Developing Phoneme-Grapheme Recognition for English as a Foreign Language: A Longitudinal Study at Japanese Primary School

Kaori Nakao[a,*], Quint W. L. Oga-Baldwin[b], Luke K. Fryer[c]

[a,*] **Corresponding Author:** Kaori Nakao, Seinan Gakuin University, Fukuoka, Japan.
E-mail: k-nakao@seinan-gu.ac.jp
ORCID: https://orcid.org/0000-0001-9015-8765

[b] Quint W. L. Oga-Baldwin, Waseda University, Tokyo, Japan.
E-mail: quint@waseda.jp?
ORCID: https://orcid.org/0000-0003-3455-6456

[c] Luke K. Fryer, The University of Hong Kong, Hong Kong, Hong Kong.
E-mail: lukefryer@yahoo.com
ORCID: https://orcid.org/0000-0001-6250-5950

Abstract

Foreign language was recently added to Japan's national primary school curriculum. Phoneme-Grapheme Recognition (PGR) skills are a critical step in both L1 and L2 development. Due to its simplistic approach and the lack of new investment in teacher training, research regarding the impact on PGR skills are important. This study tested the relative differences and growth of primary school students' (grades 3, 4, 5, 6) PGR skills across two semesters. Semester-1 (July) and Semester-3 (March) students (n = 256, female n = 130) completed the same PGR test during regular class time. ANOVA and follow-up pairwise tests assessed achievement differences between grades and across the two-semester gap. Difference testing between grades indicated substantial (R^2 = .32 & .19) and statistically significant (p < .001) differences at both time points. Pairwise follow-up tests pointed to two steps in statistically different ability (grades 3-4/5-6). Longitudinal tests suggested that the current Japanese national elementary school curriculum supports phoneme-grapheme skill development; However, the two-step ability grouping indicated that the current national curriculum is not sufficiently detailed and/or rigorous to ensure annual student improvement in PGR skills.

Keywords:

Phoneme-Grapheme Recognition; Elementary School; Foreign Language Education; Longitudinal

Introduction

Research on the foundations of learning to read has important theoretical and practical implications for elementary school education. Substantial research during the last three decades has indicated that efficient word reading is a necessary but not sufficient condition for the development of reading comprehension (Melby-Lervåg et al., 2012; Castles et al., 2018). Past studies have demonstrated that the development of phonological recognition (phonological awareness) is a prerequisite for the acquisition of literacy in alphabetic languages (Ehri et al., 2001); in fact, the importance of early phonemic awareness may extend well beyond its influence on early reading and spelling skills (Hulme et al., 2020) to broader language development.

Learning to read in a foreign language also requires some degree of mastery of the phonology of the new language (Melby-Lervåg & Lervåg, 2014). This need for phonological mastery is complicated by children's first language literacy skills, which can interfere with their acquisition of a new language (Caravolas et al., 2019). Children who fail to acquire awareness of the system of phonemes and the letter sound connections, can continuously struggle with literacy (Moll et al., 2014). As a result, early reading instruction in an alphabetic language — be the language foreign or native — needs to first focus on teaching letter to sound correspondences (Kahn-Horwitz, 2020). Phoneme-Grapheme Recognition is a critical element of foreign language acquisition for English and other alphabetic languages, despite it often not receiving the attention it deserves in many cases (Temur & Sezer, 2023). As a result, young learners of English and other alphabetic languages require standards and developmental benchmarks to establish goals for and signal improving sound-letter (phoneme-grapheme) recognition in the new language learning.

Such standards and developmental benchmarks must not only be set, but also be practical for elementary school classrooms. The present study addresses this issue of making these benchmarks practical through the longitudinal trial of a classroom test of phoneme-grapheme recognition (PGR) with Japanese elementary school students (grades 3-6). The current research assessed elements of Japanese students' phoneme-grapheme recognition for English across three semesters of curricula to create basic benchmarks for the development of sound-letter correspondences.

Literature Review

General Phonological Awareness, Phoneme-Grapheme Recognition, and Language Learning

Phonological processing of text is widely understood to be a powerful predictor of later reading comprehension (Rueckl et al., 2015). Some form of sound-letter connection is taken for granted as a universal element in learning to read (Shankweiler & Fowler, 2019). There is also evidence that learners acquire the components of literacy specific to the decoding skills necessary for that language (McBride et al., 2022). In languages with transparent phoneme-grapheme representation such as Finnish and Korean, these skills first involve recognizing and decoding individual sound units and connecting them with visual representations. For opaque logographic languages like Chinese, recognizing characters is often a more complex matter of mapping sounds onto the individual logograms. Other languages like English fit somewhere in between these two, with elements of mapping sound through memorization of lexical units as well as decoding words by semi-regular sound units.

The process of developing literacy in any language is thus a matter of recognizing the appropriate balance of skills necessary for that language, although the specific elements of this in different languages remains a subject of debate (McBride et al., 2022). One important component of learning to read alphabetic scripts is the development of sound-letter correspondences (Siegelman et al., 2020). As a step toward acquiring connections between orthography and phonology, young learners first acquire some measure of phonological awareness (Melby-Lervåg et al., 2012), defined as an understanding of how speech in a language is broken into individual sounds and how those sounds may be recombined to form words (Kahn-Horwitz et al., 2012). Learners then use this knowledge to create grapheme-phoneme correspondences, such as the recognition that in English the letter "D/d" most often represents the sound [d]. For these alphabetic languages, first- and second/foreign language phonological abilities are highly correlated (Melby-Lervåg & Lervåg, 2014), with some transfer of skills across alphabetic languages (Arfé & Danzak, 2020). Research has suggested that L1 reading can significantly predict L2 reading performance in children learning to read (Caravolas, et al., 2019). Furthermore, these elements are teachable (Castles et al., 2018), with children responding positively to instruction across diverse L1s (McArthur et al. 2018).

Elementary school foreign language curricula are designed as a way to develop the fundamentals that later lead to literacy and other language skills (Kang & Crandall, 2013). Despite a consensus regarding the central role of phonology in developing literacy in foreign language learning (Koda & Yamashita, 2019; Yamashita, 2022) and an uptick in research on the phonological contributions to reading in the L2 (Jeon & Yamashita, 2022), the process of acquiring phoneme-grapheme concordance in foreign language learning environments is an under-researched issue (Arfé & Danvak, 2020; Ikeda, 2018). Further investigations are necessary to clarify the process of letter-sound recognition within young foreign language learners and then develop tools to improve curricula to best support it.

English phonological awareness in the context of Japanese native speakers

In the case of the Japanese language specifically, phonological awareness refers to the ability to understand how mora (i.e., syllable-like sound units) make up words: for example, 猫 (neko, cat), ne-ko is two moras. Since Japanese characters directly correspond to a single syllabic unit (mora) (Kubozono & Honma, 2002), when teaching English to Japanese children, it is necessary not only to ensure that students recognize each English phoneme, but also to understand that some English characters do not directly correspond to only one sound (Siegelman et al. 2020). Furthermore, it should be taken into

account that when some characters are combined or arranged differently, their sounds can change. These complexities can also impact English speaking and listening (i.e., a considerable portion of English, like other languages, can only be fully understood through literacy) as well as reading fluency making their direct instruction important (Caravolas et al., 2019). For these reasons, it is necessary to support school curricula in developing phonemic and broader phonological awareness ability during primary school L2 English education.

While acknowledging the increasing amount of time the national curriculum has committed to foreign language instruction, it is also important to note some of the natural barriers Japanese elementary school students face when learning English. As part of the Japanese students' native language studies, students learn four different types of scripts. Kana refers to two of these: a. First, Kanji, which is a Japanese script based on Chinese characters; b. hiragana and katakana, which are types of phonetic characters, and basically one character is classified as a syllabary representing one syllable. Romaji is the fourth script, referring to the general rules for transcribing kana (Japanese script) into Latin letters (Romaji notation). Romaji is studied in grade three. For many students, this is their first formal introduction to the letters of the alphabet. The difference between Romaji and the English alphabet is that much like Kana, each letter corresponds to one sound and in some cases, it is a sound that is not represented clearly in English. This means that Japanese students can have difficulty identifying the English voiceless labio-dental fricative [f]. Japanese language does not use this exact sound, instead there is an aspirated voiceless fricative [ɸ] which is romanized using both "h" and "f;" thus characters like 服 (clothing) can be romanized as either "huku" or "fuku," though "fuku" tends to be preferred. Other phonological differences and near-similarities abound (Tsujimura, 2014).

Prior tests of phonemic awareness and sound-letter recognition have indicated cross-linguistic transfer as a potentially confounding phenomenon for Japanese learners of English. Ikeda (2018) conducted one test of elementary school learners' phonemic awareness in English, showing that Japanese children tend to process English sounds more similarly to Japanese sounds, defaulting to the use of mora units with their concordant consonant-vowel structure. Later tests of elementary students' recognition of English sounds have indicated that students are able to recognize the initial sounds of familiar words when those sounds have a clear regular correspondence to Japanese Romaji, but that performance degrades when students are presented with words with less consistent sound representations (Nakao et al., 2022). Despite the interference of the Romaji system, cross-sectional investigations of students' Phoneme-Grapheme Recognition show a clear stepwise increase in students' knowledge as they progress through elementary school (Nakao et al., 2022).

Japanese students can thus struggle to identify English sounds, both due to general differences in phonological awareness between the languages (Ikeda, 2018) and interference from the Romaji system (Nakao et al., 2022; Okada, 2005). This kind of inconsistency is exacerbated by the fact that both Romaji and English are now taught at the same curricular point (MEXT, 2017). As Japanese language is the dominant subject during elementary school, romaji is learned much more thoroughly than English phoneme-grapheme correspondences. As a result, students have learned a set of rules for alphabetic notation (i.e., one letter, one sound) that is often not correct or only partially correct for English (Okada, 2005). The fact is that many young learners struggle to acquire English orthography rules generally (Russak & Kahn-Horwitz, 2015). The added difficulty of translating from Japanese L1 to English L2 (Ikeda, 2018) only compounds this problem, making it a critical area for language learning research with young learners.

Learning English in Japan's elementary schools

For two decades foreign language study has been progressively integrated into Japan's national curriculum for elementary school students. Beginning in 2002, English oral communication became an optional part of Japanese elementary school students' comprehensive study. Beginning in 2011, one class per week for foreign language activities became compulsory for fifth and sixth grade students. In the subsequent national course of study, announced in 2017, foreign language activities were included in grades three and four, while foreign language studies were established as formal subjects for grades five and six (MEXT, 2017).

In 2017, Japanese Ministry of Education, Culture, Sports, Science and Technology [henceforth MEXT] announced that foreign language studies would become a formal subject. What had been once-a-week (35 hours a year) foreign language activity became a twice-a-week (70 hours a year) formal subject of study for fifth and sixth grade students, with third and fourth graders now getting one lesson each week (MEXT, 2017). Upper elementary school has focused on the comprehensive teaching and learning of four skills (i.e., listening, reading, speaking, and writing).

According to the current Course of Study (MEXT, 2017), students are expected to be able to write uppercase and lowercase letters in block type, and to be able to copy out familiar oral expressions. Importantly, although MEXT is clear, regarding the difficulties

Japanese students face in acquiring literacy, while students are expected to identify letters (i.e., A or C, /ei/ or /si:/), they are not yet expected to recognize the variety of sounds these letters might represent (e.g., /æ/ for A as in bag, both /s/ and /k/ for C in circle; etc) (MEXT, 2017, p. 78). Despite this curricular direction, the overwhelming evidence continues to lean towards a need to teach decoding, including phonological awareness and phoneme-grapheme correspondence (Melby-Lervåg & Lervåg, 2014). Despite the need for appropriate instruction into sound-letter correspondences, due to current MEXT policies, no standard tools or methods have been put forth to assess students' acquisition of phoneme-grapheme correspondences for use in the Japanese curricular environment.

The aims of the national curriculum are clear: Japanese elementary school students are expected to develop the fundamentals of English (MEXT, 2017). However, the national curriculum does not include a description of expected teaching methods or means of evaluation, including methods for bridging the gap between basic letter recognition and word reading. Additionally, there are few cross-sectional and longitudinal empirical studies on Phoneme-Grapheme Recognition with L2 English students in Japanese elementary schools. While the national programme was initiated in 2017 and guidelines/textbooks were created for the widespread start in 2020, there has been no nationally recognized training or retraining curriculum for future or current elementary school teachers (Machida, 2016; Nakao et al., 2019). Despite being required to teach English, this has left many teachers lacking confidence in teaching the new subject (Yonezaki et al., 2016). Many current Japanese elementary school teachers do not have even basic knowledge and skills related to language and/or language teaching; many hold no specific licenses or recognized training in language pedagogy. (Butler, 2015; Nakao et al., 2019).

The current study

Phonology is thus a central issue for native and foreign language literacy skill development. Despite its importance, direction regarding how the phonological aspects of learning might be supported or assessed is not substantively addressed in the current Japanese national curriculum or teaching approaches. A few persistent issues further exacerbating the difficulty in addressing this aim are: a). a lack of re/training on the part of current and future elementary school teachers; b). a lack of a framework and measurement for students' developmental stages of phonemic awareness learning; and c). insufficient national and international research in the area of L2 PGR with elementary school students.

Seeking to begin to address some of the issues highlighted, the present study built on a cross-sectional pilot of a practical classroom test of basic PGR (Nakao et al., 2022). While there are various levels of phonological knowledge (sound isolation, combination, blending, etc.), the most basic testable level rests on the ability to identify sound-letter connections based on the curriculum. For this reason, the present study's measurement focuses on the initial connection between phonemes and their regular graphemic representations (i.e., /k/ represented by the letter k but not irregulars such as c). The initial pilot study (Nakao et al., 2022) identified a clear pattern of cross-sectional growth across grade levels, with older children performing better than younger children. The present work sought to extend the work beyond cross-sectional findings to confirm a continued trend in the same direction over time.

In the current longitudinal examination of Japanese elementary school students' development of basic PGR, students' basic letter-sound connections were assessed twice (same students), across two semesters of school, seven months apart. Analyses tested differences between the participating grades (cross-sectionally and longitudinally) and students' ability across time.

Aims

Building on the foundation of Phoneme-Grapheme Recognition research and a previous pilot study in the same school context, this study aimed at extending our understanding of Japanese elementary school students' general and specific Phoneme-Grapheme Recognition across their four years of English instruction (grades three to six). This study therefore addressed two research questions and tested three hypotheses.

Research Question One (RQ1): Does Phoneme-Grapheme Recognition (PGR) present significant differences at each of two time-points, with a seven-month gap? Hypothesis One (H1): Based on Nakao et al., (2022) incremental differences were expected between grades, but only between grades three and four, three and five, three and six, four and six.

Research Question Two (RQ2): Does students' PGR increase across two semesters of instruction? (a): Does the overall score go up? (b): Does the PGR increase for each grade? Hypothesis Two (H2a): an overall increase in elementary school students' PGR was expected across the two semesters of the study. (H2b). Based on results from previous studies (Nakao et al., 2022), PGR was expected to increase for each grade but the differences in increase for grades five and six might not be significantly different.

Methods

Sample & Ethics

We present cross-sectional and seven months of longitudinal data from a study of public elementary school students aged eight to twelve. Elementary school students from Western Japan participated in this study. Time 1 (T1) (Total *n* = 261, female *n* = 138) in July 2020, and Time 2 (T2) (Total *n* = 256, female *n* = 130) in March 2021. Students in each grade were taught under the same curriculum. All elementary schools in Japan were closed from March to May in 2020 due to -19. The new national curriculum began from 2020 April (MEXT, 2017). The 3rd and 4th grade students engaged in foreign language activities class once a week, while 5th and 6th grade students attended twice a week.

A Japanese teacher of English who worked at this school collaborated with the researchers to conduct tests. All participating classes were taught by this Japanese teacher of English.

Data collection occurred at a small-to-medium-sized school in the suburbs of a large Japanese city in western Japan. The assessment was administered in the third semester of the 2020 school year during regular foreign language classes. Ethical oversight was included in the review process for JSPS Grant-in-aid for Scientific Research and approved by schools and boards of education. All procedures were in accordance with the ethical standards of the national research committee.

Methodology

The sound-letter recognition test used in this study consisted of fifteen items, selected from the third grade assigned Ministry textbook (MEXT, 2017). An optical mark reader (OMR) sheet was used for students to report their answers, with letters A to Z, and ? for "I don't know" presented as choices (test sheet included in appendices). Capital letters were used to increase students' recognition and ease of reporting. Previous studies using this testing format have indicated students with a stronger recognition of initial sounds have better PGR (Sodoro et al., 2002). The instrument in question is a criterion-referenced test designed by the researchers to investigate the specific phonological needs of Japanese elementary school students (Sodoro et al., 2002). The test had been successfully piloted in a study conducted seven months prior (Nakao et al., 2022).

The test was administered by the native Japanese teacher of English who taught at the elementary school where the study took place during regular class time. The participating teacher pronounced each of the fifteen words twice with no specific emphasis on any syllable. The teacher participated in a brief training on how to read the alphabet words by the first author. The students selected the letter on the test sheet (appendix) after hearing each of the fifteen target words. The teacher distributed and collected test papers for all classes. Students took between six to nine minutes to complete the test.

Tests were scanned and converted to data using OMR software. The first author reviewed each scan to address marking errors and input mistakes. The finalized data were imported to JMP14.1 (SAS Inc., 2019) for analysis.

Analyses

Descriptive statistics and test normality

This study's analyses proceeded in four stages. To begin with, students' scores for the 15-item test were calculated by summing their correct answers. Students' summed scores and their standard deviation were reviewed. To assess the normality of students' test scores, skewness and kurtosis were calculated. Based on George and Mallery's (2010) suggested heuristics for assessing distribution normality, if skew values are between -1 and 1, that skewness is deemed to be approximately symmetric. Based on the same heuristics, if kurtosis values are between -2 and 2, they were deemed to be within reasonable limits. Next, reliability (Cronbach's Alpha) for tests at each time point were calculated.

Overall and grade test longitudinal results

Addressing RQ1 and H1, analysis of variance (ANOVA) was conducted for both T1 and T2. For this analysis, the four (third, fourth, fifth, and sixth) participating grades were used as the independent variable (IV), and test scores as the dependent variable (DV). This analysis was followed by pairwise difference testing (Tukey-Kramer Honest Significance Difference; Tukey-Kramer HSD) to examine differences between the students' year of study. For all tests, statistical significance was set at $p < .05$. Where feasible, effect sizes (Cohen's d or R2) were reported.

Addressing RQ2 and H2, *t*-tests were conducted to test for differences between T1 and T2. First, overall changes between T1 and T2 were tested. Following this initial broad test, *t*-tests were conducted to test for differences in each of the four grades included in the study. In addition to p-values, confidence intervals were reviewed. For all *t*-tests Bonferroni correction was employed to account for multiple significant tests: As a result, statistical significance was set at $p < .007$ (7 tests) for all tests.

Results

Both T1 and T2 tests presented sufficiently normal distribution of scores; skew was approximately symmetric (-1 to 1) and kurtosis within reasonable limits (-2 to 2; George & Mallery, 2010). Table 1 presents the descriptive statistics with skew and kurtosis for each grade at T1 and T2. At both T1 and T2, the test presented reasonable reliability (.90 and .84 respectively). Please see Table 1 for Means and SDs for T1 and T2 tests for each grade.

A one-way between-subjects ANOVA (H1) compared the effect of year of study (IV) on test score (DV). For T1 test, there was a significant effect of year of study (IV) on test score (DV) at the $p<.05$ level for four conditions [$F(3, 257) = 40.97$, $p < .0001$]. Given the simple nature of the test, year of study accounted for a considerable amount of the variance in students' test scores ($R^2 = .32$). For T2 test, there was a significant effect of year of study (IV) on test score (DV) at the $p < .05$ level for four conditions [$F(3, 252) = 19.64$, $p < .0001$]. Given the simple nature of the test, students` grade accounted for a reasonable amount of the difference in students' test scores ($R2 = .19$).

Table 1.
Descriptive Means and SDs by Time 1 and Time 2 test score

| Grade | Time 1 | | Time 2 | |
	M	SD	M	SD
3	2.84	2.77	6.31	3.90
4	7.06	3.74	8.25	3.60
5	8.64	3.93	10.00	3.31
6	9.18	4.19	10.87	3.51

A Tukey-Kramer HSD post hoc test (H1) was significant ($q* = 2.59$, Alpha = .05) for the T1 test. Results suggested that there was no statistically significant difference between grades 4 and 5, and grades 5 and 6. Grade 3 scored lower than grades 4, 5, and 6 ($p < .0001$). Grade 4 scored lower than grade 6 ($p < .001$). For T2 test, there was no statistically significant difference between grades 5 and 6. Year 3 scored lower than grades 4 ($p < .05$), 5 ($p < .0001$), and 6 ($p < .0001$). Grade 4 scored lowered than grades 5 and 6 ($p < .0001$).

Tables 2 and 3 present the pairwise differences, confidence intervals, and p-values for T1 and T2 tests respectively.

Table 2.
T1 Tukey-Kramer HSD post hoc test for test scores by each grade

Grade	Grade	Difference	Std Err Dif.	Lower CL	Upper CL	p-Value
6	3	6.34	0.63	4.70	7.98	<.0001
5	3	5.80	0.65	4.13	7.47	<.0001
4	3	4.22	0.63	2.59	5.85	<.0001
6	4	2.12	0.64	0.47	3.77	0.0056
5	4	1.58	0.65	-0.11	3.26	0.0750
6	5	0.54	0.65	-1.15	2.23	0.8404

Table 3.
T2 Tukey-Kramer HSD post hoc test for test scores by each grade

Grade	Grade	Difference	Std Err Dif.	Lower CL	Upper CL	p-Value
6	3	4.56	0.64	2.89	6.22	<.0001
5	3	3.69	0.65	2.00	5.38	<.0001
4	3	1.94	0.63	0.30	3.58	0.0127
6	4	2.62	0.65	0.95	4.29	0.0004
5	4	1.75	0.65	0.06	3.44	0.0398
6	5	0.87	0.66	-0.85	2.59	0.5615

A Tukey-Kramer HSD post hoc test (H1) was significant ($q* = 2.59$, Alpha = .05). Results suggested that there was no statistically significant difference between grades 5 and 6, but that students' Phoneme-Grapheme Recognition increased across longer time spans (i.e., grades 3 compared to 5 and 6, grade 4 compared to 6).

Addressing RQ2, a pairwise t-test and examination of confidence intervals confirmed H2a. Between T1 and T2, test scores increased significantly: the matched t-test, $t(245) = 12.499$, $p < .007$, $d = .80$. Addressing H2b, a statistically significant difference was found across T1 and T2 for grades three, four, five and six: Grade three, $t(64) = 8.81$, $p < .007$, $d = .95$; Grade four, $t(64) = 4.10$, $p < .007$, $d = .32$; Grade five, $t(57) = 5.91$, $p < .007$, $d = .36$, and Grade six, $t(59) = 7.20$, $p < .007$, $d = .42$.

Discussion

The present longitudinal study builds on cross-sectional pilot research undertaken in the same school context as a previous pilot study (Nakao et al., 2022). The previous pilot established the basic measurement and practicality of the current testing format, while prior studies have established a clear link between Phoneme-Grapheme Recognition and literacy (Allen-Tamai, 2019). On this foundation, the current study attempted to begin to map the development of basic Phoneme-Grapheme Recognition (PGR) for Japanese elementary school students (grades three

to six) across two semesters of academic study. The current study focused on two research questions and tested three hypotheses. First, does students' sound-letter recognition significantly differ (based on their grade) at each of two time-points separated by seven months (RQ1)? Based on Nakao et al., (2022) pilot findings, incremental differences were expected between the grades at both time points; however, statistically significant differences were expected just between years three and four and then two-grade gaps (H1). Statistically significant differences between grades at both time points were found. The hypothesis was partially supported by the pairwise difference testing Tukey's Honest Difference results. T1 test results presented statistically significant differences between all grades except 4-5 and 5-6. T2 test results presented statistically significant differences between all grades except 5-6.

Longitudinal tests addressed RQ2: Does students' phoneme-grapheme recognition increase across two semesters of instruction? Analyses aimed to test whether students' overall score, and therefore students' PGR, increased in each grade. Hypothesis Two (H2a) proposed that an overall increase in elementary school students' PGR was expected across the two semesters of the study. Furthermore (H2b), based on Nakao et al. (2022), PGR was expected to increase for each grade across the seven-month gap, but the increases for years five and six were not expected to be significantly different. Addressing H2a, pairwise t-test and examination of confidence intervals confirmed that the increase in test between T1 and T2 was statistically significant. Addressing H2b, statistically significant differences in students' PGR were found between T1 and T2 for each of the four grades included in the study.

Theoretical implications

Do very simple foreign language curricula, with superficial focus on PGR provide sufficient support students' PGR skills?

Both the pilot study (Nakao et al., 2022) and the current study suggest that even limited curricula and instruction provide some support for elementary school students in developing basic PGR. Results are consistent with prior indications that language exposure can positively impact skills like basic phonemic awareness and grapheme recognition (Hulme et al., 2020; Silverman et al., 2020). This finding supports the government's efforts to initiate the national curriculum and make foreign language a formal part of elementary school education. Even with very basic curricular guidelines, simplistic textbooks and teachers who often cannot substantively speak the language of instruction, students' PGR does improve.

Do students at different grades have meaningfully different phonemic awareness?

With a cross-sectional cohort study, and using the same type of test, Nakao et al. (2022) demonstrated that Japanese elementary school students' PGR improves cross-sectionally by grade (progressing from grade 3 to 6). The stepwise difference was not always statistically significant, suggesting that particularly in the later grades, students might not be sufficiently challenged by curricula and instruction. Across two PGR assessments at the same school, separated by seven months, the current study supports those findings. Presumably during the earlier years of elementary school, curricula are sufficient to support students in gaining very basic sound-letter recognition. However, for more comprehensive PGR skill development, students might need more challenging curricula and/or more intensive instruction to continue to improve.

Can simple classroom tests capture phoneme-grapheme recognition (PGR) gains?

One large barrier to improving elementary school students' PGR is practical and effective assessment. The current study confirms that a short, pen-and-paper assessment of initial word sounds can be conducted by classroom teachers with students from grades three through six. Furthermore, such an assessment can provide reliable information about students' current and growing PGR. The test used here suggested the development of PGR during the school year, indicating suitability for use as a standard-based academic ability test (Sodoro et al., 2002). These findings can be built on, creating regular assessment or a wider variety of sounds, providing direction for students and educators about the language learning process across early learning years.

Practical implications

What are the benefits of basic foreign language instruction and how might they be enhanced given the limitations inherent in Japanese elementary schools?

One clear outcome of this study and its pilot is the fact that even simple engagement with a foreign language has benefits. Even without clearly structured curricula and explicit instruction, students' PGR does see a meaningful increase across elementary school experiences. Findings confirm the fact that children's language skills improve over time and are thus likely teachable with appropriate interventions (Hulme et al., 2020).

Currently, explicit phonological instruction is not yet part of the curriculum of foreign language education in Japanese primary school education (MEXT, 2017). Even though the current test indicates some potential

for growth without explicit instruction, the need to teach the connection between sounds and letters that are not in their mother tongue (Arfé & Danzak, 2020; McBride et al., 2022) is also well-established. Greater focus on this in the future is likely to increase teachers' impact on students' PGR, and thus improve decoding skills and prevent literacy problems that arise later (Coulson et al., 2013).

What is the potential of simple classroom tests for supporting instruction and curricula in Japanese elementary schools going forward?

A second important outcome of the present study is the fact that a very simple classroom test of PGR is a useful and reliable tool for classroom teachers. This classroom PGR assessment is a first step in the right direction, opening the door to a wider variety of assessments, undertaken regularly across elementary school foreign language classes. They have the potential to not only estimate PGR growth but also set goals/benchmarks for each year and even expectations for basic language fluency prior to the critical transition to secondary school.

How do we address weaker learning gains for grades five and six?

One issue of immediate practical value for educators that arose from the current study is the indication that upper elementary school children are being underserviced. The diminishing returns of the national curriculum are seen most poignantly in their weaker gains in PGR. While the national curriculum is less likely to immediately adjust materials and guidelines, prefectures, school districts or even single schools might make adjustments to address this issue. Given the fact that many teachers are not proficient in the foreign language and have little formal language teaching training, the simplest adjustment is to expose students to a wider variety of language content through multimedia. They might also expand students' independent studies, encouraging students to engage in individual or group projects which draw on books or the internet. Since 2020, all students have been issued personal tablets and internet access at schools (MEXT, 2021). Taking advantage of this new affordance will be critical to expanding the means by which students engage with foreign language education. In addition to offering personalisation and gamification of the language learning experience, tablets might well offer opportunities for independent learning at home as well during free periods during school.

Limitations and Future Discussion

This study used a simple (initial letter sound) PGR test to assess students' cross-sectional and longitudinal abilities. The results of this study indicate basic gains in students' sound-letter knowledge, though greater depth of testing is needed. Information about the developmental trajectories of PGR, however, would necessitate both a wider variety and a larger number of assessments. For this kind of assessment to be feasible, inside and outside the classroom, intermittent online assessment is necessary (Butler, 2022).

Conclusions

This study demonstrates the potential of a simple test of phoneme awareness in young learners in foreign language for Japan's elementary school education. The test can be easily and effectively conducted by a non-English teacher with minimal training. Findings from the longitudinal study presented here demonstrate the significance of the current early foreign language curricula in Japan. Current cross-sectional and longitudinal tests of students' PGR across four grades suggest potential developmental trends and point to minimum amounts of input required to construct PGR. To build on current foreign language PGR development in Japanese elementary schools in the near future, building PGR assessment into curricula is an important first step. A wider variety of language engagement opportunities, including self-directed learning and game-based interventions, might offer greater learning affordances. At a curricular level, integrating explicit instruction in curricula is likely needed to improve sound-letter recognition ability; this will require appropriate support systems for teachers, including comprehensive training for teachers and efficient material delivery systems. As noted, not all letters and sounds will require the same attention. By targeting the most necessary and difficult phoneme-grapheme representations, instruction will likely have maximum impact.
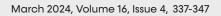
Acknowledgements

References

Allen-Tamai, M. (2019). Literacy teaching for elementary school students. Tokyo Shoseki.

Arfé, B., & Danzak, R. L. (2020). The influence of first language spelling and response inhibition skills on English-as-an-additional-language spelling. *Cognitive Development*, *56*, 100952. https://doi.org/10.1016/j.cogdev.2020.100952

Butler, Y. G. (2015). English language education among young learners in East Asia: A review of current research (2004-2014). *Language Teaching*, *48*(3), 303-342.

Butler, Y. G. (2022). Language education in the era of digital technology. *JALT Journal, 44*(1), 137-152. https://doi.org/10.37546/jaltjj44.1-7

Caravolas, M., Lervåg, A., Mikulajová, M., Defior, S., Seidlová-Málková, G., & Hulme, C. (2019). A cross-linguistic, longitudinal study of the foundations of decoding and reading comprehension ability. *Scientific Studies of Reading, 23*(5), 386-402. https://doi.org/10.1080/10888438.2019.1580284

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*(1), 5–51. https://doi.org/10.1177/1529100618772271

Coulson, D., Ariiso, M., Kojima, R., & Tanaka, M. (2013). Difficulties in reading English words: How do Japanese learners perform on a test of phonological deficit? *Vocabulary Learning & Instruction, 2*(1), 57-63. http://doi.org/10.7820/vli.v02.1.coulson.et.al

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research, 71*(3), 393–447. https://doi.org/10.3102/00346543071003393

George, D. & Mallery, P. (2010). *SPSS for Windows Step by Step (10th Ed.)*. Pearson.

Hulme, C., Snowling, M. J., West, G., Lervåg, A., & Melby-Lervåg, M. (2020). Children's language skills can be improved: Lessons from psychological science for educational policy. *Current Directions in Psychological Science, 29*(4), 372-377. http://doi.org/10.1111/1460-6984.12339

Ikeda, C. (2018). Phonological awareness of Japanese-L1 elementary school children: The influence of L1 on phoneme manipulation tasks. *JES Journal, 18,* 52–67.

Jeon, E. H., & Yamashita, J. (2022). L2 reading comprehension and its correlates. In E. H. Jeon & Y. In'nami (Eds.), *Understanding L2 Proficiency*, 29–86. https://doi.org/10.1075/bpa.13.03jeo

Kahn-Horwitz, J. (2020). 'I didn't even know one of the conventions before': Explicit EFL spelling instruction and individual differences. *Cognitive Development*, *55*, 100880. https://doi.org/10.1016/j.cogdev.2020.100880

Kahn-Horwitz, J., Sparks, R., & Goldstein, Z. (2012). English as a foreign language spelling development: A longitudinal study. *Applied Psycholinguistics, 33*(2), 343-363. doi:10.1017/S0142716411000397

Koda, K., & Yamashita, J. (2019). *Reading to learn in a foreign language: An integrated approach to foreign language instruction and assessment.* Routledge.

Kubozono, H. & Honma, T. (2002). *Onsetsu to mora.* Kenkyushya shuppan.

McArthur, G., Sheehan, Y., Badcock, N. A., Francis, D. A., Wang, H.-C., Kohnen, S., Banales, E., Anandakumar, T., Marinus, E., & Castles, A. (2018). Phonics training for English-speaking poor readers. *Cochrane Database of Systematic Reviews*, *2018*(11). https://doi.org/10.1002/14651858.cd009115.pub3

Machida, T. (2016). Japanese elementary school teachers and English language anxiety. *TESOL Journal, 7*(1), 40-66. https://doi.org/10.1002/tesj.189

McBride, C., Pan, D. J., & Mohseni, F. (2022). Reading and writing words: A cross-linguistic perspective, *Scientific Studies of Reading, (26)*2, 125–138, https://doi.org/10.1080/10888438.2021.1920595

Melby-Lervåg M., Lyster, S. A., & Hulme, C. (2012). Phonological skills and their role in learning to read: a meta-analytic review. *Psychological Bulletin. 138*(2), 322-52. https://doi.org/ 10.1037/a0026744.

Melby-Lervåg, M., & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first- and second-language learners. *Psychological Bulletin, 140*(2), 409–433. https://doi.org/10.1037/a0033890

MEXT (2017). *Shogakkou gakushuu shidou youryou kaisetu: Gaikokugokatsudou / gaikokugohen* [Elementary school curriculum guidelines: Foreign language activities / foreign language]. MEXT. https://www.mext.go.jp/content/20201029-mxt_kyoiku01-100002607_11.pdf

MEXT. (2021). *GIGA sukuru koso-no jitsugen-ni tsuite* [The realization of the GIGA School Concept]. https://www.mext.go.jp/a_menu/other/index_00001.htm

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., et al. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction, 29*(C), 65-77. http://doi.org/10.1016/j.learninstruc.2013.09.003

Nakao, K., Oga-Baldwin, W. L. Q., & Fryer, L. K. (2019). Expanding Japanese elementary school English education: Native and nonnative speaking team-teachers' perspectives on team-teaching quality. Waseda University Education Bulletin, 29, 17-32.;

Nakao, K., Oga-Baldwin, W. L. Q., & Fryer, L. K. (2022). Phonemic awareness as fundamental listening skill: A cross-sectional, cohort study of elementary foreign language learners. *Asia TEFL*, *19*(2), 609-618. http://dx.doi.org/10.18823/asiatefl.2022.19.2.13.609

Okada, T. (2005). A corpus-based study of spelling errors of Japanese EFL writers with reference to errors occurring in word-initial and word-final positions. In V. Cook & B. Bassetti (Eds.) *Second Language Writing Systems. Pp. 164*–https://doi.org/10.21832/9781853597954-008

Rueckl, J. G., Paz-Alonso, P. M., Molfese, P. J., Kuo, W.-J., Bick, A., Frost, S. J., Hancock, R., Wu, D. H., Mencl, W. E., Duñabeitia, J. A., Lee, J.-R., Oliver, M., Zevin, J. D., Hoeft, F., Carreiras, M., Tzeng, O. J., Pugh, K. R., & Frost, R. (2015). Universal brain signature of proficient reading: Evidence from four contrasting languages. *Proceedings of the National Academy of Sciences*, *112*(50), 15510–15515. https://doi.org/10.1073/pnas.1509321112

Russak, S., & Kahn-Horwitz, J. (2015). English as a foreign language spelling: comparisons between good and poor spellers. *Journal of Research in Reading, 38*(3), 307–330. https://doi.org/10.1111/JRIR.12009

SAS Inc. (2019). *JMP®, Version 14.1.* Cary, NC, 1989-2019.

Shankweiler, D., & Fowler, C. A. (2019). Relations between reading and speech manifest Universal Phonological Principle. *Annual Review of Linguistics*, *5*(1), 109–129. https://doi.org/10.1146/annurev-linguistics-011718-012419

Siegelman, N., Rueckl, J. G., Steacy, L. M., Frost, S. J., van den Bunt, M., Zevin, J. D., Seidenberg, M. S., Pugh, K. R., Compton, D. L., & Morris, R. D. (2020). Individual differences in learning the regularities between orthography, phonology and semantics predict early reading skills. *Journal of Memory and Language, 114*, 104145. https://doi.org/10.1016/j.jml.2020.104145

Silverman, R.D., Johnson, E.M., Keane, K., & Khanna, S. (2020). Beyond decoding: A Meta-Analysis of the effects of language comprehension interventions on K–5 Students' language and literacy outcomes. *Reading Research Quarterly*, *55*(1), 207– 233. https://doi.org/10.1002/rrq.346

Sodoro, J., Allinder, R. M., & Rankin-Erickson, J. L. (2002). Assessment of phonological awareness: Review of methods and tools. *Educational Psychology Review, 14*(3), 223-260. https://doi.org/10.1023/A:1016050412323

Temur, T., & Sezer, T. (2023). The agenda of the reading teacher journal on reading and reading skills: a corpus analysis in the last decade. *International Electronic Journal of Elementary Education*, *15*(4), 357–369. Retrieved from https://www.iejee.com/index.php/IEJEE/article/view/2047

Yamashita, J. (2022). L2 reading comprehension. In E. H. Jeon & Y. In'nami (Eds.), *Understanding L2 Proficiency*, 5–28. https://doi.org/10.1075/bpa.13.02yam

Yonezaki, M., Tara, S., & Tsukuda, Y. (2016). Primary school teachers' anxiety about teaching English as a compulsory subject and teaching english to middle-grade students: structuring and transition. *JES Journal, 16,* 132-146. https://doi.org/10.20597/jesjournal.16.01_132

Appendices

Name:

学年：

組：

出席番号：

先生が言う単語の最初の音をえらんで、その○をぬりつぶしてください。　例）APPLE → Ⓐ