

**LATAM Revista Latinoamericana de Ciencias Sociales y
Humanidades, Asunción, Paraguay.**

ISSN en línea: 2789-3855, 2025, Volumen VI

Thunk: the cognitive processes behind language change

Thunk: los procesos cognitivos detrás de cambio en el lenguaje

Juan Carlos Rodríguez Burgos

juarodriguez@uqroo.edu.mx

<https://orcid.org/0009-0005-6366-0211>

Universidad Autónoma del Estado de

Quintana Roo

Cozumel - México

DOI: <https://doi.org/10.56712/latam.v6i4.4543>

Artículo recibido: 06 de junio de 2025

Aceptado para publicación: 20 de
septiembre de 2025.

Conflictos de Interés: Ninguno que declarar.


Redilat
Red de Investigadores
Latinoamericanos

NÚMERO

DOI: <https://doi.org/10.56712/latam.v6i4.4543>

Think: the cognitive processes behind language change

Think: los procesos cognitivos detrás de cambio en el lenguaje

Juan Carlos Rodríguez Burgos¹

juarodriguez@uqroo.edu.mx

<https://orcid.org/0009-0005-6366-0211>

Universidad Autónoma del Estado de Quintana Roo

Cozumel – México

Artículo recibido: 06 de junio de 2025. Aceptado para publicación: 20 de septiembre de 2025.

Conflictos de Interés: Ninguno que declarar.

Abstract

This study investigates the cognitive processes underlying the emergence and usage of *think* as a non-standard past participle and, less frequently, past tense form of *think* in present-day English. A corpus-based qualitative approach was adopted, and data were drawn from the Corpus of Historical American English (COHA), the Corpus of Contemporary American English (COCA), and the Oxford English Dictionary (OED). A total of 71 relevant instances were identified, analyzed for diachronic development, frequency, and contextual distribution. Findings reveal that while *think* historically appeared in both past simple and participial contexts, its modern usage is restricted to the *would have + participle* and *woulda + participle* constructions, especially in colloquial expressions such as *Who would have think...?*. As a result, the analysis considers three cognitive mechanisms (categorization, analogy, and automatization) as potential explanations for this shift. Categorization links *think* to phonologically similar strong verbs (e.g., *drink – drunk*), which enables analogical innovation. However, evidence from corpus data suggests that automatization, driven by high-frequency co-occurrence and phonetic reduction in predictable contexts, plays a central role in the entrenchment of *think*. These findings highlight the importance of usage-based processes in language change and suggest that *think*'s persistence is shaped more by routinized discourse patterns than by systematic morphological reanalysis. Future research should explore phonetic realizations in spontaneous speech and constructional frameworks to assess its integration into contemporary English.

Keywords: cognitive linguistics, analogy, automatization, corpus linguistics, language change

Resumen

El siguiente estudio analiza los procesos cognitivos que sustentan la aparición y uso de *think* como una forma no estándar de pasado participio y, en menor frecuencia, de la conjugación en pasado del verbo *think* en el idioma inglés contemporáneo. Se adoptó un enfoque cualitativo basado en corpus lo que derivó en la extracción de información del corpus histórico del inglés americano (COHA), del corpus del inglés americano contemporáneo (COCA), y del diccionario de inglés de Oxford (OED). Se identificó un total de 71 ejemplos relevantes para la investigación, dando pie a su análisis de desarrollo diacrónico, de frecuencia, y de distribución contextual. Dentro de los resultados se puede resaltar que mientras que la forma *think* aparece históricamente en contextos donde se requiere su forma de pasado participio y pasado simple, su uso contemporáneo, por el contrario, se ve confinado a

¹ Autor de correspondencia.

construcciones del tipo *would have + participio* y *woulda + participio*, especialmente en expresiones coloquiales como *Who would have thunk...?*. Consecuentemente, este análisis considera tres mecanismos cognitivos que explicarían este cambio: la categorización, la analogía y la automatización. La categorización vincula fonológicamente al verbo *think* con verbos de frecuencia alta, *drink – drunk*, lo que permitiría una innovación analógica; sin embargo, los datos de los corpus sugieren que la automatización juega un papel importante en el afianzamiento de la forma *thunk*, debido a la coocurrencia de alta frecuencia y la reducción fonética en contextos predecibles. Estos hallazgos subrayan la importancia de procesos basados en el uso dentro del cambio en el lenguaje y sugiere que la persistencia de la forma *thunk* está moldeada por patrones de discurso rutinizado en lugar de un reanálisis morfológico sistemático. Se espera que futuras investigaciones exploren las ejecuciones fonéticas en el habla espontánea y marcos constructivos para evaluar su incorporación en el inglés contemporáneo.

Palabras clave: lingüística cognitiva, analogía, automatización, lingüística de corpus, cambio en el lenguaje

Todo el contenido de LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades, publicado en este sitio está disponibles bajo Licencia Creative Commons. 

Cómo citar: Rodríguez Burgos, J. C. (2025). *Thunk: the cognitive processes behind language change. LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades* 6 (4), 3708 – 3721.
<https://doi.org/10.56712/latam.v6i4.4543>

INTRODUCTION

Languages evolve in unique and often unexpected ways: irregularities can gradually enter a language, with users often unaware that previously standard grammatical or lexical forms are being replaced. But how do these irregularities find their way into the language?

Two major perspectives can be identified in addressing this question: the dictionary-and-grammar view of language (Taylor, 2012) and the constructionist perspective of linguistic knowledge (Goldberg, 2003). The former is concerned with a clear-cut distinction between grammar, a set of specific rules on any given language, and a second element in charge of comprising all irregularities and vocabulary. In contrast, the latter proposes that no distinction is made, rather knowledge of a language emerges through cognitive processes involved in language use. Consequently, generalizations of constructions are recognized subconsciously by language users and these are stored in what is known as the construct-i-con (Hilpert, 2014)

This paper aims to investigate the cognitive processes underlying the use of *thunk*, a non-standard past tense and participle form of *think*, in present-day English. The phenomenon is examined as an instance of language change driven by cognitive mechanisms.

The paper first contextualizes the case within existing literature on frequency effects and cognitive linguistics. It, then, explores the historical emergence of *thunk*, and examines dictionary attestations, such as the Oxford English Dictionary (OED), and corpus data to determine its emergence, frequency and contextual usage. Finally, it discusses the cognitive processes that may explain its development, including categorization, analogy, and automatization. Conclusions are drawn regarding the nature of the change and potential areas for future research.

METHODOLOGY

This study employs a corpus-based qualitative analysis to trace usage patterns of the non-standard verb form *thunk* in present-day and historical English. As a result, two corpora were consulted: the Corpus of Historical American English (COHA) and the Corpus of Contemporary American English (COCA). Each occurrence was manually examined, so a total of 58 relevant instances were extracted from COCA and 13 from COHA. Additionally, the Oxford English Dictionary (OED) was consulted to check whether it had already been established in the language as well as to trace the diachronic development and lexical status. Finally, findings were interpreted through the lens of cognitive linguistic theory, particularly in relation to categorization, analogy, and automatization. No statistical tests were applied, as the analysis was exploratory and interpretive in nature.

DEVELOPMENT

Previous studies

Frequency effects

“(…) Frequency of occurrence is an important determinant of linguistic structure and language use. (...) Frequency has an impact on the comprehension, production, and emergence of linguistic categories and rules (Diessel, 2007, p.109)”. For this reason, corpus-based analyses are necessary to determine whether a linguistic change is influenced by frequency.

Gahl and Garnsey (2004) investigated the frequency and probabilistic effects of /t, d/ deletion in environments that were biased towards specific syntactic structures. The results were categorical: the pronunciation of regular past tense verbs (e.g., *confirmed*, *suggested*, etc.) was affected by the syntactic context. Verbs that typically occurred with a direct object (DO) or a sentential complement

(SC), as shown by corpus data, exhibited greater phonetic reduction in the past tense when followed by a DO rather than an SC. Thus, it can be assumed that frequency of collocational patterns influences phonetic realization.

Hooper (1976) also highlighted how phonological reduction is affected by lexical frequency. In her research, schwa deletion occurred more frequently in common lexical items (e.g., memory, salary, summary and nursery), especially when preceded by a sonorant sound (preferably /r/ or /l/), compared to less frequent items (e.g., mammary, artillery, summery and cursory). Frequency counts provided by Francis and Kucera (1982) confirmed this pattern: for example, every appeared 492 times, while artillery only 11.

Bybee (1985, 1988) further argued that frequency of use influences the way words are stored and accessed. High-frequency items become deeply entrenched in memory, which leads to faster retrieval. Moreover, irregular word forms are maintained in the language as long as they exhibit high token frequency; for example, was/were. In contrast, irregular forms that occur infrequently undergo regularization over time; for example, burnt → burned.

Cognitive processes

Bybee (1995) compares three models of morphological representation: the dual-processing model, the Connectionist model and the Network model. Morphology is, thus, conceptualized as an emergent property of associative networks based on lexical, morphological and semantic connections, where frequency also plays a central role. Bybee notes that “sets of words having similar patterns of semantic and phonological connections reinforce one another and create emergent generalisations describable as schemas” (1995, p. 430). Therefore, when a new or infrequent item is introduced or retrieved, it will be compared and fitted into one of these schemas as long as it aligns with the schema’s properties and strength (i.e., the more items a schema possesses, the stronger it is).

De Smet and Fischer (2016) highlight the importance of analogy in language change as it is linked not only to categorization but also to broader domains such as cognitive science and language acquisition. They highlight the role that frequency plays in analogical extension and how it is connected to categorization, which other authors have already discussed (Bybee & Beckner, 2015). Their analysis of the grammaticalization of have to and as good as illustrates how supporting constructions influence analogical spread.

Bybee (2003) adds automatization as a crucial mechanism in grammaticalization. According to usage-based theory, frequent repetition of structures such as going to lead to phonological reduction and chunking, ultimately resulting in lexicalized units (e.g., gonna). Automatization reduces motor effort and leads to opacity; that is, the original components lose their transparency and are no longer perceived as compositional.

Thesis

Thunk is most frequently used as a past participle and is primarily restricted to the context of would have + participle. Its development is likely the result of cognitive processes, categorization and analogy, which are involved in language use.

RESULTS

Emergence of thunk

To determine whether thunk has been formally established in English, the Oxford English Dictionary (OED) was consulted. It lists four entries for thunk, of which only one is relevant to this study:

“Joc. variant of ‘think n.’ (OED: thunk, n.1)

“A sound of an impact, either dull or plangent. Also int. or as adv. (OED: thunk, n2)

“Dial. and joc. pa. tense and pa. ppl. of ‘think v.’”(OED: thunk, v1.)

“intr. To make a thunk; to fall or land with a thunk” (OED: thunk, v2)

The first, second and fourth definitions were excluded as they do not pertain in the verbal alternation under analysis. The third entry, which concerns the verbal form, is marked with frequency band 3 in the OED. According to the dictionary’s frequency guide, this represents relatively rarity, but not obsolescence.

The earliest attested use, dated to 1876, describes thunk as being more common in the past tense than in the participle form:

Think. (Thuongk). The last form is less employed participially than in the past, in which tense it is of constant occurrence.” (C.C. Robinson, 1876, as cited in Davies, 2010)

This observation challenges the central hypothesis of the present study, which posits a higher frequency of thunk in participial contexts. Unfortunately, the limited number of examples in the OED (only four related to this usage: two in the past tense and two in the participle) is insufficient to confirm or reject the hypothesis conclusively.

If the earliest attested form is considered, a different historical trajectory must be acknowledged, one that suggests thunk underwent semantic or functional shifts over time.

COHA

The Corpus of Historical American English (COHA) provides 135 occurrences of thunk between 1810 and 2000. However, not all instances reflect the verbal form under investigation. A detailed examination was conducted to identify only those cases where thunk functioned as either a past tense or past participle of think:

“(…) I’d be a butterfly born in a bower, Where roses and – ” Fitz. Fitz James Why, who’d have thunk it – so good and yet so bad (…)” (The Lady of the Lake, 1861, as cited in Davies, 2010)

“(…) but she leans up to him purty sharp, I reckons. Dar was one time Ise thunk Marsa Hampton had’ em all fixed; but sence he’s done gone to de Debil of de Forest, Marsa Colbrum’ peers to hev it all cl’ar wedder (…)” (The Phantom of the Forest, 1868, as cited in Davies, 2010)

These early instances confirm that thunk historically appeared in both grammatical roles. Notably, the participial form was not only restricted to the canonical structure would have + participle, but it was also used in past perfect constructions but to a lesser degree:

“(…) I knew when my Reginald staggered into the door-yard that he was on the Die, but if I’d only thunk to ask him about them mules ere his gentle spirit took flight, it would have been four thousand dollars in our pockets, and no mistake! (…)” (Artemus Ward; His Travels, 1865, as cited in Davies, 2010)

Over time, such variation diminished, and thunk has become almost exclusively associated with the would have + participle structure in contemporary usage. This restriction is further supported by the latest examples in COHA, which reflect a narrowing of contexts as opposed to early instances that could be either as a past form or a participle.

There is a time gap when it comes to the penultimate and last usage of *thunk* in the past simple form. However, this might just reflect the potential lack of information that could help track its history:

“(…) ‘He wouldn’t dare to sell them out.’ Well, I *thunk* some more *thunks* just as punky as that, and then we settles it that I’m to hike over and take a squint, anyway. (…)” (Torchy, 1911, as cited in Davies, 2010)

“(…) After you gone to Cambry this morning I *thunk* on it and I know it I bes have another look at things in Fork Stoan (…)” (Riddley Walker, 1980, as cited in Davies, 2010)

By analyzing the occurrences longitudinally, a noteworthy trend arises: COHA data shows a peak in participial usage during the 1880s (10 instances), followed by a marked decline. From 1890 to 1970, the corpus registers only a handful of relevant instances, often just one per decade, with none in the 1940s. The 1980s and 1990s show a modest resurgence, with three and two occurrences respectively, and only one case is found in the 2000s. Moreover, COHA started recording *thunk* as a noun shortly after example number (4). This might have contributed to its less frequent usage in the past simple form as this lexical overlap may have led to some confusion or underrepresentation of the verbal form:

“(…) The howling on the banks of the little stream was now almost deafening, and every second there came the *thunk* of arrows against the boat (…)” (Boy Scouts in Airship, 1912, as cited in Davies, 2010)

In summary, COHA provides useful diachronic evidence which supports the idea that *thunk* was indeed present in both past tense and past participle in the history of the language, but it fails to capture its modern-day usage. To provide more current data, it is necessary to consult a contemporary corpus, which is addressed in the following section.

Frequency of usage: COCA

A final corpus study was conducted to corroborate the usage of *thunk* in the OED. As previously mentioned, the OED classifies its frequency as belonging to band 3, which indicates relative rarity.

The results showed that its frequency is indeed quite low, with only 289 occurrences found. Upon closer examination, it became evident that these instances included not only the verbal form in question but also nominal and onomatopoeic uses. Therefore, an exhaustive analysis was carried out to isolate only the relevant tokens. A total of 58 examples were identified as legitimate verbal instances of *thunk*, which confirms that it is a highly infrequent alternation of thought.

Despite the low frequency, COCA helped determining the contexts in which *thunk* is used. Interestingly, it contradicts the OED’s claim that *thunk* functions as both past tense and past participle. It is worth noting that, in COCA, *thunk* appears to be predominant in the structure would have + participle, and more colloquially, woulda + participle.

Context of *thunk*

The OED describes *thunk* as both a past simple and a participial form; however, it provides insufficient evidence to determine whether it shows preference for one over the other. Nonetheless, the definition serves as a useful foundation for further exploration in trusted corpora. Notably, the participial form appears to be restricted to the context of would have + participle and, even more specifically, to the reduced form woulda + participle, as only such instances are shown in the online dictionary. This challenges the original thesis of this paper, as it suggests that the form may be the result of a different cognitive process not previously considered: automatization (see section 7.3).

Historical corpus data offers a broader view of *thunk*’s contextual usage. The variant appears in five main constructions: 1) would have + participle, 2) past perfect, 3) woulda + participle, 4) past simple, and 5) musta + participle.

COHA suggests that *thunk* preferred the past simple context, with 13 relevant instances in the corpus. However, when all examples which involve the participial form are taken into account, *thunk* shows a strong preference for the participle, particularly in combination with *would have*. Moreover, when comparing the *would have* + participle and *woulda* + participle variants, a surprising pattern arises: *thunk* appears more often in the reduced form context (*woulda*, 10 instances) than in the full form. Interestingly, *would* is the preferred modal, with only one example that involves a different modal verb found in 1981:

(..) So proud, she must a *thunk* she was in Longchamps (...) (Grown Ups, 1981, as cited in Davies, 2010)

This suggests a strong tendency for *thunk* to occur in contexts in which *have* becomes reduced and it is preceded by a modal verb. If this pattern is observed in a contemporary corpus data, its evolution pattern can be assessed.

COCA displays different trends in the usage of *thunk*. First, there were no instances involving modal verbs other than *would*, unlike COHA. This implies that while *thunk* briefly extended to new modal contexts in the past, these innovations were not sustained. Had it been successful, we would expect a higher number of instances with alternative modals such as *must*, *could*, or others.

In contrast to COHA, COCA shows five occurrences of *thunk* in simple past, which contradicts previous findings. This discrepancy might be due to iconicity, since the same form functions as a noun representing a sound. As Wang (2010, p.177) explains:

“An icon image is a single sign that resembles its referent with respect to its characteristics (...). In a word, iconicity as a semiotic notion refers to a natural resemblance or analogy between the form of a sign (...) and the object or concept (...) it refers to in the world or rather in our perception of the world.”

In other words, once *thunk* emerged as a noun describing a sound, the language may have avoided using it in the simple past to reduce ambiguity, thus favoring its use in the participial context. The auxiliary *have* reinforces the link with the base verb *think*, which helps maintain clarity. The overlap between the nominal and verbal forms may have contributed to the decline of the past simple usage. Particularly, COCA does not show *thunk* in other tenses, unlike COHA. As stated previously, languages evolve in peculiar ways, and those changes are motivated by cognitive processes that cannot be controlled.

There is now little doubt that *thunk* most frequently appears in the *would have* + participle structure. However, the question remains: is it more frequent in full *would have* + participle form or in the reduced *woulda* variant? The corpus data yield contradictory results. While COHA shows more examples of the reduced form (10 vs. 2), COCA displays the opposite: 32 instances of *would have* + *thunk* and 18 of *woulda* + *thunk*. This makes it difficult to draw firm conclusions. One explanation is that the reduced form was more common in earlier periods, while the full form gained ground over time. Another possibility, and perhaps the more credible, is that this discrepancy reflects differences in genre and time coverage between the corpora.

Despite this, the overall pattern is clear: *thunk* is typically found alongside *would have*, especially *Who would have thunk...?* As previously proposed, *thunk* likely emerged through cognitive processes that are involved in everyday language use. Nonetheless, the analysis has revealed an additional pathway of change that was not considered at the beginning. One sub-thesis, though, has been supported: *thunk* is more frequent in participial contexts than in the past simple.

Cognitive processes involved in the development of *thunk*

Categorization

Structuralist and generativist views of linguistic knowledge differ from the notion of the construct-i-con in that “material contained in rules does not also appear in the lexicon and vice versa” (Bybee, 2001, p. 20). Therefore, these models assume that a dedicated mental module handles words and idiosyncracies, while another independent module manages grammar. In contrast, cognitive linguistic approaches maintain that there is a close and dynamic relationship between lexicon and grammar, which makes the latter view more plausible and cognitively economical.

Langacker (1987) discusses the example of the -ed suffix: a child must be exposed to a number of instances (e.g., spilled, played, talked) to form generalizations over concrete material. In other words, “our brains record even the predictable pieces of the experience” (Bybee, 2001, p. 20). This reflects a subconscious capacity that enables speakers to recognize similarities across words and phrases. Language users can store new material based on recognized features that were derived from previous language use, whereas completely new features may create new memories or generalizations.

Bybee (2001, p. 21) provides an analogy: “if our memories for dogs excluded all the predictable features (two ears, a muzzle, fur, a tail, wet nose, etc.) what is left would not be a recognizable or coherent entity.” Consequently, speakers’ knowledge of a word or phrase includes all those predictable features that can be cognitively retrieved. When language users are exposed to the verb think, for example, they can associate it with other words that share phonological features (e.g., wink, link; three, theme) or semantic associations (e.g., believe or consider). Additionally, the aim is not to dismiss the concept of regularity, but to acknowledge that categorization allows similarities to surface naturally.

If think is analyzed at a morphological level, language users should be capable of recognizing phonological similarities in think that are comparable with verbs such as sing, drink and ring. Hence, a reinterpretation of its past forms could be prompted, given the partially similar patterns. These similarities can be visualized as follows:

Figure 1

The rise of morphology: identifying morphological similarities between words

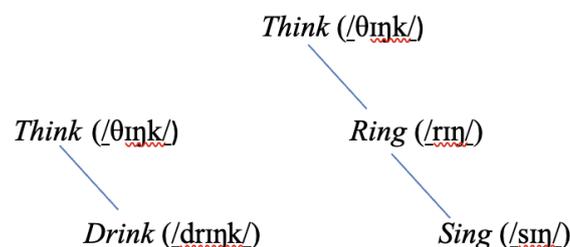


Figure 1 conceptualizes what Bybee (1985) considers to be an improvement over the idea presented in the Phonetic refinement theory (Pisoni, Nusbaum, Luce, & Slowiaczek, 1985) as “ (...) identity relations are represented by connecting lines. (...) These lexical connections can be phonological or semantic (...)” (Bybee, 2001, p. 22). As previously noted, these connections resemble what other morphological theories would describe as prototypes. Hence, there is no need for an independent module to manage irregular morphology; categorized material can be stored as schemas, even if these representations

lead to redundancy, as words that can be easily described by predictable morphemes could also be stored independently.

But why is this process of categorization important? Not only does it provide a cognitively plausible account of morphological structure, but it also offers insight into the dynamics of change. Words that are stored closely together in the mental lexicon may influence one another, which could potentially lead to reanalysis or analogical innovation.

Analogy

McColl Millar & Trask (2015, p. 99) provide a useful example to illustrate the logic behind analogy: imagine being tested on the highly infrequent nouns *ziff*, *zo* and *zax* to write their plural forms. The answer would likely be *ziffs*, *zos* and *zaxes*, respectively. Speakers, thus, are able to produce such forms through analogy; that is, speakers apply familiar patterns, in this case, the regular plural suffix *-s*, to new or unfamiliar items.

Analogy, therefore, is the process in which speakers can create new forms by drawing parallels with a number of existing forms, regardless of being a large or small quantity. As a result, analogy is considered “a powerful mechanism for language change” (McColl et al, 2015, p. 99), especially in morphological developments. Put it in simple terms, a speaker is likely to be familiar with the language pattern A-B, and once they are exposed to a new or an infrequent word in everyday speech, such pattern is readily available because of its high degree of productivity or because the word possesses traits that resemble another.

There are two primary types of analogy: extension and levelling. The former involves applying a known pattern to new or rare items. McColl Millar and Trask (2015) provide an example regarding loanwords: there is a number of nouns in English that were borrowed from Latin, whose plural form follows the *-i* suffix source pattern (e.g., *cactus-cacti*, *radius-radii*, *succubus-succubi*). When *octopus*, a Greek borrowing, entered English, speakers noticed the *-us* ending and erroneously extended the Latin plural pattern, resulting in *octopi*, despite the correct Greek plural, *octopode*.

Analogical levelling, on the other hand, describes a situation in which two or multiple morphological variants of a word coexist, but eventually one form predominates and replaces the others. Haspelmath (2002) illustrates this process through the Old French verb *trouver* and its two stems alternations. When the verb was stressed on the final syllable (as it is with the pronouns *nous* and *vous*), the stem took the form *ou*. When the stem itself was stressed (as with pronouns *je*, *tu*, *il*, *elle*, *ills*, *elles*), it appeared as *eu*. This alternation no longer exists in modern French, so *ou* became the default. Overall, analogical levelling eliminates stem alternations allowing the language to show greater consistency.

Both processes can be illustrated in the form of equations. In the case of *think*, this analogy might look like the following:

Figure 2

Proportional equation representing the formation of think

Drink : Drunk = Think : X

X = Think

As shown, think could have emerged by analogy with strong verbs like drink - drunk. However, it would be misleading to attribute the change solely on one trigger word such as drink. Instead, the change likely stems from shared patterns, word-schemas, across semantically and phonologically similar verbs (Haspelmath, 2002).

Additionally, think does not conform to the full strong-verb- pattern A-B-C (e.g., sing – sang – sung or drink – drank – drunk). There is no evidence for a hypothetical past tense form *think in the corpora examined, only think-think-think had been registered in the OED as previously stated. This likely reflects a cognitive preference for clarity and distinctiveness, as languages prefer to be as transparent as possible; a concept known as iconicity (Haspelmath, 2008). If there were, indeed, examples where *think was used as the simple past form of think, this would create confusion with words already in the lexicon but with complete different meanings (e.g., thank; meaning of gratitude). This is a peculiarity that languages tend to avoid.

McColl Millar and Trask (2015, pp. 100-101) give a phonological example: English lost the /w/ sound after /s/ and before /o/; as a result, sword would receive a silent /w/ sound, but the spelling remained unaffected. The same change was expected with swore and swollen, but it retained the /w/ sound. This may have happened either because their base forms (swear and swell) influenced their pronunciation to preserve morphological transparency, or because the /w/ sound was restored analogically. In any case, the phenomenon reinforces the idea that analogy operates not in isolation but within a network of related patterns.

Automatization

At a subconscious level, speakers tend to respond effortlessly to words or structures that occur frequently in their linguistic environment. As listeners, certain pronunciations might be perceived as 'lazy' or 'sloppy' when speakers reduce or simplify articulation. These patterns are not sign of slackness, but the result of a cognitive process that allows speakers to communicate more efficiently with reduced articulatory effort; a concept known as automatization, also referred to as routinization.

As Hilpert (2014, p. 144) explains, "you are more likely to articulate your words carefully in situations that are formal (...) or when there is ambient noise that impairs auditory comprehension." In other words, speakers tend to adjust their speech to maintain clarity and meet social expectations in formal settings. However, there are other factors, linguistic and contextual, which contribute to speech reduction.

Crucially, frequency and predictability are key linguistic variables that influence this process (Hilpert, 2014, p. 144). He illustrates this with the word and, a highly frequent conjunction that, as a result of its frequency, undergoes phonetic reduction. Moreover, given the predictability of certain collocations, listeners can infer and in its reduced form. For example, when exposed to the phrase 'gin ____ tonic',

most people can effortlessly predict the missing word. Therefore, frequency and predictability are both important factors in phonetic reduction.

Why was it necessary to define the context of *thunk* at the beginning of this study? As shown in the examples from the OED, the original hypothesis fell short: specifically, a case of automatization emerged upon analyzing the recorded instances. As it had been clarified in previous sections, *thunk* is most commonly present in the structure *would have + participle*, replacing the standard form *thought*. Additionally, corpus data indicate that *thunk* frequently appears in the structure *who would have thunk...?*

DISCUSSION

Pathways of change

On the one hand, the emergence of *thunk* can be the result of both categorization and analogy, as initially proposed. The way speakers access and use language in everyday situations allows *think* to be associated with characteristics shared by certain strong verbs. This cognitive alignment may have led *think* to shift from its standard past tense and participle form (*thought*) to resemble the pattern of verbs such as *sing*, *drink*, and *ring*.

Nevertheless, one potential counterargument is that its past tense did not change independently, as the hypothetical pattern *think-*thank-thunk* never materialized. However, this does not invalidate the analogical process. Even though the pattern A-B-C (e.g., *sing – sang – sung*) was not the driving force, it is more likely that the base and participial form of such schema may have influenced the change. Consequently, the language adopted the pattern *think – think – think* as an analogical variation. Another possibility is that the original A-B-B pattern (*think – thought – thought*) was so deeply entrenched that the simple past resisted change, or the analogical pressure was insufficient to alter both forms.

On the other hand, automatization might offer another explanation. In the historical form *thuonk* recorded in the OED (see section 5), *thuonk* had a more frequent use in the past form than as a participial. Given this increased usage, the time required to produce the form decreased, which may have led to its phonological reduction to eventually become routinized. Simultaneously, the reduction of *would have to woulda* could have facilitated the chunking of *woulda + thunk*, especially since these elements frequently co-occur in discourse (*Who woulda thunk...?*) as previously discussed.

This co-occurrence of *would*, *have* and *thunk* in highly predictable contexts has led to the automatization of the full sentence. The fact that this process can be grounded in frequency and cognitive processes, it is assumed that the emergence of *thunk* is through usage-based mechanisms rather than grammatical reanalysis.

Although the original thesis emphasized categorization and analogy, the analysis ultimately suggests that automatization plays a central role in the development of *thunk*. Additionally, further research should focus on phonetics and sociolinguistics of contemporary spoken data to determine how *thunk* behaves in real-time language use.

CONCLUSION

Kiparsky (1995) refers to language change as often being the result of analogical processes as linguistic forms shift from one pattern to another, eventually spreading to a broader set of items. In this view, language change is inherently connected to the principles of cognitive linguistics.

This paper has argued that changes in language, specifically the emergence of *thunk*, can be accounted for through cognitive processes such as categorization, analogy, schema formation, and automatization. The thesis, initially, focused on the influence of categorization and analogy in reshaping the morphological pattern of *think*. However, as the analysis progressed, particularly through corpus evidence and dictionary attestations, it became clear that automatization may offer a more compelling explanation for the evolution of *thunk*.

As speakers use language in daily interactions, we start recognizing patterns that help us store linguistic material that is either new or already in existence. The form *thunk* may have originated through analogy with other strong verbs, but its integration into everyday speech appears to have been strongly influenced by chunking, especially in constructions such as *who would have thought...?*, resulting in the form *who woulda thunk?*. This suggests that high-frequency co-occurrence and phonetic reduction played key roles in the form's entrenchment.

An area for further research would be to apply Hilpert's (2014, pp. 14-22) framework for identifying constructions, to determine whether *thunk* qualifies as part of a larger constructional schema. Additionally, it would be worth analyzing speaker pronunciation in spontaneous speech, particularly of the phrase *Who would have thunk...?*, as it could reveal how the segment would have contributed to the routinization and phonetic shaping of *thunk* in discourse.

Regardless of the precise pathway, whether analogy, automatization, or a combination of both, this paper has sought to demonstrate that *thunk* can be plausibly explained through the lens of cognitive processes involved in language use. In this respect, language change should be examined not only synchronically, in terms of current use, but also diachronically, tracing the historical and cognitive developments that shape linguistic evolution.

REFERENCES

- Bybee, J. (1985). *Morphology: A study of the relation between meaning and form*. John Benjamins.
- Bybee, J. (1988). Morphology as lexical organization. In M. Hammond & M. Noonan (Eds.), *Theoretical morphology: approaches in modern linguistics* (pp. 119-141). Academic Press.
- Bybee, J. (1995). Regular morphology and the lexicon. *Language and cognitive processes*, 10(5), 425-455.
- Bybee, J. (2001). *Phonology and Language Use*. Cambridge University Press.
- Bybee, J. (2003). Cognitive processes in gramaticalization. In M. Tomasello (Ed.), *The New Psychology of Language* (Vol. 2, pp. 145-167). Lawrence Earlbaum Associates.
- Bybee, J. (2012). Domain-general processes as the basis for grammar. In M. Tallerman & K. R. Gibson (Eds.), *The Oxford Handbook of Language Evolution* (pp. 528-536). Oxford University Press.
- Bybee, J., & Beckner, C. (2015). Language use, cognitive processes and linguistic change. In C. Bowers & B. Evans (Eds.), *The Routledge Handbook of Historical Linguistics* (pp. 503-518). Routledge.
- Davies, M. (2008-). The Corpus of Contemporary American English (COCA): 520 million words, 1990-present. Retrieved from <http://corpus.byu.edu/coca/>
- Davies, M. (2010-) The Corpus of Historical American English (COHA): 400 million words, 1810-2009. Retrieved from <http://corpus.byu.edu/coha/>
- De Smet, H., & Fischer, O. (2016). The role of analogy in language change: supporting constructions. In M. Hundt, S. Mollin, & S. Pfenninger (Eds.), *The changing English language: Psycholinguistic perspectives* (pp. 240-268). Cambridge University Press.
- Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New ideas in psychology*, 25(2), 108-127. <http://www.personal.uni-jena.de/~x4diho/Frequency%20effects.pdf>
- Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage: Lexicon and grammar*. Houghton Mifflin.
- Gahl, S., and Garnsey, S. M. (2004). Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language*, 80(4), 748-775.
- Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, 7(5), 219-224.
- Haspelmath, M. (2002). *Understanding Morphology*. Arnold.
- Haspelmath, M. (2008). Frequency vs. Iconicity in explaining grammatical asymmetries. *Cognitive Linguistics*, 19(1), 1-33.
- Hilpert, M. (2014). *Construction Grammar and its Application to English*. Edinburgh University Press.
- Hooper, J. B. (1976). Word frequency in lexical diffusion and the source of morphophonological change. In W. Christie (Ed.), *Current progress in historical linguistics*, (pp. 96-105). North-Holland.

Kiparsky, P. (1995). Explanation in Phonology. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 640-670). Blackwell.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar: Theoretical Prerequisites* (Vol. 1). Stanford University Press.

Lindblom, B., MacNeilage, P., & Studdert-Kennedy, M. (1984). Self-organizing processes and the explanation of language universals. In G. Butterworth, B. Comrie, & Ö. Dahl (Eds.), *Explanations for language universals* (pp. 181-203). Walter de Gruyter.

McColl Millar, R., & Trask, L. (2015). *Trask's Historical Linguistics* (3rd ed.). Routledge.

Oxford University Press. (n.d.). Key to frequency: calculating frequency. Oxford English Dictionary Online. <http://public.oed.com/how-to-use-the-oed/key-to-frequency/>

Oxford University Press. (n.d.) Thunk. Oxford English Dictionary Online. http://www.oed.com/search?searchType=dictionary&q=think&_searchBtn=Search

Pisoni, D. B., Nusbaum, H. C., Luce, P. A., & Slowiaczek, L. M. (1985). Speech Perception, Word Recognition and the Structure of the Lexicon. *Speech Communication*, 4(1-3), 75-95. [https://doi.org/10.1016/0167-6393\(85\)90037-8](https://doi.org/10.1016/0167-6393(85)90037-8)

Saussure, F. (1986). *Course in General Linguistics* (C. Bally & A. Sechehaye, Eds.; R. Harris, Trans.). Open Court. (Original work published 1916).

Taylor, J. R. (2012). *The Mental Corpus: How Language is Represented in the Mind*. Oxford University Press.

Wang, X. (2010). Language Iconicity from Sociolinguistic Perspective. *Asian Social Science*, 6(7), 176-179.

Todo el contenido de **LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades**, publicados en este sitio está disponibles bajo Licencia [Creative Commons](#) .