Current Issues in English Education for Young Learners in East Asia

Yuko Goto Butler
(University of Pennsylvania)


Recognizing English as an increasingly powerful lingua franca, many governments in East Asia have implemented English-language education at the primary school level. The zeal for teaching English to younger and younger learners extends to both within and outside of the formal education system. While there are many unique local issues, a number of challenges and issues apply across the East Asian region. In this article, I focus on three such topics: (a) the widespread belief in the increased benefit of starting English-language learning when children are very young; (b) the emergence of new instructional models in early English teaching; and (c) the growing gaps in English-language achievement by children’s socio-economic status (SES). In discussing these issues, I draw on examples from previous studies as well as my own applied linguistics research in East Asia. I conclude by offering suggestions for future research directions that will inform pedagogy and policies for early English education.

**Key words:** Early English education, instructional models, socio-economic disparities

1. INTRODUCTION

With English having been called “a basic skill” for success in the twenty-first century (Graddol, 2006, p. 72), the learning of English in East Asia has become far more than merely the learning a foreign language (FL). Considering the potential positive impact that knowledge of English has on a person’s future studies and careers, as Zou and Zhang (2011) rightly described, “English is more than just a school subject; it permeates into many aspects of social life” (p. 191).

The quest for learning English has intensified in recent years. Many governments in East Asia have introduced English in primary schools, and the zeal for learning and teaching
English prevails even among children. Each government and region has its own unique issues and challenges regarding early English education; however, many issues are observed across East Asia. In this article, I focus on three such cross-cutting issues, namely, (a) the widespread belief in the increased benefit of early English learning, (b) the emergence of new instructional models in early English teaching, and (c) the growing gaps in English achievement by children’s socio-economic status (SES). These issues are particularly important to address because there seem to be substantial gaps between research findings and understanding among decision-making stakeholders such as policy makers and practitioners. Therefore, my aim is to bridge the gap between research and practice. In discussing these issues, I draw on examples from existing studies as well as my own applied linguistics research in East Asia.

Following a convention in the field, I use young learners to refer to children of primary school age (approximately 5-12 years old). East Asia covers China (including Hong Kong), Japan, South Korea, and Taiwan, reflecting the body of research conducted on early English education in this region.

2. THE EFFECT OF EARLY ENGLISH LEARNING

2.1. The Critical Period Hypothesis and FL Learning

“The younger the better” in language learning is a widely held notion among the general public, and it is often used as a rationale for implementing early FL programs for children. In East Asia, English has been introduced as part of the formal education system at progressively younger grade levels. Various types of early English programs outside of the schooling are also on the rise and have attracted many parents of young learners. However, empirical evidence available to date does not necessarily support that “the younger the better” is true for FL learning.

The notion of “the younger the better” is often associated with the critical period hypothesis (CPH) in language acquisition. In terms of language acquisition, the CPH suggests that there is a limited window in the course of one’s development during which the acquisition of “native-like” proficiency is possible. The original idea of the critical period came from ethology, which is the study of animal behavior, and the idea was later applied to human language acquisition. Based on clinical observations, Penfield and Roberts (1959) claimed that there is “a biological clock” in the brain that stops sometime around 9 years of age (p. 237). Similarly, Lenneberg (1967), who formulated the CPH in language acquisition, suggested that the window closes before puberty, when the brain loses its plasticity. It is important to note, however, that Lenneberg proposed the CPH in
the context of first language (L1), and he had little to say about the critical period for second-language (L2) acquisition.

Since the formulation of the CPH in language acquisition, a substantial number of studies have been conducted in L2 acquisition. It is beyond the scope of this article to review studies on the CPH in L2 in detail (for such reviews, see Birdsong, 2006; Dekeyser, 2013; Singleton, 2005, etc.), but the results of these studies are inconclusive. Even among researchers who support the existence of a critical period in L2 acquisition, they differ in terms of the exact age of the critical period (when the critical period begins and ends), and the linguistic domains affected by the CPH (e.g., phonology, morphosyntax, lexical, and other domains). What counts as evidence for or against the CPH differs among the researchers as well (Butler, forthcoming-a).

Importantly, the studies on the CPH in L2 have been conducted in immersion contexts (e.g., immigrants who learn their L2 in a community where the target language is predominantly used). Therefore, whatever the result of the CPH in L2 acquisition may be, we cannot simply assume that the same result will be obtained in FL contexts, in which the amount of input and meaningful interaction in the target language is usually substantially limited compared with immersion contexts. In other words, applying the CPH in an FL context is potentially misleading and inappropriate.

What do we know, then, about the relationship between the age of onset of FL learning and the learning outcome in non-immersion contexts? Do earlier starters have advantages in FL instructional settings? Huang (2014) conducted a meta-analysis based on 42 studies, the majority of which were conducted either in the United States or Europe, dealing with European language combinations (e.g., English speakers learning French, Spanish speakers learning English, etc.). Huang analyzed studies examining both short-term and longer-term effects. The former type of study offered short-term treatments to students with various age groups, and examined immediate age effects; in other words, these studies examined the efficiency of learning within a limited timeframe. The latter type of study was either cross-sectional or longitudinal, and examined the effect of age of onset of FL learning in the long run. While there were some exceptional cases, by and large, Huang found that older learners, not young learners, have advantages in the short term; namely, older learners were more efficient in FL learning. For the mid-/long-term effects, after controlling for the amount of instruction, older learners again tended to have an advantage over younger learners.

That being said, some studies from Spain (e.g., García Mayo & García Lecumberri, 2003; Muñoz, 2006) indicate that younger learners may catch up with older learners after

---

1 There were four exceptions in Huang (2014): a case of Japanese children learning English as their FL, a case of English speakers learning Russian, and two cases of English speakers learning Mandarin. The four studies all examined the age effect on short-term outcomes.
receiving approximately 700 hours of FL instruction. In addition, morphosyntax learning was boosted at around the age of 12, while auditory learning was not affected by the learners’ age but was affected by the amount of instruction they received. In other words, although morphosyntax learning might be affected by learners’ cognitive maturity, auditory learning might not be. This result suggests that early FL instruction should focus on auditory learning, while also ensuring a sufficient number of instructional hours.

2.2. The Relationship Between Age and Outcome in East Asia

In East Asia, a number of studies have examined the effect of English as FL instruction during primary school (referred to as EFLPS hereafter). Some of these studies specifically addressed the relationship between the age (or the onset of FL instruction) and the learning outcome. However, the results of such studies remain inconclusive, as I detail below.

In South Korea, Kwon’s large-scale evaluation study ($N = 8,062$) indicated the effect of EFLPS (Kwon, 2006). The study compared high school students in 2006 who received EFLPS with a 2003 cohort who did not receive EFLPS. The results indicate that the 2006 group outperformed their counterparts in terms of both language attainment and affective domains, such as their motivation to learn English. Kwon speculated that, in addition to the EFLPS itself, the increase of English-language learning outside of formal schooling after the implementation of EFLPS greatly influenced the results.

An abundance of research on age effects in English learning has taken place in Japan, perhaps partially due to the fact that the Japanese government has yet to decide when to start English education as an academic subject in the formal educational system. (As of 2014 in Japan, English is not taught as a formal academic subject; instead, it is an exploratory program taught once a week at the 5th- and 6th-grade levels. While most students have minimal exposure to English at school, there is variability both in terms of the amount and the content of the English instruction at school across Japan; some schools begin teaching English in younger grade levels and/or as an academic subject as opposed to an exploratory experience.) The results of the studies on the effects of EFLPS in Japan have been mixed. Some studies have shown positive results, while others have shown negative or no effects of EFLPS. The domains where the effects were found differed across studies (see Butler, forthcoming-b, for details). Such mixed results are not too surprising, however, considering the variability of EFLPS in Japan as well as the inconsistent measurements used across these studies.

One of the challenges that researchers face in evaluating the effect of EFLPS in relation

---

2 The Ministry of Education, Culture, Sports, Science, and Technology in Japan (MEXT) plans to introduce English as an academic subject at the 5th and 6th grade levels starting from 2020.
to age is that the age of first exposure (AOE) to FL is usually confounded with the hours of instruction (HOI); in other words, earlier starters tend to receive longer accumulated hours of instruction. Because the potential impact of HOI on the learning outcome is substantial, we need to consider the confounding factor seriously. After statistically controlling for HOI, Larson-Hall (2008) found that college students in Japan who started receiving English instruction earlier (AOE < 12, \(N = 61\)) had an advantage in a phonemic task than later starters (AOE > 12, \(N = 139\)), but that the effect size was small. No advantage was found in a grammatical judgment task.\(^3\) Butler and Takeuchi (2008), in their large-scale study \((N = 6,541)\) found that HOI could positively predict the STEP Silver test scores (a proficiency test for young learners) among 5\(^{th}\)- and 6\(^{th}\)-grade students but not among 3\(^{rd}\)- and 4\(^{th}\)-grade students. This result led the authors to suggest that “the one-hour block of instruction commonly taught might have exerted a different influence on students’ performance on the Silver Test” (p. 88). Thus, we should pay close attention not only to the amount of instruction (HOI) but also to the quality or content of instruction when conducting research on age effects.

In the last 10 years or so, we have seen a growing number of neurolinguistic studies that shed light on the relationship between age and learning outcomes in East Asia. Two major techniques have been used to capture neuronal functions: functional magnetic resonance imaging (fMRI) and event-related potentials (ERPs). fMRI is better suited to obtaining special information, while ERPs are better at capturing real-time brain activities. Most recently, researchers have started using functional near-infrared spectroscopy (fNIRS). This newer brain imaging technique is considered more suitable for analyzing children’s brain activities than fMRI.

Neurolinguistic studies on the effect of age in East Asian contexts have shown that even minimal exposure to FL seems to influence children’s processing. Although it is hard to draw a conclusion yet due to the limited information available (and considering that most neurolinguistic studies have been focused on semantic processing), HOI or general proficiency levels appear to be more influential factors than the AOE itself when it comes to FL semantic processing. For example, Ojima, Matsuba-Kurita, Nakamura, Hoshino, and Hagiwara (2011), in a 3-year longitudinal project among Japanese primary school students \((N = 350)\), found that the students with longer HOI showed larger N400 in ERPs (an indication of semantic processing) and had higher scores in a STEP test (an English proficiency test) regardless of AOE. After controlling for HOI, later starters had larger

\(^3\) We need to keep in mind, however, that the participants in Larson-Hall (2008) belonged to an elite national university, and we can assume that they all passed a college entrance exam in English in which learners’ grammatical knowledge was heavily tested. Thus, Larson-Hall’s finding (no difference in the grammar task between the earlier and late starters) could be largely attributed to the result of their test preparation conducted after primary school rather than to their AOE.
N400 and scored higher in the STEP test. Takahashi et al. (2011) also used ERP to examine the effect of FL in the L1 semantic processing among 78 Japanese preschool children (age 4-5). They found that, while there was no difference among 4-year-olds (who had 15 minutes of exposure to English per day), there were significant differences in ERP wave forms among 5-year-olds (who had 45 minutes of exposure to English per day). The result led the authors to conclude that “[t]he development and processing ability for four- and five-year olds in their native language is vulnerable to exposure to non-native language, depending on the length of exposure” (p. 251).

Sugiura et al. (2011), using fNIRS, found that Japanese primary school students (age 6-10, N = 484) depended more heavily on right-hemispheric supra-segmental processing rather than left-hemispheric supra-segmental processing when they processed unfamiliar FL words, suggesting that the children processed unfamiliar or infrequently used FL words as if they were non-word auditory stimuli. The authors suggested that as children’s lexical knowledge increased, there might have been a shift in laterality from right to left in the inferior parietal region. Such bilateral activation found in this study may be advantageous for FL learning, but it is not clear if the result can be attributed to the participants’ age, proficiency levels, or the specific task effect. Hidaka et al. (2012), also using fNIRS, examined preschool children’s brain activities (ages 3-4 and 4-5) when they listened to a story in L1, familiar FL (English), and unfamiliar FL (Chinese). While the researchers found no differences among the younger group, they found differences in the activities in the bilateral frontal areas among the older group (as well as among the adult control group). The brain activities were higher when the children processed L1 and familiar FL than when they processed unfamiliar FL, suggesting that the greater exposure to FL indeed mattered. (See Butler forthcoming-b for a more detailed review of neurolinguistic studies.)

Neurolinguistic research is an emerging field and has significant potential for uncovering the relationship between age and language acquisition. So far, however, it is hard to draw a conclusion on AOE and its relation to learning outcomes using such studies. Many neural studies to date have failed to sort out confounding factors (e.g., age and HOI are confounded), and the link between brain activities and behavioral outcomes (i.e., language performance) is not yet totally clear. Furthermore, this type of investigation so far has largely been limited to semantic processing.

In summary, research findings (both behavioral and neural studies) so far do not provide evidence supporting the “the earlier the better” assumption in FL language learning. On the contrary, a number of studies show the advantage of beginning FL learning among older learners, at least in some domains. The studies also indicate a greater influence of the amount of instruction than of the age of onset itself. The relationship among HOI, AOE, and learning outcomes remains largely inconclusive. Moreover, researchers need to pay closer attention to the quality of instruction as well as the quantity of instruction.
3. INSTRUCTIONAL MODELS

EFLPS can be a good opportunity to try out new instructional methods and strategies because it typically involves a new group of teachers (i.e., primary school teachers) who might have not been constrained by traditional, secondary-school-level English-teaching methods such as instruction focused on grammar or translation. But at the same time, primary school teachers are relatively new to English-teaching professions, and they often face a number of challenges. In this section, I focus on the following major topics related to instruction: (a) communicative language teaching (CLT) and task-based language teaching (TBLT); (b) teaching English through English (TETE); and (c) content-based instruction (CBI) and immersion programs.

3.1. Communicative Language Teaching (CLT) and Task-based Language Teaching (TBLT)

In response to repeated criticism of traditional English teaching for being ineffective, communicative language teaching (CLT) has been heavily promoted to English teachers in East Asia. Task-based language teaching (TBLT) also started gaining popularity among English teaching professionals in East Asia during the 1980s. Although both CLT and TBLT are widely subscribed to in EFLPS curricula in East Asia, it has been repeatedly reported that primary school teachers face difficulties practicing them in their classrooms. For CLT and TBLT to be properly implemented in East Asian primary school classrooms, it is necessary to implement them in such a way that gives sufficient consideration to children’s age and contextual factors. In other words, it is necessary to adapt them for use in this new context rather than simply adopting them wholesale from English-speaking contexts where CLT and TBLT originated.

The basic premise of CLT is to help students develop communicative competence in meaningful and authentic interactions, as opposed to focusing instruction on linguistic forms. In CLT, learners have opportunities to receive meaningful input and to use the target language through social interaction, as well as opportunities to focus on learning processes. CLT is broadly based on the theories of communicative competence and second-language acquisition, but its theoretical “broadness” has invited various interpretations and realizations of practice (e.g., the strong version and weak version of CLT). Similarly, there is no agreed-upon definition for tasks in TBLT. Long (1985), for example, emphasized that the core feature of tasks is their connection to real-world events.

---

4 TBLT can be considered as “an offset of CLT” (Kumaravadivelu, 2006, p. 66), but the relationship between CLT and TBLT is debatable.
Samuda and Bygate (2008) synthesized key features of tasks addressed by different researchers, and defined a task as “a holistic activity which engages language use in order to achieve some nonlinguistic outcome while meeting a linguistic challenge, with the overall aim of promoting language learning, through process or product or both” (p. 69). As with CLT, researchers also differ in terms of how tasks should be implemented in practice. TBLT implementation can be characterized as having strong and weak versions (referred to as task-based teaching and task-supported teaching, respectively) according to the degree of emphasis on communication versus instruction in task designs, and the degree of task engagement in instruction.

In the initial implementation of CLT in East Asia, there were a number of challenges and constraints, which can be largely classified as (a) conceptual constraints (e.g., teachers’ misunderstanding of or lack of belief in CLT); (b) classroom-based constraints (e.g., large class sizes, limited instructional hours, and difficulties in classroom management); and (c) societal-institutional constraints (e.g., pressure from exams). Similar challenges were encountered when TBLT was introduced in East Asia (Butler, 2011).

At the primary school level, a number of studies across contexts reported that CLT and TBLT were not implemented in classrooms as the policies intended; so-called “communicative activities or tasks” practiced in class were not fully communicative (e.g., Deng & Carless, 2009 in China; Carless, 2002, 2004a; Tong, Adamson, & Che, 2000 in Hong Kong; Yukawa, 2002 in Japan; K.-M. Lee, 2002 in Korea). In response to the needs for Asian teachers to have a finer-level communicative activity model, Littlewood (2004, 2007) in Hong Kong proposed a five-level model based on a communicative continuum ranging from fully meaning-focused activities to fully form-focused activities. Littlewood (2007) suggested that teachers who were used to form-focused activities could gradually incorporate more meaning-focused activities in their classrooms.

However, we should not assume that highly communicative activities are always the best, regardless of context. After observing classroom activities and conducting interviews with primary school teachers in China, Deng and Carless (2009) suggested that, depending on contextual factors, such as students’ needs, various degrees of communicative activities can be combined and used in class. Considering the strong preference for the traditional Presentation-Practice-Production (PPP) approach among teachers in Hong Kong, Carless (2009) further suggested that a combination of PPP and TBLT may even be possible.

In addition to contextual factors, when it comes to implementing CLT and TBLT in primary schools, age-related factors need to be carefully examined. Considering that young learners tend to be attracted to unexpected and imaginary events and activities (Butler, Someya, & Fukuhara, 2014), task designs may not necessarily be restricted to “real world events”: the notion of authenticity may need to be broadened for young learners. In addition, research on tasks has focused on cognitive elements with little attention to
affective elements, such as those that excite and motivate learners. Because many EFLPS programs have as their core objective to develop young learners’ motivation and confidence, we need to pay more serious attention to affective elements when designing and theorizing tasks for young learners.

Policy makers and researchers in East Asia are still in the process of searching for effective instructional models and strategies for EFLPS. We need more research on “good” practice in EFLPS in general. Meaning-based activities, such as storytelling, are increasingly popular in EFLPS (Ellis, 2014). Another exciting step for EFLPS is the increase in the use of technology as an instructional and learning tool. Since the official introduction of EFLPS in 1997 in Korea, the Korean government has been very proactive in incorporating technology in EFLPS classrooms. A growing number of studies have examined the effect of IT tools for EFLPS in East Asia, including an online reading partnership and peer assessment system called “a mobile-device-supported peer-assisted learning system (MPAL)” (Lan, Sung, & Chang, 2007) in Taiwan; a multimedia English learning system (MEL) for phonemic awareness and pronunciation learning (Lai, Tsai, & Yu, 2009), also in Taiwan; and a “massive multiplayer online role-playing game” (MMORPG) in Korea (Suh, Kim, & Kim, 2010). These intervention studies using IT tools generally have shown positive results on young learners’ English learning, as well as their motivation. In Japan, Butler, Someya, and Fukuhara (2014) examined the use of online instructional games as part of an online assessment and identified game features that were shared by games that children (age 4-12, \(N = 3,945\)) were attracted to. Such features included being cognitively challenging, incorporating elements that evoke curiosity, allowing greater player autonomy and control, and having multiple players. The researchers also found different game-playing behaviors based on players’ age, and varying relationships between the game scores and English test scores based on the game type and the difficulty level of the assessment. Technology has great potential not only as an instructional tool (e.g., it may provide children with greater opportunities to receive input), but also as an authentic means of communication. Because communication through technology is expected to be increasingly common even among young learners, we should consider technology as a major and authentic means of communication in EFLPS.

3.2 Teaching English Through English (TETE)

Another instructional method that has been strongly promoted to English teachers in East Asia is teaching English through English (TETE). The rationale for employing TETE, a monolingual (i.e., English-only) approach, is based on the assumption that increased

---

5 As of 2014, the Japanese EFLPS policy has not mentioned TETE.
exposure to English improves learners’ chances of internalizing the target language and facilitating their language learning. Since learners in EFL contexts usually have limited exposure to the target language, teachers using TETE are expected to make every effort to maximize the students’ exposure to that language. The promotion of TETE in East Asian contexts might also be a backlash against heavy use in the past of the grammar-translation method, in which learners’ L1 was used excessively in class (Atkinson, 1987).

Indeed, some empirical studies show that teachers’ greater use of the target language positively influences young learners’ proficiency, including a classical case study conducted in a primary school French education program in England and Wales (Burstall et al., 1974, cited in Turnbull, 2001). However, we have limited knowledge when it comes to TETE’s effectiveness in East Asian EFLPS.

Despite the promotion of TETE at the policy level, a number of studies have reported that TETE is not fully implemented in many EFLPS settings in East Asia. There seem to be a couple of reasons behind such “partial” implementation. The first reason concerns teachers’ own English-language proficiency. TETE requires teachers to have a certain level of proficiency in the target language, and not all EFLPS teachers in East Asia meet this requirement (Butler, 2004). Second, even if teachers have sufficient proficiency in the target language, young learners’ limited proficiency also makes implementing TETE difficult (D.-K. Kim, 2001; S.-Y. Kim, 2002). Macaro and Lee (2013) reported that 6th-grade students in Korea showed less favorable attitudes toward TETE compared with adult learners. Shin (2012) described Korean novice teachers’ difficulties implementing TETE when they socialized into a teaching culture in which individual novice teachers had little voice. Although Shin’s study was conducted among secondary school teachers, primary school teachers may encounter a similar challenge. Kang’s (2008) case study of a Korean EFLPS teacher found that the teacher adjusted her use of English according to her students’ preference for L1 use when necessary. It appears that individual teachers’ and learners’ specific experiences with TETE greatly influence their attitudes toward the teaching approach, as Carless’s (2004b) case study of a primary school teacher in Hong Kong described.

In recent years, a growing body of research, primarily from North America, has shown some positive effects of the controlled use of students’ L1 in FL classrooms (see Anton & Dicamilla, 1999; Atkinson, 1987; Cook, 1999, 2001; Harbord, 1992; Kharma & Hajjaj, 1989; Littlewood & Yu, 2011; Rolin-Ianziti & Brownlie, 2002; Schmidt, 1995; Swain & Lapkin, 2000; Turnbull, 2001; Turnbull & Arnett, 2002; also see Auerbach, 1993 for a discussion of English-only instruction in ESL contexts). The potential benefit of using learners’ L1 included: (a) facilitating language learning at the beginning level; (b) helping less-proficient students avoid falling behind; (c) reducing affective filters and making learners feel more secure; (d) facilitating thinking processing in L2; (e) promoting more
learner-centered environments and meaningful interactions; (f) helping learners develop shared perspectives in group work; and (g) saving time in certain situations. Levine (2014) suggested that students’ L1 can facilitate what Larsen-Freeman (2003, p. 24) called “grammaring.” Grammaring does not refer to acquiring fragmental pieces of rules and forms, but is defined as dynamic process of integration of forms, lexicon, and users in context in order to make up for what we understand as “grammar.”

There is no question that learners’ L1 should be used judiciously in FL classrooms. Research has found that the excessive use of learners’ L1 not only deprives them of exposure to the target language but also diminishes their motivation to pay attention to the target language, which in turn negatively influences the learning outcome (Schmidt, 1995). Atkinson (1987) suggested that the optimal ratio of L1 and the target language use for beginners should be 5% and 95%, respectively; Shapson, Kaufman, and Durward (1978) suggested a ratio of 25% and 75%. The rationale for such suggestions, however, is not clear. We can assume that the optimal ratios should be determined based on various factors, including the learners’ age and proficiency levels, the teachers’ proficiency levels, the content and objectives of the instruction, and the teaching environment (e.g., the degree of heterogeneity or homogeneity of the students’ L1 backgrounds). In Qian, Tian, and Wang’s (2009), 4-year longitudinal study of two Chinese primary school teachers’ code-switching, they found that the teachers used intersentential switching more frequently than tag or intrasentential switching, and the code-switching was used when the teachers clarified instructions and elicited learners’ responses. The researchers also found that the frequency of code-switching dropped greatly from the first to the second year, as the students’ proficiency levels increased.

As Qian et al.’s (2009) study suggested, not only is the amount of use of the target language important but so is the quality of its use. While it is critical for learners to be exposed to a wide range of language use in order to facilitate their learning, it may be challenging for some primary school teachers to use English beyond sets of common expressions, commands, and praises. For example, in D.-K. Kim (2001), primary school teachers in Korea reported that they did not use English when answering their students’ questions or when explaining the objectives and content of their lessons.

In sum, while EFLPS policies often encourage teachers to employ TETE in their classrooms, teachers often find it challenging to use TETE in East Asian contexts. The effectiveness of TETE in East Asia remains unclear. However, we do know that the controlled use of students’ L1 can be effective. Not only is the amount of target language important, but so is the quality of its use. The decision to employ TETE should be determined flexibly while considering various contextual factors.
3.3 Content-Based Instruction (CBT) and Immersion Programs

An instructional approach referred to as content-based instruction (CBT) or content and language integrated learning (CLIL)\(^6\) has gained attention in East Asia. In this approach, the instruction not only focuses on language learning but also on learning the content of subject matter such as math and science. The degree of focus between language and content learning varies across programs. Immersion programs are usually defined as a type of bilingual program in which learners receive subject matter instructions mainly or exclusively in the target language; these programs can be further classified into various sub-types depending on how the target language is introduced in the programs (e.g., partial immersion, total immersion, early immersion, late immersion, etc.). In practice, however, different types of “immersion” programs are employed under different labels in East Asia (Feng, 2007; He, 2011), and thus the relationship between the CBI/CLIL programs and immersion programs is not clear\(^7\). In China, most “immersion programs” are offered at affluent or prestigious schools and are often used as “a synonym for quality of education” (Hu, 2009, p. 51). However, the quality of immersion programs varies greatly, and it is not uncommon to find “immersion programs” that are operated on trial-and-error bases (Song & Cheng, 2011; Zhang & Adamson, 2007).

The effectiveness of CBI/CLIL among young learners is not yet clearly known. Research on CBI/CLIL has been primarily conducted at the secondary and postsecondary education levels in North America and Europe (see Nikula, Dalton-Puffer, & Llinares, 2013, for a review of the classroom interaction in CLIL studies), and it generally has shown that CBI/CLIL instruction tends to promote more negotiation of meaning, to create more varieties of occasions to use language, and thus to produce more student utterances than in regular FL programs. (It is important to note, however, that the amount and quality of interaction heavily depend on activities introduced in class.) It has also been reported that CBI/CLIL instruction may lower the “affective filter” among students. However, students in CBI/CLIL programs tend to make more linguistic “errors” in their production—and in particular more lexical-related errors—and they tend to monitor their production less than students in regular FL programs. This seems to be in part due to the fact that students in CBI/CLIL programs predominantly receive implicit corrective feedback (typically by recast or simply repeating the correct form without obstructing the communication) when they make “errors.” Such implicit feedback does not seem to lead to

---

\(^6\) The term CLIL is used primarily in Europe whereas CBI is used primarily in North America. CLIL is based on a framework composed of four elements: content, communication, cognition, and culture/community (Coyle, Hood, & Marsh, 2010).

\(^7\) Even in Europe, Cenoz, Genesee, and Gorter (2014) argued that the distinction between CLIL and immersion is not totally clear.
student uptake. With respect to content learning, Llinares and Whittaker (2010) reported that, compared with CLIL students, students who studied the same content subjects in their L1 outperformed in terms of their use of academic register.

Empirical research on CBI/CLIL at the primary school level is scarce. A longitudinal evaluation study of CLIL among Swiss primary school students (50% of the curriculum was taught in their L2, from grades 1 to 6, \( N = 27 \)) found that the young learners made steady progress in their L2 (both in oral and written domains). They caught up with their peers in the control group in their L1 (both oral and written) as well as in math, which they learned in L2 (Serra, 2007). The author attributed this result to the program’s instructional approach: the students in this CLIL program received literacy instruction in L1 and L2 in parallel, and the instructional subactivities and topics were designed in such a way that the students needed to use both languages to complete activities.

In China, Knell et al. (2007) found that children in a partial immersion program (50% of the curriculum was taught in English, \( N = 92 \)) had higher scores in vocabulary, word identification, and oral proficiency in English, compared with non-immersion children (\( N = 91 \)). As with Serra’s study, even though the immersion children received fewer hours of instruction in their L1, they were not behind in their L1 (measured by Chinese character reading). Similarly, Cheng, Li, Kirby, Qiang, and Wade-Woolley (2010) studied three immersion programs in China and found that the children in these programs (30-40% of the curriculum was taught in English, grades 2, 4, and 6, \( N = 998 \) altogether) outperformed non-immersion children in Chinese and math, as well as in English.

The positive results of immersion programs are promising; however, we need to be careful about possible selection biases in studies that compared immersion and non-immersion children. As I discuss in more detail in the following section, immersion programs usually attract children with higher socio-economic status and/or higher aptitudes; most immersion programs tend to be highly selective. In any event, the existing studies are still very limited both in number and scope, and more research is necessary in order to draw any conclusions on the effectiveness of the CBI/CLIL and immersion approaches.

4. SOCIAL-ECONOMIC DISPARITIES AND ENGLISH LEARNING

There are growing concerns about socio-economic disparities and achievement gaps among students by their socio-economic status (SES) in many parts of the world, and East Asia is no exception. SES is usually measured by household income, educational levels, and social capital (e.g., collective values brought by resource networks, etc.). English may be one of the school subjects that is most strongly influenced by parental SES in East Asia.
In an FL context, people usually have limited access to the target language, and thus if one wishes to have greater exposure to English, he or she needs extra resources to create such opportunities.

In recent years, “English fever” (J.-K. Park, 2009, p. 50), or an intense quest for learning English, has taken hold even among young learners. East Asian countries such as Japan and South Korea have substantially high private expenditures on education (OECD, 2011). And importantly, children’s English achievement is highly correlated with the amount of money that parents spend on their children’s private English education (e.g., Statistics Korea, 2010, for the case of South Korea).

English-speaking bilinguals are becoming a symbol of success—and sometimes a symbol of wealth, because it is costly to raise a child as a bilingual in an EFL context such as China, Japan, South Korea, and Taiwan. S.-J. Park and Abelm (2004) have described English education in Korea as an “inter-generational gendered project” (p. 647) because mothers are increasingly responsible for the success of their children’s English learning. The ways that mothers commit themselves to their children’s English-learning activities reflect not how much they value English but also their SES background. Mothers who want their children to be “global elites” often accompany their children abroad for study. Depending on their SES, however, children’s and their mothers’ experiences in the host countries vary (e.g., J.-H. Kim, 2010, for a case of Singapore). In a case of Chinese mothers who went to Singapore for their children’s English learning, Huang and Yeoh (2005) described difficulties that mothers with nonprivileged backgrounds encountered.

Research on the achievement gap in English by SES among young learners in East Asia is still in its infancy, and very few studies have been conducted so far. Among the few is J.-R. Kim’s (2013) meta-analysis of extensive reading in primary schools in Korea. Of particular relevance to this article is Kim’s finding that the effect of extensive reading was bigger in wealthier regions. My own studies in Changzhou, China (Butler, 2014a, 2014b, 2014c), also reported that SES influenced children’s English performance at early stages in their learning, particularly in the oral domain. I also found that parents with lower SES started losing faith in their children’s abilities to learn English successfully. Children with lower SES backgrounds also lowered their perceptions of their own competence sometime around the transition from primary school to secondary school. Both parents’ beliefs about their children’s competence and the children’s perceptions of their own competence turned out to be strong predictors for their English performance. Moreover, I found that parental SES had varying impacts on the children’s motivation to learn English. Whereas parents with higher SES tended to adjust their assistance according to their children’s changing needs, parents with lower SES tended to remain controlling and often failed to respond to their children’s changing needs to foster their self-competence and self-determined motivation. In addition, high SES parents’ abilities to provide their children with greater
opportunities to use English outside of school were increasingly advantageous for maintaining their children’s high levels of intrinsic motivation (Butler, 2014c).

Policy makers have made efforts to narrow the gaps in achievement by SES. To illustrate this point, I use Changzhou in China as an example (Butler, 2014b). In Changzhou, three innovative policies have been implemented since 2012: (a) a teacher rotation project; (b) a free tutoring project; and (c) changes in high school admission policies. The teacher rotation project aimed to allocate the teaching resources more evenly across the city by rotating select teachers to different schools, in order to avoid concentrating the most capable teachers in the most prestigious schools, as well as to share the knowledge and skills that capable teachers have with other teachers. The aim of the free tutoring project was to make education accessible for all students, regardless of their parents’ SES. Any students in the city can sign up for a free, individualized tutoring service on weekends and can receive assistance from public school teachers in the city on any academic difficulties that they may encounter. The rising cost of private education and the achievement gap by SES were major driving forces for implementing this service. Changzhou also changed the admission policy for high school admissions. Instead of accepting students purely on merit (i.e., by exam scores), prestigious high schools set a quota for accepting students from less-prestigious junior high schools. This policy aimed to encourage students to be enrolled in designated primary and junior high schools according to their residency,8 and to narrow achievement gaps by SES.

In addition to the three policy changes at the city level, some changes took effect at the Jiangsu provincial level in 2012 as well. Among various changes that the province made, the ones that are most relevant to this discussion involve the difficulty level of the curriculum and associated materials. The policy lowered the minimum requirement in the curriculum, and the province adopted easier English-language textbooks. The motivation for lowering the standard came from concerns that English learning has increasingly become a burden for many students. By reducing the minimum standards while granting greater autonomy to local schools, the policy aimed to reduce the burden on students struggling with English education.

Changzhou’s policy changes reflect conflicting attitudes about English education. On the one hand, by accepting the premise that English is important for everybody as a powerful lingua franca and as a major barometer for academic achievement in the formal education system, the policies aimed to narrow gaps by trying to ensure equal access to English education. They did so by rotating capable teachers, offering free tutoring on weekends for anybody in need of assistance, and requiring prestigious high schools to

---

8 It is not unusual for parents with resources to send their children to prestigious primary and junior high schools by relocating their residency or paying an extra fee.
accept a certain number of students from less prestigious junior high schools. On the other hand, the policies that modified the English curriculum and lowered the standards of the English-language textbooks appeared to attempt to reduce the weight of English-language study in the formal education system. By granting greater autonomy to local schools, the policies attempted to create flexibility in their English education according to the local needs. It is important to remember, however, that schools are strongly confounded by SES. Thus, greater local autonomy may eventually contribute to widening gaps by SES (Butler, 2014b).

The seemingly conflicting policy situation in Changzhou is probably not unique to this city; rather, it is likely that many other governments and regions face similar dilemmas. Unfortunately, we have yet to find a good solution to this problem.

5. CONCLUSION AND FUTURE DIRECTIONS

As early English-language education has gained popularity in East Asia and many other regions in the world, a number of issues have arisen. In this article, I focused on three issues prevalent in East Asia, namely (a) the rise in English-language education for young learners due to the widespread belief that “younger is better” when it comes to FL learning; (b) the emergence of a number of new instructional models in early English teaching; and (c) growing gaps in English achievement by children’s SES.

First, despite the prevailing belief in “the younger the better” when it comes to FL learning, we do not have strong evidence supporting the assumption, at least based on the limited studies available. Neurological studies are emerging but have not given us a clear picture of the relationship among neurological processing, actual linguistic performance, and age. Therefore, we should not blindly assume that younger introduction of EFLPS is always better. In order to inform policy and enhance instructional decisions, we need information about the conditions for optimizing learning for different student age groups. Because we can assume that introducing EFLPS earlier in the educational process tends to increase educational expenditures from a policy point of view, analyses of cost-effectiveness should also be conducted. Finally, we should also better understand how young learners approach and process FL. It is highly possible that we have not yet captured the kinds of abilities that children are developing, such as their tolerance of ambiguity and other non-linguistic abilities. Thus, creative approaches to measurements also need to be explored.

Second, with respect to instructional methods and approaches to early English teaching, we still have relatively limited information in East Asian contexts. Popular instructional methods and approaches that originated mostly in the West, such as CLT, TBLT, TETE,
and CBT/CLIL, are now strongly promoted to teachers in East Asia, but they may not work well for EFLPS in this context. To account for the different contexts, adaptation—as opposed to adoption—is necessary, with sufficient attention to both local contextual factors and age factors. Before blindly assuming that these instructional methods and approaches should work for EFLPS in East Asia, we need to better understand (a) how to maximize their effectiveness if they do, indeed, work; (b) what kinds of conditions are necessary for them to be effective (qualifications necessary for teachers, young learners’ characteristics and needs, etc.); and (c) how best to train teachers to adapt them to their teaching contexts. More classroom-based research on EFLPS in East Asia would help address these topics.

Finally, I addressed achievement gaps based on SES. Although the impact of SES on student achievement and other behavioral matters has been much studied in the general education field, the field of applied linguistics has not paid sufficient attention to the role of SES in language learning until very recently. To echo Zou and Zhang’s (2011) quote that I include in the introduction to this article, to think of English as just a school subject fails to appreciate its potential for significantly influencing one’s life. Both researchers and policy makers need to pay much closer attention to growing achievement gaps in English among learners at earlier stages of their English learning.

REFERENCES

Butler, Y. G. (2004). What level of English proficiency do elementary school teachers need to attain in order to teach EFL?: Case studies from Korea, Taiwan, and Japan. TESOL Quarterly, 38(2), 245-278.
Butler, Y. G. (2014b). Socioeconomic disparities and early English education: A case in


113-134.


Applicable levels: Elementary, secondary, tertiary

Yuko Goto Butler
Graduate School of Education
University of Pennsylvania
3700 Walnut Street, Philadelphia,
PA 19104-6216, U.S.A.
Phone: 215-898-5937
Email: ybutler@gse.upenn.edu

Received in September 2014
Reviewed in October 2014
Revised version received in November 2014