

KONRAD SZCZEŚNIAK

THERE IS MORE TO LEARNING WORDS THAN MEETS THE CONSCIOUS EYE

INTRODUCTION

It is rather ironic that some of the key and most fundamental ideas in the study of language learning are fraught with seemingly insurmountable controversy. One example is conscious attention and its role in learning. According to Schmidt's (1990, 2010) *Noticing Hypothesis*, conscious attention is a necessary condition for learning new language forms: in order to commit to memory a new word or expression found in the input, the learner must become aware of its novelty. The learner must direct what Schmidt (1990, p. 132) calls "focal awareness" to the physical form of the new word or expression. That is, he or she must consciously and intentionally "photograph" the spelling (or register the pronunciation) of the new word. What this means in practice is that merely being exposed to a new language form while listening or reading does not guarantee that that new form will be recorded in memory. The learner must notice that the new item is not part of his or her lexicon, something that requires conscious effort and, ideally, rapt attention.

Implicit in the above reasoning is the assumption that our knowledge of language comes from the input—an assumption rejected by Chomsky and nativist authors, who insist that much of our linguistic competence does not have to be *learned* in the first place because it is innate. It is no wonder then that the Noticing Hypothesis is contested by authors like Gass (1997), who argues that since a speaker's command of a language is preprogrammed, it is therefore *not* contingent on input evidence and consequently does not require attention to that input.

KONRAD SZCZEŚNIAK, PhD, Habil., Assistant Professor at the University of Silesia in Katowice and Palacký University Olomouc; e-mails: konrad.szczesniak@us.edu.pl, konrad.szczesniak@upol.cz; ORCID: <https://orcid.org/0000-0003-3669-4983>.

More seriously, the hypothesis is not only contested by nativist authors, but, as we will soon see, it does not sit well even with those approaches that do see the input as the sole source of linguistic knowledge. Specifically, as should be evident from the discussion below, usage-based theories of language (Langacker, 1987; Croft, 1991; Tomasello, 2003; Bybee, 2010; Taylor, 2012) presuppose the need to record more information than a person is capable of attending to. I will attempt to point out why the Noticing Hypothesis in its strong form cannot account for how people learn new words, whether in their first or foreign language. However, before the main problems with the Noticing Hypothesis can be detailed, it is first important to look at its claims about attention.

1. ATTENTION TO FORM

Of course, it is rather obvious that, in a banal sense, conscious attention is vital to learning. It just stands to reason that we learn by paying attention; inattentiveness practically guarantees a failure to learn. If a person is reading absent-mindedly while being distracted, little is likely to get registered. This has often been called “reading on automatic pilot” (Brown, 1980; McTavish, 2008), an experience of following the lines of text—without engaging in the content—and realizing that we do not remember the last couple of paragraphs. Clearly, it is necessary to focus on the content. However, the claims made by the Noticing Hypothesis are much stronger. It is not enough to focus on the *content*—the storyline, the plot of the text we may be reading—instead, we must direct our ‘focal awareness’ toward the *form* of new words or expressions, that is, their spelling (or pronunciation in the case of listening). Thus, reading for pleasure does not carry any additional benefits in the form of lexical gains. According to Schmidt, incidental learning—also referred to as peripheral or unconscious learning—is not possible. This rather pessimistic conclusion seems to receive support from research in psychology. Baars (2002, p. 50) avers that “unconscious learning has been debated for decades, but there appears to be no robust evidence so far for long-term learning of unconscious input.”

At first glance, the logic of the hypothesis seems unassailable and consistent with people’s personal experience of reading. The moment we come across a new word, we do a double-take; that is, we pause to inspect the new word and sometimes decide to look it up. If we hear a new word in a conversation, we may focus on it by asking the speaker what the word means. In such situations, we can be said to consciously register the linguistic form by shifting our attention from the content to the form. These acts can be taken to represent our heightened attention directed at the new element of input, precisely the way the Noticing Hypothesis envisions it.

Further, it is not just individual words that seize a learner's active attention. People also tend to notice new *combinations* of words, especially if these are imaginative, humorous or apt descriptions of the subject at hand. To take one example of lexical combinations that are almost certain to attract listeners' or readers' attention, consider aphorisms. The title of Ochs and Schieffelin's (1989) paper "Language has a heart" describes the affective nature of language so succinctly and memorably that it would take active effort to avert attention from the wording of the expression. When reading a passage by an eloquent author, we are all but guaranteed to come across a number of such unique lexical combinations prompting us to pause and reflect on them, which increases the chances of remembering them and thus making them part of our personal lexicon.

While this much seems uncontroversial, the really important question is whether all expressions are learned consciously. As we will see, speakers know considerable numbers of language forms such as *the fact that* or *black and white*, which are not particularly salient, unlikely to attract much conscious attention. Is it realistic to talk of learning such items through conscious noticing?

2. LEXICAL VS. SUBLEXICAL PROCESSING

It is important, at this point, to consider how attention is mobilized in the process of comprehension. Research on processing visually presented words indicates that while a person is busy reading, his or her attention can be distributed among a number of distinct dimensions. While it is possible to switch attention to the physical properties of the text in front of us (e.g., spelling, font, size, etc.), the default state is to keep attention primarily directed to the meaning of words and their relations to other words (Balota et al. 2000, p. 1081). What this means is that our conscious focus does *not* alight on the shape of words. Visual recognition is automatic, typically handled by those mental mechanisms that are traditionally considered to be relatively effortless and fast-acting compared to more attention-demanding mechanisms. Attentional control mechanisms are not enlisted for word recognition unless a person closely analyzes the details of form in what is termed *sublexical processing*. This involves inspecting the ordering of the characters, their shapes, the font, or indeed any other detail such as whether they are lower or upper case.

Unfortunately, our knowledge of the nature of the mechanisms in question is rather limited, as research on text processing is still ongoing and some controversy surrounds the issue of how much attention is really involved in word recognition. For example, Besner et al. (2016) argue that word reading taps into

a lot more attentional resources than was originally assumed. On the other hand, Augustinova and Ferrand (2014, p. 344) point out that newly discovered signs of heightened attention do not invalidate the automatic view of word reading.

Now, whatever the disagreements, it is safe to say that the lexical-sublexical distinction is real, even if blurred. In the course of lexical processing, words are recognized holistically, based on their general shape, *not* on the exact internal configuration, something that can be accomplished without much conscious focus (Dehaene 2009). On the other hand, in sublexical processing, the reader can consciously register the exact sequences of individual letters. This explains how in rough, holistic word recognition, readers (especially speed readers) can miss spelling mistakes, which do not become visible until words are scanned more carefully and consciously.

Conscious focus on the form of a new word or expression as it is envisioned in the Noticing Hypothesis is tantamount to sublexical processing, the rather costly, resource-demanding mode of cognitive operation. The question is how often it is realistic to expect a person to mobilize conscious processing to scan new language forms.

3. ISLANDS OF SALIENCE

New words and unusual expressions can be seen as special islands of salience. They catch our eye causing us to shift our attention from the informational content to linguistic form. If such examples were the only language points to learn, there would perhaps be little reason to question Schmidt's claim that the chances of learning depend on noticing.

However, doubts arise when we consider the much less salient textual sea around the islands. What are the odds of readers shifting their conscious attention to unassuming lexical combinations such as *one of the most* ADJ or *as the result of* NP? To be sure, the odds are not an absolute zero. Every now and then, at least some readers do pause to take note of even perfectly regular mundane combinations of words. But it would be quite beyond belief to suppose that people should attentively scan the entire stream of words, one after another, in a piece of text. People read for entertainment, ideas, pleasure, or other purposes found in the informational content of the words they see, and that is where most of their conscious attention is focused; surely Schmidt's *focal awareness* is not constantly allocated to the shapes of words (Truscott, 1998).

Another reason why readers and listeners cannot seriously be expected to attend to all words equally consciously is what Bybee (2003) calls *habituation*. The

more frequently encountered a word is, the less attention it draws to itself because “repetition itself diminishes the force of a word, phrase, or construction” (p. 157). Sequences of words that are already familiar to a person tend to go unnoticed, perceptually taken for granted.

Yet, despite the virtual impossibility of following the surface wording of expressions, people do end up learning sequences like *one of the most* and many other similarly unassuming expressions. For example, when asked which form is more frequent *yes or no* versus *no or yes*, both native and non-native speakers of English, without a moment’s hesitation, identify *yes or no* as the right wording. The example may appear dull and unworthy of discussion but that is precisely what makes it compelling here: How do people learn its form if the combination of its component words is so unlikely to attract attention?

4. SUBLIMINAL LEARNING VS. SUBLIMINAL PERCEPTION

Schmidt does admit the possibility that people may register a stimulus below conscious awareness, but he insists that “this is generally the case only for already established representations, that is, while there is subliminal perception, there is no subliminal learning” (Schmidt, 2012, p. 37). Thus, language learners cannot rely on subliminal detection for acquiring entirely new words or expressions; they may only exploit subliminal detection when exposed to words with already established representations in their minds.

It is interesting to point out that such subliminal perception would be sufficient for the purposes of learning formulaic language. After all, most new formulaic expressions are made up of words that people already know. The only element of novelty is their co-occurrence. Does registering new combinations of familiar parts require conscious attention?

5. FORGING ASSOCIATIONS

The mental mechanism widely regarded as responsible for the perception of co-occurring stimuli is associative learning. Put simply, the brain registers coinciding stimuli, a powerful learning mechanism whose purpose is to detect meaningful links between stimuli, objects or events. A classic example is how food aversions are learned: the shape and color of a new food is instantly associated with its unpleasant taste. The links established in the mind are based on temporal

or spatial proximities of stimuli, a learning mode that is “fundamental to our sense of causality and is the basis of much of our understanding of the external world” (Christian, 2010, p. 242). What is particularly remarkable about associative learning is how ubiquitous it is: it is found across the animal kingdom, including in species (e.g., snails) not known for conscious cognition (Heyes, 2012). In humans, associative learning can be both conscious and unconscious. However, it is the latter that is credited for registering and pairing most incoming stimuli (e.g., Kuldás et al., 2013; Bargh & Marsella, 2008). That is because the amount of information that the brain receives at any given moment exceeds the capacity of conscious processing. The recent reappraisal of unconscious processing is a recurring theme in many publications whose authors converge on the conclusion that “unconscious processing appears to be structurally and functionally much more sophisticated than the conscious” (Kuldás et al., 2013, p. 3).

It is safe to suppose that associative learning is an adequate mechanism sufficient to handle formulaic language. Words that are found to occur in the company of other words can be considered examples of coinciding stimuli. Their co-occurrences must be visible to mechanisms of associative learning, just like any diverse pieces of information perceived at the same time, something that is typically registered unconsciously. If these co-occurrences are then found to be recurrences—that is, for all intents and purposes, tantamount to learning.

It should now be clear why conscious attention cannot be deployed to learn formulaic patterns: Patterns are not visible until additional encounters. The learner does not know which of the lexical combinations are frequent phrases: They are not highlighted in the text to indicate their formulaic status. The only solution is to record everything wholesale, so that some of the previously seen combinations can be recognized as repetitions if they happen to be reencountered in the future. A way of accomplishing just that—recording all lexical combinations—is through unconscious perception of familiar words.

6. EVANESCENT ACTIVATIONS

One reason why Schmidt doubts that a person can learn much subliminally is that subconscious perception triggers little activity in the brain (Schmidt, 2001). He reports on studies showing that while subliminal perception does lead to cognitive activation of previously well-learned information present in long-term memory, he adds that such activation is modest and fleeting, lasting for about a mere tenth of a second. In other words, such short-lived activation is unlikely to result in learning anything new. True, the chances of a brief weak activation of a memory trace

leading to its long-term consolidation are rather slim, especially if it results from a single subliminal encounter. However, it would be controversial to dismiss such activation, however evanescent, as inconsequential or leaving absolutely no trace in memory, especially if such reactivations recur. Surely, repeated activations, even if brief, have a cumulative effect, so much so that after a sufficient number of encounters, there should be considerable learning gains in memory. This much is in fact consistent with what usage-based theories have to say about learning.

That is, one of the main assumptions of usage-based theories is that detecting instances of already established representations is actually a big part of lexical acquisition. Learners repeatedly come across uses of words and expressions they are already familiar with, and each such reencounter is believed to lead to changes in a person's lexical representations. Usage-based models assume that each time a word is seen again, its entry in the mental lexicon is not only strengthened but also updated with information about the usage just witnessed. If this assumption is correct, increases in consolidation should be visible in the learner's enhanced sense of the word's frequency. In fact, language users have been found to keep track of the frequencies of language forms (e.g., Ellis, 2002). This is evident in people's ability to rank words in terms of their frequency (Hasher & Chromiak, 1977). Also, more frequent forms become more entrenched, leading to their increased automaticity both in comprehension and production (Bybee, 2010). All this is possible if people register all instances of forms that are part of their representations.

How much of such registration of reencounters occurs outside consciousness? To be sure, not all of it. It is possible for a language user to register some re-encountered expressions consciously, but it is highly unlikely for conscious attention to be directed at most, let alone all, familiar language forms. The reason this is implausible is that, as we saw earlier, many language forms are not very salient. Lexical sequences like *as the result of* or *one of the most* are among a great number of well-entrenched language forms that most people encounter without focusing their attention on their form. It can be assumed that such expressions are not registered consciously because there is little about their form that should attract or require much attention. If people tried to focus on the form of these and other similarly regular-looking expressions, they would have to scan the form of all lexical combinations. That is, while reading, their conscious attention would have to focus on the form of expressions practically non-stop, something that most people most certainly do not do.

What makes this even more problematic for the Noticing Hypothesis is that such examples abound. The family of irreversible binomials (Malkiel, 1959) includes hundreds of combinations such as *Adam and Eve*, *black and white*, *mom and pop*, *sooner or later*, *more or less*, most of which are too bland in form to

attract attention. In fact, research in recent decades suggests that formulaicity is not only more significant than it was given credit by generative grammarians but it may well be the default mode of language use (Pawley & Syder, 1983; Wray, 2002, Christiansen & Arnon, 2017). Reliance on a diverse range of fixed expressions is part and parcel of everyday language production and the secret behind fluency and naturalness of use.

7. EVIDENCE IN FAVOR OF INCIDENTAL LEARNING

There is in fact mounting experimental evidence in favor of learning as a result of incidental exposure to input. In an ingenious experiment (Bordag et al., 2021), subjects were asked to read a passage containing fairly basic words. Then they were given another passage which looked like a copy of the original one, differing only in the use of several words. The researchers used eye-tracking equipment to observe at which moments their subjects would pause. They found that the readers focused their gaze on the wordings that differed relative to the original passage. That suggests that the readers detected the differences between the passages even though the expressions in question were ordinary combinations otherwise assumed to be unlikely to attract any degree of conscious attention. In a similar study, Gurevich et al. (2010) asked subjects to listen to a story followed by a surprise quiz about the expressions used in the story. The listeners demonstrated high degrees of accuracy of verbatim recall of specific expressions even though they had not been told they would be tested on their memory, and so they can be assumed to have paid little attention to form. For example, in a series of studies, Arnon and Snider (2010) showed that comprehenders are sensitive to the frequencies of compositional four-word phrases (e.g., *don't have to worry*) such that more frequent phrases were processed faster. All that refers to the way language forms are acquired and processed in L1, but similar data are available about L2. Szcześniak (2022) found that foreign learners of English demonstrate a better command of formulaic expressions such as *year passed* as opposed to less frequent combinations (e.g., *person passed*).

8. THE LINGUISTIC FORM IS ONLY THE TIP OF AN ICEBERG

What makes the Noticing Hypothesis particularly problematic is that the linguistic form of expressions is not the only aspect of the information to be extracted from the input. Underneath the surface form of expressions are a myriad

of other details to do with the meaning, usage, emotional tone. To get a sense of the multitude and diversity of information to be handled, consider Taylor's (2012) description of what must happen when a person comes across a word in the input:

each linguistic encounter lays down a trace in memory. The trace pertains not only to the linguistic signal as such, but also to the context in which it is encountered. The context may include the characteristics of the speaker (her accent and voice quality, for example), and features of the situation in which the utterance is encountered, as well as the presumed semantic intent of the speaker. The mental corpus is therefore vastly more rich in detail than any available text collection, in that each element is indexed for its contextual features. (Taylor, 2012, p. 3)

Each such experience involves an element of novelty. A person will keep encountering a word that is already part of their lexicon but inevitably each time something about its use is bound to be new. This is in fact how usage-based theories explain the acquisition of lexical meaning. Namely, the meanings of words and expressions are assumed to be discovered in piecemeal fashion through multiple exposures to uses, each of which reveals some component of the meaning. Language learners discover the meaning by finding which components recur across diverse situations of usage:

Additional encounters with the same word typically overlap in some ways with the earlier representation, strengthening those shared aspects, while also potentially adding contextual information that is unique to that particular experience ... the representation of a word becomes broader as aspects of each context are added. At the same time, those aspects of memory representations that overlap across multiple encounters of a word become strengthened over time, thus becoming more central to a word's meaning. (Goldberg, 2019, p. 16)

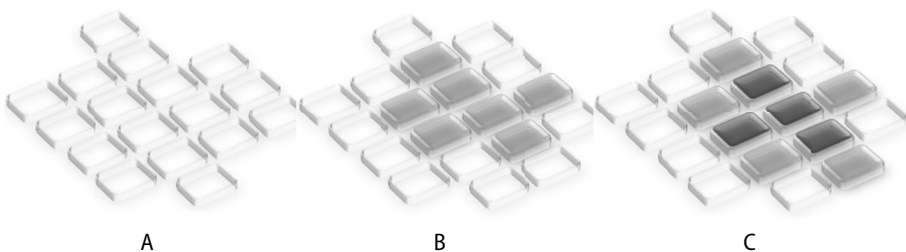


Figure 1
Schematic Visualization of the Increasing Entrenchment of a Memory Trace

Note. A is a schematic visualization of a memory trace of a word encountered in the input; the tiles represent aspects of the word's meaning found in use. B is a strengthened representation of that word after aspects have been found to recur in new contexts (the more recurrences, the darker the tiles, hence stronger representations). C is a representation of the word following more encounters (based on Goldberg, 2019, p. 16).

Thus, a person's mental picture of a word and its meaning is updated with details accumulated from each experience. A good illustration of this is how people arrive at a word's semantic prosody. An ingredient of a word's content, semantic prosody is defined as a "consistent aura of meaning" (Louw, 1993, p. 157) forming around words, absorbed from environments in which these words appear. For example, metaphoric uses of the verb *plunge* share an element of clear and strong negativity:

- (1) ... Pope Innocent III nullified the agreement, and England *plunged into internal war*.
 Every spring, China's cities are *plunged into chaos* ...
 ... when the climate was *plunged into an Ice Age*.
 Lower Manhattan was *plunged into an ominous disorientating darkness*.
 ... the world will be *plunged into a nuclear catastrophe*.
 ... energy solution will appear before our society is *plunged into crisis*.
 China was at its zenith while Europe was *plunged into the Dark Ages*.

The phenomenon of the affective aura around some words was first observed by Sinclair (1991, pp. 74–75). He pointed out that this kind of meaning is not apparent out of context. Indeed, the meaning is not apparent even in context if all a person has to go on is a single instance. It is by being exposed to multiple uses like the above, people realize that *plunge* is not synonymous with a general-purpose verb like *move*. That is, when something (e.g., a country) is plunged into a given state, that state is more specific than a "general situation". This is how an aura of negativity emerges around *plunge*: is observed to recur consistently enough to be taken to be central to the word's meaning.

What is particularly interesting here is that the diverse facts surrounding the use of a word are for the most part stimuli corresponding to "already established representations". When the learner comes across *plunge* followed by *chaos* or *depression*, these collocates tend to be familiar established concepts in the learner's mind, and therefore, according to Schmidt, available for subliminal detection. What may be unfamiliar is a new combination between the familiar concepts. These, however, are perfectly within the reach of unconscious mechanisms of associative learning.

Semantic prosody is only one case of a kind of information gathered about a word or expression. Another example of a pattern that only becomes apparent after multiple exposures is pragmatic content.

9. PRAGMATIC FUNCTIONS: MIGHT AS WELL

Pragmatic aspects of usage tend to attract little attention, but multiple exposures to recurring pragmatic functions help the learner detect a pattern. A typical case is the semi-auxiliary *might as well*, exemplified in (2):¹

- (2) a. With recent countries banning the consumption of them, I figured I *might as well* eat it before it's illegal.
 b. I thought there'll only be one show, there'll never be a repeat, so I *might as well* go for it.
 c. Cops ignore 100s of crimes every day until they get a hair up their ass about something. Usually the thought process is "I haven't done anything in a while, so I *might as well* bust this poor moron."
 d. Then once I'm wearing all the gear, I start to think, well, I *might as well* go out for a little jog. (COCA)

There is a sense that these sentences share an important element of meaning, but whatever that common denominator is, it is elusively hard to define. Just like most other pragmatic functions, it is used with effortless intuitiveness, but speakers are at a loss when asked to name the purpose it serves. (This in itself suggests that the knowledge of this pragmatic function is unconscious.)

Another indication of the subconscious nature of the pragmatic function of *might as well* is that even the definitions given in dictionaries, shown below in (3), do not exactly overlap. In fact, they seem to see the function differently. The definitions listed below come from online editions of *The Oxford English Dictionary*, *The Cambridge English Dictionary*, *Collins Cobuild*, *Macmillan*, *Longman*, and *The Merriam-Webster Dictionary*. They highlight different aspects of the function. These aspects can be thought of as typical details normally found in *might as well* situations. The lack of enthusiasm (a–d), some choice between two alternative plans one of which is better (e–f), etc. are concepts that serve as typical background features in scenarios that trigger the use of *might as well*. One more component not named in these definitions is that the choice being made is a spur-of-the-moment decision made without any delay when the speaker is faced with an alternative.

- (3) a. used to make an unenthusiastic suggestion (*Oxford Languages*)
 b. used to suggest doing something, often when there is nothing better to do (*Cambridge Dictionary*)
 c. If you say that you might as well do something, or that you may as well do it, you mean that you will do it although you do not have a strong desire to do it and may even feel slightly unwilling to do it (*Collins Cobuild*)

¹ From *The Corpus of Contemporary American English, COCA* (Davies, 2008).

- d. used to suggest doing something because you cannot think of anything better to do (*Macmillan*)
- e. used to suggest that someone should do something, because there is no good reason to do anything else (*Longman*)
- f. used to say that something should be done or accepted because it cannot be avoided or because there is no good reason not to do it (*Merriam-Webster*)

Regarding the possibility of learning the pragmatic function from usage, there are two observations to be made here. First, the features ('lack of enthusiasm', 'spontaneous decision', etc.) can be suspended. For example, in (2a–b), one can easily imagine the speaker being enthusiastic, looking forward to eating the food about to be banned or taking the last chance to see a show. It would be perfectly correct and natural to use *might as well* in such a context. What this means is that the learner exposed to such uses needs to encounter many more instances of use to eventually observe the typical features.

The second observation is that the considerable challenge is compounded by the multitude of features to keep track of. Apart from the key features, a use of the phrase can be accompanied by a host of additional details (whose relevance remains to be seen). For example, is the 'last chance' feature part of the meaning of the phrase *might as well*? In truth, it is really impossible to know ahead of time, without being exposed to more instances of use. Of course, if the 'last chance' feature turned out to recur across many uses, it would be wrong to disregard a detail like that. That being the case, the best course of action is to register and retain this and any other detail present in the use of the phrase. Now, some of these details will turn out to be sporadic and therefore likely irrelevant to the meaning of *might as well*. From the point of view of the learner, such sporadic features represent a kind of white noise that would distract a person's attention if he or she attempted to consciously scan all the features present and identify the key components. One solution to this problem is to suppose that the perception and processing of at least some of these features is handled by automatic subliminal mechanisms.

What makes subliminal mechanisms a plausible solution is that the multiple features typical of *might as well* are not novel concepts. The learner is not acquiring any new ideas beyond those that are already established representations in the mind. These are precisely those elements that are, even according to Schmidt himself, accessible to subliminal perception. One could say that no learning happens other than registering what familiar concepts are part of the meaning of *might as well*.

This involves detecting features that recur across uses. That is, the learner records each use, along with at least some of the features of context encountered to

coincide with the form *might as well*. Each such recorded use serves to update the entry for *might as well* the way illustrated in Figure 2, with the recurring features increasingly strengthening the corresponding parts of the representation.

This rather lengthy analysis is predicated on the assumption that the phrase is representative of most expressions with pragmatic content. But the case of this otherwise modest looking phrase suggests that conscious attention would be an inefficient approach to mastering its meaning. Given that language users are faced with the need to learn the details of usage of literally hundreds of thousands (Syder & Pawley, 1993; Jackendoff, 1997) of fixed expressions of various kinds, subconscious processing appears to be the only viable option.

10. CONCLUSIONS

The input carries too much information for conscious attention to handle all at once. First, we saw that many formulaic expressions are too regular and unassuming to draw much attention to themselves. Instead, they are acquired by recording all the words as they are encountered “in full flow” and then having the mental representations of recurring expressions strengthened with each subsequent re-encounter. Indeed, such a wait-and-see approach is an optimal solution given that many formulaic expressions are patterns invisible in a single contact with the input. They do not reveal themselves until later when more attestations have been gathered.

Advocates of the Noticing Hypothesis may counter that people do hold the potential to pay attention to even very commonplace lexical combinations, many more in fact than we may suspect. While this of course may be true and it is *not* impossible for a person to learn phrases like *black and white* by consciously recording the exact sequence of the component words, this would represent only the tip of a much larger iceberg to observe consciously. Underneath the surface of the spellings (or the phonetic shapes) of the words encountered in the input, there are numerous details of meaning that must be recorded in the lexical memory. These details include aspects of usage and context (such as the speaker’s attitude) that may or may not turn out to be part of the meaning. Again, like in the case of orthographic or phonetic shapes of words, the details of meaning become patterns that do not become visible until a number of encounters later. And because the learner does not know ahead of time which details matter, it is necessary to record them all and see which ones recur across attestations. It would simply stretch credulity to imagine anyone consciously keeping pace with all that diverse data. At least some of it has to be processed by unconscious mechanisms.

The skeptical opposition to unconscious operations in the service of learning is actually rather odd. While any processes that lie beyond consciousness may appear inferior to it—as the negative term ‘un-conscious’ suggests—it is in fact those very automatic unconscious mechanisms that represent the bulk of our cognition. As Wyeth (2015, p. 102) puts it, “The majority of the brain’s operations take place outside consciousness, and for information-processing only a millionth of it takes place within it.” Learners need this unconscious cognition to register and organize the profusion of stimuli, as they would overwhelm the brain’s conscious operations.

REFERENCES

- Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, 62(1), 67–82.
- Augustinova, M., & Ludovic, F. (2014). Automaticity of word reading: Evidence from the semantic stroop paradigm. *Current Directions in Psychological Science*, 23(5), 343–348.
- Baars, B. J. (2002). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, 6(1), 47–52.
- Balota, D. A., Law, M. B., & Zevin, J. D. (2000). The attentional control of lexical processing pathways: Reversing the word frequency effect. *Memory & Cognition*, 28(7), 1081–1089.
- Bargh, J. A., & Morsella, E. (2012). The unconscious mind. *Perspectives on Psychological Science*, 3(1), 73–79.
- Besner, D., Risko, E. F., Stolz, J. A., White D., Reynolds, M., O’Malley S., & Robidoux, S. (2016). Varieties of attention: Their roles in visual word identification. *Current Directions in Psychological Science*, 25(3), 162–168.
- Bordag, D., Opitz, A., Polter, M., & Meng, M. (2021). Non-native readers are more sensitive to changes in surface linguistic information than native readers. *Bilingualism: Language and Cognition*, 24(4), 1–13.
- Brown, A. L. (1980). Metacognitive development and reading. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), *Theoretical issues in reading comprehension: Perspectives from cognitive psychology, linguistics, artificial intelligence, and education* (pp. 453–482). Erlbaum.
- Bybee, Joan. (2003). Cognitive processes in grammaticalization. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (Vol. 2, pp. 145–167). Erlbaum.
- Bybee, Joan. (2010). *Language, usage and cognition*. Cambridge University Press.
- Christian, Kimberly M. (2010). Cerebellum: Associative learning. In G. F. Koob, M. Le Moal, & R. F. Thompson (Eds.), *Encyclopedia of behavioral neuroscience* (pp. 242–248). Academic Press.
- Christiansen, M. H., & Arnon, I. (2017). More than words: The role of multiword sequences in language learning and use. *Topics in Cognitive Science*, 9(3), 542–551.
- Croft, W. (1991). *Syntactic categories and grammatical relations*. University of Chicago Press.

- Davies, M. (2008–). *The Corpus of Contemporary American English (COCA)*. Available online at <https://www.english-corpora.org/coca>.
- Dehaene, S. (2009). *Reading in the brain: The science and evolution of a human invention*. Viking Penguin.
- Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188.
- Gass, Susan M. (1997). *Input, interaction, and the second language learner*. Routledge.
- Goldberg, A. (2019). *Explain me this. Creativity, competition, and the partial productivity of constructions*. Princeton University Press.
- Gurevich, O., Johnson, M. A., & Goldberg, A. E. (2010). Incidental verbatim memory for language. *Language and Cognition*, 2(1), 45–78.
- Hasher, L., & Chromiak, W. (1977). The processing of frequency information: An automatic mechanism? *Journal of Verbal Learning and Verbal Behavior*, 16, 173–184.
- Heyes, C. (2012). Simple minds: A qualified defence of associative learning. *Philosophical Transactions of the Royal Society*, 367(1603), 2695–2703.
- Jackendoff, R. (1997). *The Architecture of the Language Faculty*. MIT Press.
- Kuldás, S., Ismail, H. N., Hashim, S., & Abu Bakar, Z. (2013). Unconscious learning processes: Mental integration of verbal and pictorial instructional materials. *Springerplus*, 2(1), 1–14.
- Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. 1). Stanford University Press.
- Louw, B. (1993). Irony in the text or insincerity in the writer: The diagnostic potential of semantic prosodies. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair* (pp. 157–176). Benjamins.
- Malkiel, Y. (1959). Studies in irreversible binomials. *Lingua*, 8, 113–160.
- McTavish, M. (2008). What were you thinking? The use of metacognitive strategy during engagement with reading narrative and informational genres. *Canadian Journal of Education*, 31(2), 405–430.
- Ochs, E., & Schieffelin, B. (1989). Language has a heart. *Text*, 9(1), 7–25.
- Pawley, A., & Syder, F. G. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–225). Longman.
- Schmidt, R. (1990). The role of consciousness in second language acquisition. *Applied Linguistics*, 11(2), 129–158.
- Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 3–32). Cambridge University Press.
- Schmidt, R. (2012). Attention, awareness, and individual differences in language learning. In W. Chan, K. Chin, S. Kumar Bhatt, & I. Walker (Eds.), *Perspectives on individual characteristics and foreign language education* (pp. 27–50). Mouton de Gruyter.
- Szcześniak, K. (2022). Memory of formulaic sequences in L2 retention of chunks featuring inchoative uses of unaccusative verbs. *Linguistica Silesiana*, 43, 43–64.
- Taylor, J. (2012). *The mental corpus. How language is represented in the mind*. Oxford University Press.
- Tomasello, M. (2003). *Constructing a language. A usage-based theory of language acquisition*. Harvard University Press.

- Truscott, J. (1998). Noticing in second language acquisition: A critical review. *Second Language Research*, 14(2), 103–135.
- Wray, A. (2001). *Formulaic language and the lexicon*. Cambridge University Press.
- Wyeth, P. (2015). *The matter of vision: Affective neurobiology & cinema*. Indiana University Press.

THERE IS MORE TO LEARNING WORDS THAN MEETS THE CONSCIOUS EYE

S u m m a r y

This study focuses on the role of attention in learning new language items. These include elements of varying size and schematicity, ranging from single morphemes, through words, phrases, partially-filled sequences, to entirely schematic grammatical patterns. Following cognitive-linguistic usage-based models, it is assumed that language learning takes place by observing examples of use found in the input and whose elements must be memorised. Given the nature of linguistic use and the deluge of information to which the learner is exposed at any moment, learning is highly unlikely to involve only a continuous focus of consciousness. Thus, the present contribution argues against the main claims of Richard Schmidt's Noticing Hypothesis, under which learning is only possible when consciously attending to input elements. On the contrary, the development of proficiency is based mainly on unconscious learning mechanisms. The acquisition of phraseological compounds is particularly dependent on incidental learning.

Keywords: conscious attention; formulaic language; use-based language acquisition

NAUKA SŁOWNICTWA TO WIĘCEJ NIŻ WIDAĆ NA PIERWSZY RZUT OKA

S t r e s z c z e n i e

Niniejsze badanie koncentruje się na roli uwagi w uczeniu się nowych elementów języka. Obejmują one elementy o różnej wielkości i poziomie schematyczności, począwszy od pojedynczych morfemów, poprzez słowa, frazy, częściowo wypełnione sekwencje, aż po całkowicie schematyczne wzorce gramatyczne. Zgodnie z kognitywno-lingwistycznymi modelami bazującymi na użyciu, zakłada się, że nauka języka odbywa się poprzez obserwację przykładów użycia znalezionych w materiale wejściowym i których elementy muszą zostać zapamiętane. Biorąc pod uwagę naturę użycia języka i zalew informacji, na który uczący się jest stale narażony, jest bardzo mało prawdopodobne, aby uczenie się obejmowało jedynie ciągłe skupienie świadomości. W związku z tym niniejsza praca opowiada się przeciwko głównym twierdzeniom Hipotezy dostrzegania autorstwa Richarda Schmidta, zgodnie z którą uczenie się jest możliwe tylko wtedy, gdy świadomie zwraca się uwagę na elementy wejściowe. Wręcz przeciwnie, rozwój biegłości opiera się głównie na nieświadomych mechanizmach uczenia się. Przystawianie związków frazeologicznych w dużej mierze zależy od uczenia się incydentalnego.

Słowa kluczowe: świadoma uwaga; język formulaiczny; przyswajanie języka oparte na użyciu