UFRJ
PROGRAMA DE PÓS-GRADUAÇÃO
EM LETRAS VERNACULAS

DIADORIM

Revista Diadorim
e-ISSN: 2675-1216
v.25, n.3, e63231, 2023
DOI: 10.35520/diadorim.2023.v25n3a63231

Article

Forgetting without complete erasure: How language knowledge develops by retaining traces of all experiences

Konrad Szczesniak[1,2]

Václav Řeřicha [2]

[1] Institute of Linguistics, University of Silesia, Katowice, Poland
[2] Faculty of Pedagogy, Palacký University Olomouc, Olomouc, Czech Republic
E-mail: konrad.szczesniak@upol.cz
E-mail: vaclav.rericha@upol.cz

Autor correspondente
Konrad Szczesniak
konrad.szczesniak@upol.cz

Recebido: 11/03/2024
Aceito: 04/04/2024

Como citar:
SZCZESNIAK, Konrad; ŘEŘICHA, Václav. Forgetting without complete erasure: How language knowledge develops by retaining traces of all experiences. *Revista Diadorim,* v.25, n.3, e63231, 2023. doi: https://doi.org/10.35520/diadorim.2023.v25n3a63231

Abstract

Under the Cognitive Commitment, cognitive linguists aim to characterize the fundamental principles of human language in a way that is consistent with insights from other fields, such as philosophy, psychology, or neuroscience. While cognitive linguistic research has indeed proceeded in line with what is known about cognition, one area of research has not been consulted very closely. Namely, insights from memory engram research have rarely been considered, despite their obvious relevance to usage-based accounts of language learning and use. One reason behind the neglect is that "the gap between the behaviour of neurons and that of people was deemed too wide to bridge." (Divjak, 2019, p. 104) Engrams will be discussed at some length in this paper, but as a first approximation, they can be thought of as networks of neurons activated while a piece of information is processed in the brain. An engram is the physical substrate underlying that piece of information.

Keywords

Engrams; Lexical Memory; Usage-based Accounts of Language.

Resumo

No âmbito do Compromisso Cognitivo, os linguistas cognitivos têm como objetivo caraterizar os princípios fundamentais da linguagem humana de uma forma que seja consistente com os conhecimentos de outras áreas, como a filosofia, a psicologia ou a neurociência. Embora a investigação linguística cognitiva tenha, de facto, prosseguido de acordo com o que se sabe sobre a cognição, há uma área de investigação que não tem sido consultada com muita atenção. Nomeadamente, os conhecimentos da investigação sobre os engramas de memória raramente têm sido considerados, apesar da sua óbvia relevância para os relatos baseados no uso da aprendizagem e do uso da língua. Uma das razões por detrás desta negligência é que "o fosso entre o comportamento dos neurónios e o das pessoas foi considerado demasiado grande para ser ultrapassado." (Divjak, 2019, p. 104) Os engramas serão discutidos em pormenor neste artigo, mas, numa primeira aproximação, podem ser considerados como redes de neurónios activadas enquanto uma informação é processada no cérebro. Um engrama é o substrato físico subjacente a essa informação.

Palavras-chave

Engramas; Memória Lexical; Relatos de Linguagem Baseados no Uso.

> Everlasting layers of ideas, images, feelings, have fallen upon your brain softly as light. Each succession has seemed to bury all that went before. And yet in reality not one has been extinguished.
> Thomas De Quincey, *Suspiria de Profundis*

A central paradox

We wish to point out a serious unresolved paradox to do with memory. On the one hand, usage-based accounts posit a robust lexical memory system capable of storing fine-grained information. Such a robust system is needed for our linguistic knowledge to develop through multiple individual experiences, each one updating our memory representations (Bybee, 2010, p. 17). Under usage-based perspectives, each experience is hypothesized to leave a memory trace rich in detail (Taylor, 2012, p. 3). Rather obviously, the hypothesis of incremental gains from multiple encounters only makes sense if each resulting trace is persistent.

On the other hand, there is the recognition that the contents of human memory are subject to forgetting. This view is postulated by authors like Goldberg (2019), Dąbrowska (2009), Christiansen and Chater (2018), and Divjak (2019), as a functional feature of memory making it possible to form generalizations. Forgetting is understood in the sense of complete loss of information. Rather obviously, if details of an experience can be forgotten irretrievably, this is a problem for usage-based models.

The solution being proposed here is to redefine what forgetting actually means. We will make a case for doing away with the (otherwise reasonable) notion that forgetting is synonymous with complete erasure. As we will demonstrate, it is possible for information to be eliminated but persist in a dormant form. To understand how this is possible, it will be necessary to review the extant engram literature on the properties of the neuronal tissue.

Engrams

All forms of cognition, including language comprehension and production, involve information processing, which occurs through the simultaneous firing of interconnected neurons. Put simply, when a piece of information (e.g. image, idea, emotion, word) is entertained in the mind, a large group of neurons fires in the brain. One important consequence of such co-activations of neurons is that the synaptic connections between them become stronger, in accordance with the famous Hebbian mechanism of synaptic plasticity (Hebb, 1949), better known in aphoristic form "cells that fire together wire together." After repeated activations, the neurons involved start functioning as a consistent network. This network is the locus of the piece of information whose processing triggered the initial activation. In other words, this network is a memory trace or—as it has been termed in memory research—an engram. Each of the countless memories in the mind is stored as an engram. Each word meaning in the mental lexicon exists in engram form (Crowell, 2004).

The term *engram* was introduced over a century ago by German memory researcher Richard Semon (1921). He hypothesized that sensory stimulation leads to a lasting change in the brain tissue, an effect he compared to engraving or inscribing. Based on this conjecture, he coined *engram* (Greek for 'something inscribed') defined as "enduring though primarily latent modifications in the irritable substance triggered by a stimulus." Semon's hypothesis was beyond the technological capabilities at the time, but research in recent years has largely borne out his original conception (Schacter *et al.*, 1978; Josselyn; Frankland, 2018). Namely, studies converge on the conclusion that specific kinds of information are stored in the form of "constellations of neurons" or "neuronal ensembles" which appear to persist over time.

Latent modifications

Semon described modifications to the neuronal tissue as latent. This characteristic is also manifest today: Despite tremendous strides made in recent years, many details remain elusive. One of the unknowns is how many or which neurons comprise any given engram. It is reasonable to suppose that a single engram consists of vast numbers of cells, likely running in the millions. To identify neurons participating in an engram, researchers assume that such neurons should meet at least two of the following criteria:

a. be active during initial learning;

b. display physical changes;

c. be reactivated during memory retrieval;

d. be required or sufficient for memory retrieval (Guskjolen; Cembrowski, 2023, p. 3209).

Points (a-d) can be thought of as representing the chronology behind the development of an engram, starting with the initial activation (a), strengthening of the synaptic connections between engram neurons (b) leading to the emergence of an established engram available for retrieval (c-d). These stages correspond to what is experienced as the encoding of a memory, its consolidation, retrieval and, in some cases, forgetting.

What we do know is that memory encoding and consolidation depend on multiple factors. For example, whether or not a neuron becomes part of a neuron depends on its initial excitability. During initial learning (a), excitable neurons out-compete less excitable counterparts for allocation into the engram (Mozzachiodi; Byrne, 2010; Josselyn; Frankland, 2018).

Then, as predicted by Semon, activation has consequences for the participating neurons. This corresponds to the physical changes (b), both chemical and morphological. First, because communication between neurons occurs through the release and reception of neurotransmitters in the synaptic spaces, sensitivity levels to these chemical messengers change on the cell surfaces. Even more dramatically, some co-activations cause neurons to change shape. This involves the formation of so-called dendritic spines, which are small extensions on the surface of dendrites in contact with the axon of the next neuron. This step is associated with a piece of information being transferred to long term memory, but it requires protein synthesis—which may or may not take place.

The above is, of course, a very brief description, which can only give a rough idea of the intricacies of the processes set in motion in engram activation. Equally elusive are the exact consequences of engram activation, or what kind of trace it leaves.

Enduring modifications

If not extended, activation of a neuronal ensemble can be followed by a decay period, wherein the excitation falls back to baseline levels, weakening the engram (Gallinaro *et al.*, 2022). Even dendritic spines disappear within two weeks of their formation (Attardo *et al.*, 2015). This can be taken to be tantamount to forgetting. However, there are reasons to believe that forgetting does not mean a complete erasure of the engram.

The key point is that the activation of a neuronal ensemble affects vast numbers of participating neurons and the synaptic connections between them. The complexity and magnitude of an activation are such that the changes cannot be undone. The diverse chemical and structural alterations make it impossible to return all the neurons and connections to their pre-activation state, as if perfectly intact. Even if newly formed dendritic spines are reduced and changes in neurotransmitter sensitivity are reversed, it is certainly not the case that the entire network (of millions of neurons) can assume the exact same configuration it had prior to activation. In fact, new findings suggest that even depotentiation of the synaptic connections (i.e. separation of neurons) does not result in erasure: "Memory might be able to persist in a non-synaptic state." (Guskjolen; Cembrowski 2023, p. 3211) A trace can be held over time by other means.

Ebbinghaus' savings effect

The peculiar nature of engrams described above create paradoxical conditions in which it is possible to forget a piece of information without losing it. Even when a fact is apparently erased from memory, this erasure is not definitive. Some faint trace of it remains latent, available to be revisited, revived and eventually consolidated, should the original stimulus (word, image, etc.) recur.

This paradoxical nature of forgetting was first pointed out by Hermann Ebbinghaus (1885-1890) in his pioneering memory research, in which he experimented on himself. He was interested to determine how long it takes to learn new information and how soon such newly learned knowledge is forgotten. One of the most remarkable observations he made was that once a newly learned fact is forgotten, it then takes less time to relearn that fact than it took him to learn it at the first attempt. He dubbed this phenomenon the *savings effect*, to reflect the savings in time and repetitions necessary to learn something previously seen. In other words, we may be under the impression that a piece of information is gone from memory, but its fragmentary residue is still present in neuronal connections. Indications of such weak residual traces can be found in language users' experiences. For example, when we come across a vaguely familiar word, our ability to recognize it suggests that its memory trace persists, though it is perhaps insufficient for active recall. In fact, the well-known distinction between active and passive vocabulary can be seen as a consequence of gradual (but not definitive) forgetting. Many vocabulary items we know passively may be subserved by engrams that have been activated only sporadically (in some cases only once) and therefore not sufficiently consolidated for active use.

Each experience lays down a trace in memory

Thus, we are lead to the conclusion that forgetting does not negate persistence. This conclusion is counterintuitive for yet another reason. Namely, it is in conflict with the widely held belief that to learn a new language form effectively, a person needs to encounter it multiple times. González-Fernández and Schmitt (2017, p. 288) estimate that ideally, a couple to over a dozen encounters are necessary. Based on this, one may draw the mistaken inference that if one encounter is not enough, then memory does not retain pieces of information perceived only once. However, this reasoning suffers from a logical problem. If multiple encounters are absolutely required for memory retention (if a single encounter is insufficient), then all information would be doomed to be forgotten after the initial experience. That, in turn, would reset the counter and the next experience would appear to be the first one, which would again be insufficient for memory retention. This fact was pointed out by Bybee (2010), Goldberg (2019), and Divjak (2019, p. 151):

> If a single exposure is below the threshold where a counter begins accruing evidence, then the counter of exposures remains set to zero, and logically no experience can accumulate (Divjak, 2019, p. 151).

Clearly, to keep count of experiences with a language form, the mind has to record each single one, starting from the initial encounter, which is *not* erased even though it is below the "sufficient frequency" threshold. The idea of non-erasure is consistent with usage-based accounts, such as Taylor's *mental corpus thesis*, according to which "each linguistic encounter lays down a trace in memory." (Taylor, 2012, p. 3).

Usage-based learning

The significance of the above becomes evident in light of the heavy demands placed on memory by exposure to linguistic input as envisaged by usage-based theories. They require that the mental representations be "*continually* updated and reorganized under the influence of social interactions" (Machado Vieira; Wiedemer, 2020, p. 282, my italics). What that means is that each time we come across a word used in the input, information about the use is recorded in the mind. This way, we store considerable amounts of information associated with that word, information coming from multiple experiences—in many cases, thousands of experiences. What needs to be underscored is that usage-based learning is an operation that involves much more than a mere counter of the times a word is seen or heard. The mind keeps track of not only individual encounters of a word but also a large number of contextual details surrounding its use, as they are instrumental in constructing our idea of the meaning of the word, with all its nuances including the "semantic, discourse, pragmatic/interactional, social and cognitive values" (Machado Vieira; Wiedemer, 2020, p. 268). To give a simple example, the verb *sit through* is followed by phrases such as *a boring speech*, *a sad dud*, etc (Hunston, 2002, p. 61-62). The mental entry of the verb *sit through* includes information along the lines of 'used with (usually) long, unpleasant events.' This description is a pattern that can only emerge out of the memory of multiple experiences, where only such negative events are found with the verb; if neutral and positive events were attested, the mental definition of *sit through* would have to be revised. This is a rule, rather than an exception, for most vocabulary items. Such semantics, rich in considerable detail gleaned from the input, is true of tens of thousands of words and expressions that language speakers learn and use.

One important objective of usage-based theories is to account for how language users build their mental definitions of words, how they master "rich meanings in psychologically plausible ways" (Ungerer; Hartmann, 2023, p. 21). Incidentally, this is where a major controversy arises. On the one hand, it would make intuitive sense to predict usage-based learning should require a robust memory, immune to information loss. If forgetting entailed irretrievable loss of information derived from exposures to the input, it would undo the cumulative effects of such exposures.

On the other hand, some usage-based authors claim that input-based knowledge can in fact be acquired through a fallible memory prone to forgetting. Indeed, forgetting is considered a key mechanism that makes it possible to recognize patterns and arrive at generalizations. Quite simply, a word has a generalized meaning that is found in most of its uses because unique details (found in sporadic situations) are not retained:

> We can describe the representations of events as involving lossy compression, by which we mean simply that not all information is retained. For example, we might have a memory trace of witnessing a kumquat that is abstracted away from the color of the kitchen table upon which it sat, the tiny scratch in its surface, and the length of its stem. (Goldberg, 2019, p. 15).

In the same vein, Divjak (2019, p. 129) notes that "Forgetting is not necessarily bad: while it means loss of information, it also allows us to abstract over peculiarities and assemble general knowledge."

Schema formation

Let us consider more closely what is involved in "assembling general knowledge". An important tenet of cognitive linguistic accounts is the assumption that the mental representations underlying our linguistic knowledge are schematic in nature. In the case of word meanings, these can be seen as *generalized* descriptions of information conveyed by words. For example, the meaning of the word *ball* includes information about its round shape or about its uses for games and sports. But a schematic representation would not contain details such as specific pictorial designs imprinted on some balls. Such details are less essential to the concept's classification and are therefore assumed to be "abstracted away from reality in order to provide a more flexible representation for cognition." (Kranjec, 2018, p. 97).

How is such schematicity achieved in mental representations constructed by including information from individual experiences. The streamlined shape of mental representations is something of a paradox, given that the amount of information accumulated from experience is considerable. As mentioned in the previous section, that information includes contextual details, pragmatic purpose and indeed any aspects observed in uses that a person has been exposed to. The reason the mind records such diverse kinds of information is that this is necessary for pattern recognition. If a given detail is found to recur across many different uses of a word, it becomes a core feature that is central to its definition. On the other hand, incidental or idiosyncratic features are not retained in the schema. In the Goldberg quote above, the information about the color of the table upon which a kumquat was seen does *not* make it into the mental representation. Such one-off characteristics observed in specific instances of any concept are not central to its meaning. It has been customarily assumed that such aspects are treated by the mind as incidental and consequently disregarded in forming the schema.

The following figure (Figure 1) shows how the process of schematization takes place over time. The first exposure to a word contributes a number of features found in the usage context. These features are represented graphically by tiles. On subsequent experiences, some features are found to recur and are as a result strengthened in the mental representation, as reflected by the darker shades of the tiles. Following multiple experiences, some features emerge as the most consolidated parts of the representation. They form the core information content of a word.

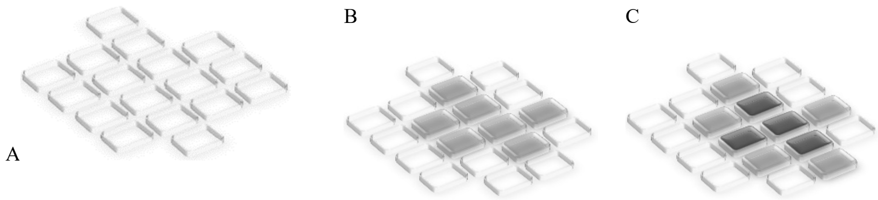

Figure 1 – (**A**) a schematic visualization of a memory trace of a word encountered in the input; the tiles represent aspects of the word's meaning found in use (**B**) strengthened representation of that word after aspects have been found to recur in new contexts (the more recurrences = the darker the tiles = the stronger the representations); (**C**) representation of the word following more encounters (Goldberg, 2019, p. 16).

Note that the mental representation shown in graphic form in the figure corresponds to a large engram. In neuronal terms, the information associated with each tile is subserved by a network of neurons activated each time that information is processed. The more frequently it is processed, the more consolidated it becomes.

Thus, frequency of experience plays a key role in determining whether or not an aspect of use is incidental. If a detail such as the color of the table (in Goldberg's kumquat example) is only observed in one experience, it is unlikely to be a central feature of the word or concept in question. That is a simple consequence of the principle that repetition strengthens memory traces and makes them more resistant to decay. Information that is encountered repeatedly or frequently is more likely to be retained in memory, while less frequently encountered details may be more prone to forgetting.

However, recall from Section 6 above that low frequency cannot be grounds for erasure. After all, nothing is frequent enough the first time it is encountered – at first, each new language form is below the "sufficient frequency" threshold. Goldberg (2019), Bybee (2010) and Divjak (2019, p. 151) all acknowledge that accumulating frequency requires some record of the initial encounter. These authors refer to experiences with words, by which they mean recording the simple fact that a given word was encountered. But note that a memory record of an encounter also entails a memory of all the specific details observed in the word's use.

> 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).

The ability to accumulate evidence should operate at all levels of analysis, including word forms as well as their semantic composition. The fact that the mind can keep count of subsequent experiences suggests that the trace left by the original experience—including the accompanying contextual information—remains, "waiting" to be strengthened by the next encounters. Thus, it makes sense to suppose that even incidental aspects of an experience must be recorded.

One important advantage of retaining incidental features is that if they are found to recur more often, they may become more central to that experience's mental representation. Of course, this may not seem like a common situation. If a person is exposed to a kumquat dozens of times, and only once is it seen on a green table, it is unlikely that from that moment on kumquats should start consistently appearing on green tables. However, some initially incidental features can in fact increase in frequency and be promoted to central positions in the schematic representation.

Promotion of an incidental feature

For example, the expression *pearls of wisdom* is used to refer to astute statements or valuable advice[1].

1. a. *Help Me Live* is beautifully written and offers extraordinary pearls of wisdom on love, on hope, on survivorship.

b. *Falling Upwards* is a spiritual text with pearls of wisdom that transcend religion.

c. Now, on the positive side, the book is full of pearls of wisdom and clever insights into the world of translating. (COCA)

A person encountering examples of use like the sentences above will eventually build a mental picture of the phrase typically used to describe wise observations on profound issues. Given that a person is likely to come across many such uses sharing

[1] Examples marked (COCA) come from the *Corpus of Contemporary American English* (Davis, 2008).

the semantic aspect of 'valuable advice', the resulting mental entry for *pearls of wisdom* will feature this as a core element of meaning. This is also true of the Portuguese equivalent *pérolas de sabedoria*, which is typically used in this reverential sense.

But what if that person encountered the expression in an ironic use designed to ridicule someone's opinion or idea, as in (2)?

2. But then possibly your pearls of wisdom are beyond my comprehension. (COCA)

Because irony is the use of a word in its opposite sense, it can theoretically be applied to any word or expression. It is therefore tempting to suppose that sentences like (2) are one-off, exceptional cases, and *not* to be incorporated into the mental definition of the expression. In fact, a look at the history of the phrase shows that at first it was not used ironically. When it originated in the 18th century, it expressed deference and admiration. Speakers of English were only exposed to such uses and constructed their mental definitions accordingly. Any ironic uses could be dismissed as incidental noise to be forgotten.

However, it is interesting to note that ironic uses have become rather frequent, so much so that some dictionaries mention the ironic use as a separate sense of the phrase. Merriam-Webster lists it as a conventionalized sense that is "often humorous". Longman Dictionary lists it as the main sense with the explanation "used especially when you really think that someone's remarks are slightly stupid." In fact, a glance at corpus data shows that such ironic uses abound:

3. a. Kanye: "It takes a god to recognize another god". Oh Kanye, there is just no shortage of brilliance to fall from your lips. Your pearls of wisdom would rival that of Aristotle and the likes.

b. Spare me the geek squad and their pearls of wisdom.

This is a case of a formerly incidental feature becoming frequent. The inclusion of the ironic sense in dictionaries is a reflection of this formerly incidental feature being promoted to core feature status. That happened because at one point, ironic uses gained momentum. This poses a problem for any theory assuming that infrequent details are forgotten. If there were no remaining mental record of a previously encountered ironic use, each next ironic use would be forgotten too because it would appear like the first one. Without some record of earlier encounters, there is no way for such uses to accumulate in memory.

The upshot is that schematicity is not achieved by forgetting incidental features. While language users do abstract away details found to be irrelevant or at best peripheral, this is not to say that these details are discarded from memory. Instead, it is more plausible to suppose that such low-frequency details are recorded as engram networks, but because of lower frequencies, their levels of activation are also lower and, as a result, those parts of the engram are weaker. That is, they are stored as latent traces, too faint for active recall (so that the person is not aware of them). Still, as faint as they may be, such latent traces represent enough of a foundation to build on when the corresponding stimulus is received.

Frequency Effects

More generally, such a non-erasure modus operandi is necessary for a lexical memory sensitive to frequency effects. Language users have been shown to be able to estimate relative frequencies at which individual verbs appear in different tenses or whether those verbs are more likely to appear in active or passive structures (Ellis, 2002, p. 144). Such knowledge is accumulated by experience with the language. People know that a verb like *repute* is more likely to appear in passive uses because they have come across such uses more often than active uses. This means that unconsciously they keep count of individual experiences with the verb and accumulate evidence of the construction in which it appears, the tense, most frequently co-occurring words. And obviously, all this requires retaining the initial encounters, no matter how infrequent they may be. Seen another way, the knowledge behind the word *repute* is a product of associations between the features observed in multiple experiences. This associative learning gives rise to our lexical knowledge:

> it is very clear that although people do not go around consciously counting features, they nevertheless have very accurate knowledge of the underlying frequency distributions and their central tendencies. We are very good at this association learning—it is the raw basis of human cognition (Ellis, 2002, p. 147).

As stressed above, this applies at all levels of analysis. The associations can be between the word and potentially any types of information (grammatical structure, accompanying words, semantic features). Some of these will turn out to be consistently frequent, and they will become entrenched as a result of repeated activations of the engram networks underlying them.

While it makes sense to imagine our lexical memories keeping track of frequencies of key recurring features, it is important to bear in mind that lexical memory has no way of knowing ahead of time which features encountered in the input will turn out to be frequent. This suggests that our mental representations must be sensitive not only to consistently frequent features but also to those that are observed some of the time. It makes sense for the representations to remain "open minded" about all the encountered features, just in case some of them may be encountered again – when that happens, the initial records are in place for subsequent to accumulate evidence.

This is evident in variable use of constructions. To take one example, the *leave* NP ADJ pattern (*The experience left me speechless*) allows both negatively and positively valued adjectives (*happy, refreshed, depressed,* or *upset*). However, its usage history in the Google Books N-grams corpus (Michel *et al.*, 2011) reveals that the construction has been acquiring an increasingly negative semantic prosody, defined as "a consistent aura of meaning with which a form is imbued by its collocates" (Louw, 1993, p. 157). It has attracted predominantly negative adjectives (e.g. *sad*, *disappointed, depressed*, and *uneasy,* among the most frequent examples.) (Figure 2).

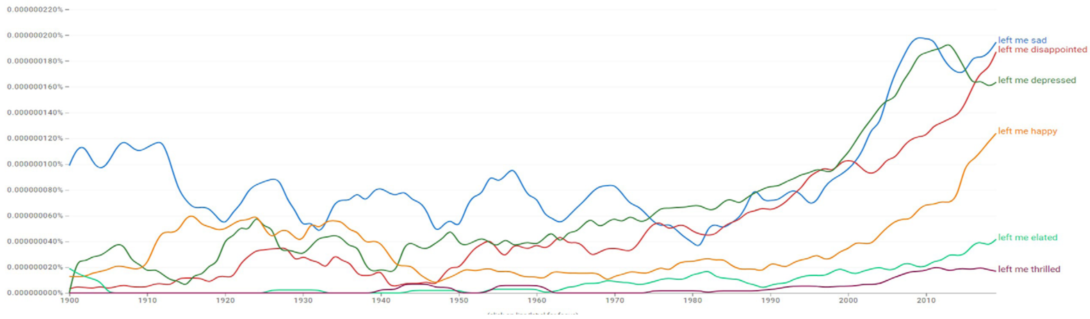

Figure 2. Usage frequencies of the *left me* ADJ sequence per the Google Books N-grams corpus

This tendency is also found in the construction's subvariant *leave* NP *feeling* ADJ (Figure 3).

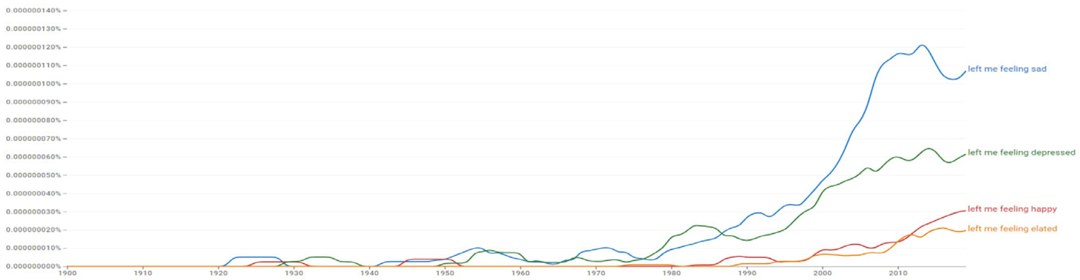

Figure 3. Usage frequencies of the *left me feeling* ADJ sequence per the Google Books N-grams corpus

The negative preference is even more visible when the adjective is modified by an intensifier (e.g. *rather exhausted*). Prosodically neutral (e.g. *very*) or negative (e.g. *rather, utterly*) intensifiers appear only before negative adjectives in the construction, intensifiers with relatively positive prosodies, such as *perfectly* (Partington, 2004), appear before more negative than positive adjectives (Figure 4):

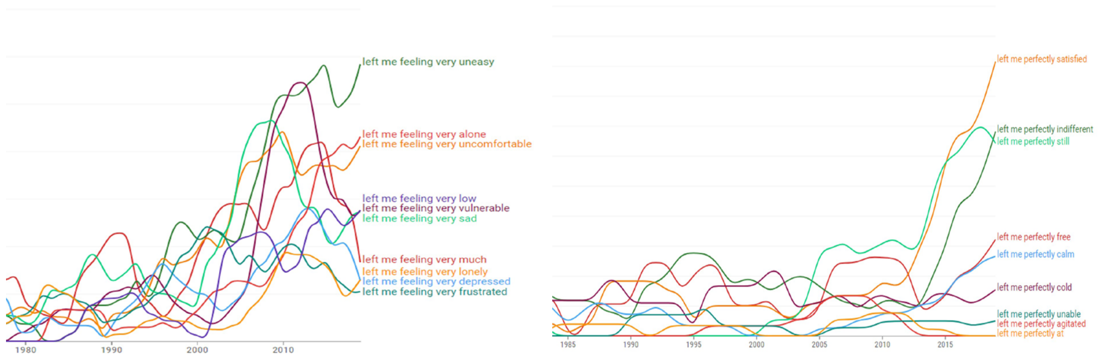

Figure 4. Usage frequencies of the sequences *left me very* ADJ and *left me perfectly* ADJ per the Google Books N-grams corpus

What the frequencies show is that speakers of English have a sense of the construction's relative preference for negative prosody uses. That is, language users know that the construction is predominantly negative but does not rule out positive prosody uses. In terms of core and incidental features, the less frequent positive prosody values represent incidental features. According to Goldberg's lossy memory view, such features, being infrequent, should be treated as irrelevant and forgotten. If that were the case, the usage statistics shown in the figures would include no evidence of variation.

The *leave* NP ADJ pattern is only one example of a construction exhibiting variable use patterns. Variability is a common property typical of most (if not all) constructions. At least some of that variability is a result of mental representations containing less entrenched features—that is, features encountered less often, but still recorded, and certainly *not* forgotten.

Conclusions

One goal of the present review is to make sense of this paradox. On the one hand, it is assumed that information that does not undergo consolidation through frequent repetition is forgotten. On the other hand, forgetting would make it impossible to learn new information by exposure to the input. That is, forgetting rules out frequency-driven entrenchment, one of the cornerstones of usage-based accounts. In this contribution, we argued that forgetting should not be understood in the sense of complete erasure but a weakening of memory traces. This allows for both usage-based learning from every encounter and the ability to generalize and form broader concepts. We contest the idea of "lossy memory" where infrequent aspects encountered in the input are forgotten. Instead, the hypothesis being proposed is that all features are retained, but with varying degrees of entrenchment based on frequency.

References

ATTARDO, A.; FITZGERALD, J. E.; SCHNITZER, M. J. Impermanence of dendritic spines in live adult CA1 hippocampus. *Nature*, v. 523, p. 592-596, 2015.

BYBEE, J. *Language, usage and cognition*. Cambridge: Cambridge University Press, 2010.

CHRISTIANSEN, M. H.; CHATER, N. *Creating language:* Integrating evolution, acquisition, and processing. Cambridge, MA: MIT Press, 2018.

CROWELL, S. E. 2004. The neurobiology of declarative memory. *In:* SCHUMANN, J. *et al.* (ed.). *The neurobiology of learning:* Perspectives from second language acquisition. Mahwah: Lawrence Erlbaum, 2004. p. 75-109.

DĄBROWSKA, E. Words as constructions. *In:* EVANS, V.; POURCEL, S. (ed.). *New directions in cognitive linguistics*. Amsterdam: John Benjamins, 2009. p. 201-223.

DAVIS, M. *Corpus of contemporary American English*. Brigham Young University, 2008.

DIVJAK, D. *Frequency in language:* Memory, attention and learning. Cambridge: Cambridge University Press, 2019.

EBBINGHAUS, H. *Memory:* A contribution to experimental psychology. Translated by Henry A. Ruger and Clara E. Bussenius. New York: Teachers College, Columbia University, 1885-1890.

ELLIS, N. C. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, v. 24, n. 2, p. 143-188, 2002.

GALLINARO, J. V. *et al*. Homeostatic control of synaptic rewiring in recurrent networks induces the formation of stable memory engrams. *PLOS Computational Biology*, v. 18, n. 2, e1009836, 2022.

GOLDBERG, A. E. *Explain me this:* Creativity, competition, and the partial productivity of constructions. Princeton: Princeton University Press, 2019.

GONZÁLEZ-FERNÁNDEZ, B.; SCHMITT, N. 2017. Vocabulary acquisition. *In:* LOEWEN, S.; SATO, M. (ed.). *The routledge handbook of instructed second language acquisition*. New York: Routledge, 2017. p. 280-298.

GUSKJOLEN, A.; CEMBROWSKI, M. S. Engram neurons: Encoding, consolidation, retrieval, and forgetting of memory. *Molecular Psychiatry*, v. 28, p. 3207-3219, 2023.

HEBB, D. *The organization of behavior*. New York: Wiley & Sons, 1949.

HUNSTON, S. *Corpora in applied linguistics*. Cambridge: Cambridge University Press, 2002.

JOSSELYN, S. A.; FRANKLAND, P. W. Memory allocation: Mechanisms and function. *Annual Review of Neuroscience*, v. 41, p. 389-413, 2018.

KRANJEC, A. 2018. Representational biases in space and language. *In:* HUBBARD, T. (ed.). *Spatial biases in perception and cognition*. Cambridge: Cambridge University Press, 2018. p. 94-106.

LOUW, B. 1993. 'Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies', in M. Baker, G. Francis and E. Tognini-Bonelli (eds), *Text and Technology: In Honour of John Sinclair*. Amsterdam: John Benjamins, pp.157–175, 1993.

MACHADO VIEIRA, M. S.; WIEDEMER, M. L. 2020. A variação no modelo construcionista da Linguística Funcional-Cognitiva. *In:* BRESCANCINI, C. R.; MONARETTO, V. N. O. (ed.). *Sociolinguística no Brasil:* Textos selecionados. Porto Alegre: ediPUCRS, 2020. p. 265-304.

MICHEL, J.-B., SHEN, Y. K., AIDEN, A.P., VERES, A., GRAY, M.K., PICKETT, J.P., HOIBERG, D., CLANCY, D., NORVIG, P. ORWANT, J. PINKER, S., NOWAK, M.A. & AIDEN, E.L. Quantitative Analysis of Culture Using Millions of Digitized Books. *Science* 331, 176-182, 2011. http://books.google.com/ngrams

MOZZACHIODI, R.; BYRNE, J. H. More than synaptic plasticity: Role of nonsynaptic plasticity in learning and memory. *Trends in Neuroscience*, v. 33, n. 1, p. 17-26, 2010.

PARTINGTON, A. Utterly content in each other's company: Semantic prosody and semantic preference. *International Journal of Corpus Linguistics*, v. 9, n. 1, p. 131-156, 2004.

SCHACTER, D. L. *et al*. Richard Semon's theory of memory. *Journal of Verbal Learning and Verbal Behavior*, v. 17, p. 721-743, 1978.

SEMON, R. *The Mneme*. Translated by L. Simon. London: George Allen and Unwin, 1921.

TAYLOR, J. R. *The mental corpus:* How language is represented in the mind. Oxford: Oxford University Press, 2012.

UNGERER, T.; HARTMANN, S. *Constructionist approaches:* Past, present, future. Cambridge: Cambridge University Press, 2023.