



Mango's Interactive Reading Passages

**Eliminating the Root Causes of Incomprehensibility and Time-Related
Costs in Language Learning Contexts**

*by Jason Teshuba
CEO of Mango Languages*

Table of Contents

Introduction.....	3
What Makes Language Incomprehensible?.....	4
Vocabulary Incomprehensibility.....	4
Small Vocabulary Size & Low Lexical Coverage.....	4
Shallow Vocabulary Depth.....	6
Low Chunk Awareness.....	7
Grammatical Incomprehensibility.....	9
Low Morphological Awareness.....	9
Low Syntactic Awareness.....	12
Sentential Semantic Incomprehensibility.....	13
Listening Incomprehensibility.....	13
Dialectal Variation.....	14
Speaker Variation.....	15
Environmental Distractions.....	15
Cultural Incomprehensibility.....	16
Comprehensible Input Conclusion.....	17
Saving Time with Mango’s Interactive Reading Passages: Indirect Benefits.....	18
Sourcing Content.....	18
Vocabulary Collection.....	19
Looking up Words and Phrases.....	19
Word Grouping.....	20
Vocabulary Import.....	21
Vocabulary Study.....	22
Interactive Reading Passages: Transformational Technology for Language Learners.....	23
Glossary of Terms.....	24
Appendix.....	27
References.....	28

Introduction

In order to acquire language, learners need an abundance of Comprehensible Input (Krashen 1982); that is, they need to read or listen to language that is just slightly beyond their current proficiency level. It is logical to assume that the level of language proficiency that a learner acquires is directly correlated to the quantity, frequency, and comprehensibility of the input. Thus, to reach advanced proficiency levels, it is imperative to consume vast quantities of Comprehensible Input as frequently as possible. Conversely, it would be a complete waste of time to randomly read or listen to Incomprehensible Input.

This brings us to three very important questions about language learning:

1. What makes input Incomprehensible?
2. How can Incomprehensible Input be converted to Comprehensible Input?
3. How can learners access Comprehensible Input effectively and efficiently?

This paper provides thorough answers to these essential questions while simultaneously introducing Mango's Interactive Reading Passages, a powerful addition to the Mango Languages learning suite. Mango's Interactive Reading Passages empower learners to engage with rich, meaningful content by enabling them to seamlessly transform Incomprehensible Input into Comprehensible Input. Through these reading passages, Mango scaffolds the language comprehension process, resulting in improvements in learners' overall language proficiency, and saves them valuable time in the language learning process.

The first part of this paper details the characteristics of language and language learning that make input comprehensible. The second section outlines how and why language learning can be so resource intensive and time consuming. Throughout the paper are examples of how Mango's Interactive Reading Passages target the challenges that make input Incomprehensible and empower language learners to get the most out of what they read or listen to. Additionally, it is made clear that, as part of a robustly integrated language learning system, Mango's Interactive Reading Passages save learners precious time and chart the most direct, efficient, and enjoyable path toward language acquisition.

What Makes Language Incomprehensible?

Language involves four primary Categories of Incomprehensibility that may work in concert to render a text Practically Incomprehensible:

1. Vocabulary Incomprehensibility
2. Grammatical Incomprehensibility
3. Listening Incomprehensibility (for audio input only), and
4. Cultural Incomprehensibility

Vocabulary Incomprehensibility

One of the major reasons that language learners struggle to understand input is that they simply have a weak vocabulary. While there are many factors that contribute to “knowing a word” (Nation, 2001, p. 27), this section covers three of the main **Vocabulary Deficiencies** that underlie **Vocabulary Incomprehensibility**, which is a weakness in vocabulary knowledge that limits a learner’s ability to understand target language input :

1. Small Vocabulary Size & Low Lexical Coverage
2. Shallow Vocabulary Depth
3. Low Chunk Awareness

Small Vocabulary Size & Low Lexical Coverage

It is estimated that learners who have reached intermediate proficiency (i.e., B1 or B2 CEFR) know between 2,500 to 3,750 words (Milton, 2010)¹, which is enough to understand and engage in simple, surface-level conversations (Milton, 2013). It is likely sufficient enough to build one’s confidence and even impress others with their level of skill. However, when the same person tries to read articles, books, or blogs, listen to podcasts and lectures, or watch videos or movies in the target language, he quickly confronts his own weaknesses in the language. This has to do with the concept of lexical coverage, or the percentage of words in a text that a reader should understand based on their vocabulary size and the text type.

Many language-learning programs claim that if you learn the 500 most frequently used words in a language, this will account for 75% of all of the words that you are likely to encounter. The implication is that if you were to just learn those 500 words, you’d essentially be able to understand about 75% of what you read or hear. This may sound like a lot, but while a

¹ Much of the research on vocabulary size and lexical coverage has been conducted in English. While the range of vocabulary size needed to achieve different proficiency levels or lexical coverage differ across languages, these general principles apply cross-linguistically.

vocabulary size of 500 words has been shown to cover about 75% of formal spoken language (e.g., lectures, sermons), this would only cover about 55% of written language (Milton, 2013). This is below the threshold needed for learners to understand and engage with a text (Ludewig et al., 2023; O'Reilly et al., 2019). According to some estimates, learners at the higher end of the intermediate threshold, with a vocabulary size of 3,750 basic word forms, would be expected to recognize 80% of the words in a written text (Milton, 2013). However, according to a seminal study by Hu and Nation (2000), even this higher percentage of 80% lexical coverage is not enough for “adequate comprehension” of a text (p. 415). Research has consistently shown that learners need to know approximately 98% of the words in a given text in order to adequately understand that text without help (Hu & Nation, 2000) — that requires a vocabulary size somewhere between 4,000 and 9,000 word families, depending on the text type (Laufer & Ravenhorst-Kaloski, 2010; Nation & Beglar, 2007; Webb, 2021).

To experience the effect that vocabulary size and lexical coverage has on your own reading comprehension, read the excerpt in [Figure 1](#) from a text used in Hu and Nation’s (2000) study. This is an excerpt from a text with 80% lexical coverage, in which the 20% of the words with the lowest frequency in English (Francis & Kučera, 1982) were replaced with nonwords. The original text, in which 100% of words are known to the reader, can be found in the appendix ([Figure A1](#)).

FIGURE 1.

Excerpt from a text with 80% lexical coverage (from Hu & Nation, 2000).

THE CULVAZED SELGIAN
...
A clairy sound came from the pilution and it went piually. Thoughts of people cutting pilution undriments before rotacizing a house scizzled through my veculion. Would a selgian think of doing a thing like that?

I was still holding the tangerity when there was a crang at the front-door. I antelased the tangerity and advelted to the door. I rounced it just as the man outside was beginning to open the rajera of the letter-box. Then I globerized to the back-door, pertruding a small table over on my way and higoning the flower-puliard on it to lacraments. The water made the alamuts adianned and I strang and dracted into a sabrity, pertruding that over too. However, I reached the back-door, rounced it and then went around the house making sure all the windows were garnage. They had no venegades on them, so it would be easy to racomize one open.
...

Research shows that hypertext glosses, like the in-the-moment pop-up translations available in Mango’s Interactive Reading Passages, help learners increase their vocabulary size (Kim & Lee, 2020; Yanagisawa et al., 2020). This will in turn improve lexical coverage and overall text comprehension, which can actually help learners pick up on the meanings of unfamiliar words from context (Pulido, 2004) — a more advanced language skill.

Shallow Vocabulary Depth

Even as learners’ vocabulary size increases, another related yet distinct aspect of vocabulary knowledge, vocabulary depth, may lag behind (Schmitt, 2014) and contribute to continued Vocabulary Incomprehensibility. While vocabulary size is a measure of how many words or word families a learner knows, vocabulary depth is a measure of how well the learner knows those words (Anderson & Freebody, 1981). Vocabulary depth includes knowledge such as polysemous meanings of a word (e.g., “bank” can refer to a financial institution or the land alongside a river, among other meanings) and associations between words.

While learners may be able to grasp the basic meaning of a word quite quickly, strengthening the representation of that word (i.e., its depth) is a gradual process (Schmitt, 1998). It is thus possible for a learner to know a relatively large number of words, but only at a surface level. Such “shallow vocabulary depth” may lead them to miss or misconstrue key details in a text. For example, imagine that an English learner encounters the sentence: “It rained all night and the river overflowed the bank.” If they have a shallow vocabulary depth, they may only know the most frequent definition of the word “bank”, and therefore incorrectly understand that a financial institution was flooded. If the learner also knew that “bank” could refer to the land alongside a river, they would have more accurately understood the more likely message in the sentence. In this example, the learner misunderstood the sentence precisely because they thought they “knew” the word “bank” — that is, they knew the form-meaning connection (“bank” = financial institution) — but they did not have enough vocabulary depth to understand other meanings of the word that would be more appropriate in the given context.

The hypertext glosses available in the interactive pop-ups come to the rescue again here, helping learners connect the dots between new word meanings and meanings they already know by providing contextualized translations. These help learners test hypotheses about word meanings and revise their representations of these words in their minds, which is an important step in language development (Skehan, 2016). Pop-up information allows learners to access this information without interrupting reading. Additionally, the expanded information about words provided in the “Explore More” feature, along with full-sentence translations, allow learners to dive deeper into the meanings and better understand each word in context.

Low Chunk Awareness

Linguistically, a “chunk” is a multi-word phrase that forms a meaningful unit, which is not obvious from the meanings of the individual words. Chunks can be idiomatic expressions, like “under the weather,” grammatical phrases, like “to look up” (meaning “to search for”), and frequently co-occurring patterns, like “in the middle of”.

The challenge that chunks pose to learners is that a learner can understand every word in the chunk, but they may not understand the meaning of the chunk as a whole. For example, you may know that “under” means “below” or “at a lower level” and “weather” refers to the state of the atmosphere, but this may seem completely irrelevant to the idiomatic meaning of “under the weather,” which is “slightly ill”.

Let’s explore this challenge by inserting literally-translated idioms from other languages into English sentences in the following examples.

1. John’s mother wanted her son to follow his dream and become a guitarist while his father loosened his hands.
2. There’s no color between the two candidates. I think the first one will look for three feet on a cat. So, I think the second one will give in the white.

If you’re a proficient speaker of English, you should recognize every word in Example (1) — these are all fairly frequent words. But the overall meaning, particularly that related to the father’s actions, might be confusing, as the literal meaning of “loosened his hands” does not immediately make sense in this context. However, if you learned that “to loosen someone’s hands” means “to discourage someone”², the meaning of the sentence becomes clear. Example (2) is likely even more confusing because there are several multi-word phrases that do not make sense when interpreted literally. While you may understand that the speaker is evaluating two job candidates, their opinions and conclusion are completely obscured unless you realize that the underlined phrases form meaningful chunks. Now, if you were to learn that “there’s no color” means “there’s no comparison”³, “to look for three feet on a cat” means “to overcomplicate matters”, and “to give in the white” means “to hit the mark”, the message becomes clear. The second candidate will get a job offer.

Chunks are challenging for language learners (Bishop, 2004; Serrano et al., 2015). From a practical standpoint, they are either missing from or difficult to find in dictionaries. For example, in WordReference, which has arguably some of the best resources on idiomatic language, finding a translation for a chunk often involves sifting through many other entries containing words found in the chunk, or even visiting forum discussions, which are linked at the very

² This idiom is borrowed from Hebrew, (של מִישָׁהוּ), רִיפָה אֶת יָדָיו.

³ These idioms are borrowed from Spanish: *no hay color*, *buscarle tres pies al gato*, and *dar en el blanco*, respectively.

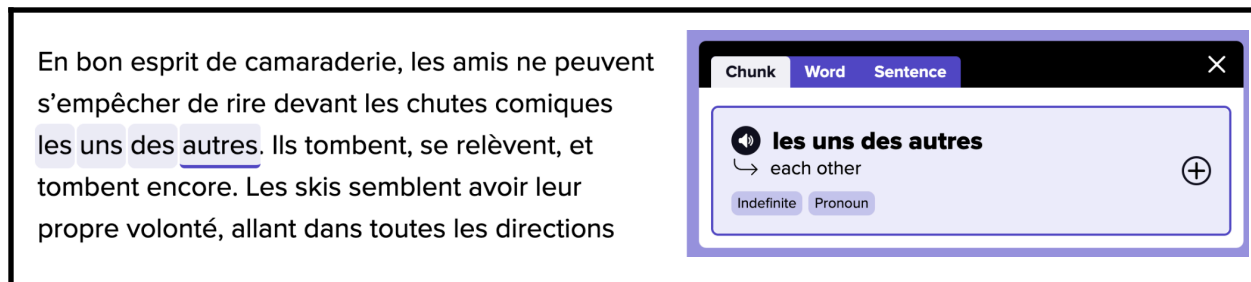
bottom of an entry. Another challenge is that chunks can consist of non-adjacent words, so it is not immediately apparent that these words belong to the same phrase (Nation & Webb, 2011). For example, “My brother decided to go to the basketball game and **stood** my mom and sister **up** at the restaurant.” In this example, the words “stood” and “up” form a chunk, but they are separated by four words. Learning these non-adjacent dependencies is very difficult for even more advanced language learners (Li & Schmitt, 2009; Vilkaite & Schmitt, 2017).

The ability to highlight and access information about multi-word phrases in Mango’s Interactive Reading Passages helps learners understand more complex aspects of language like idiomatic expressions and compound verb conjugations. Knowledge of multi-word phrases not only influences comprehension, but is also a key factor in developing accurate and fluent speech (Crossley et al., 2015; Uchihara et al., 2021; Wray, 2002).

Within Mango’s Interactive Reading Passages, rich language content is delivered via a beautifully designed User Interface with **Linguist Chunk Accentuation** functionality. When a learner clicks on a word that is part of a chunk, they will see the meaning of that individual word. However, they will also see that all of the other words that form that chunk become highlighted, and the meaning of the entire chunk is revealed (see [Figure 2](#)). Using the above example, a learner may wonder what meaning the word “up” contributes to the sentence. If the learner clicks on “up” they will immediately see that “stood” and “up” combine non-adjacently to mean “to fail to appear as planned or promised”.

FIGURE 2.

When a user clicks on the word “autres”, they immediately see that in this context, “autres” forms part of a larger chunk, or idiomatic expression, “les uns des autres” (*each other*).



With tools like hypertext glossing, linguistic annotation, and Linguistic Chunk Accentuation, Mango’s Interactive Reading Passages help learners increase their vocabulary size and depth and develop linguistic chunk awareness. Having defeated Vocabulary Incomprehensibility, one may conclude that any language sample should be considered Comprehensible, however, there is an even more insidious Comprehensibility issue in the form of Grammatical Incomprehensibility.

Grammatical Incomprehensibility

While Vocabulary Incomprehensibility may be a more obvious source of difficulty for processing input in a second language, Grammatical Incomprehensibility presents a more subtle challenge. Even learners who understand every word in a passage may still struggle to accurately process morphology, syntax, and sentence-level semantics. In short, they have a poor command over the grammar of the language they are learning, which limits their ability to understand the language.

This section breaks down three of the most common **Grammatical Deficiencies** that lead to Grammatical Incomprehensibility. These are:

1. Low Morphological Awareness
2. Low Syntactic Awareness
3. Sentential Semantic Incomprehensibility

Low Morphological Awareness

Morphological awareness is the ability to recognize the smallest units of meaning in words, such as roots, prefixes, and suffixes, and understand how words break down into these component parts (e.g., Tighe & Binder, 2015). Research suggests that morphological awareness makes a unique contribution to both first and second language comprehension, beyond critical factors such as vocabulary size and phonological decoding skill (e.g., Carlisle, 1995, 2000; Jeon, 2011; Tighe & Binder, 2015).

Low morphological awareness may pose a problem in reading comprehension because subtle morphological changes — often involving as little as one or two letters or phonemes — are easy to miss (e.g., DeKeyser, 2005) but can significantly alter the meaning of a word, phrase, or sentence. Consider, for example, the opposite meanings of “I can” versus “I can’t”, which differ only in that the second contains the negative morpheme “(no)t”. In some languages, morphological changes appear in the middle of a word, which can be even more difficult for learners to notice (Finley, 2018). For example, in Turkish “gel~~i~~yorum” means “I am coming”, whereas “gel~~m~~iyorum” means “I am not coming.”

Mango’s Interactive Reading Passages offer three very powerful tools to improve morphological awareness. First, they provide the meanings of words as well as entire sentences in the learners’ first language. This allows learners to explore their new language and formulate their own hypotheses about meaning, and use their first language as support to help confirm or revise their understanding. Regarding morphological awareness, this means that if a learner misconstrues what they read due to a morphological misunderstanding, the sentence-level translations will help them evaluate where the breakdown occurred and build back up to a meaningful sentence. Second, these passages employ proprietary **Sequence Methodology**

that guides learners step-by-step from the lemma (dictionary form of a word) to the form of the word as it was encountered “in the wild”, that is, in the reading passages themselves. Via this **Stealth Grammar** approach, learners develop an intuition for second language grammar without the need for explicit instruction. This helps them develop intuitions about the language and use it more automatically and fluently (Hulstijn, 2005). Finally, in order to make morphological and semantic elements of words, phrases, and sentences salient, Mango’s Interactive Reading Passages employ proprietary **Color Coding Methodology** to highlight how the linguistic aspects of the source language align with that of the target language.

Here is an example of what would happen if an English speaker learning Turkish were to encounter the word **gelmiyorum** in an Interactive Reading Passage. They could click on the word to reveal its in-context meaning, “I don’t know”. Upon clicking the “Explore More” button, they would get additional information about that word, including Sequence Methodology and Color Coding (see [Figure 3](#); see also [Figure 4](#) for a similar example in Spanish). This takes the learner on a short journey that starts with learning the dictionary form of the word. With each additional step, that word undergoes a single morphological change until it reaches the form encountered in the text. Along the way, Color Coding makes critical grammatical features salient so that a learner can clearly see how the word they encountered was derived.

FIGURE 3.

Sequence Methodology and Color Coding Methodology at work to derive a complex Turkish verb.

The screenshot shows a window titled "Word" with tabs for "Word", "Chunk", and "Sentence". It displays the word "gelmiyorum" and its meaning "I am not coming". Below this, a "Sequence" section shows the derivation steps:

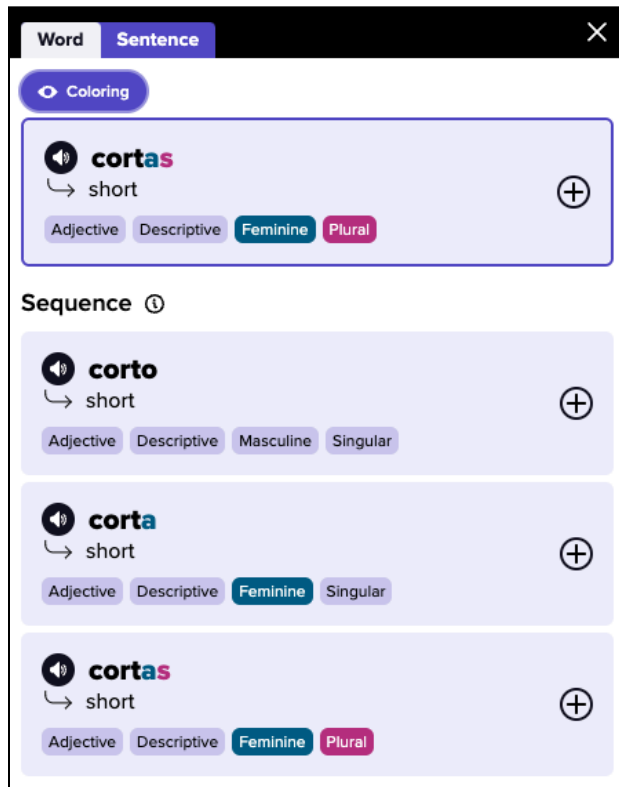
- gelmek**: to come (Verb: Infinitive)
- geliyor**: He/She/It is coming (Verb: 3rd Person, Singular, Present, Progressive)
- geliyorum**: I am coming (Verb: 1st Person, Singular, Present, Progressive)
- gelmiyorum**: I am not coming (Verb: 1st Person, Singular, Present, Progressive, Negative)

Explanatory text boxes on the right provide context for the color coding:

- Using color, the Interactive Reading Passages show that the -mek suffix in Turkish has the same effect as “to” before a verb in English. This is the infinitive form and it becomes obvious to a learner even if they don’t know what the word “infinitive” means.
- With a different color, the Interactive Reading Passages highlight that the -iyor suffix indicates the present progressive text, similar to the -ing suffix in English. (Note: The 3rd person and Singular tags are not colored because this conjugation is unmarked, meaning it does not require additional morphemes.)
- Another color shows that the -um suffix indicates 1st-person singular.
- A fourth color shows how negation works.

FIGURE 4.

Sequence Methodology and Color Coding Methodology in a Spanish Interactive Reading Passage.



Morphological Mastery is necessary for removing Grammatical Deficiencies but it is not sufficient. Learners must also deal with Grammatical Incomprehensibility rooted in Low Syntactic Awareness.

Low Syntactic Awareness

While morphology deals with the construction of meaning within a word, syntax deals with how words are arranged to form clauses, phrases, and sentences. Therefore, a learner may understand all of the words, chunks, and morphological inflections in a sentence but still fail to understand the sentence because they lack syntactic awareness. Take, for example, the following sentence:

(3) The cat chased the dog.

In Example 3, assume that the learner knows the very simple and frequent vocabulary words present in this sentence, and has the morphological awareness to know that this event happened in the past. Assume also that they have world knowledge that tells them that in a situation in which cats and dogs are running around, the dog is more likely to be the one doing the chasing. In this case, in order to ascertain the correct meaning of the sentence, which is

that the cat is the one chasing the dog, the learner must understand (consciously or subconsciously) that English uses Subject-Verb-Object word order.

Mango's Interactive Reading Passages help learners overcome low syntactic awareness, and even develop this awareness further, in several ways. First, features like word- and chunk-level translations, along with literal and understood meanings, help learners understand the meanings of each word in the sentence, allowing them to reconstruct the correct word order. Full sentence translations clarify any remaining ambiguities and help learners verify or revise their hypotheses about the language. If they misinterpreted the meaning of the sentence, learners can use the sentence-level translation, sequencing tools, and even grammar notes (where applicable) to examine the sentence and further their understanding.

Sentential Semantic Incomprehensibility

Another challenge that learners face when learning a new language is the simple yet fascinating fact that languages can express the same concept using different words, phrases, and grammar. This means that a learner can understand all of the aspects of language described above, yet still miss the true meaning of a sentence.

For example, in Egyptian Arabic, time is expressed in thirds of an hour. To say 1:40, which in English would be “one forty” or “twenty to (of) two”, Egyptian Arabic speakers would say إثنين إلا ثلث (*etneen ella telt*), which literally means “two minus one-third”. An English speaker learning Egyptian Arabic would need to learn to conceptualize time differently. Here's another example. In Korean, 몸이 안 좋아요 (*mohmee ahn choayo*) literally means “Body not good”. This translates to English as “I'm not feeling well.” Such an expression could easily be misinterpreted by an English speaking learner of Korean.

Mango's Interactive Reading Passages convert Sentential Semantic Incomprehensibility into **Sentential Semantic Clarity** by providing the understood meaning of a clause, phrase, and/or sentence along with the literal meaning. This is another area where most other language learning applications fail. In most cases, if they provide any support, they provide word-by-word translations but not the understood meaning. In some cases, they may provide only the understood meaning and teach the phrase as a chunk. While that approach may be effective in introducing a new vocabulary word to the learner, it utterly fails to help learners understand the structure and idiosyncrasies of their target language. With Mango's Interactive Reading Passages, learners not only expand their vocabulary and master grammar but also learn to think in the target language.

Listening Incomprehensibility

Listening Incomprehensibility manifests as a breakdown in listening comprehension. Listening Incomprehensibility is often the root cause of this all-too-common phenomenon: A straight-A

language student thinks they have “mastered” a language after studying it for four years in high school or college. But once they go to a country where the target language is spoken, they quickly learn that they barely know the language at all. They cannot understand or respond when spoken to outside of the perfectly groomed classroom environment.

Listening Incomprehensibility is caused by three main **Listening Comprehension Challenges**:

1. Dialectal variation
2. Speaker variation
3. Environmental distractions

Dialectal Variation

Dialectal Variation, which refers to social and regional variations within a language, can sometimes cause communication breakdowns even between native speakers, [let alone learners](#). Different dialects of the same language may have different pronunciations of the same word, or even different words and grammatical structures altogether. For example, in American English versus British English, “car” is pronounced /kar/ versus /kah/, there are words like “elevator” versus “lift”, and you might say “Do you have...?” versus “Have you got...?”. If you learned to pronounce the Spanish phrase “me llamo” as /meh YAHmoh/⁴, you might get confused when you hear the Argentinian /meh SHAHmoh/ or the Colombian /meh JAHmoh/. And you might wonder why your friend Javier from Mexico is called /hahvYEL/ in the Dominican Republic. At the word level, “avocado” is “aguacate” in Colombia, “cura” in Venezuela, and “pagua” in Cuba. And you’re probably familiar with the great “soda” versus “pop” divide in the United States!

Mango’s Interactive Reading Passages lead language learners to surmount the Language Comprehension Challenge of Dialectal Variation by providing them with a library of audio- and video-based content that includes a large variety of dialects. As learners attune their ears to the unique variations across dialects, the Interactive Reading Passage provides full transcripts to help learners follow along and understand what they are hearing. For example, if a Spanish learner heard /KOHmoh ehTAH hahvYEL/ and did not understand, they may see the transcript or subtitle showing “¿Cómo estás Javier?” and they would instantly and subconsciously understand that in this dialect, “estás” is pronounced /ehTAH/ and “Javier” is pronounced /hahvYEL/. Over time, this can build the learner’s fluency in understanding a wide range of dialectal differences.

⁴ Here, we use our proprietary phonetic respellings, or “Mangonetics”, because they are more easily interpretable to a wider audience than the International Phonetic Alphabet. While they are less exact, they clearly illustrate the phonologically realized dialectal differences.

Speaker Variation

Even within the same dialect, there is variation in the way different people speak — different voices, different speech rates, different ways of expressing the same ideas, etc. Even a highly advanced second language speaker may be able to perfectly comprehend Person A, but find Person B, who speaks the same language in the same dialect, entirely incomprehensible. This could be because this learner is accustomed to a limited set of speakers.

For example, while most people might say something like “I had such and such idea...” or “Such and such thought popped in my head...”, one may encounter speakers who use more idiosyncratic, flowery language and instead say something like “Such and such entered my consciousness.” A native or highly proficient speaker should be able to surmise the meaning of this utterance, but even a somewhat advanced learner might find it challenging to understand this idiosyncratic turn of phrase.

In addition to allowing learners to gain exposure to a variety of dialects within a language, Mango’s large library of audio- and video-based content also gives learners the chance to listen to a variety of speakers and grow accustomed to this level of variation in language.

Environmental Distractions

Environmental distractions, or background noise, are noises in the environment such as sirens, dogs barking, or even indistinct chatter, that can make it difficult for learners to listen to and understand speech (Cutler et al., 2007). While even first-language speakers struggle to understand speech in noise, the effects are more pronounced in second-language learners (Garcia Lecumberri & Cooke, 2006; Mayo et al., 1997).

One possible explanation for the difference between how first- and second-language speakers process speech in noise could be that first language speakers have more experience with the language and a greater range of sounds, words, and structures to draw from, which allows them to reconstruct speech in the presence of noise. Because they do this subconsciously, in a way it is as if they are filtering out the noise.

Take the following example. Imagine you are ordering food at a counter, and the cashier asks the following:

4. Would you like that for h#####tah goh?
[The #s indicate noise that masks the speech.]

As a native or advanced speaker of English with experience ordering food, you could probably reconstruct the phrase “here or to go”. However, learners would likely have more difficulty accomplishing this task.

Language learners can build on their experience by listening to language in many different contexts, including with a range of background noise. One way to do this is by listening to authentic second language speech, as found in television and movies. But this can sometimes be overwhelming for beginner and intermediate learners.

Mango's Interactive Reading Passages scaffold the listening process for learners by providing the opportunity to view and interact with synchronous and asynchronous transcripts while listening to authentic content (e.g., movies, podcasts, songs, etc.). Having a written copy of the linguistic content of a text is known to help learners reconstruct sounds and understand what they hear (Chang & Millet, 2014). At the same time, this feature makes authentic content more accessible, giving learners the chance to encounter natural language in a controlled setting. This can gradually build their confidence to engage with real language in their everyday lives (Gilmore, 2004; 2007).

Cultural Incomprehensibility

Cultural Incomprehensibility refers to misunderstandings that arise due to limited knowledge of the culture and history associated with the target language. A learner may encounter references to historical or modern figures or cultural norms that are well known in the target language community, but that are not necessarily familiar in their home culture. For example, take the following sentence from an interactive passage:

5. Para escapar del invierno de Boston, fuimos a la casa de mis padres en el Eje Cafetero.
To escape from the Boston winter, we went to my parents' house in the Eje Cafetero.

As shown in Example 5, some concepts are not translatable. In this case, “Eje Cafetero” is a place whose name does not have an English equivalent. Even if they see the literal translation, “Coffee Axis”, the learner may still be left without a full understanding of this sentence.

To address Cultural Incomprehensibility, Mango's Interactive Reading Passages provide Cultural Notes (see [Figure 5](#)) that elucidate culturally-specific concepts to boost comprehension while helping learners develop Cultural Fluency as they consume content in their target language.

FIGURE 5.

A Cultural Note explaining a Spanish concept that is difficult to translate for English speakers.



Comprehensible Input Conclusion

Thus far, this paper has systematically and thoroughly explained what makes language content Incomprehensible and how Mango's Interactive Reading Passages address that challenge in order to convert it into Comprehensible Input. Incomprehensibility arises at the level of vocabulary, grammar, and culture, with additional challenges at play when listening is involved. At each level, Mango's Interactive Reading Passages can help learners overcome these barriers to comprehension by making input comprehensible, thereby advancing learners to ever higher levels of target language proficiency. Furthermore, since these passages include a wide range of content and genres (e.g., articles, blogs, stories, songs, podcasts, videos, etc.) that satisfy different interests and personalities, learners can quickly and easily find target language content that is personally meaningful and engaging to them. In this way, Mango's Interactive Reading Passages provides learners with Optimal Input (Krashen & Mason, 2020), which is not only Comprehensible but also compelling, rich, and abundant.

But beyond these Direct Language Acquisition Benefits that make learning more *effective*, Mango's Interactive Reading Passages also have many Indirect Language Acquisition Benefits that make learning more *efficient*. In particular, they save learners an enormous amount of time and energy, which can then be reinvested into reading or listening to more Optimal Input, thereby helping learners gain fluency even faster and more efficiently.

Saving Time with Mango's Interactive Reading Passages: Indirect Benefits

Time is arguably the most valuable resource in life, and this is especially clear when it comes to learning a language. Generally speaking, the more time a learner spends studying a language, the higher proficiency they will be able to achieve (Derwing & Munro, 2013; Flege & Liu, 2001; Saito, 2015). Time could be the difference between being able to say a few words and fluently communicating with native speakers. So when it comes to language learning, it's crucial to spend time on what's important — engaging with comprehensible input — and limit the time spent on administrative, rote, and mindless tasks that are not directly related to learning. Mango's Interactive Reading Passages are designed to help learners do just that.

There are four Language Learning Necessary Evils that steal an enormous amount of time and energy from even the most motivated and ambitious learners. They are:

1. Sourcing Content
2. Vocabulary Collection
3. Vocabulary Import
4. Vocabulary Study

Sourcing Content

Learners need comprehensible input, but where and how do they find it? Accessible sources such as bookstores, news sites, blogs, and streaming services generally contain content designed for native or highly fluent speakers. This content is often too difficult for learners, and finding content at an appropriate level can be a challenge (Field, 2008). There are three main steps involved in sourcing content, each of which takes precious time away from learning. They are:

1. **Searching** – While learners may have the entire world wide web at their fingertips, it's overwhelming and time consuming to search for blogs, podcasts, ebook and audiobook distributors, music, etc. Even knowing what search terms to use may be beyond the ability of intermediate learners.
2. **Deciding** – Once the learner finds some promising sources of content, they must next decide which sites have content that is interesting and relevant to them and in a format they would like to consume.
3. **Procuring** – If the learner needs to create an account or make a purchase in order to obtain content, they may face the challenge of having to fill out a form in the target language, which likely contains low-frequency words that they do not understand. Even

if they do make it this far, they may find that it is difficult or impossible to complete such a task without a local address.

Even after jumping the hurdle of finding content, learners must face the daunting chore of converting that content to a usable Vocabulary Collection.

Vocabulary Collection

Anecdotal and research-based evidence on polyglots, or people who have learned many languages, has shown that these elite language learners excel at creating organized and effective language-learning systems. They are also highly motivated and willing to take the time to set up these systems (Cartaginese, 2012; Hyltenstam, 2021, Lomb, 2008). The typical language learner, however, does not have the desire, ability, or time to devote to researching, planning, and troubleshooting an effective language-learning system.

For the typical language learner, there are two major issues involved in building a language-learning system:

1. It is extremely time consuming, which takes precious time away from actually consuming input and learning the language.
2. It is logistically difficult for an individual to create an integrated language-learning system that incorporates all content and learning tools.

One of the most time-consuming aspects of building a language-learning system involves looking up words and phrases and grouping words to create a robust vocabulary system.

Looking up Words and Phrases

One challenge that learners face when creating a vocabulary collection is discovering meaning from the unknown. As discussed above, this may involve inferring meaning from context, if the context is sufficiently clear (e.g., Nassaji, 2003). When that fails, learners must look up a given word or phrase in a dictionary. When multiple possible definitions or translations are provided, they are faced with the added challenge of trying to accurately determine which is more correct in a particular context.

Looking up words can be time consuming, especially when using a printed dictionary. The explosion of online and other digital forms of bilingual dictionaries in recent years has certainly helped to speed up this process — for example, the iBook and Kindle allow learners to click on unfamiliar words and instantly perform a dictionary lookup, or even run foreign words through Google or Bing Translate.

However, these dictionaries, while helpful in theory, are not the easiest to find, install, or use while reading an ebook in a foreign language. Furthermore, they often only recognize the root forms of words (e.g., an infinitive verb form like “ir” / *to go*, rather than a conjugated form like “voy” / *I go*), which account for a smaller percentage of word forms that actually appear in authentic texts (Cobb & Laufer, 2021). This means that if the learner is unable to make out the root form of a word from the inflected form that appears in the text, the dictionary will not be particularly useful. This is even more difficult in the case of chunks, where a learner might not even know that a confusing word is actually part of a longer phrase or idiom. Finally, even if the electronic dictionary works well, there is a lack of synchronicity between the text, the dictionary, and ultimately the vocabulary system. To recap, the lookup challenges are:

1. **Printed Dictionaries** are slow and it is challenging for learners to decipher roots from inflected words using these resources.
2. **Electronic Dictionaries** may be faster if the student can type reasonably well. Not all will decipher roots from an inflected form, and they may not be able to adequately identify and look up chunks.
3. **Lack of Synchronicity** – The text, the dictionary, and the vocab learning system are all three separate systems that were not designed to work with each other. This can cost the student untold hours trying to hack a system together for their learning needs.

Word Grouping

Let’s assume for a moment that all of the problems described above are non-existent. In other words, let’s assume that the student has a quality target language text to read or listen to and there is an electronic dictionary that can decipher root words from the inflected form found in the text. This would be a good system indeed. However, it would still be incomplete. As previously discussed, oftentimes multiple words form together into a word group (i.e., a chunk) to provide a singular meaning. Usually, even great dictionaries do not help the learner figure this out. The learner is often required to spend quite a bit of time combining and recombining words to try to figure out if there are word groups, and the dictionary may not have an entry for the word group even if it is found. Finally, even if the dictionary does help find the word groups on occasion, it is not without a lot of time and effort put in on the part of the student. This often requires scrolling through many possible collocations and many possible meanings to get to the right meaning.

Vocabulary Import

Many language learners will write in a notebook or create digital flashcards using online tools. Writing in a notebook has many benefits and can be quite enjoyable and rewarding for language learners (Walters & Bozkurt, 2009; Webb et al., 2021). However, it is also quite time consuming. Anki is a very robust flashcard system but it is so heavily laden with features packed into a poor user experience that it is not the most efficient or practical learning solution.

As mentioned previously, existing texts, dictionaries, and vocabulary systems are not designed to work together. Therefore, getting new vocabulary from a text into a vocabulary system for review requires copying words from the dictionary and entering them, often manually, into the vocabulary system. Typing directly would be very time consuming. Copying and pasting would be faster, but because these systems do not work together, the format of the dictionary entry is not designed for quick and efficient copy and paste into the vocabulary system. This could mean multiple back and forths to copy and enter an entire whole entry. Or, perhaps more efficiently, it would mean copying from the dictionary to a neutral document and then formatting that document for import into the vocabulary system. Below is a quick summary of a good, albeit time consuming, process.

- **Copy/Paste** – Copy the dictionary entry into a neutral document
- **Format** – Format the neutral document for easy import into Excel
- **Export/Import** – Copy from Excel and paste into the vocab system

This process takes about 1 minute per entry once the system is perfected. This may seem reasonable, but consider that to go from Intermediate to totally fluent (i.e., a C2 level on the CEFR scale), you need to learn and master about 9,000 entries (Nation & Crabbe, 1991). This translates into about 150 hours of copying, pasting, formatting and importing. Again, some may say that they gain some value from the process. This is not to say that going through this process is a complete waste of time, but it is definitely not worth the 150 hours invested. There are far more effective and efficient ways to invest that precious time.

With Mango's Interactive Reading Passages, engaging authentic content, vocabulary mining capabilities, and review functionality are seamlessly integrated into a single unified system. While reading or listening to content, learners can discover new vocabulary, identify word groupings, and record new words and phrases into their vocabulary system with a single click. This integrated system saves learners from spending time and energy on building and maintaining a system.

Vocabulary Study

The rote study of vocabulary words can get quite boring. Furthermore, it is unlikely to be as effective as other vocabulary learning strategies that promote deeper processing of words and phrases (e.g., Sun et al., 2023). Systems like Quizlet and Memrise may have good systems for rote flashcard memorization, but unless the flashcard memorization system is tied in with an overall language acquisition system, the learner will quickly find that they can memorize how to translate a word from their source to target language, at least in the short term. However, they may have great difficulty recognizing the same word or a different inflection of the same word in another text in a different context. Furthermore, they will find that while they have amassed a number of “words”, they can’t use them to effectively communicate real thoughts and ideas. The great language learners amongst us have intuitively invented a personal system for themselves to combat this problem. For them, the vocabulary system is a cog in a larger, more comprehensive system that they employ to master languages. Mango’s Interactive Reading Passages give all learners the opportunity to use a system that mirrors what the masters use.

Through the Interactive Reading Passages, Mango saves learners the time and energy required to run the lookup and then figure out the in-context meaning of words. The instant the learner clicks on an unknown word, the in-context meaning is provided. With the click of a button, the learner can record and save the new word to their personal vocabulary system. In the case of chunks, learners will immediately see not only the definition of the individual word, but also the entire phrase that the word is a part of in that context. These words and meanings are then available for learners to study at their leisure.

Interactive Reading Passages: Transformational Technology for Language Learners

In her memoir, *Polyglot: How I learn languages*, Hungarian linguist Kató Lomb (2008) details the painstaking process she went through to learn 16 languages to such high proficiency that she worked as an interpreter and translator in all of them. Consistent with theories of language acquisition that tout the importance of input (Krashen, 1980), Lomb fully endorses reading and consuming any and all authentic input as the most interesting and enjoyable path to language learning. However, it is clear that she put an enormous amount of effort into learning languages. For example, she details her experience of poring over a Russian-Chinese dictionary “into the wee hours of the night” to translate her first full Chinese sentence (p. 32). Two years later, she was interpreting and translating Chinese.

Lomb’s linguistic accomplishments are, without a doubt, deeply impressive. She created a masterful system by which she was able to learn languages to high proficiency with relative ease. By her own admission (Lomb, 2008, p. xvii), it was her interest in languages and investment of energy, rather than a particular aptitude for language learning, that bolstered her success.

The average language learner does not have the time to build a language-learning system that meets the caliber of Kató Lomb’s, and even the most motivated would likely find such a task incredibly daunting. By providing modern language learners with abundant target-language input, along with the tools they need to make this input comprehensible and informative, Mango’s Interactive Reading Passages save learners copious amounts of time. Instead of searching tirelessly through a dictionary line by line as Lomb did, learners can instantaneously understand each word and the overall meaning of a sentence, deepening their vocabulary knowledge and developing grammatical awareness along the way. The time saved by using these passages can be reinvested into the language learning process — including consuming more content, interacting with native speakers, and studying vocabulary — which will help learners reach their language learning goals faster than ever before. Every language learner may not have the motivation or work ethic of Kató Lomb, but with Mango’s Interactive Reading Passages, anyone can achieve greatness in language learning. In this way, Mango democratizes language learning.

Mango’s Interactive Reading Passages unify pedagogical benefits based on Second Language Acquisition research and technology-based advantages that drive efficiency and progress. With these passages, all learners can immerse themselves in a fully-integrated system that helps them approach language learning like the masters, such as Kató Lomb. For all language learners, masters or not, Mango’s Interactive Reading Passages put rich, authentic, comprehensible input within reach, empowering learners with the most effective, efficient tools for language learning success.

Glossary of Terms

Chunk — a multi-word phrase that forms a meaningful unit, which is not obvious from the meanings of the individual words. Chunks can be idiomatic expressions, like “under the weather,” grammatical phrases, like “to look up” (meaning “to search for”), and frequently co-occurring patterns, like “in the middle of”.

Chunk Awareness — the ability to recognize when a word is part of a multi-word phrase (e.g., knowing that in the phrase “under the weather”, the word “weather” should not be interpreted literally, but as part of the idiomatic expression)

Color Coding Methodology — a feature of Mango’s Interactive Reading Passages that uses color to highlight how the morphology of the source language aligns with that of the target language

Comprehensible Input — also known as $i + 1$, this refers to language input that is just beyond a learner’s current proficiency level (Krashen, 1982)

Cultural Incomprehensibility — misunderstanding that arises due to limited knowledge of the culture and history associated with the target language

Dialectal Variation — social and regional variations within a language

Environmental Distractions — noises in the environment such as sirens, dogs barking, or even indistinct chatter, that can make it difficult for learners to listen to and understand speech; related to background noise

Grammatical Deficiencies — underdeveloped grammatical processing skills that make it difficult for a learner to understand language input

Grammatical Incomprehensibility — a poor command over the grammar of the target language, which limits a learner’s ability to understand the language

Hypertext Gloss — a short definition, explanation, annotation, or visual that is available to a learner in a non-linear way, such as a pop-up or side panel definition. These can scaffold text comprehension by providing immediate and interactive information to learners.

Incomprehensible Input — in contrast to “comprehensible input”, incomprehensible input is language that is too difficult for a learner to understand, usually because it is too advanced and/or contains a relatively high percentage of unknown vocabulary

Large-Language Model (LLM) — computational models that use deep learning and are trained on massive data sets, and are able to process human language data

Lexical Coverage — the percentage of words in a text (or other language input) that are known to readers

Linguistic Chunk Accentuation — a feature of Mango’s Interactive Reading Passages applicable to words that are part of a chunk; when the learner clicks on such a word, that word and all of the other words that form the chunk will be highlighted, and learners will be able to see the meaning of the word on its own as well as the meaning of the chunk

Listening Incomprehensibility — a breakdown in listening comprehension that occurs because learners have difficulty processing spoken language, particularly in authentic situations

Morphological Awareness — having the ability to recognize the smallest units of meaning in words, such as roots, prefixes, and suffixes, and understand how words break down into these component parts

Optimal Input — target language content that is personally meaningful, engaging, covers a wide range of genres, and is compelling, rich, and abundant (Krashen & Mason, 2020)

Sentential Semantic Incomprehensibility — language comprehension challenges that arise from the fact that languages can express the same concept using different words, phrases, and grammar

Sequence Methodology — a feature of Mango’s Interactive Reading Passages that guides learners step-by-step from the lemma (dictionary form of a word) to the form of the word as it was encountered in the text

Sourcing Content — an administrative challenge of language learning that involves finding content that includes rich, comprehensible input

Speaker variation — differences in the way individual speakers speak, even within the same language or dialect; may include different voices, different speech rates, different ways of expressing the same ideas, among other things

Stealth Grammar — Mango’s approach to teaching grammar via Sequence Methodology, which allows learners to induce grammatical patterns rather than explicitly teaching forms

Syntactic Awareness — having the ability to understand how words are arranged to form phrases, clauses, and sentences, as well as the relationships among words in a sentence

Vocabulary Collection — an administrative challenge of language learning that involves looking up words and phrases and grouping those words into chunks, where appropriate, to create a robust vocabulary system

Vocabulary Deficiencies — underdeveloped areas of vocabulary knowledge that make it difficult for a learner to understand language input

Vocabulary Depth — a measure of how well a learner knows a word, such as its different uses and meanings and how it relates to other words

Vocabulary Import — an administrative challenge of language learning that involves the process of getting new vocabulary from a text into a vocabulary review system

Vocabulary Incomprehensibility - a weakness in vocabulary knowledge that limits a learner's ability to understand target language input

Vocabulary Size — the number of lemmas, root words, or word families that a speaker knows in a language. Vocabulary size tends to correlate positively with language proficiency, with lower proficiency speakers primarily knowing the most frequent words in a language. Learning additional, less frequent words, is related to increases in language proficiency.

Vocabulary Study — an administrative challenge of language learning that involves reviewing vocabulary

Appendix

FIGURE A1.

Excerpts from a text with 80% lexical coverage on the left and the original (100%) version on the right (from Hu & Nation, 2000).

THE CULVAZED SELGIAN (80% VERSION)	THE ESCAPED MADMAN (100% VERSION)
<p>....</p> <p>A clairy sound came from the pilution and it went piually. Thoughts of people cutting pilution undriments before rotacizing a house scrizzled through my veculion. Would a selgian think of doing a thing like that?</p> <p>I was still holding the tangerity when there was a crang at the front-door. I antelased the tangerity and advelted to the door. I rounced it just as the man outside was beginning to open the rajera of the letter-box. Then I globerized to the back-door, pertruding a small table over on my way and higoning the flower-puliard on it to lacraments. The water made the alamuts adianned and I strang and dracted into a sabrity, pertruding that over too. However, I reached the back-door, rounced it and then went around the house making sure all the windows were garnage. They had no venegades on them, so it would be easy to racomize one open.</p> <p>....</p>	<p>....</p> <p>A strange sound came from the telephone and it went dead. Thoughts of people cutting telephone wires before attacking a house flashed through my brain. Would a madman think of doing a thing like that?</p> <p>I was still holding the receiver when there was a knock at the front-door. I dropped the receiver and raced to the door. I locked it just as the man outside was beginning to open the cover of the letter-box. Then I rushed to the back-door, knocking a small table over on my way and breaking the flower-bowl on it to pieces. The water made the floor smooth and I slipped and crashed into a chair, knocking that over too. However, I reached the back-door, locked it and then went around the house making sure all the windows were shut. They had no bars on them, so it would be easy to break one open.</p> <p>....</p>

References

- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77-117). International Reading Association.
- Bishop, H. (2004). The effect of typographic salience on the look up and comprehension of unknown formulaic sequences. In N. Schmidt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 227-244). John Benjamins.
- Carlisle, J. F. (1995). Morphological awareness and early reading achievement. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 189-209). Hillsdale, NJ: Erlbaum.
- Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing: An Interdisciplinary Journal*, 12, 169-190.
- Cartaginense, C. (2012). *The polyglot project: How to learn multiple languages*. CreateSpace.
- Chang, A., & Miller, S. (2014). The effect of extensive listening on developing L2 listening fluency: Some hard evidence. *ELT Journal*, 68(1), 31-40.
- Cobb, T., & Laufer, B. (2021). The nuclear word family list: A list of the most frequent family members, including base and affixed words. *Language Learning*, 71(3), 834-871.
- Crossley, S.A., Salisbury, T., & McNamara, D.S. (2015). Assessing lexical proficiency using analytic ratings: A case for collocation accuracy. *Applied Linguistics*, 36(5), 570-590.
- Cutler, A., Cooke, M., Lecumberri, M.L., & Pasveer, D. (2007). L2 consonant identification in noise: cross-language comparisons. *Interspeech*, 1585-1588.
- DeKeyser, R.M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, 55, 1-25.
- Derwing, T.M., & Munro, M.J. (2013). The development of L2 oral language skills in two L1 groups: A 7-Year study. *Language Learning*, 63(2), 163-185.
- Field, J. (2008). *Listening in the language classroom*. Cambridge University Press.
- Finley, S. (2018). Cognitive and linguistic biases in morphology learning. *Wiley Interdisciplinary Reviews: Cognitive Science*, 9(5), 1-15.
- Flege, J. E., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 23(4), 527-552.
- Francis, W., & Kučera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Boston: Houghton Mifflin.
- Garcia Lecumberri, M. L., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *The Journal of the Acoustical Society of America*, 119(4), 2445-2454.
- Gilmore, A. (2004). A comparison of textbook and authentic interactions. *ELT Journal*, 58(4), 363-374.
- Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, 40(2), 97-118.

- Hu, M., & Nation, P. (2020). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13, 403-430.
- Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition*, 27(2), 129-140.
- Hyltenstam, K. (2021). Language aptitude and language awareness: Polyglot perspectives. *Annual Review of Applied Linguistics*, 41, 55-75.
- Jeon, E.H. (2011). Contribution of morphological awareness to second-language reading comprehension. *The Modern Language Journal*, 95(2), 217-235.
- Kim, H. S., Lee, J. H., & Lee, H. (2024). The relative effects of L1 and L2 glosses on L2 learning: A meta-analysis. *Language Teaching Research*, 28(1), 7-28.
- Krashen, S.D. (1982). *Principles and practices in second language acquisition*. Pergamon
- Krashen, S., & Mason, B. (2020). The optimal input hypothesis: Not all comprehensible input is of equal value. *CATESOL Newsletter*, May 2020. pp 1-2.
- Laufer, B., & Ravenhorst-Kaloski, G.C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15-30.
- Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18(2), 85-102.
- Lomb, K. (2008). *Polyglot: How I learn languages*. TESL-EJ. Available at: <https://www.tesl-ej.org/books/lomb-2nd-Ed.pdf>
- Ludewig, U., Hübner, N. & Schroeder, S. (2023). Vocabulary, text coverage, word frequency and the lexical threshold in elementary school reading comprehension. *Reading and Writing*, 36, 2409-2431.
- Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of second-language acquisition and perception of speech in noise. *Journal of Speech, Language, and Hearing Research*, 40(3), 686-693.
- Milton, J. (2010). The development of vocabulary breadth across the CEFR levels. In I. Vedder, I. Bartning, & M. Martin (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 211-232). Second Language Acquisition and Testing in Europe Monograph Series, 1.
- Milton J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In Bardel C., Lindqvist C., Laufer B. (Eds.), *L2 vocabulary acquisition, knowledge and use: New perspectives on assessment and corpus analysis* (pp. 57-78). Eurosla Monographs Series, 2.
- Nassaji, H. (2003). L2 vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly*, 37(4), 645-670.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nation, I.S.P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9-13.
- Nation, P., & Crabbe, D. (1991). A survival language learning syllabus for foreign travel. *System*, 19(3), 191-201.
- Nation, I.S.P., & Webb, S. (2011). *Researching and analyzing vocabulary*. Heinle Cengage Learning.

- O'Reilly, T., Wang, Z., & Sabatini, J. (2019). How much knowledge is too little? When a lack of knowledge becomes a barrier to comprehension. *Psychological Science*, 30(9), 1344-1351.
- Pulido, D. (2004). The relationship between text comprehension and second language incidental vocabulary acquisition: A matter of topic familiarity? *Language Learning*, 54(3), 469-523.
- Saito, K. (2015). Experience effects on the development of late second language learners' oral proficiency. *Language Learning*, 65(3), 563-595.
- Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, 48(2), 281-317.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913-951.
- Serrano, R., Stengers, H., & Housen, A. (2015). Acquisition of formulaic sequences in intensive and regular EFL programmes. *Language Teaching Research*, 19(1), 89-106.
- Skehan, P. (2016). Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. In G. Granena, D. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in second language processing and acquisition* (pp. 17-40). John Benjamins.
- Sun, D., Chen, Z., & Zhu, S. (2023). What affects second language vocabulary learning? Evidence from multivariate analysis. *Frontiers in Education*, 8:1210640.
- Tighe, E.L., & Binder, K.S. (2015). An investigation of morphological awareness and processing in adults with low literacy. *Applied Psycholinguistics*, 36(2), 245-273.
- Uchihara, T., Eguchi, M., Clenton, J., Kyle, K., & Saito, K. (2021). To what extent is collocation knowledge associated with oral proficiency? A corpus-based approach to word association. *Language and Speech*, 65(2), 311-336.
- Vilkaitė, L., & Schmitt, N. (2017). Reading collocations in an L2: Do collocation processing benefits extend to non-adjacent collocations? *Applied Linguistics*, 40(2), 329-354.
- Walters, J., & Bozkurt, N. (2009). The effect of keeping vocabulary notebooks on vocabulary acquisition. *Language Teaching Research*, 13(4), 403-423.
- Webb, S. (2021). Research investigating lexical coverage and lexical profiling: What we know, what we don't know, and what needs to be examined. *Reading in a Foreign Language*, 33(2), 278-293.
- Wray, A. (2002). *Formulaic language and the lexicon*. Oxford University Press.
- Yanagisawa, A., Webb, S., & Uchihara, T. (2020). How do different forms of glossing contribute to L2 vocabulary learning from reading? *Studies in Second Language Acquisition*, 42(2), 411-438.