Memory of Formulaic Language[1]

Konrad Szcześniak
Palacký University, Olomouc; University of Silesia, Sosnowiec
Jana Kořínková
Palacký University, Olomouc

This contribution revisits the issue of differences between children and adults in their respective levels of language attainment. We start by summarizing how the differences have been explained in nativist and empiricist accounts. Our comparison of the two approaches focuses mainly on the role of input in building linguistic representations in lexical memory. We look at memory encoding at the neuronal level: we review our current knowledge of engrams, the networks of neurons participating in the formation of a new memory. The persistence of a memory (whether non-linguistic or lexical) depends on the strength of its underlying engram, and we list a number of factors that contribute to the strengthening of the connections between participating neurons. Apart from a number of factors identified in the extant literature, which we review here, we present one new candidate factor, previously unconsidered in the language acquisition literature. We propose that the strength of a memory trace depends on the learner's belief in the validity of a newly encountered idea. We argue that this factor may account for age effects in language attainment.

Keywords: *input, lexical representations, engrams, memory persistence*

1 Introduction

The glaring disparity between children and adults in their respective language aptitudes may have been evident and obvious to anyone throughout history, but it wasn't until the middle of the twentieth century that the issue was formally addressed by researchers. Why is it that all children succeed in mastering their mother tongue, whereas practically all adults *fail* to attain the native command of a foreign language they attempt to learn? Scholars working on the problem have taken two main tacks.

On the one hand, nativist scholars stressed the importance of innate predispositions. Children have been hypothesized to possess the inborn foreknowledge of some fundamental facts about language, a kind of genetic head-start reducing their workload of acquisition. According to the nativist school of thought, only part of the knowledge of language comes from learning, the other part being preprogrammed, coming from internal processes. A prime assumption of the nativist account is that these internal processes are only active during the critical period (Lenneberg 1967), after which whatever remarkable intuitions children may enjoy atrophy with age, becoming effectively unavailable for adults learning a foreign language.

The other approach, represented mainly by cognitive linguistics, assumes no division between knowledge from the outside (input) and from the inside (innate linguistic predispositions). Cognitive authors insist that all linguistic knowledge is gained through experience observing linguistic input. Under the cognitive linguistic view, input serves as a model of how words and expressions are and should be used (as is reflected in the name "usage-based models" to which cognitive linguists subscribe). To extract linguistic knowledge from

[1] We wish to thank Sabina Dubiel and our three anonymous Reviewers for their insightful comments, which helped us rethink a number of key ideas and improve the quality of the original manuscript.

input, children rely on general cognitive mechanisms (Bybee 2010) such as generalization, schematization, implicit learning—or in other words, the same mechanisms that are available to adult learners. Beyond these general mechanisms, usage-based theories postulate no innate linguistic predispositions necessary for language development. The assumption made by most cognitive linguists is that whatever the differences between first and foreign language acquisition, they do not result from the atrophy of innate domain-specific mechanisms. Thus, the difference between language development in childhood and adulthood may not be as qualitative as Chomsky's nativism originally assumed. In this paper, no distinction will be made between first and foreign language acquisition.

2 Input and Formulaicity

The difference between nativism and the cognitive linguistic empiricism can be said to come down to how we look at input. While it is a crucial element of usage-based models, in nativist models it is neither necessary nor sufficient as a condition of successful acquisition. The nativist dismissal of input rests on two arguments: one is that input is often fragmentary, if not downright ungrammatical, and yet children are able to master their mother tongue despite these obstacles.

The other argument is that input serves an at best marginal purpose. When it comes to mastering syntax, a child only needs minimal amounts of input "as a trigger" to set the parameters of her language. An oft cited example of what a child can discover by observing input is whether or not a subject must be overtly present in a sentence, a choice governed by the "PRO-drop parameter" (Chomsky 1981). A child does need some exposure to input to discover that English requires overt subjects, whereas based on Czech input, a child will discover that the subject can be omitted in a Czech sentence. However, the important point is that such parameters are present in the child's language faculty ahead of time, and all that is left for the child is to discover the correct setting. To do so, the child does not need to keep track of countless examples of sentence subjects to update the parameter; a handful of attestations will do.

Input may play a more important role as a source of lexical knowledge, but Chomsky downplays the lexicon, which he views, following Bloomfield (1933: 274), as "an appendix of the grammar, a list of basic irregularities." These irregularities, which include stock phrases, idioms, proverbs, irregular forms of verbs and nouns, collocational preferences, *etc.*, have come to be referred to as "periphery"[2] – a clear indication of its low-priority status, secondary and largely irrelevant to "core" syntax. Under this view, syntax and the lexicon are two separate realms, the peripheral lexicon having no real effect on the shape of core syntax.

This syntactocentric view of language and the peripheral position of the lexicon were accepted relatively unanimously for some time. However, the last three decades have seen advances in corpus research showing that the amount of idiomaticity and of all "irregular matter" is too large to be swept under the rug of periphery. What was underappreciated by nativist approaches was the formulaic component of language. Proficient speakers make use of great numbers of fixed expressions that go far beyond classic idioms and include phrases like *more*

[2] "phenomena that result from historical accident, dialect mixture, personal idiosyncrasies, and the like" (Chomsky 1995: 20).

or less, *black and white*, *be that as it may*, and the like. These and other expressions have been included under the term formulaic sequence, defined as

> *a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar. (Wray 2002: 9)*[3]

How many such stock phrases are there to learn? It is impossible to know for sure (no dictionary exists listing all of them), but their numbers are estimated to be at least equal to the number of single words (Jackendoff 1997), running in the tens of thousands. The significance of this fact is that these expressions are what makes a speaker proficient. It would be difficult to consider someone a native speaker of a language if that someone didn't know the extant collocations or formulaic expressions (and instead used his or her own variants like ?*white and black* or ?*less or more*).

It is obvious enough that reliance on input is inevitable in the case of idiomatic expressions such as *kick the bucket*, *give no quarter*, or other opaque phrases. However, while not all fixed expressions are as opaque as these examples, perhaps no formulaic phrases are entirely transparent either. Each fixed expression can be shown to express aspects of meaning that cannot be deduced with absolute certainty or confidence based on its form alone, without recourse to input. For example, the meaning of *thanks for nothing* includes associations of slight insult on the part of the speaker disappointed at her request not being answered as expected. At first glance, one may be under the impression that this meaning follows straightforwardly from the presence of the words *thanks* and *nothing*. However, there are other theoretically possible (but wrong) interpretations of the phrase, such as for instance, 'the right thing to say when being forgiven, as in being let off of a speeding ticket,' but this sense is not attempted by proficient speakers. How do people select the correct sense? It would be beyond belief that speakers detect the correct meaning through some robust underlying innate logic. A more plausible explanation is that people have experience coming across uses where the conventional interpretation was intended and manifestly demonstrated. The diverse peculiarities of usage cannot be figured out based on the form *a priori*; they are established *a posteriori*, after the fact of exposure to input.

In a brief summary like the above, there is only room for a handful of examples to illustrate the idiosyncrasy inherent in formulaic language. However, studies after studies conducted by cognitive linguists in recent years (e.g. Goldberg 2006) converge on the conclusion that few, if any, fixed expressions exist without any trace of idiosyncrasy at all. They all seem to harbor something or other about their form, usage, or meaning that is <u>*not*</u> predictable through innate predispositions. These facts about formulaic language do not necessarily put in question the notion of innateness (after all, inborn predispositions may help children extract the right information from input more efficiently), and here we will not argue the merits of innateness one way or another. However, formulaicity does suggest that input is a much more central to language attainment than nativism allowed. Language users rely on

[3] To take another influential definition, Wood (2015: 3) proposes "The general consensus on a definition of formulaic language seems to be that the items will be:
1 Multi word
2 Have a single meaning or function
[3] Be prefabricated or stored and retrieved mentally as if a single word"

input for aspects of use more extensively than merely treat it as a trigger. And the knowledge thus extracted is much more important than was previously appreciated.

Some argue that it is formulaic language, the kind that is learned through exposure to input, that is behind a proficient speaker's command and fine intuitions of usage (Dąbrowska 2009). That is, competent speakers have a sense of naturalness of usage which they owe to their knowledge of fine details such as frequencies of expressions (Ellis 2002; Bybee 2010) or collocational preferences of words (Dąbrowska 2009), none of which could be deduced through innate intuitions, without experience with input. Further, it has been demonstrated that reliance on formulaic phrases enhances a speaker's fluency (Bybee 2010; Wray 2002). That makes it possible to "speak by stitching together swatches of text that we have heard before" Becker (1975: 60). Consider the following sentence (taken from a fictional speech).

(1) The point is, ladies and gentlemen, that greed, for lack of a better word, is good. (Stephen Cherry, *The Dark Side of the Soul*)

Producing an utterance like this would require looking up sixteen words in the mental lexicon and arranging them according to the rules of English syntax. This would impose a considerable cognitive load on the speaker's working memory, likely affecting her fluency. However, it is possible to build this sentence out of the longer chunks *the point is that, ladies and gentlemen*, and *for lack of a better word*. Because each such expression functions as a previously rehearsed unit, its production is automatic and much faster than if it were put together out of individual words. Though not all sentences are made up of as many stock phrases, at least some prefabricated material is present in most sentences.

Last but not least, knowledge of formulaic language is the secret behind native-like usage. The reason why (1) above sounds natural and idiomatic is thanks to phrases like *for lack of a better word* (and not say, ?*due to deficit of a better word*). This account of natural usage is in contrast with the nativist view of the native speaker's intuitions. As mentioned earlier, according to Chomsky, native proficiency and natural-sounding usage were a matter of following the combinatorial rules of syntax. Under this view, a generative syntax allowed speakers to creatively combine words in novel, native-like ways. However, as has become increasingly clear, many theoretically possible word combinations do not sound anything like what native speakers would say:

> *native speakers do not exercise the creative potential of syntactic rules to anything like their full extent, and . . . indeed, if they did so they would not be accepted as exhibiting nativelike control of the language. The fact is that only a small proportion of the total set of grammatical sentences are nativelike in form – in the sense of being readily acceptable to native informants as ordinary, natural forms of expression, in contrast to expressions that are grammatical but are judged to be 'unidiomatic', 'odd', or 'foreignisms'.* (Pawley & Syder 1983:193)

One conclusion emerging from the above is that language proficiency involves much more memorization than nativist authors assumed. Chomsky dismissed memorization, idiosyncrasy and formulaicity as "periphery," and placed creative generative syntax at the core of knowledge, but periphery turns out to be much larger and more significant than was originally thought. As important as core syntax undoubtedly is, people's familiarity with the vast store of the

formulaic and idiosyncratic is at least as impressive and critical to proficiency. In light of this, in accounting for the differences in attainment between children and adult learners, we should bear in mind the formulaic component and the demands it places on memory. This issue will be addressed directly in this contribution.

3 Paradox

However, for all the merits of the cognitive linguistic emphasis on input and usage-based knowledge of language, it is not without problems. Namely, given the usage-based notion that the cognitive mechanisms needed for language are present in children and adults alike, how can the disparity in learning outcomes be explained? Clearly, the explanation should involve at least one strong discriminator, a factor that sets apart the two groups. A possible candidate is motivation. It has been proposed that children are motivated to fit in: they want to be a part of their peer group and, more generally, of society (e.g. Tomasello 1992; Over 2016). This explanation is dubious for several reasons. For one, while social integration and ultimate attainment may be low on many adults' priority lists, it is unlikely that all adults are indifferent to these goals. Indeed, if anything, mindful motivation is where some adults excel, but even then, this does not necessarily correlate with increases in proficiency.

Another possibility is one that was dismissed by Chomsky as banal, uninteresting and ultimately false: the fact that children enjoy a lot more access to input than adult learners (Conklin & Schmitt 2012; Taylor 2012). On the assumption that a solid command of a language depends on the knowledge of large numbers of words and fixed phrases, native speakers hold an outsized advantage over foreign learners in terms of the amount of exposure to new input. A typical native speaker will have come across greater numbers of vocabulary items in diverse uses. Of course, the total numbers of words children hear depends on factors such as their socioeconomic background, and there is considerable variability in this respect. The number of words in caregiver speech may range from 670 to 12,000 words per day (von Tetzchner 2022: 93). Later in adulthood, native speakers continue to be exposed to great amounts of text, both spoken and written. Ellis et al. (2008) estimate that native speakers are exposed to academic English for over 10 years at a rate of 30,000 words per day. By comparison, non-native speakers can be estimated to enjoy input rates of at most 10,000 words per day (p. 390). Needless to say, non-native speakers tend to receive little or no foreign language input in their childhood, so they do not come even close to the minimum of 670 words per day of native speaking children in low socioeconomic status families. What that means is that many lower-frequency expressions that are common knowledge to native speakers may be entirely unfamiliar to most foreign learners.

One problem with the insufficient input solution is that it is hard to find examples of individuals who enjoy access to large amounts of input in adulthood and *succeed* in taking advantage of it. Even those whose professional lives revolve around language are known to admit inadequacies of their foreign language skills. Novelists writing in a foreign language have typical problems such as heavy foreign accents and heavy reliance on native proofreaders, as was the case of Polish-born authors Joseph Conrad and Jerzy Kosinski. Perhaps more interestingly, they admit having insufficient intuitions about whether a given novel usage sounds natural, a sentiment voiced by another Polish author Gustaw Herling Grudzinski, who wrote for Italian magazines. Grudzinski compared writing in Italian to "touching this language through thick gloves, not with your bare skin." This is unusual, since people of letters have the

benefit of deep immersion in the input, exceeding on aggregate what a native teenager will have amassed, but unlike that teenager, they evidently fail to exploit their intensive exposure. Clearly, children must benefit from some advantage other than large amounts of input. There must be something about *how* they approach input that allows them to draw from it more efficiently than do adult learners. If it is correct to assume that proficiency is the product of a knowledge of vast numbers of uses, it then follows that native-speaking children's advantage must consist in successfully committing to memory more than adults learning a foreign language.

4 Engrams

In this connection, it should be noted that up until quite recently, surprisingly little importance was attached to memory retention in theories of language knowledge. For example, although Chomsky did address the issue of memory in *Aspects of The Theory of Syntax*, he was only interested in the limitations of working memory (Chomsky 1965: 14), not in how language forms are stored in long-term memory. His characterization of the Language Acquisition Device made no explicit mention of the neurological processes responsible for memory retention. One reason behind this was likely that, quite simply, back then formal research in the area had only just begun. Although much about memory remains elusive today, our understanding of the neurological processes has progressed enough to inform our discussion of the conditions under which language is successfully acquired.

Central to our focus is what happens in the brain cortex when a memory is formed. At the neuronal level, each and every memory is recorded in the form of an engram, which can be defined as a network of neurons activated whenever a person considers a thought, image, word use, or indeed any mental event. Equivalent to "memory trace", the term engram was introduced over a century ago by German psychologist Richard Semon (1904, 1909), whose pioneering insights into the functioning of memory have largely been corroborated by recent research. For example, studies on engram cells confirm that each single experience leaves a trace in the brain, observable as a specific constellation of neurons dispersed across diverse brain regions, firing together when reactivated (e.g. Tonegawa et al. 2015; Roy et al. 2022). It is these networks of millions of co-activated cells that are the neuronal substrate of mental experiences. Once established, an engram is subject to processes of consolidation and pruning (Kim et al. 2016; Ortega-de San Luis & Ryan 2022; Hong et al. 2016), which determine its strength. The exact mechanisms behind consolidation and pruning are still only partially understood, it is well established that each activation of an engram has the potential to consolidate the synaptic connections between the component neurons. This, in short, can essentially be understood as the formation of a memory. The more often an engram is activated, the more long-lasting it becomes. Although the original activation decays (Rao-Ruiz et al. 2020), once formed, an engram persists in a latent state, available to be reactivated at a later time. This much follows from Semon's definition of an engram as "an *enduring* change which is initially imperceptible"[4] (Semon 1909: 138, our italics). This notion of lasting memory traces is implicit in cognitive linguistic accounts of language knowledge, such as exemplar models, which assume that "memory for linguistic forms is represented in exemplars, which are built

[4] "eine bleibende Veränderung zurückgelassen, die ihrerseits allerdings zunächst nicht wahrnehmbar ist" (Semon 1909: 138).

up from tokens of language experience" (Bybee 2010: 7). The permanence of engrams is also assumed in Taylor's (2012: 3) proposal that "each linguistic encounter lays down a trace in memory". Indeed, various studies show that listeners retain fine-grained phonetic details such as the speaker's accent or voice quality (Goldinger 1996, 1998; Pierrehumbert 2003; Quam & Creel 2021), suggesting that episodic experiences with language uses are stored permanently.

The significance of this fact is that while some traces of forms encountered in the input may appear to have vanished from memory, they nevertheless remain present. It then follows that subsequent encounters with those forms can serve to consolidate the residual traces, which represents a gain compared to having to lay down such traces anew. This can be seen in a familiar effect first pointed out by Ebbinghaus (1885/1890), who noted that learning something entirely new takes longer than relearning it. He demonstrated that by experimenting on himself in a task that involved memorizing syllables. Ebbinghaus found that after having apparently forgotten some syllables, it took him much less time to learn them again than it did at the first attempt. Because relearning previously encountered information can be seen to save time, he called this the *savings effect*.

Not all networks are created equally strong and not all networks are equally likely to persist. Following its formation, an engram is subject to procedures that will either eliminate it (by pruning the connections between participating neurons) or consolidate it (by strengthening the connections within it). While research focusing on the exact mechanisms of elimination and consolidation is still ongoing (e.g. Szcześniak to appear), we know that memory retention is selective in the sense that persistence is conferred on those "promising" engrams that seem to exhibit signs of future utility. Persistence is achieved via Long Term Potentiation (LTP), a process involving enduring changes in synaptic strength of pre-existing connections (Poo et al. 2016: 2). Briefly, the stronger synaptic connections between neurons, the more likely it is that signal transmission will occur between these neurons, so that the activation of one neuron leads to the activation of the associated neurons in the engram. This mechanism is known as the Hebbian (Hebb 1949) rule "neurons wire together if they fire together" (Löwel & Singer 1992: 211).

Assuming that lexical memory is subject to the same (or similar) LTP procedures that determine what is recorded permanently and what is allowed to decay, the task of retaining new language forms involves exploiting factors that increase the chances of consolidation. Some of these factors are presented below.

5 Factors Contributing to Persistence

5.1 *Frequency*

One such factor is increased frequency. If an engram is reactivated after it was formed, it is more likely to become consolidated in line with the Hebbian logic that "coactivated neurons strengthen their connections" (Pulvermüller 1996: 319). Repeated activation contributes to the strengthening of the synaptic connections underlying a given memory: "Thus, if LTP is involved in memory formation, it too should be strengthened through repetition … through repeated exposure to the … stimulus" (Shors & Matzel 1997: 603). Ideally, successful learning depends on multiple exposures to a language form, with estimates of required frequency ranging from a couple to over a dozen encounters (e.g. González-Fernández & Schmitt 2017: 288). However, it would be a mistake to infer that one encounter "is not enough" or that

memory does not retain pieces of information perceived only once. As Bybee (2010) and Goldberg (2019) point out, deleting the trace of a single encounter would make it impossible to keep count of any subsequent exposures. Without some record of the first experience, each repetition would appear to be the first one. Memory evidently retains a provisory trace of even a single experience, an engram persisting in a subdued state. This follows from a rather paradoxical fact. Namely, even though the potentiation involved is subject to extinction ("destruction of what was originally learned"), "there is ample evidence that much of the original learning survives extinction" (Bouton & Woods 2008: 151). In the case of language learning, a newly learned expression may be so faint that it may appear absent to the learner. However, because engram decay may have no endpoint of complete erasure, it is fair to suppose that each lexical trace laid after a single experience is still there, biding its time before the next activation.

The need for enduring traces reinforced by repeat encounters is especially evident in the case of formulaic expressions, especially those of the *black and white* kind, whose form is rather predictable, unassuming and thus unlikely to attract much attention. Most people do not focus their conscious attention on such expressions, so their acquisition must involve weak traces laid down after an incidental exposure, followed by repeat exposures. The reason why people are unlikely to focus consciously on the form of such expressions is that any lexical sequence is a potentially formulaic sequence. The problem is that the learner has no way of knowing which expressions are formulaic until they are found to recur. This means that it is necessary to hold provisory traces for considerable numbers of encountered expressions. Indeed, studies like Gurevich et al. (2010) or Bordag et al. (2021) suggest that people do remember elements of input even without having apparently attended to their surface forms.

5.2 *Emotional Involvement*

Another obvious factor is emotional involvement, evident most clearly in what has been termed flashbulb memories (Brown & Kulik 1977). These are people's detailed recollections of stirring, usually dramatic events witnessed directly. Typical examples are the attacks of 11/9, JFK's assassination, and other events of global significance, but flashbulb memories can also be more positive and personal, such as running into a friend in an unexpected place. Fascinatingly, flashbulb memories are of an associative character, whereby people typically claim remembering not only a specific event in question but also what they themselves were doing at the time, regardless of whether they participated in that event. Thus, what is being remembered is a combination of (at least) two pieces of new information conjoined in one memory record, and the relationship between those pieces can be entirely coincidental and arbitrary. Apart from the event at the center of a flashbulb memory and the activity in which a person was involved, what is also being retained are details like the time of the day, the weather, as well as the persons' mood before and after the event. It should be noted that research following Brown and Kulik's seminal work showed that emotional memories are not always entirely accurate. For example, Neisser (1982) found that when memories are frequently rehearsed, they become more elaborate, and the resulting accounts tend to be inaccurate or inconsistent. This is also evident in what has been termed the Mandela Effect (Broome 2010), people's tendency to form consistently similar false memories of a well-known event. The phenomenon is named after Broome's observation that many people, including herself, believed that Nelson Mandela died in prison in the 1980s, before he became president of South Africa. The phenomenon has been confirmed experimentally (e.g. Prasad & Bainbridge 2022)

and although it is obvious that emotionality does not guarantee perfect accuracy, it nevertheless remains true that strong emotional involvement does correlate with memory persistence.

Here it is worth noting that this associative nature of flashbulb memories is characteristic of lexical memory: lexical entries in memory also comprise at least two logically unrelated pieces of information: the semantic content and the word's form (pronunciation and/or spelling). This lack of logical relation is, after all, the essence of Saussure's (1916/1966) arbitrariness inherent in linguistic signs. In some cases, what is also being recorded is the context in which the lexical item was encountered, along with a host of details including the affective tone, the speaker's mood, the word's effect and the like. The likelihood of such co-occurring details being recorded together, seared as one, depends on the degree of emotionality perceived by the listener / learner. That is because strong emotionality underlies "the rapid stress-induced enhancement of hippocampal LTP. … *hippocampal mechanisms of memory storage are rapidly engaged, rather than suppressed, by an arousing and stressful experience.*" (Diamond et al. 2007: 9, original italics)

5.3 *Personal Relevance*

A closely related factor is personal relevance. A word or expression may stand out by virtue of being "highly significant for personal reasons" (Dąbrowska 2009: 207), directly relevant to the learner by filling an expressive gap in his or her lexicon when encountering "just the thing" needed. Dąbrowska's supposition finds support in studies confirming that memory encoding is enhanced when people process new information in a self-referential fashion. That is when they focus on how it affects them personally or how it makes them feel, that information is then more likely to leave an enduring trace in memory (Macrae et al. 2004). Similarly, people's memory of an event appears to be more resistant to forgetting if they participate in the event directly (Pezdek 2003). This perception of usefulness and personal significance may or may not be a conscious observation, but the point is that words that tie in with people's prior knowledge and experiences are more likely to be accommodated into their engram networks.

What the above three factors have in common is their role in increasing the strength of the synaptic connections between neurons participating in an engram. Synaptic strength is a widely variable parameter, playing a decisive role in what can be imagined as a triage of newly formed engrams. As mentioned earlier, some activations are relatively weak and are thus less likely to persist, being either pruned or allowed to decay. On the other hand, synapses of higher strength levels may be consolidated in the course of replay, the secondary activation of the engram set in motion "in the background" when a person is involved in another unrelated activity, awake or asleep (Squire et al. 2015). Thus, the initial strength of activation is key to memory persistence. Activations are known to be inordinately strong in flashbulb memories. Similarly, activation strength is correlated with personal relevance and it also benefits from frequent encounters. However, these three factors do not explain why children and adults differ in attainment. Children and adults are almost certainly equally capable of experiencing strong emotional states (flashbulb memories), relating to the content of new expressions (personal relevance), and coming across multiple occurrences of a word (the frequency factor). Even though differences occur between children and adults in terms of these three factors, they are mere differences of degree. The factor discussed next is behind a difference of kind: it involves a crucial property that one group has and the other undoubtedly lacks.

5.4 *Belief in Validity*

One last factor instrumental in memory retention is the learner's sense of strong conviction vis-à-vis his or her understanding of a new lexical item. Faced with new information – of any kind, not only linguistic – people normally evaluate it in terms of validity, as they are "motivated to hold correct attitudes." (Petty & Cacioppo 1986: 127) That is because beliefs are consulted in "problem solving, decision making, goals setting as well as in maneuvering in the environment." (Seitz & Angel 2020: 5) This can be understood as a restatement of an otherwise well-known fact of the selective nature of human memory designed to retain only some information while protecting itself against overload with incorrect maladaptive content (Petty & Briñol 2010). Because there is little point in remembering something that one has reasons to doubt, memory "allows in" only that information which appears valid.

In the case of language development, faced with a new lexical item, the learner entertains a hypothesis about what that item may mean. People's hypotheses come with varying degrees of conviction, ranging from practically complete certainty to serious doubt. It stands to reason that if a learner does not find his or her hypothesis very compelling, it is then unlikely to become a permanent addition to lexical memory. In neuronal terms, a lack of conviction decreases the strength of the corresponding synaptic connections. One mechanism responsible for increased memory persistence is that information perceived to be valid is accompanied by affective loading, which involves arousal and greater neural activation (Seitz & Angel 2020). In other words, belief may contribute to greater emotional involvement discussed in Section 5.2.

One recent study (Szcześniak & Sitter 2021) yielded results supporting this prediction. Briefly, in one group of participants, a manipulation was introduced to encourage the subjects' perceptions of their understanding of new lexical items, whereas in the other group, the subjects were given clear reasons to doubt their preliminary hypotheses. In the "strong belief" group, the subjects were asked to describe in their own words what they thought a new expression meant and their definition was followed by a clear emphatic gesture of approval from a native speaker present. In the "serious doubt" group, the subjects' guesses were not endorsed by the native speaker, who offered comments such as "I'm not sure" or "That's not how I'd use it." One week later, the subjects were given a surprise quiz to test their memory of the newly encountered items. As expected, the belief group outperformed the doubt group.

While the study did not focus on children, the results make it possible to speculate why children do not experience the same difficulty as adults when it comes to remembering newly encountered forms. To wit, children are not hampered by the adult-specific impulse to second-guess themselves. Instead, they readily accept as valid their interpretation of the meaning of a new language form. This point requires a brief excursus into how belief has been viewed by philosophers and psychologists. In an uncannily prescient insight, Spinoza (1677/1982) proposed that a newly perceived idea is always automatically believed before it can be questioned. In the words of William James, "*The primitive impulse is to affirm immediately the reality of all that is conceived.*" (James 1890: 319, original italics). Crucially, this first step of "default belief" is a precondition to understanding what one is being told; without first accepting a new fact as true, it would be impossible to interpret what that fact means. Then, after the initial acceptance, one is free to question the idea and reject it (i.e. "unbelieve" it) but this represents the next step, an extra effort over the initial spontaneous belief. Remarkably, Spinoza's claim has been borne out by recent research. It has been shown that under conditions of cognitive depletion (e.g. sleep deprivation or distraction), people may not take the second

step of questioning and instead continue believing otherwise false, even implausible ideas (e.g. Asp & Tranel 2013).

Furthermore, and more relevant to our focus, critical evaluation is not only a mental operation that comes after the initial belief, but it is also a developmentally later operation. That is, the ability to question newly acquired information takes children some time to develop fully. As a result, children tend to accept new knowledge uncritically by default because strong critical thinking skills depend on the prefrontal cortex, which is among the last regions of the brain to mature. Children are known to take a considerable amount of time to learn to question the validity of any new information they encounter:

> *Children are especially credulous, especially gullible, especially prone toward acceptance and belief—as if they accepted as effortlessly as they comprehended but had yet to master the intricacies of doubt.* (Gilbert 1991: 111)

It is important to sound a note of caution here. Gilbert's depiction of children's readiness to trust new information is a strong claim at odds with experimental findings to the contrary. For example, children have been demonstrated to reject information from unreliable informants (e.g. Koenig & Harris 2005), so clearly children are capable of doubt. However, there is reason to suppose that they are more likely than adults to accept new information as true. Some scholars have suggested that children's tendency to believe newly encountered information may be an evolutionary adaptation, allowing them to learn information quickly and efficiently (e.g. Dawkins 1993). This makes sense given that in their early years, children are surrounded by informants that intend to provide accurate information.

Thus, the difference between children and adults may not be as categorical as portrayed by Gilbert; after all, some ability to question new information is clearly present even in infants. Rather, the difference may be one of degree. Because examining a claim represents a cognitive effort, children may not make that effort as spontaneously or as often as adults. The upshot is that children do seem to be more biased toward accepting new information than adults.

Belief appears to be especially relevant when it comes to acquiring formulaic sequences. These are especially challenging because they do not typically lend themselves to conscious focus. While people take advantage of conscious attention (Schmidt 1990) to notice a new language form as information one is not yet familiar with, there are limits to how often this noticing mode can be employed. Learners do occasionally direct their conscious attention to new individual words as a way of increasing the chances of successful learning, but that is not an option in the case of learning formulaic sequences. That is because there are simply too many lexical sequences to pause and focus on each single one. Some of them are loose one-off sequences and some may be formulaic expressions to remember. The problem is that the learner does not know ahead of time which ones are formulaic until they are found to recur (e.g. Szcześniak 2022). Now, in order to encounter a recurring sequence and recognize it as such, the learner must commit it to memory on the first encounter. Keeping track of a great number of word sequences (i.e. retaining relatively persistent traces in memory dedicated to encountered sequences) could not be performed consciously, in Schmidt's noticing mode. Instead, memory persistence is likely to benefit from belief, in two ways. First, learners form a belief regarding their understanding of the meaning behind a sequence of words. Second, perhaps more importantly, it is necessary to perceive a new lexical sequence as a potential lexical unit and entertain a belief that it may be such a unit rather than a loose combination of words. Interestingly, this is where foreign learners are found to be lacking. Adult learners have

been known to analyze formulaic sequences into individual words and not treat them as lexical units the way children do (e.g. Wray 2008: 276)[5].

The above is not to suggest that children are superior to adults when it comes to accuracy. Children's interpretations of newly encountered language forms can of course be incorrect: they can misinterpret or overextend the meanings of new words, as is evident in their use of *daddy* referring to any adult male. As a consequence, mental representations that do not match those of other speakers have to be revised. But that is not as calamitous as it may seem and there is certainly more profit in forming an imperfect unfinished representation than not retaining a trace. That is because people's knowledge is in general developed through trial and error. People's mental representations of meanings are shaped by exposure to diverse uses of expressions, where "each new experience serves to update and extend our previous knowledge by strengthening or adding to existing connections and representations, or by weakening prior connections"[6] (Goldberg 2019: 52).

What matters is not whether children guess the meaning or use, but whether they trust their idea of the new form. If children are not yet in the habit of questioning newly acquired information, it is fair to suppose that their initially accepted hypothesis about the meaning of a new language form is upheld and stored as such. A child's lexical memory benefits from the relative absence of doubt, which is, in the case of adult learners, an almost constant obstacle. Once the skill of critical judgment develops in adolescence (Piaget 1928/2002), the impulse to doubt is always "on standby", available the moment an adult learner takes a guess about the meaning of a new word or expression. While it is possible that children can occasionally revise an idea (following corrective feedback), adults are more likely to reject ideas, which in language study is an unfortunate habit, a signal to memory to not form strong synaptic connections. Then the resulting trace in memory is likely very weak and while it can still be consolidated through frequent encounters, foreign learners may not have sufficient exposure to usage to achieve that.

6 Conclusions

In this contribution, we have presented the question of language attainment as a memory problem. We have been mainly concerned with how memory stores new language forms that a person encounters in the input. We have listed factors that increase the chances of successful encoding, but for reasons of space, we could focus on only some factors identified in recent cognitive linguistic and psychology research. That is because our main aim was not to compile a comprehensive list, but to propose a new candidate factor instrumental in lexical memory storage, that of critical assessment of newly encountered information. The hypothesis being proposed here is that memory retention depends on new information being labeled as true. On the other hand, the moment a person questions the validity of that information, the activation

[5] Given that formulaic sequences are instantiations of more general grammatical patterns (Culicover et al. 2017), their enhanced retention can contribute to a person's command of syntax.

[6] This should also make it possible to revise inaccurate representations resulting from L1 interference. Under Jiang's (2000) L1 lemma mediation hypothesis, L2 representations are based on the learner's L1 semantic structure, which is a simple consequence of insufficient experience with input. That is, when the information contained in the lexical representations is incomplete, it is complemented by reference to the closest L1 equivalent, which may not match the L2 definition. But through continued exposure to input, the L2 representation can be revised and updated to approach the target lexical entry.

strength of the underlying memory engram decreases, ultimately decreasing the chances of its persistence. While the role of belief has been a burgeoning field of research in cognitive science (e.g. see Porot & Mandelbaum 2020 for a review), it has not, to the best of our knowledge, been considered in language research.

This is surprising especially in light of the persistent puzzle of age effects in language attainment. As it turns out, insights from neuroscience offer promising clues as to what may be responsible for the stark differences between children and adults. Before any information is encoded in memory for extended storage, it is filtered for properties like perceived importance and usefulness. Since information perceived to be inaccurate is of little usefulness, it can be assumed to be filtered out. This filter functions differently in children and adults. Children are known to be rather uncritical believers, who are developmentally unable to question or reject new information, including their understanding of the meaning and function of new language forms, and they are therefore more likely to retain them in memory. Conversely, adults tend to question their idea of the meaning of a new word or expression, which is a strong signal for memory to filter it out.

Funding Acknowledgement
This contribution was supported by grant 914106111UVV 2021-*Fond děkanky Pedagogické fakulty Univerzity Palackého v Olomouci.*

References

Asp, Eric & Tranel, Daniel. 2013. False Tagging Theory. In Stuss, Donald T. & Knight, Robert T. (eds.), *Principles of Frontal Lobe Function (Second Edition)*, 383–416. New York: Oxford University Press.

Becker, Joseph D. 1975. The phrasal lexicon. In Webber, Bonnie & Schank, Robert (eds.), *Theoretical issues in natural language processing*, 70–73. Cambridge, MA: Association for Computational Linguistics.

Bloomfield, Leonard. 1933. *Language*. London: Allen & Unwin.

Bordag, Denisa & Opitz, Andreas & Polter, Max & Meng, Michael. 2021. Non-native Readers Are More Sensitive to Changes in Surface Linguistic Information than Native Readers. *Bilingualism: Language and Cognition.* 1–13.

Bouton, Mark E. & Woods, Amanda M. 2008. Extinction: Behavioral Mechanisms and Their Implications. In Byrne, John H. (ed.), *Learning and Memory: A Comprehensive Reference*, 151–171. Amsterdam: Elsevier.

Broome, Fiona. 2010. *Nelson Mandela died in prison? The Mandela Effect*. https://mandelaeffect.com/nelson-mandela-died-in-prison/

Brown, Roger & James Kulik. 1977. Flashbulb Memories. *Cognition* 5(1). 73–99.

Bybee, Joan. 2010. *Language, Usage and Cognition*. Cambridge: Cambridge University Press.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Dordrecht: Foris Publications.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Conklin, Kathy & Schmitt, Norbert. 2012. The Processing of Formulaic Language. *Annual Review of Applied Linguistics* 32. 45–61.

Culicover, Peter W. & Jackendoff, Ray & Jenny Audring. 2017. Multiword Constructions in the Grammar. *Topics in Cognitive Science*: 1–17.

Dąbrowska, Ewa. 2009. Words as Constructions. In Evans, Vyvyan & Pourcel, Sophie (eds.), *New Directions in Cognitive Linguistics*, 201–223. John Benjamins Publishing Company.

Dawkins, Richard. 1993. Viruses of the mind. In Dahlbom, Bo (ed.), *Dennett and his critics: Demystifying mind*, 13–27. Oxford: Blackwell.

Diamond, David M. & Campbell, Adam M. & Park, Collin R. & Halonen, Joshua & Zoladz, Phillip R. 2007. The Temporal Dynamics Model of Emotional Memory Processing: A Synthesis on the Neurobiological Basis of Stress-Induced Amnesia, Flashbulb and Traumatic Memories, and the Yerkes-Dodson Law. *Neural Plasticity* 2007(60803). 1–33.

Ebbinghaus, Hermann. 1885/1890. *Memory: A Contribution to Experimental Psychology* (Transl. by Ruger, Henry A. & Bussenius, Clara E.). New York, NY: Teachers College, Columbia University.

Ellis, Nick. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition* 24. 143–88.

Ellis, Nick, Rita Simpson-Vlach & Maynard, Carson. 2008. Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly* 42. 375–96.

Gilbert, Daniel T. 1991. How Mental Systems Believe. *American Psychologist* 46. 107–119.

Goldberg, Adele. 2006. *Constructions At Work. The Nature of Generalization in Language*. Oxford: Oxford University Press.

Goldberg, Adele. 2019. *Explain Me This. Creativity, Competition, and the Partial Productivity of Constructions*. Princeton, NJ: Princeton University Press.

Goldinger, Stephen D. 1996. Words and Voices: Episodic Traces in Spoken Word Identification and Recognition Memory. *Journal of Experimental Psychology* 22(5). 1166-1183.

Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105. 251–279.

González-Fernández, Beatriz & Schmitt, Norbert. 2017. Vocabulary Acquisition. In Loewen, Shawn & Sato, Masatoshi (eds.), *The Routledge Handbook of Instructed Second Language Acquisition*, 280–298. New York, NY: Routledge.

Gurevich, Olga & Johnson, Matthew A. & Goldberg, Adele E. 2010. Incidental verbatim memory for language. *Language and Cognition* 2(1). 45–78.

Hebb, Donald Olding. 1949. *The Organization of Behavior: A Neuropsychological Theory*. New York: Wiley & Sons.

Hong, Soyon & Dissing-Olesen, Lasse & Stevens, Beth. 2016. New insights on the role of microglia in synaptic pruning in health and disease. *Current Opinion in Neurobiology* 36. 128–134.

Jackendoff, Ray. 1997. *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

James, William. 1890. *The Principles of Psychology*. New York, NY: Henry Holt and Company.

Jiang, Nan. 2000. Lexical Representations and Development in a Second Language. *Applied Linguistics* 21(1). 47–77.

Kim, Ji-Il & Cho, Hye-Yeon & Han, Jin-Hee & Kaang, Bong-Kiun. 2016. Which Neurons Will Be the Engram - Activated Neurons and/or More Excitable Neurons? *Experimental Neurobiology* 25(2). 55–63.

Koenig Melissa A. & Harris, Paul L. 2005. Preschoolers mistrust ignorant and inaccurate speakers. *Child Development* 76(6). 1261–1277.

Lenneberg, Eric Heinz. 1967. *Biological Foundations of Language*. New York, NY: Wiley.

Löwel, Siegrid & Singer, Wolf. 1992. Selection of Intrinsic Horizontal Connections in the Visual Cortex by Correlated Neuronal Activity. *Science* 255(5041). 209–212.

Macrae, C. Neil & Moran, Joseph M. & Heatherton, Todd F. & Banfield, Jane F. & Kelley, William M. 2004. Medial Prefrontal Activity Predicts Memory for Self. *Cerebral Cortex* 14. 647–654.

Neisser, Ulric. (1982). Snapshots or benchmarks? In Neisser, Ulric (ed.), *Memory observed: Remembering in natural contexts*, 43–48. New York: W. H. Freeman Co.

Ortega-de San Luis, Clara & Ryan, Tomás J. 2022. Understanding the physical basis of memory: Molecular mechanisms of the engram. *Journal of Biological Chemistry* 298(5). 1–23.

Over, Harriet. 2016. The Origins of Belonging: Social Motivation in Infants and Young Children. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371. 1–8.

Pawley, Andrew & Syder, Frances Hodgetts. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In Richards, Jack C. & Schmidt, Richard W. (eds.) *Language and communication*, 191–226. New York; Longman.

Petty, Richard E. & Cacioppo, John T. 1986. The Elaboration Likelihood Model of Persuasion. *Advances in Experimental Social Psychology* 19. 123–205.

Petty, Richard E. & Briñol, Pablo. 2010. "Attitude Change." In Finkel, Eli J. & Baumeister, Roy F. (eds.), *Advanced Social Psychology: The State of the Science*, 217–259. Oxford: Oxford University Press.

Pezdek, Kathy. 2003. Event Memory and Autobiographical Memory for the Events of September 11, 2001. *Applied Cognitive Psychology* 17. 1033–1045.

Piaget, Jean. 1928/2002. *Judgement and Reasoning in the Child* (translated by M. Warden). London: Routledge.

Pierrehumbert, Janet B. 2003. Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech* 46(2-3). 115–154.

Poo, Mu-ming & Pignatelli, Michele & Ryan, Tomás J. & Tonegawa, Susumu & Bonhoeffer, Tobias & Martin, Kelsey C. & Rudenko, Andrii & Tsai, Li-Huei & Tsien, Richard W. & Fishell, Gord & Mullins, Caitlin & Gonçalves, J. Tiago & Shtrahman, Matthew & Johnston, Stephen T. & Gage, Fred H. & Dan, Yang & Long, John & Buzsáki, György & Stevens, Charles. 2016. What is memory? The present state of the engram. *BMC Biology*. 1–18.

Porot, Nicolas & Mandelbaum, Eric. 2020. The Science of Belief: A Progress Report. *WIREs Cognitive Science* 12(2). 1–17.

Prasad, Deepasri & Bainbridge, Wilma A. 2022. The Visual Mandela Effect as Evidence for Shared and Specific False Memories Across People. *Psychological Science*. 1–18.

Pulvermüller, Friedemann. 1996. Hebb's concept of cell assemblies and the psychophysiology of word processing. *Psychophysiology* 33(4). 317–333.

Quam, Carolyn & Creel, Sarah C. 2021. Impacts of acoustic-phonetic variability on perceptual development for spoken language: A review. *WIREs Cognitive Science* 12(5). 1–21.

Rao-Ruiz, Priyanka & Visser, Esther & Mitrić, Miodrag & Smit, August B. & van den Oever, Michel C. 2021. A Synaptic Framework for the Persistence of Memory Engrams. *Frontiers in Synaptic Neuroscience* 13. 1–15.

Roy, Dheeraj S. & Park, Young-Gyun & Kim, Minyoung E. & Zhang, Ying & Ogawa, Sachie K. & DiNapoli, Nicholas & Gu, Xinyi & Cho, Jae H. & Choi, Heejin & Kamentsky, Lee & Martin, Jared & Mosto, Olivia & Aida, Tomomi & Chung, Kwanghun & Tonegawa, Susumu. 2022. Brain-wide mapping reveals that engrams for a single memory are distributed across multiple brain regions. *Nature Communications* 13(1). 1–16.

Schmidt, Richard. 1990. The Role of Consciousness in Second Language Learning. *Applied Linguistics* 11. 129–158.

Seitz, Rüdiger J. & Angel, Hans-Ferdinand. 2020. Belief formation – A driving force for brain evolution. *Brain and Cognition* 140. 1–8.

Semon, Richard. 1904. *Die Mneme als erhaltendes Prinzip im Wechsel des organischen Geschehens*. Leipzig: Wilhelm Engelmann. (Transl. *The Mneme.* London: Allen & Unwin Limited).

Semon, Richard. 1909. *Die mnemischen Empfindungen*. Leipzig: Wilhelm Engelmann. (Transl. *Mnemic Philosophy*.) London: Allen & Unwin.

Saussure, Ferdinand De. 1916/1966. *Course in General Linguistics* (Transl. by W. Baskin). New York, NY: McGraw-Hill.

Shors, Tracey J. & Matzel, Louis D. 1997. Long-term potentiation: What's learning got to do with it? *Behavioral and Brain Sciences* 20. 597–655.

Spinoza, Baruch. 1677/1982. *The Ethics and Selected Letters*. Indianapolis, IN: Hackett.

Squire, Larry R. & Genzel, Lisa & Wixted, John T. & Morris, Richard G. 2015. "Memory consolidation." *Cold Spring Harbor Perspectives in Biology* 7(8). 1–21.

Szcześniak, Konrad. 2022. Linguistic contrasts out of conscious control. *Lege artis. Language yesterday, today, tomorrow* 7(2). 192–207.

Szcześniak, Konrad & Sitter, Hanna. 2021. "To The Best of My Memory and Belief. Learning New Language Forms." *Australian Journal of Applied Linguistics* 4(1). 1–17.

Szcześniak, Konrad. To appear. Attention and Memory. In Fuyin, Li Thomas (ed.), *Handbook of Cognitive Semantics*. Leiden: Brill.

Taylor, John R. 2012. *The Mental Corpus. How Language Is Represented In The Mind*. Oxford: Oxford University Press.

Tomasello, Michael. 1992. "The Social Bases of Language Acquisition." *Social Development* 1(1). 67–87.

Tonegawa, Susumu & Liu, Xu & Ramirez, Steve & Redondo, Roger. 2015. Memory Engram Cells Have Come of Age. *Neuron* 87. 918–931.

Von Tetzchner, Stephen. 2022. *Typical and Atypical Child and Adolescent Development 5: Communication and Language Development*. London: Routledge.

Wood, David. 2015. *Fundamentals of formulaic language*. London/New York: Bloomsbury.

Wray, Alison. 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Wray, Alison. 2008. *Formulaic Language: Pushing the Boundaries*. Oxford: Oxford University Press.

Konrad Szcześniak
Jana Kořínková
Pedagogická fakulta Univerzity Palackého
Žižkovo nám. 951
779 00 Olomouc
Czech Republic
konrad.szczesniak@upol.cz
jana.korinkova@upol.cz

In SKASE Journal of Theoretical Linguistics [online]. 2022, vol. 19, no. 2 [cit. 2022-12-12]. Available on web page http://www.skase.sk/Volumes/JTL51/02.pdf. ISSN 1336-782X