. Continued exposure to familiar words can strengthen retention and increase the likelihood that children will retrieve and use previously encountered vocabulary in later interactions. In this way, vocabulary learning involves more than immediate imitation because repeated and meaningful encounters can contribute to the development of longer-term lexical knowledge and spontaneous language use (Ellis, 2002; Nation, 2001).

Audiovisual input plays an important role in supporting young children's second language vocabulary acquisition because it combines spoken language with meaningful visual representations. Visual information helps children establish word–referent relationships by connecting unfamiliar English words with concrete objects or actions displayed on the screen, while repeated exposure across videos strengthens vocabulary retention. Songs and recurring story patterns further reinforce vocabulary learning by presenting familiar linguistic forms in predictable contexts. Previous studies have suggested that educational audiovisual media facilitates vocabulary development when language input is engaging, repetitive, and appropriate for children's developmental level (Gowenlock et al., 2024; Linebarger & Vaala, 2020).

Although digital media provides valuable language input, children's vocabulary development is also shaped by social interaction with more knowledgeable individuals. Family members play an essential mediating role by selecting appropriate educational content, responding to children's emerging English expressions, providing corrective feedback, encouraging repetition, and extending newly learned vocabulary into everyday communication. Such interactions enable children to practice vocabulary beyond the immediate viewing context and support the gradual development of productive language use. From a sociocultural perspective, language learning is facilitated through supportive interaction and guidance, in which adults scaffold children's developing abilities until they are able to use the language more independently (Syahriani, 2024; Vygotsky, 1978).

Previous research has examined the use of digital media and YouTube in English language learning among children. Most studies, however, have focused on school-aged learners, classroom implementation, or the effectiveness of digital learning media (Anggraini et al., 2022; Dewi & Anggraeni, 2023). Comparatively fewer studies have explored the natural process of English vocabulary acquisition among very young children who are exposed to English-language YouTube content in their everyday home environment. Moreover, limited attention has been given to how children retain and use newly acquired vocabulary in daily communication and how family support facilitates this process. This gap highlights the need for further investigation into second language vocabulary acquisition during early childhood within natural home settings. Furthermore, previous studies have predominantly emphasized formal educational settings rather than natural home environments, leaving limited evidence regarding informal second language acquisition among toddlers (Saville-Troike & Barto, 2017).

Therefore, this study investigates how a three-year-old child acquires English vocabulary through YouTube exposure in a natural home environment. Specifically, the study

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Keysi Muretya, et al.

aims to explore the child's vocabulary acquisition process, examine vocabulary retention and use in daily communication, and investigate the role of family support in second language vocabulary acquisition. The findings are expected to contribute to the growing body of research on early childhood second language acquisition and provide practical insights for parents and educators regarding the effective use of YouTube as a source of English language exposure.

METHOD

This study employed a qualitative case study design to investigate how a young child acquired English vocabulary through exposure to English-language YouTube content in a home-based learning environment. A qualitative case study was selected because it allows researchers to examine a contemporary phenomenon in depth within its natural context and to draw on multiple sources of evidence (Creswell, 2014; Yin, 2018). In this study, the design enabled an in-depth examination of the child's vocabulary acquisition, retention and use of English in everyday communication, and the role of family support in the learning process. Data were collected through naturalistic observation and semi-structured interviews and were analyzed using thematic analysis (Braun & Clarke, 2006).

Participants

The primary participant was a three-year-old female child whose first language was Indonesian and who was learning English as an additional language primarily through informal exposure to English-language YouTube content. To protect the child's identity, she is referred to as Child A throughout this study. Child A had been regularly exposed to English-language YouTube videos since the age of two and received approximately 30–45 minutes of English exposure per day. The exposure was distributed across several short viewing sessions according to the family's daily routine, including approximately 10 minutes during mealtimes and 15 minutes before bedtime. The remaining exposure time occurred during other routine activities and varied according to the child's daily schedule. YouTube therefore functioned as an informal source of English input rather than as part of formal classroom instruction.

Child A was selected purposively because she met two criteria relevant to the research focus: regular exposure to English-language YouTube content and observable spontaneous use of English vocabulary in everyday communication. Her mother and primary caregiver were also included as interview participants because they were directly involved in the child's daily language environment and could provide contextual information about her English exposure, vocabulary development, language use, and family support. Their accounts were used to complement the observational data and provide a broader understanding of the child's English-learning experiences in the home environment.

Data Collection

Data were collected through nine naturalistic observation sessions and semi-structured interviews with the child's mother and primary caregiver. The observations were conducted in the child's natural home environment and focused on the child's responses to English-language input, recognition and use of English vocabulary, spontaneous language production, and use of previously encountered words during everyday activities. Particular attention was given to instances in which the child independently identified objects, produced English words, repeated vocabulary from previous exposure, or used English vocabulary appropriately in communication. Field notes were used to document relevant verbal and behavioral responses during the observation sessions.

Semi-structured interviews were conducted with the child's mother and primary caregiver to obtain additional information about the child's English-learning experiences. The interviews addressed the child's exposure to English-language YouTube content, patterns of vocabulary development, retention and use of English words, family involvement, and the types of interaction and feedback provided during or after viewing. The interviews complemented the observations by providing information about language-learning

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Keysi Muretya, et al.

experiences that might not have been directly observable during the scheduled observation sessions.

Data Analysis

The observation notes and interview data were analyzed using thematic analysis following Braun & Clarke's (2006) framework. The analysis began with repeated reading of the observation notes and interview data to develop familiarity with the dataset. Relevant statements and observed behaviors were then coded according to their relevance to the research focus. Related codes were subsequently grouped into broader categories and reviewed to identify recurring patterns. The categories were refined into themes representing the child's vocabulary acquisition process, retention and use of English vocabulary, and the role of family support. Patterns identified in the observational data were compared with information obtained from the mother and caregiver to strengthen the interpretation of the findings. The analysis therefore integrated direct observations of the child's language use with family accounts of her English-learning experiences.

Ethical Considerations

Because the study involved a young child, ethical procedures were given particular attention. Informed consent was obtained from the child's mother before data collection. The mother was informed about the purpose of the study, data collection procedures, the voluntary nature of participation, and the use of the data for research purposes. The child's identity was protected by using the pseudonym Child A, and identifying personal information was excluded from the report. The observational and interview data were used solely for the purposes of the study and were handled confidentially.

.

FINDINGS AND DISCUSSION

English Vocabulary Acquisition through YouTube Exposure

The findings indicate that the participant acquired English vocabulary through continuous exposure to English-language YouTube content within a natural home environment. Analysis of the observation and interview data generated three interconnected themes describing the vocabulary acquisition process: (1) visual vocabulary acquisition, (2) vocabulary retention and use in daily communication, and (3) family support in English vocabulary acquisition. Together, these themes illustrate that English vocabulary learning was facilitated through consistent audiovisual input, meaningful language use, and continuous support from family members.

Theme 1: Visual Vocabulary Acquisition

The first theme identified in this study is **visual vocabulary acquisition**, which refers to the participant's ability to acquire English vocabulary through visual information presented in YouTube videos. Throughout the observation sessions, the participant consistently recognized and named familiar objects by relying on visual cues without requiring direct translation or explicit vocabulary instruction. This pattern was particularly evident when the participant watched ASMR videos that contained visual demonstrations without narration or subtitles.

The observation findings showed that the participant successfully identified a variety of familiar objects, including fruits, vegetables, animals, and household items, immediately after they appeared on the screen. For example, during Observation Session 6, the participant watched a toy fruit-cutting ASMR video without narration or subtitles for approximately 35 minutes. While watching the video, she spontaneously produced English words such as apple, kiwi, lemon, and strawberry immediately after the fruits appeared on the screen, despite receiving no verbal prompts or translations from the researcher. Because the video contained no spoken language, these responses suggest that the participant relied on previously acquired vocabulary associated with visual representations rather than immediate auditory input.

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Keysi Muretya, et al.

Frequent exposure further strengthened vocabulary acquisition. This pattern became evident across multiple observation sessions. For example, during Observation Session 3, the participant spontaneously produced the expression *"Stuckie duckie, quack quack"* while watching English-language YouTube content, whereas in Observation Session 6, she independently identified fruits such as apple, kiwi, lemon, and strawberry from visual cues alone. These observations indicate that continuous engagement with English-language YouTube content enabled the participant to retrieve previously acquired vocabulary in different contexts without requiring additional explanation or prompting from the researcher. This suggests that recurring visual experiences contributed to stronger lexical representation and facilitated independent vocabulary retrieval over time. The participant no longer depended solely on hearing the words but was able to associate visual objects directly with their English labels.

The interview findings also supported these observations. Both the participant's mother and primary caregiver explained that the child frequently identified fruits, vegetables, animals, and other familiar objects in English while watching YouTube and later mentioned the same vocabulary during play activities. They also observed that repeated viewing of similar videos enabled the participant to recognize English vocabulary more confidently and accurately.

These findings suggest that visual input played an important role in the participant's early English vocabulary acquisition. Consistent with Krashen's (1985) Input Hypothesis, repeated exposure to comprehensible audiovisual input provided meaningful language input that facilitated vocabulary learning. In addition, the combination of visual representations and repeated exposure appeared to strengthen vocabulary acquisition by allowing the participant to form direct associations between objects and their English lexical items. This finding is also consistent with Mayer's (2021) Cognitive Theory of Multimedia Learning, which suggests that vocabulary learning is enhanced when verbal information is presented together with relevant visual representations.

These findings are consistent with previous studies indicating that continous audiovisual exposure facilitates children's English vocabulary development through visual cues, contextualized input, and repetition (Anggraini et al., 2022; Dewi & Anggraeni, 2023). Likewise, Gowenlock et al. (2024) argued that engaging audiovisual input supports vocabulary learning when children receive repeated exposure in engaging contexts. The present study extends these findings by demonstrating that repeated visual exposure enabled the participant to retrieve English vocabulary independently even when no spoken language was available, as observed during the ASMR viewing sessions. This suggests that vocabulary acquisition was supported not only by auditory input but also by the participant's ability to establish stable word–referent associations through repeated visual experiences. This finding further supports Krashen's (1985) Input Hypothesis and Mayer's (2021) Cognitive Theory of Multimedia Learning, emphasizing that repeated and comprehensible audiovisual input contributes to durable vocabulary acquisition in early childhood.

> *For example, the child's mother explained:*
> *["She often repeats English words that she hears from YouTube while watching the videos. Sometimes, after the video ends, she continues mentioning the same words during playtime or when she sees similar objects around the house."]*
> *Similarly, the primary caregiver noted that:*
> *["Repeated exposure to the same YouTube videos helped her remember English vocabulary more easily. She became more confident in naming familiar objects in English without being asked first."]*

Theme 2: Vocabulary Retention and Use in Daily Communication

The second theme identified in this study is vocabulary retention and use in daily communication, which describes how the participant retained previously acquired English

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Keysi Muretya, et al.

vocabulary and applied it spontaneously in everyday situations. The observation and interview findings demonstrated that English vocabulary learned through YouTube exposure was not only remembered but also transferred into everyday communication during daily routines, play activities, and social interactions.

Throughout the observation period, the participant consistently produced English words and expressions outside the YouTube viewing context. Rather than repeating vocabulary immediately after watching videos, the participant used English naturally during everyday activities. During Observation Session 4, for example, she spontaneously said *"Mommy, I'm hungry."* In Observation Session 5, she independently produced the expression "Brush my teeth." During Observation Session 7, she said *"Time to sleep,"* while preparing for bedtime, and in Observation Session 3, she enthusiastically said *"Swim, duckies!"* while watching the Stuckie Duckie, Quack Quack video. These findings indicate that the participant was able to retrieve previously acquired vocabulary and apply it appropriately without immediate audiovisual support.

The participant also demonstrated spontaneous use of English to express emotions and interact with the surrounding environment. During Observation Session 9, after watching a cartoon scene, the participant said, *"Oh sad… I'm feeling blue,"* while pretending to cry. Later, during Observation Session 8, she greeted familiar objects by saying *"Hello, Bintang"* and *"Hello, Bulan."* Although these expressions combined English and Indonesian, they reflected the participant's willingness to communicate by integrating newly acquired English vocabulary into natural interactions.

The interview findings reinforced these observations. The participant's mother explained that English expressions frequently reappeared several hours or even days after watching YouTube videos, particularly during play activities and daily routines. As the mother explained, "after watching YouTube, she often repeated English words later while playing by herself. Sometimes she suddenly said *'Brush my teeth'* or *"Time to sleep"* even though the video had already finished.". Similarly, the primary caregiver reported that the participant often remembered vocabulary, songs, and short expressions from previously viewed videos and later used them appropriately while interacting with family members. The caregiver stated *"She sometimes remembered songs and English words from previous videos and used them naturally while playing or talking with family members."*. These observations suggest that repeated exposure enabled the participant to retain English vocabulary and retrieve it when meaningful communicative situations arose.

These findings indicate that vocabulary acquisition extended beyond simple imitation. This suggests that the participant had begun to develop productive vocabulary, demonstrating the ability to retrieve and use previously learned lexical items in meaningful contexts rather than merely imitating language input (Nation, 2001; Webb & Nation, 2017).

These findings are consistent with previous research showing that repeated exposure to English-language audiovisual content promotes children's vocabulary retention and meaningful language use (Anggraini et al., 2022; Dewi & Anggraeni, 2023). However, the present study provides further evidence that vocabulary learned through YouTube exposure was retained and spontaneously retrieved during everyday activities that were unrelated to video viewing. Rather than merely imitating recently heard expressions, the participant applied previously acquired vocabulary in play, daily routines, and emotional expression. This suggests that repeated exposure contributed not only to vocabulary recognition but also to the gradual development of productive vocabulary, supporting the view that meaningful retrieval strengthens long-term vocabulary knowledge (Lightbown & Spada, 2021; Webb & Nation, 2017).

Theme 3: Family Support in English Vocabulary Acquisition

The third theme identified in this study is family support in English vocabulary acquisition, which highlights the important role of the participant's family in facilitating second language learning through YouTube exposure. Although YouTube provided the primary source of English input, the observation and interview findings demonstrated that

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Keysi Muretya, et al.

vocabulary acquisition was continuously reinforced through daily interactions with the participant's mother and primary caregiver. Across all nine observation sessions, the participant's mother and primary caregiver consistently supported English vocabulary learning. During every observation session, the researcher observed that the mother selected age-appropriate YouTube videos, responded to the participant's English expressions, and encouraged her to use English vocabulary during daily interactions. Similarly, the primary caregiver reinforced the participant's vocabulary by responding positively to English words and incorporating them into everyday conversations and play activities.

The interview findings revealed that the participant's mother and primary caregiver carefully supervised the child's YouTube viewing by selecting age-appropriate educational videos and avoiding content with inappropriate visual elements. Rather than allowing unrestricted access, they actively guided the participant's viewing choices to ensure that YouTube functioned as a learning resource. This supervision enabled the participant to receive consistent exposure to comprehensible English input while minimizing distractions from unsuitable content. The participant's mother explained, *"I always choose educational YouTube videos that are appropriate for her age. If a video contains inappropriate language or scenes, I immediately change it because I want her to learn from suitable content."* This interview evidence indicates that family supervision ensured that YouTube exposure remained appropriate for the participant's age while providing consistent opportunities for English vocabulary learning.

Family interaction also played a significant role in reinforcing vocabulary learning. The interview findings revealed that family members intentionally encouraged the participant to use English during everyday communication. The participant's mother explained that whenever the child produced English expressions, she responded in English rather than immediately switching to Indonesian. For example, the mother recalled that the participant spontaneously said, *"I'm hungry, Mommy,"* during a daily routine, and she responded with expressions such as *"Okay, wait a minute"* or *"I'll make you breakfast"* to encourage continued communication in English. Similarly, the primary caregiver reported asking simple follow-up questions such as *"Why?"* to maintain the interaction and provide opportunities for the participant to use English meaningfully. These findings indicate that responsive family interaction extended the participant's opportunities to practice newly acquired vocabulary beyond YouTube viewing and supported meaningful second language development (Lightbown & Spada, 2021; Vygotsky, 1978; Wahyuni & Tin, 2024).

In addition, the family consistently provided corrective feedback and positive reinforcement throughout the participant's language development. During Observation Sessions 2 and 6, the participant occasionally misidentified English vocabulary, such as confusing pants with skirt and corn with pineapple. In response, family members gently introduced the correct vocabulary and encouraged repetition without criticism. Positive expressions such as *"Good job!"* were also frequently used to motivate the participant and build confidence when using English vocabulary.

These findings demonstrate that YouTube exposure alone was not sufficient to support vocabulary acquisition (Brushe E. et al., 2024; Pediatrics, 2016). This finding supports previous research indicating that digital media is most effective when accompanied by active parental mediation and meaningful interaction (Pediatrics, 2016). Instead, continuous family involvement provided meaningful interaction, guided practice, and emotional support that strengthened the participant's vocabulary development.

These findings are consistent with Vygotsky's Sociocultural Theory Vygotsky (1978), which emphasizes that children's language development is facilitated through social interaction and guidance from more knowledgeable individuals. Unlike previous studies that generally emphasize the importance of parental involvement, the present study illustrates how family support was enacted through specific practices, including selecting age-appropriate YouTube content, responding to the child's English expressions, providing corrective feedback, and extending English use into everyday interactions. These findings demonstrate that family support functioned not merely as supervision but as meaningful scaffolding that enabled the child to use English vocabulary beyond the immediate viewing

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Keysi Muretya, et al.

context. As Syahriani (2024) argues, drawing on Vygotsky's concept of the Zone of Proximal Development, what is important is how such assistance contributes to the child's learning and development. Similarly, Lantolf et al. (2015) emphasized that social interaction and scaffolding play a central role in second language development. Similarly, Dewi & Anggraeni (2023) reported that family involvement encourages children to practice and maintain newly acquired English vocabulary in everyday communication. Therefore, the combination of comprehensible digital input and supportive family interaction created a supportive learning environment that promoted second language vocabulary acquisition during early childhood. This finding further supports the view that successful early second language acquisition is facilitated through the integration of meaningful input and responsive social interaction (Lightbown & Spada, 2021).

Challenges Encountered during English Vocabulary Acquisition through YouTube Exposure

The second research question explored the challenges encountered during the participant's English vocabulary acquisition through YouTube exposure. Although the findings demonstrated that YouTube provided meaningful language input, the observation and interview data revealed several factors that influenced the participant's vocabulary learning process. Three major challenges emerged from the analysis: (1) difficulties in understanding spoken language, (2) limited attention during viewing sessions, and (3) pronunciation difficulties. These challenges did not prevent vocabulary acquisition but influenced the pace and effectiveness of the participant's learning process.

Difficulty in Understanding Spoken Language

One challenge identified during the study was the participant's difficulty in understanding English spoken in videos containing continuous conversations. According to both the participant's mother and primary caregiver, videos such as Bing presented language at a speed and level of complexity that was more difficult for the participant to understand than videos emphasizing songs or visual demonstrations. As the participant's mother explained, *"In my opinion, the challenges experienced by the child depend on the type of YouTube channel she watches. For channels such as Bing or BabyBus that contain more conversations, she needs more time to understand what the characters are saying."* Similarly, the primary caregiver stated, *"I noticed that my granddaughter experienced difficulties, particularly when watching Bing videos that contained more conversational language. I believe the pronunciation used in those videos was relatively difficult for her to follow, so she required more time to imitate some of the English vocabulary."* Compared with songs or videos emphasizing visual demonstrations, conversational videos required greater listening ability, making vocabulary comprehension more challenging.

Despite these difficulties, the participant continued to benefit from repeated exposure. Both the participant's mother and primary caregiver explained that repeated viewing gradually helped the participant become familiar with English words and expressions, although complete understanding was not always achieved during the initial viewing. The participant's mother explained, *"When my child learns new vocabulary from YouTube, I usually explain its Indonesian equivalent so that she can understand both languages."* Likewise, the primary caregiver reported *"At first, my granddaughter only moved her lips while following the words she heard without producing any sound. After the video was played again, she gradually began to pronounce some of the words. After repeated viewings, she was able to sing the songs and pronounce the words from the videos."* These findings suggest that the complexity of spoken input may influence vocabulary acquisition, particularly when young children are exposed to natural conversational speech. This finding is consistent with Krashen's (1985) Input Hypothesis, which proposes that language acquisition occurs most effectively when learners receive comprehensible input appropriate to their level of understanding. Similarly, Gowenlock et al. (2024) reported that children's language development through video-based media is influenced by the accessibility and comprehensibility of linguistic input. Furthermore, Brushe E. et al. (2024) emphasized that children's language learning from digital media

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

Copyright (c) 2026 Keysi Muretya, et al.

becomes more effective when language exposure is supported by responsive adult guidance and repeated interaction.

Limited Attention During Viewing Sessions

The observation and interview findings also indicated that the participant's attention span influenced the vocabulary learning process. During Observation Session 5, while watching nursery rhyme videos from the ABC Kids TV YouTube channel, the participant gradually lost interest during the song Skip to My Lou and later became frightened by another nursery rhyme containing unfamiliar animated visual elements. As a result, she stopped watching and requested that the caregiver change the video. Similarly, during Observation Session 7, while watching a Savya ASMR toy food-cutting video, the participant left the viewing area after several minutes and continued playing in another room. These observations suggest that boredom, fatigue, unfamiliar visual stimuli, and shifting attention occasionally interrupted continuous exposure to English vocabulary.

The participant's mother explained that the participant often stopped watching videos before they ended when she became tired, bored, frightened by unfamiliar characters, or preferred to play instead. As the mother explained:

"The child stopped watching when she encountered characters that she considered unfamiliar or frightening, when she became bored or tired, when she preferred to play, or when she felt hungry, thirsty, or wanted to take a bath."

Similarly, the primary caregiver reported:

"My granddaughter stopped watching when she became bored, sleepy, or frightened by certain visual elements appearing in the videos. In those situations, she usually came to me or went into her room." These interview findings support the observations by showing that temporary interruptions during YouTube viewing were mainly influenced by the participant's emotional state and attention span rather than by a lack of interest in English itself.

These findings indicate that children's emotional responses and attention span influence the effectiveness of YouTube as a source of language input. Although repeated audiovisual exposure supports vocabulary acquisition, sustained learning depends on the child's continued engagement with the content. Therefore, active parental supervision through selecting age-appropriate videos, limiting viewing duration, and responding to the child's emotional needs plays an important role in maintaining meaningful language exposure. This interpretation is consistent with Brushe E. et al. (2024), who reported that children's screen-based learning is more effective when accompanied by responsive parental involvement and appropriate media selection.

This finding is also consistent with Linebarger & Walker (2005), who argued that children's learning from television is influenced not only by exposure itself but also by attention to educational content. Likewise, Anderson & Hanson (2017) emphasized that young children's language learning from digital media is most effective when audiovisual content successfully maintains their attention and is supported by adult guidance.

Pronunciation Difficulties

Another challenge observed during the study involved pronunciation accuracy. Although the participant successfully acquired and used many English words, some vocabulary items were occasionally pronounced inaccurately due to the participant's developing speech abilities. For example, during Observation Session 6, when identifying the vegetable broccoli from the ASMR video, the participant pronounced the word as "bokoli", omitting the English /r/ sound. Despite the pronunciation error, the participant successfully identified the intended object, indicating that the lexical meaning had been acquired even though articulation remained developmentally incomplete.

The observation further showed that family members did not interrupt or criticize the participant's pronunciation. Instead, they naturally modelled the correct pronunciation during subsequent interactions while allowing the participant to continue communicating

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

confidently. This supportive response provided repeated exposure to the correct pronunciation without discouraging the participant's willingness to use English.

These findings suggest that pronunciation difficulties are a natural part of early second language development and can be reduced through repeated language input and supportive social interaction. This finding is consistent with Vygotsky's Sociocultural Theory (Vygotsky, 1978), which emphasizes the importance of guidance from more knowledgeable individuals during children's learning process. However, adult assistance within the ZPD should not be understood simply as the presence of a more knowledgeable person who corrects or helps the child. As Syahriani (2024) emphasizes, what is important is the function of the assistance in relation to the child's learning and development.

In the present study, this can be seen during Observation Session 6, when the participant pronounced broccoli as "bokoli" while identifying the vegetable shown in the ASMR video. The omission of the /r/ sound reflects a pronunciation approximation that is common among young children whose speech development is still emerging. Rather than interrupting or criticizing the participant, the caregiver responded naturally and continued providing English input during subsequent interactions. Although the participant's pronunciation was not yet fully accurate, she consistently attempted to use English vocabulary when communicating. This finding illustrates Vygotsky's notion that adult assistance is not limited to correcting errors directly, but also involves providing meaningful support that encourages children to continue participating in language use (Syahriani, 2024; Vygotsky, 1978).

CONCLUSIONS

This study concludes that the three-year-old child acquired English vocabulary through repeated exposure to English-language YouTube content and supportive family interaction. Repeated audiovisual exposure helped the child recognize, understand, and retain new vocabulary through visual and auditory associations. Some acquired words were later recalled and used spontaneously during play, daily routines, and emotional expressions, indicating their integration into the child's everyday language repertoire. Family involvement further supported vocabulary development through guided media exposure, responsive interaction, and corrective feedback. Although the child experienced difficulties with conversational English, limited attention, and pronunciation, these challenges did not prevent vocabulary development and were addressed through continued exposure and family support. Thus, YouTube functioned not as a standalone source of vocabulary acquisition but as a resource whose effectiveness was strengthened by social interaction at home. The findings highlight the importance of combining accessible digital English input with active family involvement. Future research should examine longer periods and more diverse children, families, and digital learning environments.

ACKNOWLEDGEMENTS

The author would like to express sincere gratitude to the participant and her family for their cooperation and willingness to participate in this study. The author also extends heartfelt appreciation to Ma'am Titin Rahmiatin Rahim and Ma'am Ririn Syahriani for their invaluable guidance, continuous support, and constructive feedback throughout the research process and the preparation of this manuscript.

REFERENCES

Anderson, D. R., & Hanson, K. G. (2017). Screen media and parent-child interactions. In *Media exposure of content and context on learning and development* (pp. 173–194). https://doi.org/10.1007/978-3-319-45102-2_11

Anggraini, P. P., Fitria, L., Listiawati, L., Apriliani, N. A., Supeni, I., & Handrianto, C. (2022). The use of the cocomelon youtube channel as a medium for introducing children's English vocabulary. SAGA: Journal of English Language *Teaching and Applied Linguistics,*

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.

3(2), 81–90 https://doi.org/10.21460/saga.2022.32.137
Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101.
Brushe E., M., Hagg, D. G., Melhunish, E. C., Reily, S., & Gregory, T. (2024). Screen time and parent-child talk when children are aged 12 to 36 months. *JAMA Peditrics*, *178*(4), 369–375. https://doi.org/10.1001/jamapediatrics.2023.6790
Creswell, J. W. (2014). Qualitative inquiry and research design: Choosing among five approaches. In *Sage Publications*.
Dewi, H. C., & Anggraeni, D. (2023). Acquisition of English vocabulary for children aged 3 years through youtube social media and habit forming. Edunesia: Jurnal Ilmiah Pendi*dikan, 4(1), 197–209* https://doi.org/10.51276/edu.v4i1.295
Ellis, R. (2002). The study of second language acquisition. In Oxford University Press https://doi.org/10.1075/lald.27.12roc
Gowenlock, A. E., Norbury, C., & Rodd, J. M. (2024). Exposure to language in video and its impact on linguistic development in children aged 3–11: A scoping review. J*ournal of Cognition, 7(*1), 57–200. https://doi.org/10.5334/joc.385
Hoff, E. (2013). Language development. In *Wadsworth Cengage Learning*.
Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. Longman.
Lantolf, J. P., Thorne, S. L., & Poehner, M. E. (2015). Sociocultural theory and second language development. In *Theories in second language acquisition: An introduction* (pp. 207–226).
Lenneberg, E. H. (1967). Biological foundations of language. In Wiley https://doi.org/10.1212/wnl.17.12.1219
Lightbown, P. M., & Spada, N. (2021). How language are learned. In *Oxford University Press*.
Linebarger, D. L., & Vaala, S. E. (2020). Screen media and language development in infants and toddlers: An ecological perspective. D*evelopmental Review, 30(*2), 176–202.
Linebarger, D. L., & Walker, D. (2005). Infants' and toddlers' television viewing and language outcomes. *American Behavioral Scientist*, *48*(5), 624–645. https://doi.org/10.1177/0002764204271505
Mayer, R. E. (2021). Multimedia learning. In *Cambridge University Press*.
Nation, I. S. P. (2001). Learning vocabulary in another language acquisition. In *Cambridge University Press*.
Pediatrics, A. A. of. (2016). Media and young minds. *Pediatrics*, *138*(5). https://doi.org/10.1542/peds.2016-2591
Saville-Troike, M., & Barto, K. (2017). Introducing second language acquisition. In *Cambridge University Press*.
Syahriani, R. (2024). Revisiting Vygotsky's zone of proximal development: Misinterpretations and practical applications in education. *International Conference on Research Issues (ICORES) 2024*, 76–83.
Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. In *Harvard University Press*.
Wahyuni, N. T., & Tin, T. B. (2024). Beyond the classroom walls: Exploring parental involvement on children's interest development in EFL learning (A case from Indonesia). *Education 3–13*. https://doi.org/10.1080/03004279.2024.2340548
Webb, S., & Nation, P. (2017). How vocabulary is learned. In *Oxford University Press*.
Yin, R. K. (2018). Case study research and applications: Design and methods. In *SAGE Publications*.

OPEN ACCESS
© 2021 The Author.This article is licensed CC BY SA 4.0.
visit Creative Commons Attribution-ShareAlike 4.0 International License.