Cultural Representation and Vocabulary Accessibility in U.S. Elementary ESL Textbooks: A Two-Dimensional Analysis of Grade 4 Newcomer Units

Xie He *

Graduate School of Education, University of Pennsylvania, Philadelphia, USA

* Corresponding Author Email: hexie846866234@gmail.com

Abstract. Newcomer English learners meet their new language largely through textbooks, yet cultural representation analysis and vocabulary load profiling have rarely been applied to the same materials. This study examines the first unit of the Grade 4 student books of the two English language development series most widely adopted in Texas, *On Our Way to English* (2014) and *Reach* (2011). Cultural references, characters and selections were coded with a three-layer scheme combining Yuen's, Ndura's and Ebe's categories, the last scored for a hypothetical nine-year-old newcomer from China; every student-visible word was profiled against Nation's British National Corpus/Corpus of Contemporary American English (BNC/COCA) word-family lists. Coding was replicated by an independent large language model of a different family, with reliability reported per variable. The two units occupy opposite corners of a visibility–accessibility space: *Reach* foregrounds named non-U.S. cultures through dense, culture-specific and often unglossed lexis and reaches 95% coverage only at the 4,000-family level, whereas *On Our Way to English* narrates the immigrant experience through largely generic references, keeps 96% coverage at 2,000 families, and relies heavily on proper nouns. East Asian characters are peripheral in one series and absent in the other. Implications for materials selection are discussed.

Keywords: ESL textbooks; newcomer English learners; cultural representation; vocabulary load; lexical coverage; content analysis.

1. Introduction

English learners (ELs) made up 10.6% of U.S. public school enrolment, about 5.3 million students, in fall 2021 [1]. A subset of these students are newcomers, defined in federal guidance as foreign-born students who arrived in the United States within the previous three years and are still acquiring English [2]. For a newcomer placed in an elementary classroom, the English language development (ELD) textbook is often the most stable source of printed English available: it is present every day and, unlike teacher talk, it can be re-read. Which textbook a district uses are not incidental. In the most recent statewide adoption of ELD materials in Texas (2011–12), two series accounted for roughly two thirds of purchases, *Reach* (44%) and *On Our Way to English* (25%), and schools that bought neither scored about 0.30 standard deviations lower on EL reading outcomes, with no reliable difference between the two series themselves [3].

Two research traditions speak to what such books offer a newcomer, but they have developed apart. Cultural representation analysis asks whose cultures, products, practices and perspectives a textbook makes visible and how deeply. It has been pursued mainly on international English language teaching (ELT) coursebooks, where inner-circle, target-culture content dominates and products outnumber perspectives [4, 5]; for U.S. K–12 ESL materials the only content analysis located by the literature searches for this study remains Ndura's [6] study of six textbooks. Vocabulary load profiling asks how much of a text's running words a reader with a given vocabulary size will know. It rests on adult second-language evidence that comprehension becomes adequate at about 95% coverage and comfortable at about 98% [7–9], the relation being approximately linear rather than a sharp threshold [10]; the coverage–comprehension relation has not been established for children reading in a second language [11]. The two traditions must nevertheless be read together: with minority-language children in the Netherlands the advantage of culturally familiar content was restricted to linguistically simple

texts [12]; a replication with adult learners found no effect of cultural familiarity when a large proportion of the vocabulary was unknown [13]; and in an Iranian sample familiarity with the content had a far larger effect on comprehension than linguistic simplification of the text [14]. Proper names and culture-specific words sit at the junction of the two traditions: they are the lexical carriers of cultural visibility and at the same time the items most likely to be unknown to a newcomer and to impair comprehension [15–17].

No study located in searches of ERIC, OpenAlex and general scholarly web search has analyzed U.S. elementary ESL materials on both dimensions at once; the nearest precedent examines inclusiveness in second-language reading materials outside the United States without cross-tabulating it with lexical accessibility [18]. The present study therefore asks, for the first unit of the Grade 4 student book of each of the two most widely adopted ELD series: (RQ1) whose cultures are represented, at what depth and with what forms of bias, with particular attention to East Asian and Chinese-heritage representation; (RQ2) how lexically accessible the unit is for a newcomer who knows only the most frequent one or two thousand word families, and how much of the load is carried by proper nouns and culture-specific words; and (RQ3) how representation and accessibility co-vary across the selections of each unit. The contribution is a two-dimensional framework operationalized on adopted materials, with a scripted vocabulary pipeline and a model-replicated coding procedure with per-variable reliability.

2. Background

2.1. Cultural representation in language textbooks

Content analyses of ELT coursebooks converge on a small set of categories. Yuen [4] classified references to foreign cultures into products, practices, perspectives and persons, and found products to be the most frequently depicted aspect and perspectives the least. Shin, Eslami and Chen [5] distinguished whose culture is presented (source, target or international) from what kind of knowledge is presented (surface facts or deeper values) and reported that inner-circle cultures prevailed and that cultural content stayed at the level of information. Images carry much of a textbook's cultural meaning, so that text-only coding underestimates representation [19]; in primary-school textbooks used in China, a recent multimodal analysis found culture presented as simplified, conflict-free and implicit facts, with homogenized persons and products [20].

Within the United States, Ndura's [6] analysis of ESL textbooks used from elementary through high school remains the reference point; it identified three forms of bias, stereotyping, invisibility and unreality, the first being the portrayal of one set of people exhibiting one set of values, behaviours and roles. Two further constructs bear on newcomers as readers. Bishop's [21] metaphor of mirrors, windows and sliding glass doors captures whether a text reflects a reader's own experience or opens onto someone else's. Ebe [22] operationalized cultural relevance for English learners as eight factors, the ethnicity of the characters, the setting, the year the story takes place, the age and gender of the characters, the language or dialect used, the genre and exposure to that type of text, and the reader's background experiences, and found that English learners read texts rated higher on these factors more successfully.

2.2. Vocabulary accessibility

Lexical coverage, the percentage of running words in a text that a reader knows, is the standard proxy for the vocabulary demand of a text. Hu and Nation [7] found that adequate unassisted comprehension of fiction required about 98% coverage; Laufer and Ravenhorst-Kalovski [9] proposed 95% as a minimal and 98% as an optimal threshold; and Schmitt, Jiang and Grabe [10] found a linear relation between coverage and comprehension without a clear inflection point. These reference points are adult and second-language; they are used here as reference bands, not as claims about what a nine-year-old newcomer needs.

Two features of the profiling apparatus matter for textbooks. Word-family counting credits a reader with all inflected and derived forms of a base word, which overstates what beginners know. Proper nouns (PN) are conventionally treated as known because they are transparent in context [8]; for a newcomer, however, unfamiliar names impair listening comprehension [16], and replacing culturally distant names of persons and places with familiar ones has produced sustained gains on cloze tests [15]. Coverage is therefore reported with proper nouns treated as known and as unknown.

2.3. A two-dimensional framework

The two dimensions define a space in which a selection can be located. Cultural visibility is indexed at the reference level by the density of specific, named cultural references and at the selection level by depth (surface products versus perspectives) and by the presence of the newcomer experience. Lexical accessibility is indexed by coverage at the 2,000-family level with proper nouns treated as unknown, the most conservative reading of what a newcomer brings. A selection in the high-visibility, high-accessibility quadrant offers cultural content a newcomer can actually read; the other quadrants describe trade-offs.

3. Materials and Methods

3.1. Corpus

The corpus consists of the first unit of the Grade 4 student book of each series (Table 1). Grade 4 was chosen because a newcomer at this level is beyond the picture-book stage and is expected to read connected informational and literary text; the first unit is what a newcomer meets first. Both books were obtained as complete library scans with an optical character recognition (OCR) layer and were verified against their ISBNs, page counts and tables of contents. Pages were segmented into selections (student-facing reading texts, including model texts and the writing-project text) and apparatus (unit openers, vocabulary, comprehension and grammar pages). Text embedded in photographs, maps and posters was excluded; all other student-visible text, including margin glosses, was retained. The OCR layer was corrected page by page against the page images, with the nine pages of highest OCR error rate transcribed in full.

Table 1. The corpus.

	On Our Way to English, Grade 4	*Reach*, Level E (Grade 4)
Publisher, year	Houghton Mifflin Harcourt / Rigby, 2014	National Geographic Learning / Hampton-Brown, 2011
ISBN	978-0-544-23533-5	978-0-7362-7429-6
Unit analysed	Unit 1, *American Journeys* (pp. 6–51)	Unit 1, *Living Traditions* (pp. 2–71)
Pages	46	70
Selections coded (Layer C)	10	5
Running words, all text	4,557	8,676
Running words, selections only	3,493	4,374
Word families (1k–25k), all text	754	1,014
Share of Texas 2011–12 ELD adoption [3]	25%	44%

3.2. Coding of cultural representation

A codebook with three layers was written before coding and revised once after a pilot on the first selection of each series. Layer A codes cultural references: each explicit textual reference to the product, practice, perspective or person of a specific culture or group, following Yuen [4], with the culture named (country or group code), its specificity (specific, naming a culture or item, versus generic) and whether the text glosses the item in situ (a margin gloss, an appositive or a parenthetical explanation). Conventions for non-English words, group labels and image-only elements are given in the codebook, which is supplied as supplementary material. Layer B codes persons: every named character or character with dialogue or action in the text and every identifiable foreground person in an image, with perceived ethnicity coded from page evidence only (nationality, language, name and

visual cues), origin status (newcomer or immigrant, U.S.-born, non-U.S. resident, unknown), role, gender, age group and any first language displayed on the page. Layer C codes each selection for genre, setting, focal culture, depth (surface, when only products such as food, festivals, names and dress are involved, versus deep, when a perspective-level reference to values, identity or migration experience is thematic), Ndura's three biases with quoted evidence, whether the newcomer experience is portrayed, and Ebe's eight cultural-relevance factors [22], each scored 1–4 for a hypothetical reader, a nine- to ten-year-old fourth grader from mainland China who arrived in the United States within the previous three years. The reader was fixed because cultural relevance is reader-relative and because Chinese is the home language of about 1.8% of ELs [1]; the choice does not affect Layers A and B.

3.3. Vocabulary profiling

All student-visible text was profiled against Nation's British National Corpus/Corpus of Contemporary American English (BNC/COCA) word-family lists, version 1.0.0, which comprise twenty-five 1,000-family levels plus supplementary lists of proper nouns, marginal words, transparent compounds and acronyms [23]. Contractions were expanded, hyphenated forms split and numerals excluded. Tokens were classified by a fixed priority: a manually curated tag table (proper nouns, non-English words, marginal words and OCR corrections), then the 25 frequency levels, then the supplementary lists; anything left was counted as off-list, and after correction the off-list share was below 0.1% in both books. For each series and layer the pipeline reports running words, families, frequency-band shares, proper-noun and non-English shares, cumulative coverage with proper nouns treated as known and as unknown, and the levels at which 95% and 98% coverage are reached; Flesch–Kincaid and Dale–Chall scores are given as conventional descriptors. Every number in the tables and figures is generated by the same scripts from the same intermediate files; none is typed by hand.

3.4. Coding procedure, replication and reliability

Both coders in this study were large language models, so the procedure is described in full, following the reporting guidance of O'Connor and Joffe [24]. Primary coding was carried out with Claude (Anthropic) working from the page images and the corrected text under the codebook, in three parallel passes over disjoint page ranges whose outputs were harmonized against the codebook before merging; every code carries a page number and a quoted or described piece of evidence, and the merge script rejects any value outside the codebook's vocabulary. The second coder was Gemini 3.7 Flash (Google), a model of a different family, invoked through a command-line wrapper that records the model actually routed and checks that the output file is complete; it received the codebook, the sampled units with page numbers and identifying information, the page images and the corrected text, but not the primary codes. The sample comprised all 15 selections and a stratified random 30% of references (58) and characters (48), 121 units in total, above the minimum of 50 recommended for reliability subsamples [25]. A first run in which the second coder had not opened any page image, as its logged tool calls showed, was discarded and repeated under a protocol that required each page image to be opened and logged before a unit was coded. Krippendorff's alpha was computed for every variable, nominal for categories and ordinal for the Ebe scales [26, 27], with simple percentage agreement alongside, because chance-corrected coefficients approach zero when nearly all units fall into one category, as with the three bias flags. Alpha for the Layer A variables ranged from .35 (glossing) to .91 (specificity); for Layer B from .60 (perceived ethnicity) to .83 (gender); at the selection level depth reached .71 and the newcomer-experience code .87, while the Ebe factors varied from -.34 (genre) to .76 (year), with percentage agreement between 47% and 73%; the full table is supplied as supplementary material. Every disagreement was adjudicated against the page evidence and the decision logged; the results below use the adjudicated primary codes, and variables whose alpha fell below .67 are reported descriptively and interpreted with corresponding caution.

3.5. Cross-tabulation

For RQ3 each selection is located in the framework of Section 2.3 by the number of specific cultural references per 1,000 running words and by 2,000-family coverage with proper nouns treated as unknown, with depth and the Ebe total shown as marker attributes. With fifteen selections the analysis is descriptive; no inferential statistics are used.

4. Results

4.1. Cultural representation (RQ1)

Table 2 summarises the three layers. The two units are of comparable reference density, 17.3 and 13.3 references per 1,000 running words, but differ in kind. In *On Our Way to English* (hereafter OOWTE) perspectives are the largest category, 37 references or 47%, because the unit, *American Journeys*, is organised around migration; products (16%) and practices (19%) are few. In *Reach* the unit, *Living Traditions*, is built from products and practices (38% each), with perspectives at 17%. The series also differ in specificity: 85% of *Reach*'s references name a culture or item, against 61% in OOWTE, where many references to migration are generic ("immigrants", "homeland", "a new country") and coded as international. Among specific references, in-situ glossing is somewhat more frequent in OOWTE (50%) than in *Reach* (42%).

Table 2. Cultural representation in the two units, by coding layer. Percentages are of the layer total for the series; region counts allow multiple regions per reference.

Measure	OOWTE G4 Unit 1	Reach E Unit 1
Cultural references (Layer A), n	79	115
per 1,000 running words (all text)	17.3	13.3
Products	13 (16%)	44 (38%)
Practices	15 (19%)	44 (38%)
Perspectives	37 (47%)	19 (17%)
Persons	14 (18%)	8 (7%)
specific (named culture/item)	48 (61%)	98 (85%)
glossed in situ, among specific	24 (50%)	41 (42%)
region: EastAsia	7 (9%)	2 (2%)
region: Europe	12 (15%)	1 (1%)
region: INTL	16 (20%)	18 (16%)
region: LatAm	11 (14%)	35 (30%)
region: SouthAsia	2 (3%)	8 (7%)
region: SubSaharanAfrica	10 (13%)	34 (30%)
region: US	37 (47%)	11 (10%)
region: US-AfAm	0 (0%)	4 (3%)
region: US-NativeAm	0 (0%)	7 (6%)
region: other (Asia-generic, MENA)	3 (4%)	0 (0%)
lexical item beyond 2,000 or PN/foreign	52 (66%)	78 (68%)
Characters/persons (Layer B, selections), n	50	44
White	15 (30%)	1 (2%)
Black-AfAm	2 (4%)	1 (2%)
Latino	4 (8%)	1 (2%)
EastAsian	6 (12%)	0 (0%)
SouthAsian	1 (2%)	4 (9%)
MENA	1 (2%)	0 (0%)
NativeAm	0 (0%)	6 (14%)
SubSaharanAfrican	6 (12%)	18 (41%)
Multi-Ambiguous	9 (18%)	1 (2%)
NotShown	6 (12%)	12 (27%)
newcomer/immigrant characters	23 (46%)	1 (2%)
characters with L1 displayed	6 (12%)	11 (25%)
persons on apparatus pages, n (% ambiguous)	16 (69%)	40 (15%)
Selections coded (Layer C), n	10	5
deep (perspective-level)	6 (60%)	3 (60%)
newcomer experience portrayed	6 (60%)	0 (0%)
Ndura bias flagged: stereotype / invisibility / unreality	0 / 1 / 2	1 / 1 / 1
Ebe relevance total (8–32), mean (range)	22.8 (17–28)	18.6 (16–21)

Whose cultures are represented differs sharply. Nearly half of OOWTE's references are to mainstream U.S. culture (47%), with Europe (15%), Latin America (14%), sub-Saharan Africa (13%) and East Asia (9%) forming a second tier. *Reach* is centred elsewhere: Latin America (30%) and sub-Saharan Africa (30%) account for three fifths of its references, through a Cuban folk tale and an interview about Gabonese music, followed by South Asia (7%), a coming-of-age article set in India, and two U.S. minority cultures, Tewa Pueblo (6%) and African American (3%). East Asia is nearly absent from *Reach*'s text (2 references, both incidental mentions of the Vietnam War), although a Japanese tea scene appears as an apparatus illustration.

The person layer sharpens the picture. Of the 50 persons in OOWTE's selections, 23 (46%) are newcomers or immigrants, and six of the ten selections portray the newcomer experience: a Bulgarian boy's first weeks in an American school ("Voices"), a Sudanese refugee boy's arrival in winter ("Home of the Brave"), a new girl from Trinidad in a school play, and two poems in a newcomer's voice. The protagonists are mostly White or ambiguous, and the 6 East or Southeast Asian persons (12%) are all secondary or background figures: a classmate with a Korean name, a Filipino-heritage teacher, Vietnamese refugees in two photographs and the architect I. M. Pei in a chart. No selection has an East Asian protagonist, and no Chinese word or phrase appears anywhere in the unit. *Reach* presents 44 persons, of whom 41% are sub-Saharan African, 14% Native American and 9% South Asian; only 1 is an immigrant, and none is East Asian. First languages are displayed for 25% of *Reach*'s persons, almost all Spanish in the Cuban folk tale, and for 12% in OOWTE, mainly Bulgarian in "Voices".

At the selection level, 60% of OOWTE's and 60% of *Reach*'s selections reach perspective-level depth, so neither unit is confined to food and festivals. The biases differ in form. In OOWTE the coded instances are of unreality and invisibility: the informational selections state that Vietnamese immigrants "quickly settled into life in the United States" and that immigrants "have always worked hard and made a big contribution", success narratives without hardship or discrimination, and the historical overview stops in 1915. In *Reach* all three biases are coded in a single selection, the interview about Gabonese music, in which villagers are generalised as leading "very simple" lives, no Gabonese person speaks in his or her own voice, and life without electricity is idealised. On Ebe's factors scored for a Chinese newcomer reader, OOWTE's selections average 22.8 (17–28) of a possible 32, driven by shared age, school settings and the migration experience rather than by ethnicity or language, and *Reach*'s average 18.6 (16–21); the Cuban folk tale scores lowest (16) and the story of the Bulgarian newcomer highest (28). Given the modest reliability of several Ebe factors, these totals are treated as an ordering rather than as measurement.

4.2. Vocabulary accessibility (RQ2)

Table 3 gives the vocabulary profile and Fig. 1 the cumulative coverage curves. With proper nouns treated as known, OOWTE's selections reach 95% coverage at the 2,000-family level and 98% at 3,000, whereas *Reach*'s selections require 4,000 families for 95% and 11,000 for 98%. The difference arises above the second thousand: 1.9% of OOWTE's selection tokens lie beyond the 3,000 level against 4.9% in *Reach*, and non-English words, negligible in OOWTE (0.1%), reach 1.1% in *Reach*. The conventional readability descriptors point the same way (Flesch–Kincaid grade 4.2 versus 6.1; Dale–Chall 7.8 versus 9.0). Lexical variety is high in both: 50.1% of the word families in OOWTE's selections and 43.4% in *Reach*'s occur only once, so most new words are not met again within the unit.

Table 3. Vocabulary profile of the two units against the BNC/COCA word-family lists.

Measure	OOWTE G4 — selections	OOWTE G4 — all text	Reach E — selections	Reach E — all text
Running words (tokens)	3,493	4,557	4,374	8,676
Word families (1k–25k lists)	667	754	761	1,014
Families occurring once (%)	50.1	45.6	43.4	38.0
1st 1,000 (% tokens)	80.4	81.3	77.0	78.5
2nd 1,000 (% tokens)	5.9	6.6	6.4	7.7
3rd 1,000 (% tokens)	2.1	2.2	2.7	3.3
4,000+ (% tokens)	1.9	1.7	4.9	4.0
Proper nouns (% tokens)	8.0	6.2	6.3	4.2
Non-English words (% tokens)	0.1	0.1	1.1	0.7
Off-list (% tokens)	0.0	0.0	0.1	0.0
Coverage at 2k, PN known (%)	96.0	96.1	91.3	92.0
Coverage at 2k, PN unknown (%)	87.9	89.8	85.0	87.7
95% level (PN known / unknown)	2 / >25	2 / >25	4 / >25	3 / 22
98% level (PN known / unknown)	3 / >25	3 / >25	11 / >25	8 / >25

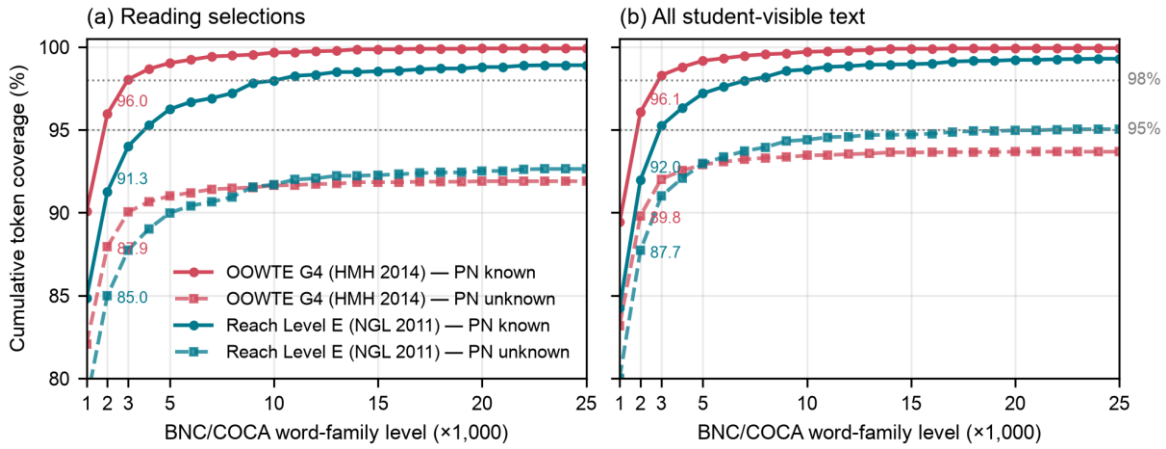


Fig. 1 Cumulative lexical coverage by 1, 000-family level for the selections (left) and all student-visible text (right) of each unit, with proper nouns treated as known (solid) and as unknown (dashed). Horizontal lines mark the 95% and 98% adult reference bands.

Proper nouns change the picture more than any frequency band. They make up 8.0% of OOWTE's selection tokens, a consequence of its historical and biographical selections, and 6.3% of *Reach*'s. Treating them as unknown lowers 2,000-family coverage from 96.0% to 87.9% in OOWTE and from 91.3% to 85.0% in *Reach*, and under this reading the selections of neither unit reach 95% coverage anywhere within the 25,000-family lists (Fig. 1a, dashed curves). For a newcomer who does not yet recognise English renderings of place and personal names, the two units are closer in demand than the conventional profile suggests, and both are demanding.

4.3. Representation and accessibility together (RQ3)

Fig. 2 locates each selection by the density of specific cultural references and by 2,000-family coverage with proper nouns treated as unknown. Two thirds of the culture-specific lexical items in both units lie beyond the 2,000-family level or are proper nouns or non-English words (66% in OOWTE, 68% in *Reach*), which is why the two coordinates are related: the selections richest in named cultural content are, with one exception, the least accessible.

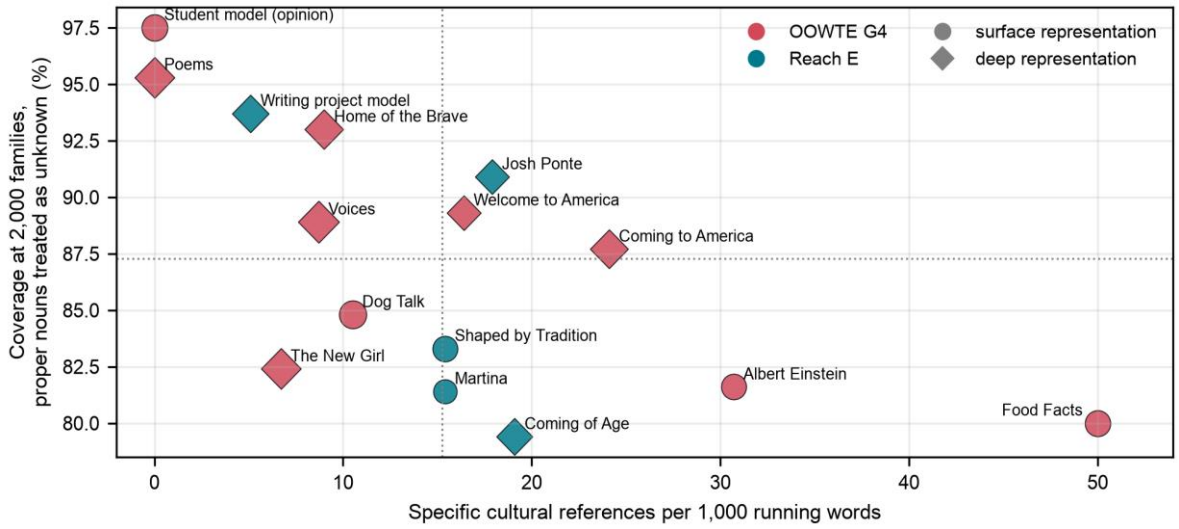


Fig. 2 Selections located by specific cultural references per 1,000 running words and by 2,000-family coverage with proper nouns treated as unknown. Colour marks the series, marker shape the depth of representation and marker size the Ebe cultural-relevance total; dotted lines are the means of the fifteen selections.

The lower-right quadrant, high visibility and low accessibility, holds OOWTE's "Food Facts" (50.0 specific references per 1,000 words; 80.0% coverage), its "Albert Einstein" biography (30.7; 81.6%), and *Reach*'s "Coming of Age" (19.1; 79.4%), "Shaped by Tradition" (83.3%) and "Martina" (81.4%), whose Sanskrit, Telugu and Spanish terms (*samskara*, *voni*, *abuela*, *cucaracha*) are only partly glossed. The upper-left quadrant, accessible but culturally sparse, holds the student model text and the poems, which carry OOWTE's highest coverage (97.5% and 95.3%) and almost no specific references. The exception is *Reach*'s interview "Josh Ponte" (17.9 per 1,000; 90.9%), which combines named Gabonese content with a conversational register and is also the selection in which all three of Ndura's biases were coded. The selections that portray the newcomer experience cluster on the accessible side with low to moderate reference density: OOWTE conveys migration through generic vocabulary rather than named cultural items. Deep representation occurs in both quadrants, so depth and lexical demand are not the same thing; what varies with demand is the number of named, culture-specific items a newcomer must decode.

5. Discussion

Three findings follow from reading the two dimensions together. First, the two most widely adopted ELD series make opposite trade-offs in their opening units. *Reach* offers windows in Bishop's sense [21], deep and specific representation of Gabonese, Cuban, Tewa and Indian cultures, but pays for it in lexis: a newcomer with 2,000 families would know roughly 85% of the running words of its selections once proper nouns are counted as unknown and would meet an unclosed culture-specific item about once every 107 running words. OOWTE offers the newcomer experience itself, the nearest thing to a mirror the two units contain and keeps its narrative selections within the 2,000-family band but concentrates cultural specificity in informational pages built from proper nouns, where coverage falls to 80.0%. By the evidence reviewed in Section 1, the first design forfeits much of the benefit of culturally rich content for a newcomer who cannot yet read it; the second confines cultural specificity to pages where its main effect is lexical load.

Second, for a Chinese-speaking newcomer neither unit offers a mirror. The migration stories in OOWTE are told through Bulgarian, Sudanese and Trinidadian protagonists, and East Asian persons appear only at the margins; no Chinese word appears on any page. *Reach* contains no East Asian person in its selections at all. This pattern is the elementary-level counterpart of the invisibility Ndura [6] documented, and it coexists with representational depth: a unit can be deep about several cultures and still leave a large group of its readers unreflected. Ebe's [22] factors make the point concrete:

OOWTE's selections score higher for a Chinese newcomer not because of ethnicity or language but because of shared age, school settings and the experience of arrival.

Third, proper nouns and culture-specific words are where the two dimensions meet. Counting proper nouns as known, the usual convention, makes OOWTE's selections look comfortably readable at 2,000 families; counting them as unknown removes 8.0 percentage points of coverage from OOWTE's selections and places both units below the 95% band throughout the frequency range. Since names of persons and places are precisely the items that culturally distant readers fail to process [15, 16], and since these units use them to carry much of their cultural content, the conventional profile understates the demand of the culturally richest pages. Two implications follow. Teachers using either series with newcomers should treat the proper nouns of informational and biographical selections as vocabulary to be pre-taught and should expect the culture-specific nouns of *Reach* to need glossing beyond what the margins supply. Publishers and adoption committees should ask of any unit not only whose culture it shows and how deeply, but at what lexical price, and should not assume that a unit about immigration is thereby relevant to every immigrant child.

Limitations. The study is descriptive and small: one unit of one edition of each series, fifteen selections, and no inferential claims. Perceived ethnicity was coded from page evidence and is a judgement about representation, not about the identity of any real person; where cues were absent it was coded as ambiguous. Both coders were large language models; although the second coder was of a different family, saw no primary codes and demonstrably opened the page images, model coders may share biases that a human pair would not, and reliability was modest for perceived ethnicity, glossing and several Ebe factors, whose results are reported as orderings. Word-family counting credits a beginner with derived forms he or she may not know, so coverage is overstated even in the conservative reading, and the 95% and 98% bands are adult reference points, not thresholds validated for children reading in a second language. Oral input, images, teacher mediation and the teacher's edition were not analyzed. The claim that no prior study combines the two dimensions on U.S. elementary ESL materials is bounded by the databases searched and the languages read.

6. Conclusion

Read on two dimensions at once, the opening Grade 4 units of the two most widely adopted U.S. ELD series make opposite bargains: one purchases deep, named representation of non-U.S. cultures with lexis that a newcomer cannot yet read, the other narrates the newcomer experience in accessible English while leaving its culturally specific content to pages of proper nouns. Neither unit gives a Chinese-speaking newcomer a character to recognize. The framework, the coding scheme and the scripted profiling pipeline used here can be applied to further units, series and grade levels; the immediate recommendation to teachers and adoption committees is to weigh cultural visibility and lexical accessibility together, because in the materials examined here the one is largely bought with the other.

References

[1] National Center for Education Statistics. English learners in public schools (Condition of Education indicator CGF). NCES, U.S. Department of Education, 2024.

[2] U.S. Department of Education (OELA/NCELA). Newcomer Toolkit (2023 edition). U.S. Department of Education, 2023.

[3] Lê QTN, Polikoff MS. Do English Language Development Curriculum Materials Matter for Students' English Proficiency? SAGE Open, 2021, 11(3).

[4] Yuen K-M. The representation of foreign cultures in English textbooks. ELT Journal, 2011, 65(4): 458–466.

[5] Shin J, Eslami ZR, Chen W-C. Presentation of local and international culture in current international English-language teaching textbooks. Language, Culture and Curriculum, 2011, 24(3): 253–268.

[6] Ndura E. ESL and cultural bias: An analysis of elementary through high school textbooks in the Western United States of America. Language, Culture and Curriculum, 2004, 17(2): 143–153.

[7] Hu H-CM, Nation P. Unknown vocabulary density and reading comprehension. Reading in a Foreign Language, 2000, 13(1): 403–430.

[8] Nation ISP. How large is vocabulary needed for reading and listening? The Canadian Modern Language Review, 2006, 63(1): 59–82.

[9] Laufer B, Ravenhorst-Kalovski GC. Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. Reading in a Foreign Language, 2010, 22(1): 15–30.

[10] Schmitt N, Jiang X, Grabe W. The percentage of words known in text and reading comprehension. The Modern Language Journal, 2011, 95(1): 26–43.

[11] Ludewig U, Hübner N, Schroeder S. Vocabulary, text coverage, word frequency and the lexical threshold in elementary school reading comprehension. Reading and Writing, 2023, 36(9): 2409–2431.

[12] Droop M, Verhoeven L. Background Knowledge, Linguistic Complexity, and Second-Language Reading Comprehension. Journal of Literacy Research, 1998, 30(2): 253–271.

[13] Klassen K. The Effect of Cultural Familiarity on Reading Comprehension: An Approximate Replication of Erten and Razi (2009). Reading Psychology, 2020, 41(3): 205–228.

[14] Keshavarz MH, Atai MR, Ahmadi H. Content schemata, linguistic simplification, and EFL readers' comprehension and recall. Reading in a Foreign Language, 2007, 19(1): 19–33.

[15] Chihara T, Sakurai T, Oller JW. Background and culture as factors in EFL reading comprehension. Language Testing, 1989, 6(2): 143–149.

[16] Kobeleva PP. Second language listening and unfamiliar proper names: Comprehension barrier? RELC Journal, 2012, 43(1): 83–98.

[17] Erten IH, Razı S. The effects of cultural familiarity on reading comprehension. Reading in a Foreign Language, 2009, 21(1): 60–77.

[18] Simsek MR. A Corpus-Based Comparison of Inclusiveness in L2 Reading Materials for Refugee Children. IAFOR Journal of Education, 2022, 10(1): 53–71.

[19] Weninger C, Kiss T. Culture in English as a Foreign Language (EFL) textbook: A semiotic approach. TESOL Quarterly, 2013, 47(4): 694–716.

[20] Tang Y, Liu L. Cultural representation in primary ELT textbooks: an approach of multimodal discourse analysis. Discourse: Studies in the Cultural Politics of Education, 2025, 47(1): 54–67.

[21] Bishop RS. Mirrors, windows, and sliding glass doors. Perspectives: Choosing and Using Books for the Classroom, 1990, 6(3).

[22] Ebe AE. Culturally relevant texts and reading assessment for English language learners. Reading Horizons: A Journal of Literacy and Language Arts, 2010, 50(3): 193–210.

[23] Nation ISP. The BNC/COCA Level 6-word family lists (Version 1.0.0) [Data set]. Victoria University of Wellington, 2017.

[24] O'Connor C, Joffe H. Intercoder Reliability in Qualitative Research: Debates and Practical Guidelines. International Journal of Qualitative Methods, 2020, 19: 1–13.

[25] Lombard M, Snyder-Duch J, Bracken CC. Content Analysis in Mass Communication: Assessment and Reporting of Intercoder Reliability. Human Communication Research, 2002, 28(4): 587–604.

[26] Krippendorff K. Reliability in Content Analysis: Some Common Misconceptions and Recommendations. Human Communication Research, 2004, 30(3): 411–433.

[27] Hayes AF, Krippendorff K. Answering the Call for a Standard Reliability Measure for Coding Data. Communication Methods and Measures, 2007, 1(1): 77–89.