Intensive or extensive reading: what the research says

Pindu has been evolving from an intensive-first approach to an extensive-first approach, but is that the right direction? 2026-11-02

Review written for Anki Add-On v1.16.1 (May 2026).

In short

- We reviewed second language acquisition research to inform whether Pindu should focus on extensive or intensive reading.
- An extensive approach is supported for developing fluency.
- Comprehension rises with vocabulary coverage, with no clean threshold. Pindu's 95% target remains a product hypothesis that needs its own validation.
- Learning new words benefits from a more intensive approach, supporting Pindu's scaffolded reader.
- Reading pays off with regular reading over time, so sustaining motivation through good design matters.

"Reading helps" is one of the least controversial ideas in language learning, but it hides harder questions: what kind of reading helps, and under what conditions? Should intermediate learners read a lot of easy material, or struggle through harder material with support? Those are at the center of an unresolved debate in second language acquisition research that the field has been arguing about since at least the 1980s. The answers matter to me as both a language learner and the builder of Pindu, which at its core facilitates reading in a foreign language. First envisioned using just my intuitions, the software's approach is now based also on the research literature. The core design question I set out to answer – should Pindu focus on extensive reading or on intensive reading – ended up being, perhaps, the wrong question to ask.

Pindu's evolution towards extensive reading

Pindu was initially conceived not as an extensive reading (ER) engine but as a particularly artificial form of intensive reading (IR). Instead of plodding through hundreds of individual flashcards with Anki, users could arrange them into a dense passage and read them in context. I thought this would be faster, more motivating, and more useful for developing reading skill than isolated card review. That intuition was sufficient to push me, as a learner who had spent 45 minutes per day in Anki, to build the prototype.

The results were… tremendously hard to read. Constructing coherent passages was technically difficult, and parsing them as a learner was so slow and halting that the experience could not plausibly develop reading skill or sustain interest. As I experimented, Pindu gradually moved away from dense IR and toward workflows closer to ER: longer passages composed mostly of familiar words, with a smaller number of target words embedded in easier surrounding context. The new premise was simple: a more readable, more natural passage would better support fluency, motivation, and repeated exposure. The contrast is easy to see in a toy example like Figure 1.

Difficult to read	Easy to read
The perspicacious yet obstreperous pedagogue □ s grandiloquent dissertation on the ubiquitous manifestations of quotidian ennui proved inordinately recalcitrant to even the most sagacious and indefatigable scholars. His loquacious and vituperative diatribe, though ostensibly magnanimous in intent, was perceived as merely perfunctory obsequiousness by the truculent assemblage of fastidious cognoscenti.	The smart but stubborn teacher wrote a long, fancy essay about how people get bored in everyday life, but even the most sagacious students couldn □ t understand it. His long, angry speech was supposed to be helpful, but the grumpy group of picky experts thought he was just showing off and didn □ t really mean it.

Figure 1: The left passage, an example of Pindu's old IR approach, is composed almost entirely of uncommon Latinate words. The right, much more readable, models Pindu's current ER approach. Both passages contain the target word "sagacious".

But this evolution was driven more by usability than by any awareness of second language acquisition (SLA) research. Pindu's text engine still supports dense IR modes as well as the ER modes emphasized on the website, so the design question remained unresolved: should Pindu primarily support extensive reading, intensive reading, or some mixture of the two? This article looks at the literature around extensive reading, its challenges, and the role more intensive methods may play in developing fluency. In the end, the review supported Pindu's current direction, though not by settling the question of "extensive vs intensive" as cleanly as I expected.

A tour of the literature on extensive reading for second language acquisition

We start with the position that extensive reading works. A good starting point in the literature to examine that claim is Stephen Krashen's comprehensible input (CI) hypothesis from the 1980s. Krashen argued that language acquisition happens primarily when learners read or listen to language (or "receive input") just beyond their current level (making it "comprehensible"), allowing them to naturally learn the unfamiliar grammar and words. He believed that CI was not just a helpful technique but the central mechanism for fluency measured over any number of dimensions: vocabulary, automaticity, comprehension, etc. (see Figure 2 for his schematic model). Reading lots of level-tuned material, as in ER, is a good method for CI (as would be conversation with a tutor who adjusts their speech to the student's level). Richard Day and Julian Bamford took this line of thinking and put it into a practical context with their 1998 book Extensive Reading in the Second Language Classroom. They gave a 10-point characterization of ER programs that has become the field's working definition. From Day's 2015 retrospective, the most important of these points are: reading a lot, reading what you want, reading varied material, and reading comprehensibly. Krashen might say: design a program or tool around these principles and you'll be set. Not quite.

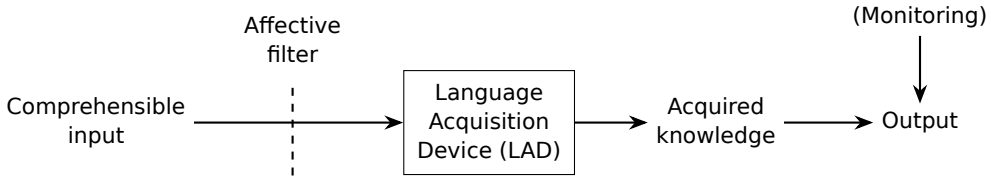


Figure 2: Krashen's Input Hypothesis Model of second-language learning and production, adapted from Krashen (1982). Comprehensible input must pass through the affective filter (low if attitudes and motivation are good) to reach the Language Acquisition Device, which produces acquired knowledge.

The ER vision for SLA runs into a practical problem almost immediately: access to material. A text cannot function as CI if it isn't comprehensible, and learners definitionally do not know many words.

They are left with unengaging children's books and a limited selection of graded readers. They become trapped in what Christine Nuttall called the "vicious cycle of the weak reader" (Figure 3): having few reading options leads to low motivation and little reading, which slows or halts progress, which maintains the same few options, etc. This chicken-and-egg problem has been part of practical teaching discourse from the start, and it is one reason for a continued focus on intensive grammar and vocabulary instruction in lower-level instruction instead of reading. If the purpose of the second language classroom is, as Krashen wrote in his 1982 book, to bootstrap students' abilities to the point that they can get CI from native texts and conversations, then an ER approach might not be practically feasible. This is not a fatal criticism, though: probably in popular languages like English there are enough graded readers; and as technology improves, access to individually-graded material is becoming increasingly available. (Pindu, for example, is enabled by massive online text corpora and generative AI.)

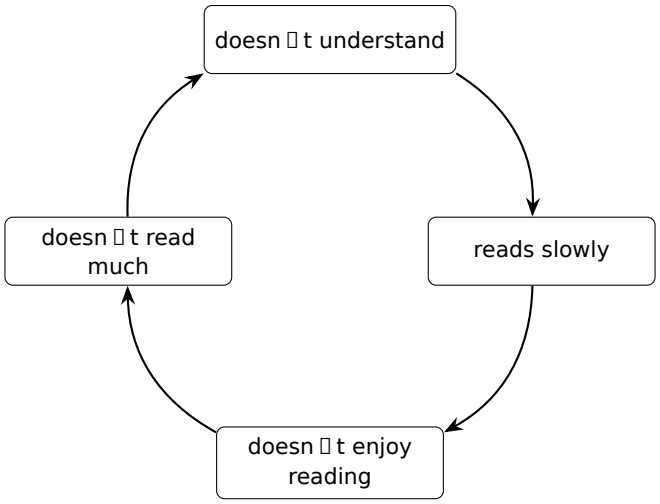
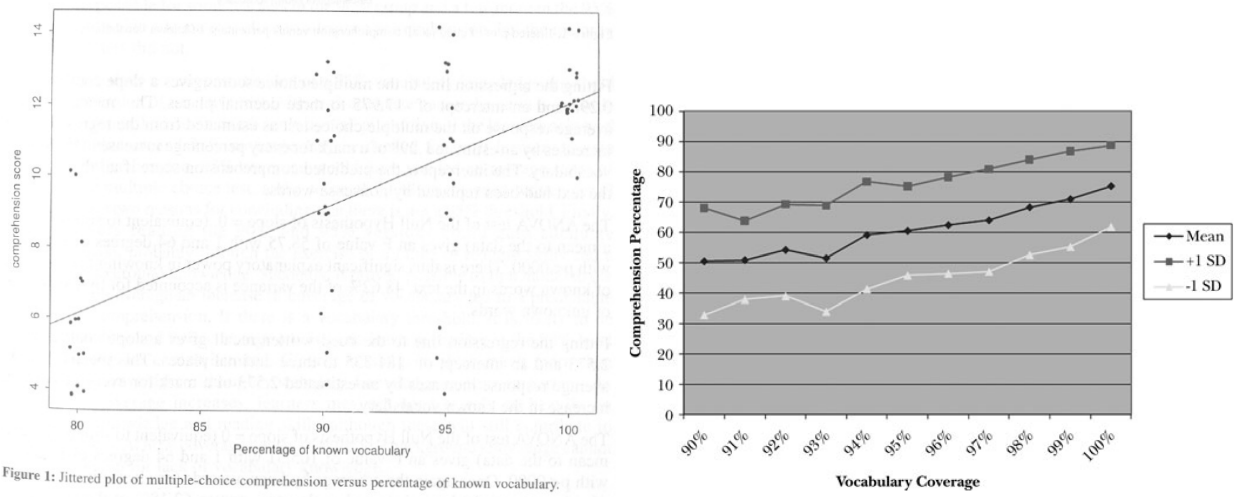


Figure 3: Nuttall's (1996) vicious cycle of the weak reader. Naïve reading of native material is not the solution to efficient second language acquisition, and neither is reading unengaging content below the reader's level.

Even assuming an adequate library of texts spanning all fluency levels, there is still another practical problem in choosing exactly what to read. Krashen wants the learner to read at the edge of their ability, but where is the edge? This is a two-part question: first, what level of comprehension is optimal for learning; second, how can the comprehensibility of a text be assessed? The second part has been better studied. In their widely-cited study from 2000, Marcella Hu & Paul Nation investigated the role of vocabulary coverage, i.e., the percentage of familiar words in a text, in determining comprehensibility. (To motivate the selection of vocabulary as the relevant text feature, they cited prior research that established vocabulary as the primary determinant of comprehension, followed by subject matter knowledge and syntactical familiarity.) Their results showed the expected increase of comprehension with increasing vocabulary coverage in the test texts, but not any convincing threshold or non-linear feature. The strongest conclusion from this study is that for near 100% levels of comprehension, near 100% levels of vocabulary coverage are needed. Even 1 in 20 words missing from the learner's vocabulary cause noticeable defects in understanding. At low coverage, reading becomes qualitatively different; it becomes, in the memorable words of Leonard Newmark, "cryptoanalytic decoding".

Despite what I have said against a threshold finding, a few of Hu & Nation's sentences like "98% coverage may be needed for most learners to gain adequate comprehension" have led some in the literature to use this study exactly as evidence for a threshold. More recent studies, like a 2011 study by Schmitt, Jiang & Grabe, confirm the cautious stance on a threshold but confirm the importance of vocabulary for comprehension. Figure 4 shows the shape of the evidence: comprehension generally rises with coverage, but not at a single clean break point.



(a) Hu & Nation (2000) **(b)** Schmitt, Jiang & Grabe (2011)

Figure 4: Key result charts from two studies of vocabulary coverage and reading comprehension. Both show a strong relationship between vocabulary coverage and comprehension but no clear threshold.

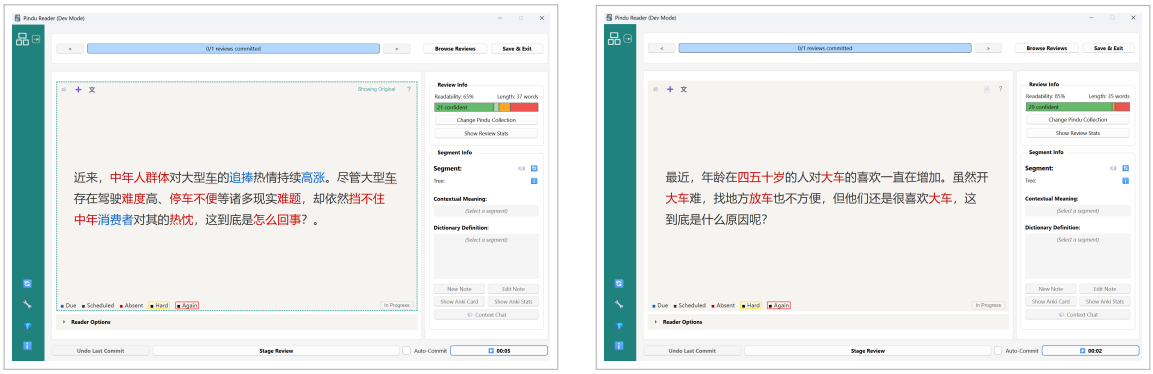
Those are the practical problems with building an ER program. But there are more fundamental concerns. Even as you read extensively, you will encounter unfamiliar words and grammar. That is rather the point: you meet new words so that you can acquire them. But is this actually an efficient way to acquire them? Krashen's model says yes, but other researchers have other ideas for how best to build. Among those, Batia Laufer is one of the most prominent. The through line of her work in the 1990s and 2000s is a model called "task-induced involvement load" that she introduced in a landmark paper with Jan Hulstijn in 2001. This basically means that learning is most effective when the learner has high "involvement" with unfamiliar words during their learning "task" or activity with the target language. Involvement includes a variety of cognitive and motivational factors, but essentially means doing "word-focused activities": looking up unfamiliar words, studying them, and reviewing them. Laufer comes through as a skeptic of the strong form of Krashen's Input Hypothesis and a foil to the ER dogmatists who say that "reading is all you need". The evidence she marshals shows instead that vocabulary acquisition, which is all-important for the ER programs that Day & Bamford outline, needs more than passive reading to work well.

To summarize the discussion so far: extensive reading is a specific kind of activity for second language acquisition that relies entirely on large amounts of vocabulary-calibrated reading material. While ER is well grounded in Krashen's Input Hypothesis as a type of Comprehensible Input, there is active debate in the researcher and practitioner community about whether unaided reading alone is optimal or even effective for vocabulary acquisition. If there is such a thing as a middle ground or consensus position in the literature and in practice, it seems to be the sort of integrated (or less politely, compromise) position proposed by Paul Nation in his 2001 book. There he puts forward a "four-strand" model for effective SLA that includes: (1) meaning-focused input like ER; (2) meaning-focused output; (3) language-focused learning like Laufer's word-focused activities; (4) fluency development. In other words: do a little bit of everything. Patricia Carrell and Joan Carson described an advanced reading program like this in their 1997 article that called for a combination of ER and IR to effectively teach their university students.

What do empirical studies say about extensive reading? They support it, but not in a way that would suggest an exclusive approach. Two large meta-analyses – Nakanishi (2015), covering 34 studies, and Jeon & Day (2016), covering 49 – find consistent positive effects of ER programs on various fluency dimensions compared to programs without ER. The effects are significant and hold across a wide variety of program designs, but the designs generally include ER only as one component in a more diversified pedagogy.

What this means for Pindu

The high-level takeaway from the above research seems to be that “extensive reading probably works, but maybe is not better than other methods, at least in isolation.” Unfortunately that does little to resolve the design tension that I have with Pindu. Pindu has two core pieces of technology: a text engine that can make IR passages, ER passages, and anything in between; and a reader interface that can support either style of reading (Figure 5). But what sort of reading should the user do, and how should Pindu be positioned in marketing? Although the research seems not to answer those questions, it’s a mistake to think it doesn’t say anything useful.



(a) Intensive mode **(b)** Extensive mode

Figure 5: Pindu can present similar reading content in either a dense intensive mode or a more readable extensive mode.

Laufer’s argument in favor of “word-focused work” is quite well supported, and it’s something that resonates with Pindu’s audience (flashcard-enthusiast Anki users). Pindu’s reader interface provides comprehensive scaffolding and glossing – translations, text-to-speech, example sentences, definitions, interactive chats, etc. – that lets users read smoothly and interact with words or sentences at will to improve comprehension. In this way, it operationalizes Laufer’s critique with Krashen’s Input Hypothesis in the same tool. That Pindu’s text engine is built with spaced-repetition at its core, and so resurfaces targeted vocabulary often, is a further feature that provides usability at this interface. My conclusion: these features are core to the product and should not be sidelined.

The empirical work from Hu & Nation and Schmitt, Jiang & Grabe, even if it does not mark a clear threshold of vocabulary coverage for extensive reading, still shows clearly that any claim to reasonable comprehension needs to be extremely close to 100%. Since this number is set by the user in Pindu, having clear guidance on the categories considered extensive (98%+, say) matters. However, an interesting wrinkle appears here: the cited research for vocabulary coverage generally treats word familiarity as binary, whereas Pindu estimates recallability probabilistically through spaced repetition. That means Pindu’s coverage setting is not identical to the coverage measures used in those studies. A 95% target in Pindu should therefore be understood not as a proven equivalent to Hu & Nation’s 98% figure, but as a product hypothesis: probabilistic familiarity may allow a lower nominal target to produce a similarly extensive reading experience. This is a place where the literature gives Pindu a useful design rationale, but where product-specific validation would still be needed.

Looking through the lens of Nation’s four-strand model, we see how both Pindu’s IR and ER modes fit in. Reading is always going to be situated in his first strand, “meaning-focused input”, but as the user modulates a passage’s vocabulary coverage, their session will blur into others as well. At and above high levels sufficient for extensive reading (above 95%, say), the session is squarely in the first strand. If the user turns down coverage, the session becomes more intensive, and the work blurs into Nation’s third strand, “language-focused learning”: it engages more intentionally with new words and structures. If the user turns up coverage to 100%, the work blurs into the fourth strand, “fluency

development". (Pindu makes no effort to occupy the second strand, "meaning-focused output".) Takeaway: both the IR and the ER modes are worth keeping and highlighting. Close reading is a useful tool, though laborious parsing is not necessarily efficient as a default mode. The 100% comprehensibility text-passage – something diametrically opposed to my original intent when building Pindu – should be considered.

And there's a real, actionable shortcoming of Pindu that the research reveals: the programmatic aspect. The benefits of ER do not accrue immediately but over time; the above meta-reviews emphasized that all of the tested programs integrated reading over the course of many weeks. To avoid Nuttall's vicious cycle, repeated practice is key. To the extent that Pindu encourages regular reading, it would work better. Yet while Pindu makes it easy to sit down for a reading session, it currently lacks an engagement flywheel that would make a reader want to return. This was another intentional feature of my original vision: Pindu would target self-motivated learners and avoid gamification. But reading the literature causes me to reflect that such a principled stance may in the end be unhelpful to the exact learners I am trying to serve. Gamification in the service of motivation, and not engagement for engagement's sake, should be entertained.

In some interesting way, Pindu's development so far mirrors the trajectory that many independent learners (myself included) follow by accident: start drilling flashcards intensively but burn out without actually using the language, begin to use graded readers and immersion methods extensively but stall out on vocabulary, and then back off to a balanced approach. This literature review has done much less to steer the future of Pindu than I originally thought it might, but it certainly has given me confidence that the product's current direction is not wrong and, on the contrary, is quite defensible. To return to the original question I asked at the top of the article, the more informed phrasing would probably be, "How can Pindu best support both intensive and extensive reading?" It's and, not or.

Sources

The analysis behind this article took the form of a brief literature review. It was not overly formal: I selected just the key works from the SLA, ER, and Vocabulary fields that seemed to matter for Pindu. I used a combination of Google Scholar, LLMs, and bibliography mining to find the sources, and I ended up with a pruned list of 17. These are not all directly cited in the article, but they all have influenced it. For fun, Figure 6 is a "Citation Map" that visualizes their mutual chronology and dependencies.

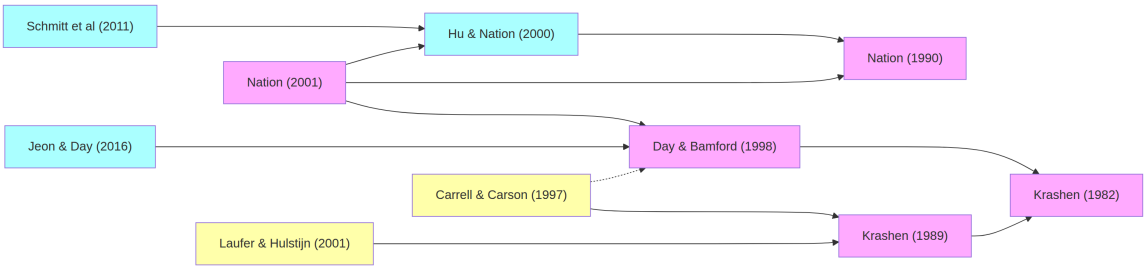


Figure 6: A stylized portrayal of how this article's sources relate. Arrows point from later works to the earlier ones they directly cite. Nodes are colored by category: conceptual foundations, empirical support, and practical considerations. Many lines are omitted for clarity, and the map is not comprehensive.

The sources I used are listed below. (Note for the curious: the Extensive Reading Foundation maintains a bibliography of ER-related works at bib.erfoundation.org.)

Carrell, Patricia L. & Joan G. Carson (1997). Extensive and Intensive Reading in an EAP Setting. English for Specific Purposes.

Day, Richard R. (2015). Extending extensive reading. Reading in a Foreign Language.

Day, Richard R. & Julian Bamford (1998). Extensive Reading in the Second Language Classroom. Cam-

bridge University Press.

Hu, Marcella & I.S.P. Nation (2000). Unknown vocabulary density and reading comprehension. Reading in a Foreign Language.

Jeon, Eun-Young & Richard R. Day (2016). The effectiveness of ER on reading proficiency: A meta-analysis. Reading in a Foreign Language.

Krashen, Stephen D. (1982). Principles and Practice in Second Language Acquisition. Pergamon Press.

Krashen, Stephen D. (1989). We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. The Modern Language Journal.

Laufer, Batia (2001). Reading, Word-Focused Activities and Incidental Vocabulary Acquisition in a Second Language. Prospect.

Laufer, Batia & Jan Hulstijn (2001). Incidental Vocabulary Acquisition in a Second Language: The Construct of Task-Induced Involvement. Applied Linguistics.

Nakanishi, Takayuki (2015). A Meta-Analysis of Extensive Reading Research. TESOL Quarterly.

Nation, I.S.P. (1990). Teaching and Learning Vocabulary. Newbury House.

Nation, I.S.P. (2001). Learning Vocabulary in Another Language. Cambridge University Press.

Nation, I.S.P. & Ming-Tzu Karen Wang (1999). Graded Readers and Vocabulary. Reading in a Foreign Language.

Nuttall, Christine (1996). Teaching Reading Skills in a Foreign Language. Heinemann.

Schmitt, Norbert, Xiangying Jiang, & William Grabe (2011). The Percentage of Words Known in a Text and Reading Comprehension. The Modern Language Journal.

Suk, Namhee (2017). The Effects of Extensive Reading on Reading Comprehension, Reading Rate, and Vocabulary Acquisition. Reading Research Quarterly.

Zhou, Jing & Richard R. Day (2023). Establishing an Extensive Reading Program in a Chinese as a Foreign Language Context. Reading in a Foreign Language.