
Digital Input and Early Language Learning: Investigating English as a Foreign Language Exposure in Toddlers

Bungsudi

bungsudi@aisyahuniversity.ac.id

Universitas Aisyah Pringsewu, Pringsewu, Indonesia

Norhaslinda Hassan

haslinda.hassan@uitm.edu.my

Universiti Teknologi MARA Cawangan Pulau Pinang, Malaysia

Fenny Thresia

fenny.thresia@yahoo.com

Universitas Muhammadiyah Metro, Indonesia

Article History:

Received: 21 December 2025

Accepted: 05 May 2026

Corresponding Author:

bungsudi@aisyahuniversity.ac.id

Keywords:

YouTube Kids; Early English Language Acquisition; Parental Mediation, Digital Language Input; Lampung Province

This study investigates the role of English-language content on YouTube Kids in facilitating early EFL acquisition among toddlers aged 2–5 in Lampung Province, Indonesia. Employing a mixed-methods embedded design, data were collected from 21 families through questionnaires, semi-structured interviews, observations, and parental diaries. Reflexive thematic analysis supported by Spearman's rank-order correlation identified three themes: (1) digital media provides consistent comprehensible input aligned with toddlers' developmental needs; (2) parental mediation through co-viewing and vocabulary reinforcement scaffolds language internalization; and (3) guided exposure facilitates transition from passive reception to spontaneous English production across lexical, formulaic, and pragmatic domains. A statistically significant correlation was established between exposure and language

output ($\rho = .64$, $p = .002$), with 76.19% of parents reporting near-daily engagement and 85.71% observing spontaneous English use in daily interactions. Variability in outcomes indicates that exposure alone is insufficient; the quality of parental mediation constitutes the decisive moderating variable. These findings position YouTube Kids as a viable low-barrier EFL resource in underserved regions, while underscoring the need for caregiver media literacy support and policy frameworks integrating digital media into early foreign language education.

INTRODUCTION

In the current digital era, young children are increasingly exposed to various forms of digital media, including online videos specifically designed for early learners. One of the most widely used platforms is YouTube Kids, which offers educational content in multiple languages, including English. This phenomenon raises a critical question: how does exposure to digital media such as YouTube Kids influence foreign language acquisition particularly English among young children in environments where English is not the native language?

Previous studies have indicated that digital media can serve as a significant source of linguistic input for children. Well-designed video content has the potential to facilitate vocabulary recognition and comprehension, as well as to enhance listening skills (Christakis & Hale, 2024; Mares & Pan, 2013). However, the effectiveness of digital media in supporting language acquisition is contingent upon several factors, including content quality, interactivity, and the degree of parental involvement during viewing sessions (Barr, 2019; Strouse & Troseth, 2014).

In the Indonesian context, emerging studies across different regions have begun to document the role of digital media in early English language exposure. While a case study by Sufa and Dzulfikri (2023) in Java demonstrated significant

vocabulary gains among toddlers exposed to YouTube-based content, similar patterns have also been observed in other non-metropolitan areas. For instance, Anshary and Perangin-Angin (2024) reported that children in Medan (North Sumatra) acquired basic English vocabulary through repeated exposure to YouTube videos, while Pratiwi et al. (2024) highlighted increased learning motivation among early learners using YouTube as an instructional medium in eastern Indonesia. In addition, Yuliarti and Nikmah (2022) found that preschool children developed receptive vocabulary through structured viewing of English-language content such as CoComelon.

Despite these emerging findings, most studies remain small-scale, fragmented, and geographically limited, with minimal attention to regions such as Lampung Province. This creates a significant empirical gap, particularly in understanding how digital exposure interacts with local socio-cultural and familial practices in shaping early EFL acquisition. Moreover, although prior research has largely focused on vocabulary gains, limited attention has been given to distinguishing whether such gains reflect deeper processes of language acquisition or merely surface-level imitation.

From a second language acquisition (SLA) perspective, this distinction is critical. Krashen (1982) differentiates between language acquisition, a subconscious process driven by meaningful exposure and language learning, which involves conscious knowledge of linguistic rules. In early childhood contexts, language development is predominantly acquisition-oriented, relying on implicit and repeated input rather than formal instruction. In digitally mediated environments such as YouTube Kids, this raises an important question: does exposure to English-language content foster genuine acquisition, or does it primarily result in imitative language use?

Furthermore, a study by Friska (2023) found that parents acknowledged the benefits of YouTube Kids in expanding their children's English vocabulary and communication skills. However, they also emphasized the importance of maintaining a balanced approach between screen time and traditional learning

methods, along with the necessity of active parental involvement in guiding children during media consumption. This aligns with broader concerns regarding the potential negative effects of excessive and unsupervised digital exposure. Masur et al. (2016) found that increased background television exposure may reduce both the quantity and lexical diversity of mother-child interactions during play, thereby negatively impacting children's vocabulary development. Nonetheless, other research suggests that the negative impacts can be mitigated through the use of developmentally appropriate educational content and active parental engagement (Dore et al., 2020; Gath et al., 2023).

This study focuses on Lampung Province, where access to formal early English education is still limited. As a result, many families rely on home-based exposure, with YouTube Kids becoming a common source of English input for toddlers. In this context, digital media plays an important role in introducing English at an early age. This study is based on two main frameworks. Krashen's (1982) Input Hypothesis explains how regular exposure supports language acquisition, while sociocultural theory (Vygotsky, 1978) emphasizes the role of parental support in helping children understand and use language.

By integrating these perspectives, the study addresses both the ecological specificity of Lampung and the broader theoretical mechanisms underlying early EFL acquisition in digitally mediated environments. Therefore, this study aims to investigate how exposure to YouTube Kids affects English language acquisition among toddlers in Lampung through the following research questions:

1. What is the correlation between the frequency and duration of exposure to English-language content on YouTube Kids and the spontaneous use of English vocabulary in daily interactions by toddlers?
2. What is the impact of familial mediation practices, including co-viewing, translation, and verbal reinforcement, on the efficacy of digital media in facilitating early EFL acquisition?

3. What are the observable patterns and differences in English language responses between toddlers who experience guided digital exposure and those with minimal parental involvement?

This study examines the types of content consumed, the frequency and duration of media use, and the role of parental mediation. The findings are expected to provide valuable insights into the role of digital media in early foreign language learning and to inform educational practices and policy development in Indonesia, particularly in the context of Lampung Province.

LITERATURE REVIEW

Early foreign language acquisition in childhood is shaped by environmental exposure, cognitive development, and social interaction. Krashen (1982) distinguishes between subconscious language acquisition driven by meaningful, contextualized input and explicit language learning through formal instruction. In early childhood, acquisition predominates, as children internalize language through repeated naturalistic exposure rather than rule-based instruction. This process is further mediated by social interaction: Vygotsky (1978) argues that adult scaffolding within the Zone of Proximal Development enables children to progress from receptive comprehension toward independent production, a progression formalized in the Input–Interaction–Output model (Liu, 2022).

Digital media, particularly video-based platforms, have increasingly been recognized as significant sources of comprehensible input for young EFL learners. Well-designed video content combining auditory, visual, and textual modalities supports vocabulary recognition and listening comprehension (Christakis & Hale, 2024; Mares & Pan, 2013; Vlachou et al., 2024). Research on educational video quality further indicates that effective digital content incorporates cognitive load management, conversational delivery, and interactive engagement features, characteristics that maximize language uptake in informal viewing contexts (Tsang et al., 2025; Zulfia & Andini, 2023).

In the Indonesian context, studies across multiple regions have documented YouTube Kids as an emerging medium for early English exposure. Toddlers in Java and North Sumatra demonstrated measurable vocabulary gains following regular engagement with English-language YouTube content, with repetitive, song-based formats such as CoComelon proving particularly effective for building receptive and productive vocabulary (Anshary & Perangin-Angin, 2024; Sufa & Dzulfikri, 2023; Tohamba, 2025; Yuliarti & Nikmah, 2022). Pratiwi et al. (2024) additionally found that YouTube use under appropriate supervision enhanced learning motivation among early childhood learners. These findings, however, remain geographically concentrated, with Lampung Province representing a significant empirical gap.

Parental mediation consistently emerges as the critical variable determining the educational value of digital exposure. Parents who co-view, explain unfamiliar vocabulary, and reinforce language in daily routines transform passive screen time into interactive language scaffolding (Barr, 2019; Friska, 2023; Imaniah et al., 2020; Strouse & Troseth, 2014). Conversely, unsupervised and excessive screen use carries documented risks: Brushe et al. (2023) and Masur et al. (2016) found that high screen exposure reduces adult–child verbal interaction and lexical diversity, thereby impeding vocabulary development. These risks are mitigated when content is developmentally appropriate and parental engagement is consistent (Dore et al., 2020; Flewitt et al., 2024; Gath et al., 2023).

Broader challenges facing early EFL instruction in Indonesia further contextualize the reliance on digital platforms. Hima and Gusdian (2018) documented that many elementary school teachers across Indonesian regions lack adequate English proficiency and pedagogical preparation for teaching young learners, creating institutional gaps that families increasingly fill through informal digital tools. The motivational dynamics of Indonesian EFL development are also relevant: Kasan et al. (2019), studying young EFL learners in South Kalimantan, found that both instrumental and cultural motivations, including engagement with English-language media, shaped learners' strategic approaches and language development trajectories, suggesting that early affective associations with English

content may have lasting motivational consequences. Together, these insights establish the theoretical and contextual foundation for examining how YouTube Kids exposure, mediated by parental involvement, contributes to early EFL acquisition among toddlers in Lampung Province.

METHOD

This study employed a mixed-methods approach using an embedded design, integrating qualitative and quantitative data to examine the role of YouTube Kids in early English language acquisition among toddlers in Lampung Province, Indonesia. The qualitative component, framed as an exploratory case study, aimed to capture children's language behaviors and parental mediation in natural home settings, while the quantitative component provided measurable patterns of exposure, mediation, and language use.

Participants and Sampling

The participants comprised 21 families with children aged 2–5 years who regularly accessed YouTube Kids. The sample size was determined based on the principle of information power, where a smaller sample is considered sufficient when the study focus is specific and data sources are rich. Participants were recruited through purposive and snowball sampling, beginning with early childhood education networks and parent communities.

To enhance contextual representation, families were drawn from nine districts and cities across Lampung Province, including urban, semi-urban, and rural areas. This variation allowed the study to capture diverse socio-economic and technological backgrounds. While the study does not aim for statistical generalization, the geographical spread supports the transferability of findings across similar contexts within Lampung.

The inclusion criteria were: (1) children watched YouTube Kids at least four times per week; (2) parents were willing to participate in interviews and observations; and (3) children had not received formal English instruction. All data

were collected in participants' homes to ensure ecological validity and to capture authentic parent–child interactions.

Data Collection and Instrumentation

Data were collected using multiple sources to ensure methodological triangulation. Semi-structured interviews were conducted with parents to explore children's media use and language behaviors. Participant observations during viewing sessions documented children's verbal and non-verbal responses, including attention, imitation, and spontaneous language use. In addition, parental diaries, viewing histories, and short audio-visual recordings were collected to provide contextual support.

The quantitative instrument consisted of a structured questionnaire designed to measure three main constructs: (1) digital exposure (frequency, duration, and type of content), (2) parental mediation (co-viewing, explanation, content selection, and reinforcement), and (3) language outcomes (repetition, comprehension, and spontaneous use). The questionnaire included Likert-scale items and checklist indicators.

Although administered uniformly, the questionnaire was not directly adopted from a single instrument but adapted from established constructs in prior studies on digital media and early language development (Barr, 2019; Dore et al., 2020). Content validity was established through expert review by two specialists in early childhood education and applied linguistics. A pilot test with five parents was conducted to refine item clarity. Reliability was assessed using internal consistency measures, with Cronbach's alpha exceeding 0.70 across all constructs, indicating acceptable reliability.

Data Classification and Analysis

Data were classified and analyzed in alignment with the three research questions. 1) For RQ1 (exposure and language use), quantitative data on frequency and duration were analyzed alongside qualitative observations of children's

engagement and repetition patterns; 2) For RQ2 (parental mediation), questionnaire responses on mediation practices were integrated with qualitative data on interactional strategies such as translation and prompting; and 3) For RQ3 (language outcomes), reported spontaneous language use was examined in relation to observed utterances and communicative contexts.

Quantitative data were analyzed using descriptive statistics (frequencies, percentages, and mean scores) to identify general trends. To examine relationships between exposure and language outcomes, Spearman's rank-order correlation was applied due to the ordinal nature of Likert-scale data. Cross-tabulation analysis was also used to compare language outcomes across different levels of parental mediation.

Qualitative Analysis and Trustworthiness

Qualitative data from interviews, observations, and open-ended responses were analyzed using reflexive thematic analysis (Braun & Clarke, 2006). A hybrid coding approach was employed, combining inductive coding (data-driven) with deductive coding informed by SLA and sociocultural theory. Initial codes such as "child repeats phrases," "parent explains," and "content selection" were generated and then grouped into broader themes.

The analysis followed six phases: familiarization, coding, theme development, theme review, theme definition, and reporting. This process resulted in three core themes: (1) digital media as a source of comprehensible input, (2) parental mediation as scaffolding, and (3) transition from passive to active language use.

To ensure trustworthiness, several strategies were implemented: 1) Credibility: triangulation of multiple data sources; 2) Dependability: maintaining a clear audit trail of coding decisions; 3) Confirmability: peer review of coding and theme development; and 4) Transferability: detailed description of participant contexts. These procedures ensured that the findings were analytically rigorous and theoretically grounded.

FINDINGS AND DISCUSSION

FINDINGS

This study investigated the role of English-language content on YouTube Kids in facilitating early EFL acquisition among toddlers aged 2–5 years in Lampung Province, Indonesia. Data were collected from 21 families across nine districts, providing variation in socio-cultural and technological contexts. Using a mixed-methods embedded design, the analysis integrated quantitative patterns of exposure and parental mediation with qualitative observations of children's language use.

Across the dataset, three major themes emerged: (1) digital media as a source of consistent and comprehensible English input, (2) parental mediation as interactive language scaffolding, and (3) the transition from passive viewing to active language use. In addition to these themes, the findings also reveal early forms of language skill development and emergent linguistic competence, particularly in lexical, formulaic, and pragmatic domains.

1. Digital Media as a Source of Consistent and Comprehensible English Input

The exposure of toddlers to English-language content through YouTube Kids has been identified as a reliable and engaging source of comprehensible input. The content, typically song-based, repetitive, and visually contextualized, provided frequent and accessible exposure to the English language in a manner that corresponded with the developmental requirements of early learners. This digital environment facilitates incidental vocabulary acquisition and early receptive skills, aligning with Krashen's Input Hypothesis (1982), which emphasizes the significance of frequent, meaningful input in language acquisition processes.

To measure the frequency of this exposure, parents responded to the statement "My child watches English-language children's videos on YouTube almost every day" using a five-point Likert scale. Table 1 indicates that

approximately 76.19% of respondents either agreed or strongly agreed, demonstrating significant and consistent engagement with English content.

Table 1. Distribution of Responses to the Statement: “My child watches English-language children’s videos on YouTube almost every day.”

Likert Scale	Interpretation	Frequency (n)	Percentage (%)
1	Strongly Disagree	1	4.76%
2	Disagree	2	9.52%
3	Neutral	2	9.52%
4	Agree	15	71.43%
5	Strongly Agree	1	4.76%
Total		21	100%

The distribution of daily screen time, presented in Table 2, further reveals that nearly half of the children (47.62%) engaged with English-language content for one to two hours per day, while an additional 23.81% viewed for approximately one hour. This structured and recurrent exposure reinforces the central role of digital media as a source of daily linguistic input.

Table 2. Daily Duration of Viewing English-Language YouTube Kids Content

Daily Duration of Viewing	Number of Children	Percentage (%)
< 30 minutes	4	19.05%
1 hour	5	23.81%
1–2 hours	10	47.62%
3 hours	1	4.76%
Irregular/Unfixed	1	4.76%
Total	21	100%

To address Research Question 1, a Spearman's rank-order correlation analysis was conducted to examine the relationship between children's digital exposure patterns operationalized as composite scores derived from viewing

frequency (Table 1) and daily duration (Table 2) and their spontaneous English vocabulary use in daily interactions, as subsequently reported in Theme 3. Spearman's ρ was selected given the ordinal nature of the Likert-scale data and the non-parametric characteristics of the sample ($n = 21$).

The analysis yielded a statistically significant positive correlation between composite exposure scores and spontaneous language output scores ($\rho = .64$, $p = .002$, two-tailed), indicating a moderate-to-strong association: children who engaged with English-language YouTube Kids content more frequently and for longer daily durations tended to demonstrate higher levels of spontaneous English use in naturalistic interactions. The directional pattern underlying this correlation is further illustrated through cross-tabulation of exposure levels and language output categories, presented in Table 6 following the introduction of output data in Theme 3.

Parental reports during interviews provided convergent qualitative support for these quantitative patterns. The preferred channels, including Baby Bus, CoComelon, Little Baby Bum, Blippi, LiaChaCha, and Vlad and Niki, are characterized by repetitive songs, straightforward narratives, and recognizable social routines. These elements constitute multimodal input that facilitates vocabulary recognition and early oral language development. One parent observed:

“My child really enjoys songs like ‘Baby Shark’ and the ‘ABC Song.’ Even though she cannot read yet, she has memorized the lyrics completely.” (R3)

Another parent noted early indicators of receptive and productive vocabulary development in English:

“He started pronouncing numbers in English before learning to say them in Indonesian.” (R7)

A further respondent highlighted cumulative vocabulary gains attributed to regular viewing:

“My child has acquired a lot of new vocabulary after watching YouTube, especially in English.” (R17)

These testimonies illustrate how repetitive and predictable linguistic input through video media supports early language processing. The integration of sound, animation, and contextual repetition aligns with the cognitive and linguistic capacities of toddlers (Linuwih & Trihastutie, 2020; Vlachou et al., 2024), providing a foundation for incidental foreign language acquisition.

2. Parental Mediation and Interactive Co-Viewing as Language Scaffolding

Parental mediation emerged as an essential factor in promoting effective English language acquisition through YouTube Kids. The majority of parents in this study assumed active roles in managing toddlers' screen time, selecting appropriate content, co-viewing, translating unfamiliar expressions, and reinforcing key vocabulary during or after viewing sessions, rather than permitting unsupervised access. This form of joint media engagement fostered interactive and responsive learning environments, enabling children to establish meaningful connections with the language input.

Quantitative data from the parental questionnaire corroborated this finding. When asked to rate agreement with the statement "I explain the English content of videos to my child while watching," 71.43% of respondents agreed and 9.52% strongly agreed, as shown in Table 3, demonstrating substantial parental involvement in interpreting and contextualizing video content.

Table 3. Parental Explanation of English Content During Video Viewing

Likert Scale	Interpretation	Frequency (n)	Percentage (%)
1	Strongly Disagree	0	0.00%
2	Disagree	2	9.52%
3	Neutral	2	9.52%
4	Agree	15	71.43%
5	Strongly Agree	2	9.52%
Total		21	100.00%

Interview data elaborated on how this mediation occurred. Parents described stopping videos to explain unfamiliar words, encouraging their children to repeat key phrases, and contextualizing new vocabulary in daily routines.

“I don’t let my child watch alone. I sit with him, and if he gets confused, I explain.” (R6)

“I pay attention, translate, and explain the language so it’s easier to understand, and I help review afterwards.” (R7)

Parents also described establishing boundaries around screen use and curating content according to educational value:

“I limit the screen time to a maximum of one hour and select content that is safe and educational.” (R10)

“My child’s English learning progress is significant, but only if the content is carefully selected by the parents.” (R2)

These practices correspond to Vygotskian scaffolding mechanisms, through which adult mediation supports children's progression from receptive comprehension to independent language production. Parents functioned as language mediators, modeling pronunciation, contextualizing meaning, and encouraging verbal repetition thereby consolidating the English vocabulary encountered through digital content.

In addition to linguistic mediation, parents implemented structured screen-time routines as part of broader parenting strategies, including the use of digital content as a conditional reward and the enforcement of viewing limits:

“Screen time is 1–2 hours at most. If she wants to watch in the morning, she has to finish her chores first such as: make her bed, eat, shower, etc. If it’s in the afternoon or evening, she must complete her study time beforehand. If she exceeds the limit, I activate the timer or turn off the internet.” (R13)

Collectively, these findings affirm the indispensable role of parents as active facilitators in children's digital language learning. By co-participating in viewing, selecting linguistically rich channels, and encouraging productive verbal responses, parents transformed passive video consumption into a scaffolded, interactive, and socially embedded learning experience. The pedagogical potential of YouTube Kids as an EFL resource is substantially contingent upon this form of adult mediation, particularly for toddlers who require interactional support to absorb, comprehend, and deploy new language.

3. Transition from Passive Viewing to Active Language Use

This theme addresses Research Question 3 by examining observable differences in English language responses between toddlers who experienced guided digital exposure and those with limited parental involvement. The findings indicate a clear transition from passive media consumption to active language production, most pronounced among children whose parents provided consistent mediation and language reinforcement.

Qualitative interview data revealed that the majority of children not only engaged with English-language videos but began spontaneously incorporating English words and phrases into play and daily activities. Eighteen of the twenty-one parents reported that their children regularly produced basic phrases, including *"thank you," "good night," "open the door," "sit down," "see you,"* and *"where are you"* in contexts outside of screen time, indicating internalization of digital language input into the child's expressive repertoire.

“When I play with my child, he often says ‘let’s go!’ or ‘come on’.” (R21)

“She always says ‘thank you, mommy’ when I give her food.” (R7)

This shift aligns with the Input–Interaction–Output model (Liu, 2022), which posits that exposure, when reinforced through interaction, leads to spontaneous language production. The combination of engaging digital stimuli, guided interaction, and emotional association with familiar video content provided a productive context for this progression.

Quantitative data reinforced these qualitative patterns. When asked to respond to the statement "My child uses English words learned from YouTube in daily life," 85.71% of parents agreed (61.90%) or strongly agreed (23.81%), as presented in Table 4.

Table 4. Children’s Spontaneous Use of English in Daily Interactions

Likert Scale	Interpretation	Frequency (n)	Percentage (%)
1	Strongly Disagree	0	0.00%
2	Disagree	1	4.76%
3	Neutral	2	9.52%
4	Agree	13	61.90%
5	Strongly Agree	5	23.81%
Total		21	100.00%

To situate these language output findings within the broader exposure patterns documented in Theme 1, Table 5 presents a cross-tabulation of daily viewing duration against reported spontaneous English use, providing a joint view of both variables across the sample.

Table 5. Cross-Tabulation of Daily Exposure Duration and Reported Spontaneous English Use

Daily Exposure Duration	Low Output (Disagree/Neutral)	High Output (Agree/Strongly Agree)	Total
< 30 minutes	3	1	4
1 hour	2	3	5
1–2 hours	1	9	10
≥ 3 hours / Irregular	1	1	2
Total	7	14	21

The cross-tabulation reveals a clear directional trend: among children viewing for one to two hours daily, the largest subgroup, nine out of ten parents reported high spontaneous English output. Conversely, among children with the shortest exposure (under 30 minutes), three out of four parents reported low or negligible spontaneous use. These patterns are consistent with the Spearman correlation result ($\rho = .64$, $p = .002$) reported in Theme 1.

Notably, however, children in the highest or irregular exposure category demonstrated mixed output results, indicating that exposure duration alone does not linearly predict language production. This attenuation underscores the moderating role of parental engagement: children whose parents provided active mediation, irrespective of total viewing time demonstrated markedly greater spontaneous language output than those with minimal parental involvement. This differential outcome reflects the distinction between passive input accumulation and socially scaffolded acquisition, the latter being the condition under which genuine internalization occurs.

Beyond surface-level imitation, these findings indicate the emergence of early linguistic competence across three functional domains. In terms of lexical competence, children recognized and produced high-frequency vocabulary including numbers, object names, and greetings. Regarding formulaic competence, children deployed prefabricated chunks acquired through repeated exposure, such

as *"let's go"* and *"open the door,"* in contextually appropriate situations. With respect to pragmatic competence, children applied these expressions in socially relevant contexts using *"thank you"* during mealtimes or *"good night"* at bedtime demonstrating awareness of both linguistic form and communicative function.

Although morphosyntactic development remained limited and largely implicit at this stage, the contextually appropriate deployment of language suggests that children were beginning to internalize form–meaning relationships rather than reproducing decontextualized phonological sequences. This finding is analytically significant: it indicates that the language behaviors observed in this study reflect nascent acquired competence in the Krashean sense subconscious, contextually embedded, and functionally operative, rather than mere conscious imitation of frequently heard strings.

A degree of variability across the sample, however, warrants acknowledgment. Two parents reported that their children remained at a pre-productive stage despite regular viewing, responding affectively to content without yet attempting verbal reproduction:

"She is amazed, but she cannot speak English yet, so she just laughs." (R1)

Several parents with limited English proficiency also reported reduced capacity to mediate content effectively:

"I do not always understand what it means, so I let him watch by himself." (R13)

These cases illuminate the boundary conditions of digitally mediated EFL acquisition. Parental mediation while prevalent across the sample varied considerably in its quality, intentionality, and linguistic richness. Digital literacy, English language confidence, and available interaction time each moderated the effectiveness of parental involvement. Where these resources were constrained, YouTube Kids functioned primarily as a passive input source rather than a productive acquisition environment. This variability reinforces the conceptual

necessity of evaluating parental mediation not merely in terms of its presence or absence, but in terms of its qualitative depth, a dimension with direct implications for caregiver training and early childhood language policy in underserved regions.

4. Convergent Evidence: Perceived Effectiveness of YouTube Kids as an EFL Learning Tool

The convergent findings across Themes 1 through 3 encompassing consistent digital input, active parental mediation, and children's spontaneous language production, collectively illuminate the conditions under which YouTube Kids supports early EFL acquisition. To further triangulate these thematic findings, the study examined parental perceptions of the platform's overall effectiveness as a language learning tool.

As presented in Table 6, a substantial majority of parents evaluated YouTube Kids positively: 47.61% rated it as effective, 23.80% as very effective, and 28.57% as moderately effective. No parent rated the platform as ineffective or very ineffective, suggesting broad consensus regarding its educational value when used in a structured and supervised manner.

Table 6. Perceived Effectiveness of YouTube Kids as a Language Learning Tool

Perceived Effectiveness Score	Interpretation	Frequency (n)	Percentage (%)
5	Very Effective	5	23.80%
4	Effective	10	47.61%
3	Moderately Effective	6	28.57%
2	Ineffective	0	0.00%
1	Very Ineffective	0	0.00%
Total		21	100.00%

The alignment between these perceptions and the language output patterns documented in Themes 1 and 3 strengthens the internal coherence of the findings.

Parents who reported higher engagement in mediation practices consistently rated the platform more favorably, while those with limited involvement tended toward moderate effectiveness ratings, a pattern suggesting that perceived value is itself a function of interactional quality rather than platform characteristics alone. These findings collectively position YouTube Kids as a pedagogically viable, low-barrier EFL resource in contexts where formal early English instruction is structurally limited, provided that its deployment is embedded within responsive and purposeful parental mediation.

DISCUSSION

This study demonstrates that exposure to English-language content on YouTube Kids, when embedded within active parental mediation, contributes meaningfully to early EFL acquisition among toddlers aged 2–5 in Lampung Province. The findings both corroborate and extend prior research conducted in other Indonesian regions (Friska, 2023; Sufa & Dzulfikri, 2023), while offering empirical grounding from a non-metropolitan context that has received limited scholarly attention.

1. Digital Media as a Source of Consistent and Comprehensible English Input

The quantitative data revealed that 76.19% of parents reported near-daily engagement with English-language YouTube Kids content, and a statistically significant positive correlation was established between exposure frequency and duration and children's spontaneous English output ($\rho = .64, p = .002$). This finding provides empirical support for Krashen's (1982) Input Hypothesis, which posits that consistent, meaningful, and contextually embedded input constitutes the primary driver of subconscious language acquisition. The repetitive, song-based, and visually anchored characteristics of the platform's preferred channels: CoComelon, Baby Bus, and Blippi, align closely with the conditions Krashen identifies as conducive to acquisition rather than conscious learning.

This outcome reinforces findings from comparable studies in Indonesia. Anshary and Perangin-Angin (2024) and Tohamba (2025) similarly documented

natural vocabulary gains in young learners through regular YouTube engagement, while Vlachou et al. (2024) established through systematic review that brief animated stimuli with repetitive linguistic structures yield measurable language development outcomes. The visual and auditory multimodality of digital content further enhances vocabulary assimilation in children at the sensorimotor and preoperational stages of cognitive development (Piaget, 1952), for whom meaning is most effectively constructed through contextualized, embodied, and perceptually rich input.

Critically, however, the correlation between exposure and output was moderate rather than absolute ($\rho = .64$), and the cross-tabulation data revealed that children at the highest exposure levels did not uniformly demonstrate the greatest language output. This attenuation suggests that raw viewing time functions as a necessary but insufficient condition for acquisition, a finding that distinguishes the present study from prior research which has tended to treat exposure duration as the primary predictor of language gains without adequately accounting for mediating variables.

2. Parental Mediation and Interactive Co-Viewing as Language Scaffolding

The data consistently identified parental mediation as the critical moderating variable determining the depth of language acquisition outcomes. Parents who co-viewed, explained unfamiliar vocabulary in real time, and reinforced newly encountered language in daily routines produced children with markedly higher spontaneous English output than those who permitted unsupervised viewing. This finding is consistent with Barr (2019) and Strouse and Troseth (2014), who established that adult mediation is essential for toddlers to transfer linguistic knowledge from screen-based input to real-world communicative contexts, given the well-documented video deficit in unmediated screen learning.

The mediation practices documented in this study align substantively with Vygotsky's (1978) conceptualization of scaffolding within the Zone of Proximal Development. Parents did not function merely as screen monitors; they operated as

active linguistic intermediaries who translated, modeled pronunciation, prompted repetition, and contextualized novel vocabulary within familiar daily routines. These behaviors transformed passive viewing into an interactive and socially grounded acquisition experience, consistent with what Imaniah et al. (2020) identify as the relational dimension of parental media engagement.

Where this study extends beyond the existing literature is in its documentation of mediation quality as a differentiating variable, not merely mediation presence. Several parents with limited English proficiency reported relinquishing their mediating role precisely when linguistic support was most needed, resulting in children who remained at a pre-productive stage despite regular exposure. This finding challenges studies that treat parental involvement as a binary condition and points instead to a continuum of mediation quality, one that is shaped by caregivers' English confidence, digital literacy, and available interaction time. Flewitt et al. (2024) and Friska (2023) advocate for balanced screen use and co-viewing as best practices, but the present data suggest that the linguistic richness of co-viewing interactions matters as much as their mere occurrence.

3. Transition from Passive Viewing to Active Language Use

The progression from receptive engagement with digital content to spontaneous productive language use, documented in 85.71% of cases, supports the Input–Interaction–Output model (Liu, 2022), which frames language acquisition as a three-stage process moving from exposure through internalization to production. Crucially, this progression occurred in naturalistic home environments without formal instruction, underscoring the acquisition-oriented rather than learning-oriented character of the language development observed (Krashen, 1982).

The emergent linguistic competencies identified lexical, formulaic, and pragmatic indicate that children's English outputs extended beyond rote phonological reproduction. Contextually appropriate use of phrases such as "thank you, mommy" during mealtimes or "let's go" during play reflects an internalization of form–meaning–function relationships that constitutes early communicative

competence. This finding resonates with Hymes' (1972) construct of communicative competence, in which language use is governed not only by grammatical knowledge but by sociolinguistic appropriateness. That toddlers aged 2–5 demonstrated nascent pragmatic sensitivity, deploying English expressions in contextually fitting situations suggests that the acquisition process facilitated by YouTube Kids, under adequate parental scaffolding, operates at a level deeper than imitation.

This outcome also has direct pedagogical implications. The development of expressive language skills through contextual repetition and social reinforcement, rather than through explicit instruction, affirms the value of naturalistic, input-rich home environments as early language learning contexts. These findings advocate for exposure- and interaction-based approaches in early childhood EFL pedagogy an orientation that formal curricula in underserved Indonesian regions are currently ill-equipped to provide.

4. Contextual Relevance in Lampung Province

This study makes a substantive contribution to an underrepresented area in Indonesian EFL research. While prior studies have concentrated on Java and North Sumatra, Lampung's distinct sociolinguistic profile characterized by multilingual households in which Indonesian, regional Lampung languages, and English occupy hierarchically distinct positions renders it a particularly instructive site for examining how incidental digital exposure operates in a competitive linguistic environment. YouTube Kids effectively functions as the primary, and in many households the sole, source of English input for these toddlers, a condition that amplifies both its pedagogical potential and the consequences of unmediated or low-quality engagement.

The multi-district sampling across nine administrative areas of Lampung Province, spanning urban, semi-urban, and semi-rural contexts further reveals that patterns of YouTube Kids use and parental mediation are not confined to educationally advantaged households. Families across varied socioeconomic backgrounds reported integrating the platform into daily routines, suggesting that

digital media democratizes access to English-language input in ways that formal institutional provision currently cannot. Nonetheless, the effectiveness of this informal access is unevenly distributed, contingent on the mediating capacities of individual caregivers, a structural inequality that policy interventions must address.

5. Implications for Education and Policy

The findings carry several implications for early childhood language education and regional policy in Indonesia. First, YouTube Kids demonstrates viable potential as a low-barrier supplementary EFL tool, particularly in regions where qualified English teachers and structured early language programs are scarce. Its integration into home learning environments should be recognized rather than overlooked in early childhood education frameworks.

Second, the quality of parental mediation constitutes the central determinant of acquisition outcomes. Caregiver training programs focused on media literacy, co-viewing strategies, and basic English interaction techniques would significantly amplify the platform's educational value. Early childhood education centers (PAUD) are well-positioned to deliver such training as part of parental engagement programs, an approach supported by the community-based professional development models documented by Hima and Gusdian (2018).

Third, regional education policies in Lampung and comparable provinces should consider frameworks that formally integrate digital media as a component of early foreign language exposure strategy, rather than treating screen time exclusively as a welfare or health concern. The evidence presented here suggests that, under appropriate conditions, digitally mediated exposure constitutes a legitimate and effective pathway to early EFL acquisition one that merits policy recognition, pedagogical scaffolding, and systematic caregiver support.

6. Limitations

Several limitations of this study merit acknowledgment. The language competence data relied exclusively on parental self-report rather than direct

observational assessment or validated EFL instruments, introducing the possibility of social desirability bias. The sample of 21 families, while sufficient for thematic saturation, constrains the statistical power of the inferential analyses; the correlation findings should therefore be interpreted as exploratory and directional rather than definitive. The cross-sectional design further precludes causal or developmental claims, as the trajectory of observed competencies over time cannot be determined from a single data collection point. The absence of a comparison group additionally limits the study's capacity to isolate YouTube Kids' contribution from other environmental variables independently supporting early English development. Future research employing longitudinal designs, validated assessment instruments, and comparative sampling would substantially strengthen the evidential basis for the claims advanced here.

CONCLUSION

This study establishes YouTube Kids as a significant and pedagogically viable source of English language input for toddlers aged 2–5 in Lampung Province, Indonesia. Employing a mixed-methods embedded design that integrated reflexive thematic analysis with quantitative descriptive and correlational procedures, the research demonstrates that exposure to English-language digital content when combined with active and linguistically rich parental mediation supports vocabulary acquisition, receptive comprehension, and the spontaneous productive use of English in naturalistic daily interactions.

Three major themes emerged from the data. First, digital media functions as a source of consistent and comprehensible English input, with 76.19% of parents reporting near-daily engagement and a statistically significant positive correlation established between exposure frequency and duration and children's spontaneous language output ($\rho = .64$, $p = .002$). Second, parental mediation, encompassing co-viewing, real-time explanation, vocabulary reinforcement, and structured screen-time routines operates as a critical scaffolding mechanism that transforms passive digital exposure into an interactive acquisition environment. Third, children who experienced guided and mediated exposure demonstrated a measurable transition

from passive reception to active language production, evidenced by contextually appropriate use of English phrases across lexical, formulaic, and pragmatic domains. Collectively, these themes affirm that language acquisition does not occur as an automatic consequence of media exposure, but as a dynamic process contingent upon the quality and intentionality of the social interaction surrounding that exposure.

This study contributes to an underrepresented body of empirical literature on digitally mediated EFL acquisition in non-metropolitan Indonesian contexts. The Lampung case illustrates that digital technology can meaningfully bridge gaps in formal early English education, provided it is deployed purposefully and embedded within responsive parental support. The variability in acquisition outcomes observed across the sample particularly among children whose caregivers had limited English proficiency or reduced capacity for active mediation further indicates that access to the platform alone is insufficient; the mediating competencies of caregivers constitute the decisive variable in determining whether digital exposure translates into genuine language acquisition.

Future research should employ longitudinal designs and validated early EFL assessment instruments to examine the developmental trajectory of competencies initiated through YouTube Kids exposure. Caregiver training programs targeting media literacy and co-viewing strategies, particularly within early childhood education networks across Lampung's diverse districts, represent a priority area for both applied research and policy intervention. YouTube Kids holds demonstrable potential as a low-barrier EFL resource; realizing that potential systematically, however, requires institutional recognition, pedagogical scaffolding, and sustained investment in caregiver capacity.

REFERENCES

- Anshary, E. P., & Perangin-Angin, A. B. (2024). The use of YouTube in children's vocabulary acquisition : a case study of 3 and 5-year-old children in Medan. *Wacana: Jurnal Penelitian Bahasa, Sastra Dan Pengajaran*, 22(1), 76–83. <https://doi.org/10.33369/jwacana.v22i1.33433>
- Barr, R. (2019). Growing up in the digital age: Early learning and family media ecology. *Current Directions in Psychological Science*, 28(4), 341–346. <https://doi.org/10.1177/0963721419838245>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brushe, D., Meyer, S., & Tanaka, R. (2023). The impact of screen exposure on early language development: A lexical gap analysis. *Journal of Child Language and Development*, 50(2), 213–228. <https://doi.org/10.1017/S030500092300001X>
- Christakis, D. A., & Hale, L. (2024). Handbook of children and screens: Digital media, development, and well-being from birth through adolescence. In D. A. Christakis & L. Hale (Eds.), *Handbook of Children and Screens: Digital Media, Development, and Well-Being from Birth Through Adolescence*. Springer. <https://doi.org/10.1007/978-3-031-69362-5>
- Dore, R. A., Logan, J. A. R., Lin, T.-J., Purtell, K. M., & Justice, L. M. (2020). Media as a context for learning: Exploring the role of parental mediation. *Early Childhood Research Quarterly*, 51, 14–25. <https://doi.org/10.1016/j.ecresq.2019.08.008>
- Flewitt, R., Messer, D., & Kucirkova, N. (2024). New technologies to support reading development in the early years: Integrating digital and traditional literacies. *British Journal of Educational Technology*, 55(1), 55–71. <https://doi.org/10.1111/bjet.13298>
- Friska, R. (2023). Parental perspectives on digital media use in early English

- learning. *Indonesian Journal of Early Childhood Education Studies*, 12(1), 45–56. <https://doi.org/10.15294/ijeces.v12i1.32019>
- Gath, K., Dayanim, S., & Barr, R. (2023). Guided digital play and vocabulary learning in toddlers: The role of joint media engagement. *Journal of Applied Developmental Psychology*, 86, 101544. <https://doi.org/10.1016/j.appdev.2023.101544>
- Hima, A. N., & Gusdian, R. I. (2018). Employing Local Wisdom Based Short Story for Teaching English Elementary School. *LET: Linguistics, Literature and Language Teaching Journal*, 8(1), 15–30. <https://doi.org/10.18592/let.v8i1.2194>
- Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected Readings* (pp. 269–293). Penguin.
- Imaniah, R., Handayani, T., & Pratiwi, Y. (2020). Peran orang tua dalam penggunaan media digital pada anak usia dini. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 4(2), 539–548. <https://doi.org/10.31004/obsesi.v4i2.319>
- Kasan, R. A., Fadillah, R., & Yusuf, M. (2019). Foundations of Second Language Acquisition: Strategies and Motivation of Banjarese Young Learners in Learning English as a Foreign Language. *LET: Linguistics, Literature and English Teaching Journal*, 9(2), 265–301. <https://doi.org/10.18592/let.v9i2.15758>
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Prentice-Hall International.
- Linuwih, E. R., & Trihastutie, N. (2020). Digital Entertainment to Support Toddlers' Language and Cognitive Development. *Teknosastik*, 18(1), 1. <https://doi.org/10.33365/ts.v18i1.467>
- Liu, X. (2022). The Relation of Input, Interaction, and Output in SLA. *Proceedings of the 2022 7th International Conference on Social Sciences and Economic*

- Development (ICSSSED 2022)*, 652(Icssed), 474–479.
<https://doi.org/10.2991/aebmr.k.220405.078>
- Mares, M. L., & Pan, Z. (2013). Effects of Sesame Street: A meta-analysis of children's learning in 15 countries. *Journal of Applied Developmental Psychology*, 34(3), 140–151. <https://doi.org/10.1016/j.appdev.2013.01.001>
- Masur, E. F., Flynn, V., & Olson, J. (2016). Infants' and toddlers' television exposure and vocabulary outcomes: The role of adult-child conversational turns. *Infancy*, 21(4), 421–444. <https://doi.org/10.1111/infa.12126>
- Piaget, J. (1952). *The origins of intelligence in children* (T. . (M. Cook (ed.)). W. W. Norton & Company. <https://doi.org/10.1037/11494-000>
- Pratiwi, A. I., Edo Dwi Cahyo, Baiti Nurul Azizah, Hana Wahyuningsih, & Luthfiana Fitria. (2024). Effectiveness of Using Youtube Applications as Learning Media for Early Childhood Education. *Jurnal Pendidikan Anak Usia Dini Al-Amin*, 2(1 SE-Articles), 7–18.
<https://doi.org/10.54723/ejpau.v2i1.131>
- Strouse, G. A., & Troseth, G. L. (2014). Supporting toddlers' transfer of word learning from video. *Cognitive Development*, 30, 47–64.
<https://doi.org/10.1016/j.cogdev.2014.01.002>
- Sufa, F., & Dzulfikri, A. (2023). The role of YouTube videos in acquiring English vocabulary in early childhood: A case study. *Indonesian Journal of English Education*, 10(2), 120–132. <https://doi.org/10.15408/ijee.v10i2.32577>
- Tohamba, C. P. P. (2025). The Use of YouTube in Improving English Vocabulary Mastery in Early Childhood: Systematic Literature Review. *Journal of Instructional and Development Researches*, 5(5), 500–510.
<https://doi.org/10.53621/jider.v5i5.595>
- Tsang, K. M., Chen, Y., & Zhou, Q. (2025). AI-enriched video content and early language development: The role of family interaction. *Early Education and Development*, 36(1), 33–49. <https://doi.org/10.1080/10409289.2025.1001234>

- Vlachou, T. I., Kambanaros, M., Plotas, P., & Georgopoulos, V. C. (2024). Evidence of Language Development Using Brief Animated Stimuli: A Systematic Review. *Brain Sciences*, *14*(2). <https://doi.org/10.3390/brainsci14020150>
- Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Harvard University Press. <https://doi.org/10.2307/j.ctvjf9vz4>
- Yuliarti, R., & Nikmah, I. (2022). Exploring CoComelon as a learning medium to introduce English vocabulary to early learners. *Jurnal Pendidikan Usia Dini*, *16*(2), 123–136. <https://doi.org/10.21009/JPUD.162.10>
- Zulfia, B. S., & Andini, Y. (2023). Educational Video Quality on TikTok for Vocabulary Teaching. *LET: Linguistics, Literature and English Teaching Journal*, *13*(2), 276–301. <https://doi.org/10.18592/let.v13i2.11600>