COMENIUS UNIVERSITY IN BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS, AND INFORMATICS

EXAMINING COLOMBIAN SPANISH LANGUAGE SPEAKERS' PROCESSING OF MORPHOLOGICALLY COMPLEX WORDS

Diploma Thesis

Martha Juliana Aponte Niño

Bratislava, 2023

COMENIUS UNIVERSITY IN BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS, AND INFORMATICS

EXAMINING COLOMBIAN SPANISH LANGUAGE SPEAKERS' PROCESSING OF MORPHOLOGICALLY COMPLEX WORDS

Diploma Thesis

Study programme:	Cognitive Science
Field of study:	Computer Science
Supervising Department:	Department of Applied Informatics
Supervisor:	Bojana Ristić, PhD.

Bratislava, 2023 Martha Juliana Aponte Niño

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Martha Juliana Aponte Niño
Study programme:	Cognitive Science (Single degree study, master II. deg., full time form)
Field of Study:	Computer Science
Type of Thesis:	Diploma Thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: Examining Colombian Spanish language speakers' processing of morphologically complex words

Annotation: We will examine how Colombian Spanish language speakers process morphologically complex words with derivational affixes, and whether different processing stages are detected, as previously shown for different languages. Using a lexical decision task and the pseudoword paradigm, we will add new evidence from a yet uninvestigated language - Colombian Spanish. We will also investigate suffix productivity (measured through frequency), and the category of the stem that the suffix attaches to (noun vs. verb). Finally, a comparison will be made between these results and those obtained in other languages.

Aim:
1. Distinguish post-decomposition stages in processing complex (suffixed) words in Colombian Spanish by using online psycholinguistics methods such as lexical decision task to tap into the language processing.
2. Compare results with the results from other languages (English, Greek, Slovenian) and see if the pattern holds cross-linguistically
3. Add new insights to the existing research line by looking at the differences between affixes that attach to the nouns and affixes that attach to verbs, and looking at the affixes of different productivities and frequencies.

Literature:
Manouilidou, C., & Stockall, L. (2014). Teasing apart syntactic category vs. argument structure information in deverbal word formation: A comparative psycholinguistic study. Italian Journal of Linguistics, 26(2), 71-98.

Supervisor:	Bojana Ristic, PhD.
Consultant:	doc. RNDr. Martin Takáč, PhD.
Department:	FMFI.KAI - Department of Applied Informatics
Head of department:	doc. RNDr. Tatiana Jajcayová, PhD.
Assigned:	23.02.2023
Approved:	23.02.2023

prof. Ing. Igor Farkaš, Dr.
Guarantor of Study Programme

.....
Student

.....
Supervisor

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Martha Juliana Aponte Niño
Študijný program: kognitívna veda (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: informatika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Examining Colombian Spanish language speakers' processing of morphologically complex words
Skúmanie spracovania morfologicky zložitých slov v kolumbijskej španielčine

Anotácia: Budeme skúmať, ako hovoriaci kolumbijskou španielčinou spracúvajú morfologicky zložité slová s derivačnými afixmi a či sa zistili rôzne fázy spracovania, ako sa už predtým ukázalo pri rôznych jazykoch. Pomocou lexikálnej rozhodovacej úlohy a paradigmy pseudoslov pridáme nové dôkazy z doteraz neskúmaného jazyka - kolumbijskej španielčiny. Budeme tiež skúmať produktivitu sufixov (meranú prostredníctvom frekvencie) a kategóriu kmeňa, ku ktorej sa sufix pripája (podstatné meno vs. sloveso).

Cieľ: 1. Rozlíšiť postdekompozičné štádiá pri spracovaní komplexných (príponových) slov v kolumbijskej španielčine pomocou online psycholingvistických metód, ako je lexikálna rozhodovacia úloha, ktorá umožňuje preniknúť do spracovania jazyka.
2. Porovnať výsledky s výsledkami z iných jazykov (angličtina, gréčtina, slovinčina) a zistite, či tento vzorec platí naprieč jazykmi.
3. Skúmať rozdiely medzi afixmi, ktoré sa pripájajú k podstatným menám, a afixmi, ktoré sa pripájajú k slovesám, ako aj afixmi s rôznou produktivitou a frekvenciou.

Literatúra: Manouilidou, C., & Stockall, L. (2014). Teasing apart syntactic category vs. argument structure information in deverbal word formation: A comparative psycholinguistic study. Italian Journal of Linguistics, 26(2), 71-98.

Vedúci: Bojana Ristic, PhD.
Konzultant: doc. RNDr. Martin Takáč, PhD.
Katedra: FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry: doc. RNDr. Tatiana Jajcayová, PhD.

Dátum zadania: 23.02.2023

Dátum schválenia: 23.02.2023 prof. Ing. Igor Farkaš, Dr.
garant študijného programu

.....
študent

.....
vedúci práce

Declaration: I hereby declare that this thesis is written by myself, and investigated by myself, with help of relevant literature that had been properly cited and quoted where needed.

Bratislava, 2023

…….

Bc. Martha Juliana Aponte Niño

Acknowledgments:

I sincerely thank my supervisor, Bojana Ristić, for her exceptional mentorship and patience. She guided me step by step and provided space for independent learning. I am fortunate to have such a fantastic mentor. To Professor Christina Manouilidou for giving me the research opportunity. To Martin Takac, who greatly influenced my understanding of interdisciplinarity. I am deeply thankful to Miguel Maia for his support and belief in me. Lastly, I thank my family and friends for understanding my absence during this time.

Abstract:

This thesis investigates the distinct sub-processes during the processing of morphologically complex words that occur after decomposition, namely licensing and semantic composition, Building upon previous research that has evidenced these different post-decomposition sub-processes, this study aims to explore whether these sub-processes can be differentiated among native speakers of Colombian Spanish when the suffix attachment rules of the language are violated. Drawing on the framework proposed by Schreuder and Baayen (1995), which posits licensing (checking stem category) and composition (checking semantic compatibility of stem+affix combinations) as separate processes after decomposition, this research seeks to extend the understanding of these processes. The investigation focuses on the effects of violating suffix attachment rules and examines the influence of suffix productivity and base category on the resulting acceptability, processing speed, and accuracy. Furthemore, in contrast to previous studies that primarily examined verb-attaching affixes, this research incorporates noun-attaching affixes as well. By employing a 2 x 2 factorial design, the experiment manipulates suffix productivity (low vs. high) and base category (nouns vs. verbs). Pseudowords with semantic violations are created by attaching low-productivity suffix -udo to object nouns and high-productivity suffix -ano to concrete nouns. Pseudowords with category violations are generated by attaching the suffixes to verbs. Similarly, high-productivity suffix -ble and low-productivity suffix -dizo are used to construct semantic violations by adding them to semantically non-fitting verbs, and category violations by attaching them to nouns. The experimental setup also includes grammatically correct words, fillers, and non-words, resulting in a comprehensive lexical decision task. An acceptability judgment task was also administered before the lexical decision task. Preliminary analysis, employing linear mixed-effects models, reveals an expected pattern for the suffixes -ble and "-dizo," with semantic violations eliciting higher acceptability ratings, longer response times, and increased error rates. However, no significant effects are observed for the suffixes -ano and "-udo." These findings suggest that the observed pattern may be contingent upon base category but not affix productivity. Specifically, violations of category and violations of semantics exhibit comparable effects in terms of acceptability, processing speed, and accuracy with noun-attaching affixes.

Further research is required to investigate whether this observed pattern stems from differential processing mechanisms between noun- and verb-attaching affixes or specific semantic restrictions explored here. By deepening our understanding of the licensing and semantic composition processes in morphologically complex word processing, and adding evidence from yet unattested language (Spanish) and factors (base category and productivity) to this research line, this study contributes to the broader field of psycholinguistics and sheds light on the cognitive mechanisms underlying language comprehension and word formation.

Abstrakt:

Táto práca skúma odlišné čiastkové procesy pri spracovaní morfologicky zložitých slov, ku ktorým dochádza po dekompozícii, a to licencovanie a sémantickú kompozíciu. Vychádzajúc z predchádzajúcich výskumov, ktoré preukázali existenciu týchto odlišných čiastkových procesov po dekompozícii, je cieľom tejto štúdie preskúmať, či je možné tieto čiastkové procesy rozlíšiť od seba medzi rodenými hovoriacimi kolumbijskej španielčiny, keď sú porušené pravidlá pripájania prípon v jazyku. Na základe rámca navrhnutého Schreuderom a Baayenom (1995), ktorý predpokladá licencovanie (kontrola kmeňovej kategórie) a skladanie (kontrola sémantickej kompatibility kombinácií kmeň+afix) ako samostatné procesy po dekompozícii, sa tento výskum snaží rozšíriť chápanie týchto procesov. Výskum sa zameriava na účinky porušenia pravidiel pripojenia sufixu a skúma vplyv produktivity sufixu a základovej kategórie na výslednú prijateľnosť, rýchlosť spracovania a presnosť. Okrem toho na rozdiel od predchádzajúcich štúdií, ktoré skúmali predovšetkým afixy pripájajúce slovesá, tento výskum zahŕňa aj afixy pripájajúce podstatné mená. Experiment využíva 2 x 2 faktorový dizajn, ktorý manipuluje s produktivitou prípon (nízka vs. vysoká) a základnou kategóriou (podstatné mená vs. slovesá). Pseudoslová so sémantickým porušením sa vytvárajú pripájaním sufixu s nízkou produktivitou -udo k objektovým podstatným menám a sufixu s vysokou produktivitou -ano ku konkrétnym podstatným menám. Pseudoslová s porušením kategórie sa vytvárajú pripojením prípon k slovesám. Podobne sa vysokoproduktívna prípona -ble a nízkoproduktívna prípona -dizo používajú na tvorbu sémantických porušení ich pridaním k sémanticky nezhodným slovesám a kategoriálnych porušení ich pripojením k podstatným menám. Experimentálna zostava zahŕňa aj gramaticky správne slová, plnovýznamové slová a neslovenské slová, čím vzniká komplexná lexikálna rozhodovacia úloha. Pred úlohou lexikálneho rozhodovania bola administrovaná aj úloha posudzovania prijateľnosti. Predbežná analýza s využitím lineárnych modelov so zmiešanými efektmi odhalila očakávaný vzorec pre prípony -ble a -dizo, pričom sémantické porušenia vyvolávajú vyššie hodnotenia prijateľnosti, dlhší čas odpovede a vyššiu chybovosť. V prípade prípon -ano a -udo však neboli pozorované žiadne významné efekty. Tieto zistenia naznačujú, že pozorovaný vzorec môže byť podmienený základnou kategóriou, ale nie produktivitou afixu. Konkrétne, porušenie kategórie a porušenie sémantiky vykazujú porovnateľné účinky z hľadiska prijateľnosti, rýchlosti spracovania a presnosti pri afixoch s príponou

podstatného mena. Je potrebný ďalší výskum, aby sa zistilo, či tento pozorovaný vzorec vyplýva z rozdielnych mechanizmov spracovania medzi afixmi pripájajúcimi podstatné mená a slovesá alebo zo špecifických sémantických obmedzení, ktoré sa tu skúmali. Tým, že táto štúdia prehlbuje naše chápanie procesov licencovania a sémantického skladania pri spracovaní morfologicky zložitých slov a pridáva do tejto výskumnej línie dôkazy z ešte neprebádaného jazyka (španielčiny) a faktorov (základná kategória a produktivita), prispieva k širšej oblasti psycholingvistiky a vrhá svetlo na kognitívne mechanizmy, ktoré sú základom porozumenia jazyka a tvorenia slov.

Contents

List of figures

List of tables

List of plots

1. Introduction

Morphological processing involves the analysis and comprehension of words based on their morphological structure. This study examines the effects of affix productivity, base category, and semantic well-formedness on morphological processing, specifically focusing on accuracy and reaction times associated with different suffixes. By investigating these factors, we seek to contribute to the existing literature and enhance our understanding of the complex processes involved in morphological analysis.

We also expand the scope of research by investigating an under-studied dialect of Spanish, namely Colombian Spanish. This allowed us to explore how these factors operate within a specific linguistic context and shed light on the processing of morphological structures in this dialect. To achieve our objectives, we adopt a paradigm that compares pseudowords violating categorial rules and semantic rules of affix attachment.

Previous studies in this research line have shown that violations of semantic constraints have a greater impact on accuracy and reaction times compared to violations of categorial constraints (Manouilidou, 2007; Manouilidou & Stockall, 2014). We observe a comparable pattern in our study, supporting the cross-linguistic nature of this effect.

In addition, our findings reveal an intriguing pattern in the processing of suffixes attached to nouns. While violations of semantic constraints lead to prolonged processing times and reduced accuracy, we find no significant differences in processing times or accuracy between the semantic and categorial violation conditions for noun-attaching suffixes. This new piece of evidence suggests that the underlying base category and its associated semantic restrictions play a more significant role in understanding and processing derived words in this context.

Furthermore, we investigate the influence of affix productivity on morphological processing. Contrary to our expectations, our study does not reveal a significant effect of affix productivity in the processing of morphological structures within Colombian Spanish. This deviation from previous research suggests that other factors, such as syntactic and semantic constraints, may be more influential in determining the integration and compatibility of morphological constituents.

THEORETICAL FRAMEWORK

2. Morphological derivation

Morphological derivation, the process of forming new words by adding affixes, has intrigued linguists and language researchers for years. According to Bauer (1983), affixes like prefixes and suffixes are attached to existing words, altering their syntactic category or adding meaning. For instance, "un-" in "unhappy" changes "happy" to "not happy," exemplifying derivational morphology with a prefix. Another example is the formation of nouns like "driver" from verbs such as "drive" or "teacher" from "teach," where a suffix is added, changing the syntactic category.

In response to this interest, psycholinguistic research has delved into studying morphological processes, unraveling how speakers process and comprehend words formed through these processes. Taft & Forster (1975) found that the transparency of morphological structure influences word recognition speed. Additionally, Joanisse & Seidenberg (1999) highlighted the role of morphological awareness in reading development. Another influential study by Rastle & Davis (2008) investigated the influence of morphological structure on visual word recognition and lexical access. They found that visual word recognition and lexical access are influenced by rapid orthographic analysis. Thus, these studies have contributed to our understanding of how morphological processes influence language processing and have provided insights into the intricate mechanisms involved in recognizing and understanding derived words.

However, the process of morphological derivation is more complex than simply adding affixes, and it involves a range of factors that can affect the well-formedness of complex words and subsequently their processing.

2.1 The syntactic argument vs. semantic argument of suffixation in derivational word formation.

The syntactic argument versus semantic argument in suffixation for derivational word formation is a long-standing debate among linguists and cognitive scientists (Bauer, 2001; Booij, 2010; Plag, 2003; Pustejovsky, 1991). This debate revolves around the question of whether the choice of a suffix is primarily guided by syntactic considerations or semantic considerations. On the one hand, the syntactic argument proposes that the syntactic category of the base word determines the choice of a particular suffix. In contrast, the semantic argument claims that the base word's meaning determines the suffix's choice. The semantic argument suggests that suffixation primarily serves a semantic function by adding meaning to words and creating new lexical categories. For example, the addition of the suffix "-ness" to the adjective "happy" creates the noun "happiness," which refers to the state of being happy. The question of the importance of syntax and semantics in suffixation elicits divergent viewpoints among scholars, and a consensus has not been reached yet. The importance and impact of each factor can vary based on aspects such as the specific language being studied, the morphological system being analyzed, and the theoretical framework adopted by researchers. However, some scholars argue for a more integrated perspective that considers both syntactic and semantic factors in the process of suffixation (Bauer, 2015; Kiparsky, 1982). They propose that the choice of a suffix is influenced by a combination of syntactic and semantic factors, suggesting that the two approaches are not mutually exclusive. This view acknowledges the interplay between syntax and semantics in word formation and recognizes that multiple linguistic factors can influence the choice of a suffix. By considering both the syntactic category of the base word and its meaning, this perspective offers a more comprehensive understanding of how suffixation operates in language.

In this study, we adopt the perspective put forth by Manouilidou & Stockall (2014) that both syntactic and semantic approaches to morphological analysis recognize the importance of morphemes in constructing words and the role of word structure in language comprehension and production. Both approaches also acknowledge the significance of the mental lexicon in language processing. Furthermore, they acknowledge that linguistic and extralinguistic factors can affect the variability of morphological structures, such as a

word's frequency of use, semantic context, and phonological properties. When investigating morphological processing with a lexical decision task, as in the current study, it is crucial to take all these aspects into consideration, as well as the question of lexical access.

3. Lexical access

Extensive research in psycholinguistics has examined the impact of morphological structure on the access of complex words, leading to a long-standing debate regarding whether these words are accessed through decomposition or as whole units. Proponents of the strict decomposition perspective (Taft & Forster, 1975; Rastle & Davis 2008) argue that complex words are accessed by breaking them down into constituent morphemes, while other advocates of the whole-word perspective (starting with Butterworth 1983) propose that complex words are stored and accessed as integrated units. Computational models of lexical access have been instrumental in investigating these perspectives. Morphological decomposition models simulate the cognitive processes involved in accessing complex words through morphological rules or statistical patterns, aligning with the decomposition perspective (Coltheart et al., 2001; Taft & Forster, 1975). As in previous psycholinguistic studies relevant for the current research (Manouilidou, 2007; Manouilidou & Stockall, 2014), we will base our investigation on the model by Schreuder & Baayen (1995). Thus, we will analyze this model to gain a more comprehensive understanding of its stages and approach to morphological processing.

3.1 The Model by Schreuder and Baayen (1995)

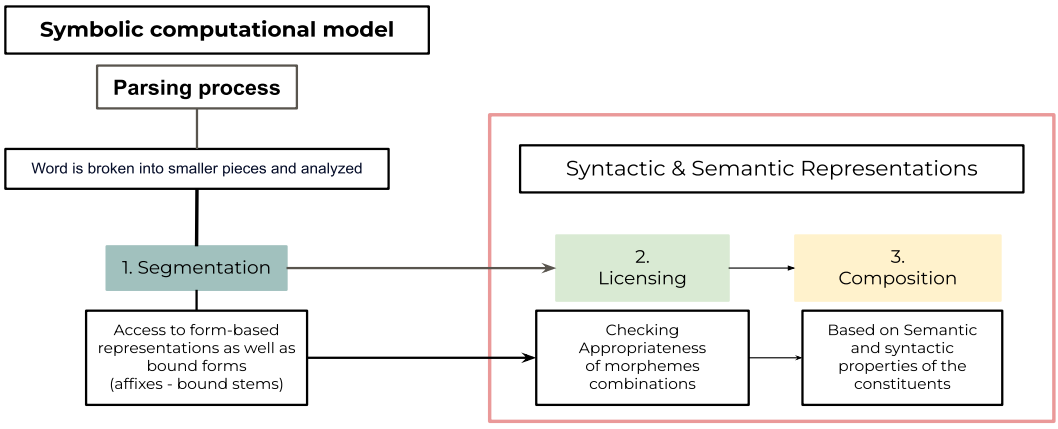


Fig. 1: Adaptation of the Model from Schreuder and Baayen (1995)

Schreuder and Baayen (1995) proposed a morphological processing model involving two subprocesses after segmentation, namely, licensing, and composition. These subprocesses describe distinct stages that dictate how morphemic constituents are recombined and interpreted in accessing and recognizing morphologically complex words.

Assumptions of the model

The two stages after decomposition are the licensing stage, and the semantic composition stage. These are two post-decomposition stages relevant for our current study. In the licensing stage, (the lexeme lookup stage) the constituents are accessed in the mental lexicon and are evaluated to determine if they can be combined. This stage ensures that the co-active representations are integrated correctly. To accomplish this, the relevant information associated with each constituent, such as semantic and syntactic properties, and frequency of occurrence in the language, is retrieved. By extracting this information, the model gains access to the specific linguistic knowledge tied to each morpheme.

Then, in the semantic composition, the model computes the lexical representation of a complex word by assessing the semantic compatibility of the stem+affix combinations. By combining the extracted information from the morphological constituents, a comprehensive representation of the word's meaning is constructed. This process allows the model to understand the semantics of the complex word and establish its relation to the surrounding context.

Bringing back the debate surrounding the syntactic and semantic arguments in suffixation for derivational word formation, the model proposed by Schreuder and Baayen (1995) offers valuable insights. While the model itself does not explicitly support either the syntactic or semantic argument approach in suffixation (see section 2.1), it provides a framework that incorporates aspects of both perspectives. By considering the sequential subprocesses of segmentation, licensing, and composition, the model integrates the form, meaning, and syntactic properties of morphemes in the processing of complex words. This integration allows for a more comprehensive understanding of how syntax and semantics interact in suffixation and derivational word formation. Therefore, the model sheds light on the interplay between syntactic and semantic factors in the formation and comprehension of morphologically complex words.

4. Previous studies in this research line

As mentioned before, Schreuder & Baayen's 1995 model distinction between the licensing and semantic composition stages is particularly important for the current study. Substantial research has supported the distinction between the licensing and semantic composition subprocesses in word recognition experiments. A study by Manouilidou (2007) aimed to determine if native speakers distinguish between different types of constraint violations in pseudowords containing suffixes. The study found that subjects differentiated between pseudowords violating different constraints. Thematic constraints, where suffixes were added to verbs of mismatching argument structure (**orimastis* 'maturer') were found to be more violable than categorial constraints, where suffixes were added to nouns (**karekla-tis* 'chair-er') and the processing of pseudowords with thematic violations incurred a higher processing cost compared to categorial violations. These results suggest that thematic processing is a distinct part of word recognition for deverbal pseudo-words. Additionally, the study suggested that different types of constraints have varying levels of violability and apply at different stages of word formation. The findings have implications for the theory of lexical access, indicating that features might not be processed simultaneously during word recognition.

Furthermore, the study by Manouilidou & Stockall (2014) explored categorial and thematic constraints in deverbal word formation in Greek and English following the same paradigm as Manouilidou (2007). Results from this study showed that speakers of both languages differentiated between pseudo-words that violated Categorial (**potiritís* 'glasser'), and Thematic (**trék-simos* 'runable') constraints regarding acceptance rates and processing time. Taking together results from both languages, the study made claims about the structured mental representation of deverbal derivatives. It suggested that these derivatives have various properties that can be accessed through distinct operationsduring word processing. This implies that processing morphological well-formedness of morphologically complex words involves multiple subprocesses and operations.

Overall, the behavioral psycholinguistic research and the studies by Manouilidou (2007) and Manouilidou & Stockall (2014) provide empirical evidence for the immediate processing of morphological information and support Schreuder & Baayen's model. The

findings enhance our understanding of morphological processing, highlighting the role of decomposition into morphemes, linguistic rules, and semantic properties in word recognition. They also shed light on the psycholinguistic aspects of lexical access and word formation, emphasizing the intricate nature of processing morphologically complex words. The studies underscore the importance of thematic constraints in accessing deverbal pseudo-words and their distinct role in word formation in comparison to the category/syntactic constraints.

Besides, neuroimaging studies using the pseudoword paradigm have supported the proposal of temporally separate processing stages. Researchers used magnetoencephalography (MEG) to measure brain activity during a lexical decision task. Neophytou et al. (2018) conducted a Greek experiment, replicating the difference between category and argument structure violation pseudowords. They found that category violation pseudowords elicited more activity in the 200-300 ms time window, associated with licensing effects, in the ventral and posterior regions of the temporal lobe. On the other hand, thematic/argument structure violation pseudowords showed increased activity in the orbitofrontal region in the 300-500 ms time window. Stockall et al. (2019) conducted an English experiment using prefixes and obtained similar results. Although the effect in the orbitofrontal cortex was less statistically robust, the combined findings of these studies provide spatiotemporal evidence for the distinction between the earlier licensing/subcategorization stage and the later semantic composition stage.

Recent studies within the same research line have expanded the languages and dimensions over which the investigation into the distinction between licensing and semantic composition is conducted. Moreover, these studies have further explored the processing of derivations from other categories, with a particular emphasis on nouns. Ongoing studies in Bangla (Moitra et al., 2022) have examined the differences between morphologically complex Bangla nominals, specifically investigating whether they follow similar processing stages as observed in other languages. Surprisingly, these studies have found that in some cases, prefixes associated with category violations are rejected more frequently than those associated with semantic violations. Additionally, category violation prefixes take longer to be rejected and result in lower accuracy compared to semantic violation prefixes. These findings suggest that the processing of morphologically complex words in Bangla exhibits unique characteristics, deviating from the patterns observed in

other languages. Furthermore, Ristić et al. (2022) investigated prefixes in two South Slavic languages, Slovenian and BCS, where category violation pseudowords were made by attaching prefixes to nouns, while semantic violation pseudowords were created by attaching prefixes to verbs whose semantics does not match these specific prefixes (i.e. they do not present unstable or reversible states). The expected pattern was found for both languages and across the measures. Unlike previous studies, where semantic violations were based on argument structure violations, this study presents evidence from purely semantic violations, confirming the existence of a separate sub-process for semantic well-formedness, that is different from the category evaluation.

5. Current study

Building upon previous research, this study aims to determine if we can differentiate between two post-decomposition processes when processing suffixed complex words in Spanish. The pseudoword paradigm, previously employed by Manouilidou (2007), Manouilidou et al. (2015), and Manouilidou & Stockall (2014), will be utilized in the current study. However, unlike the previous studies, and following Moitra et al. (2022) and Ristić et al. (2022), we will compare category restrictions violations to semantic restrictions violations in Spanish suffixes, rather than solely focusing on argument structure restrictions. This approach will provide additional, unambiguous evidence pertaining to the semantic composition stage, as examining argument structure restrictions alone may not fully capture the semantic evaluation processes. Namely,argument structure encompasses both syntactic and semantic aspects (Bresnan, 2015). Therefore, we aim to explore a violation of a more purely semantic suffix attachment rule to address this gap in knowledge. By adopting this approach, we hope to obtain clearer evidence regarding the sub-process of semantic composition in word processing.

To add novelty to this study, we will incorporate two additional aspects that have not been extensively explored thus far. Firstly, we will consider the influence of suffix productivity, distinguishing between low and high-productivity. Previous research, such as the work by Baayen and Lieber (1991) Plag and Baayen (2009) Burani and Thornton (2003), and Lazaro et al. (2015), have demonstrated the impact of productivity measured through token and type frequency on lexical access with morphologically complex words and

pseudowords. By investigating the role of productivity, we aim to gain a deeper understanding of its effect on the processing of suffixed complex words in Spanish in relation to the licensing and semantic composition subprocesses. Secondly, we will introduce a new base/stem category for comparison, specifically nouns. This addition expands the scope of our study, allowing us to examine potential variations in the post-decomposition processes across different base/stem categories. By comparing the processing mechanisms of noun bases/stems with those of verb bases/stems, as in the previous studies in this research line, we can uncover potential differences in how these words are processed. Specifically, the present study aims to investigate the effects of semantic violations and categorial violations in pseudowords on response times (RTs) and error rates.

Our hypothesis posits that pseudowords with semantic violations will lead to longer RTs and higher error rates compared to pseudowords with categorial violations. Furthermore, we anticipate, based on Lazaro et al. (2015) and similar research, that pseudowords with low-productivity suffixes will elicit shorter RTs and lower error rates when compared to pseudowords with high-productivity suffixes, which might annul the relevant (semantic vs. categorial) contrast for the low-productivity suffixes. However, we predict no significant differences in RTs and error rates with respect to the relevant contrast between noun-attaching and verb-attaching affixes, suggesting that the post-decomposition processing develops equally regardless of the stem/base category.

6. Productivity and frequency of suffixes in lexical access

The role of productivity and frequency of suffixes in lexical access is an important area of study within morphology and language processing. Researchers aim to gain insights into the mechanisms underlying word formation and comprehension by examining how suffix productivity and frequency impact the recognition and retrieval of derived words. Several studies have investigated this relationship in different languages. Baayen and Lieber (1991) conducted a corpus-based study on English derivation and found a positive correlation between the productivity and frequency of suffixes. Additionally, Burani and Thornton (2003) focused on Italian and showed that highly productive suffixes are accessed more quickly, highlighting the influence of productivity on lexical access. Plag and Baayen

(2009) demonstrated in their investigation of German that more frequent suffixes are accessed faster during word recognition tasks.

In the case of Spanish, Lazaro et al. (2015) found that words with productive derivational suffixes were recognized faster than those with unproductive suffixes. The opposite trend was observed with pseudowords. When it comes to pseudowords, the presence of highly productive suffixes actually hinders the processing. These suffixes activate various potential morphological options, leading readers to consider the pseudoword as a possible real word. However, since pseudowords lack valid entries in the lexicon, combining the stem with a highly productive suffix does not create actual words. As a result, readers have to actively suppress their tendency to classify the pseudoword as a real word, which leads to inhibition. On the other hand, for real words, highly productive suffixes have a facilitating effect. They are connected to a wide range of valid lexical entries, making it easier for readers to recognize and process the word as a legitimate part of the language. Consequently, productivity acts as a facilitator for real words, enhancing their processing, while it acts as an inhibitory factor for pseudowords.

In a subsequent study by Lazaro et al. (2016), 50 Spanish derivational morphemes were examined to evaluate their frequencies in terms of types and tokens. The results showed a positive correlation between suffix productivity and type frequency (number of words containing that suffix), indicating that highly productive suffixes tend to have higher type frequencies. However, no significant correlation was found between suffix productivity and token frequency (the frequency of occurrence of these words), suggesting that more than frequency alone is needed to explain suffix productivity. It was also shown that both suffix productivity and frequency contribute to lexical access. Highly productive and frequently used suffixes were associated with faster lexical access, indicating that words formed with these suffixes are more easily recognized and retrieved by speakers. These findings highlight the importance of productivity in determining word formation potential and emphasize the significance of morphological knowledge in language comprehension.

However, productivity alone does not fully account for facilitating lexical access. It interacts with other variables such as word length, suffix frequency, or the frequency of words containing the suffix. The specific influence of these variables and whether they are independent or not remains to be determined, necessitating further research. A separate study by Lázaro (2012) examined the role of suffix type frequency in lexical decision

tasks. It was found that morphologically complex words formed by high-frequency suffixes were recognized more quickly than those formed by low-frequency suffixes.
The paper by Lazaro et al. (2016) was used as a guide to select four suffixes, comprising two high-frequency and two low-frequency ones. Subsequently, we conducted an investigation of these suffixes within the corpora of Sketch Engine, which encompasses an American Spanish web corpus sourced from South American country code domains, along with a specific subcorpus focused on the Colombian domain. This careful selection and analysis were undertaken to ensure the significance and applicability of the variables within the Colombian Spanish context.

7. Derivation Rules in Spanish

Regarding the derivation rules in Spanish, they involve various processes for creating new words. The following are some of Spanish's most common derivation rules:

- Prefixation: Adding a prefix to a base word to create a new word. For example, the prefix "des-" can be added to the base word "hacer" (to do) to make "deshacer" (to undo).
- Conversion: Changing a base word from one part of speech to another without adding any affixes. For example, the base word "bailar" (to dance) can be converted into a noun by changing it to "el baile" (the dance).
- Compounding: Combining two or more base words to create a new word. For example, the base words "casa" (house) and "grande" (big) can be combined to create "Casagrande" (mansion).
- Suffixation: Adding a suffix to a base word creates a new word. For example, the suffix "-ción" can be added to the base word "conversar" (to talk) to make "conversación" (conversation).

However, this research aims to investigate the derivation by the suffixation process, how Spanish speakers process derived words with suffixes, including deverbal and denominal adjectives. As in other languages, in Spanish, the formation of adjectives through suffixation is a typical process that involves adding a suffix to a word from another grammatical category to form a derived adjective. This strategy of adjective formation is productive in Spanish and is used to express a wide range of semantic nuances. For

example, the suffix "-oso" is a productive suffix in Spanish that is added to nouns to form adjectives that indicate the abundance or presence of something. Thus, from the noun "ruido" (‘noise’), the adjective "ruidoso" (‘noisy’) is formed. Similarly, from the verb "cansar" (‘to tire’), the adjective "cansado" (‘tired’) is formed, indicating that someone or something is in a state of fatigue or exhaustion.

By examining the processing of complex words composed of multiple morphemes, including bases/stems and suffixes, we aim to gain insights into how speakers of Spanish perceive and interpret these words. For this purpose, understanding the different types of suffixes and their usage in Spanish is crucial to comprehend the meaning and function of these new words as well as to gain a more comprehensive overview of different affix + stem combination rules.

7.1 Affix selection

In addition to selecting a particular class of words by their category, the suffixes take into account other more specific aspects of these. Some suffixes are sensitive to certain 'subcategorical' stem features, which we call this way because they divide a given grammatical category into subclasses, such as a specific type of verb or noun.

Therefore, although syntactic and morphological constraints exist, there are also constraints related to semantics that govern the meanings and ideas that can be expressed. These semantic limitations dictate the suitability or well-formedness of a specific suffix when combined with a base word. We will then provide separate descriptions for verb-attaching and noun-attaching suffixes, with a focus on semantic constraints. However, in order to present a comprehensive understanding of their attachment behavior, we will also discuss the argument structure of -ble and -dizo, suffixes added to verbs, even though our primary restrictions were based on semantics when creating the pseudowords.

7.1.1 Suffixes attached to verbs

Deverbal word formation is a process that involves creating new words from verb stems. According to Manouilidou (2007), this process is subject to various linguistic constraints, including phonological, syntactic, and semantic constraints. However, additional constraints must be considered when forming deverbal words, which arise from the Argument Structure (AS) properties of both the verb stem and the suffix. These constraints

play an important role in determining which suffixes can be attached to a given verb stem and thus have significant implications for the formation of new words in general. This chapter will explore the AS properties, as well as semantic properties of verb stems and suffixes and how they interact to influence deverbal word formation in Spanish.

7.1.1.1 The suffix -ble

The suffix -ble plays an essential role in Spanish language study because of its frequent usage and distinctive properties. It is added primarily to transitive verbs with an agent subject. According to Oltra-Massuet (2014), as not all verbs are suitable for forming words with -ble, to be eligible for this suffix, a verb must fulfill two key conditions: the presence of an internal argument (not always explicitly stated) and an implication of an originator such in the following example:

<Internal Argument<Originator>

1. *pasar (to pass(* → *pasable (passable)*
   <Originator<Internal Argument > <Internal Argument<Originator>

Thus, unaccusative or unergative verbs generally do not fulfill these criteria and are not considered suitable for forming words with -ble. Instead, according to Albano et al. (2021), unaccusative verbs have a deep object (patient subject), and their thematic role is either theme or patient. Thus, the thematic roles scheme for the intransitive verbs would be as follows:

<Patient (Internal Argument)<Originator (unexpressed)>

Let's consider the following examples:

2. *a. brotar (to sprout)* → **brota-ble "sproutable"*
   <Internal Argument (patient)>
   b. expirar (to expire) → **expira -ble "expirable"*
   <Internal Argument (patient)>

In the case of the verb "brotar" (to sprout), it is an intransitive unaccusative verb, which means it does not take an object. As such, "brotar" does not have an internal argument in the same way that transitive verbs do. Instead, the internal argument (the sprout itself) represents the sprouting or budding entity. This verb does not have any external agent or originator either. In case *b.* again the internal argument of "expirar" is the patient subject of the sentence, which refers to the thing that is expiring or coming to an end of its validity period but again, it is implied. Therefore, none can be considered a valid derived word.

Other examples of these verbs are *"ocurrir" (happen)* or "*crecer" (grow),* which have an internal argument but do not imply an originator, while unergative verbs, like "*dormir" (sleep)* or "*estornudar" (sneeze),* have an originator but lack an internal argument. Therefore, the ability of -ble to form words from these types of verbs is not expected either.

As stated in Ortega,et al. (2005), we do not form "*tener" (have) *tenible (hav-able)* or "*ver" (see) *veible (see-able)* because, although the base verbs *tener* and *ver* are transitive, their subject is not an agent and their direct object is not affected by the verbal action.

Finally, another type of verbs which seem not to attach to -ble are the antiaccusative verbs. These verbs are a type of intransitive verb that have an additional argument that is not an agent or a causer of the action, a verb whose direct object is the same as its subject. In the case of Spanish, reflexive verbs would belong to this category, for instance: *derretirse (melt), dormirse (fall asleep), aburrirse (get bored)* etc..

Exceptions:

However, we find some intransitive verbs, which have a transitive equivalent, known as ergative verbs, such as congelar 'freeze' or romper 'break'. With those verbs -ble can be used.

3. congelar (to freeze) → congelable (freezable)
<Originator<Internal Argument > <Internal Argument >

The verb *"congelar"* means *to freeze*, and it is an ergative verb because it can be used transitively, as in "congelar la comida" (freeze the food). In this case, "la comida" (the food) is the direct internal argument of the verb but also there is an agent or originator of the action. Therefore, the adjective "congelable" can be formed to indicate that something is "freezable" or capable of being frozen. Besides, as stated in Albano et al. (2021) the participles of unaccusative verbs, like those of transitive verbs, can function as modifiers of nouns, resembling adjectives, and like transitive verbs, they can form constructions with the neuter article "lo". We will see the importance of this when we analyze the results of the processing of some pseudowords.

In our current study, we would like to highlight the complexity surrounding the attachment of the suffix -ble due to numerous exceptions and the absence of straightforward rules. With that in mind, we have established certain guidelines based on semantic constraints for our study design. Specifically, we propose that verbs denoting existence or appearance (e.g., acaecer "to happen," acechar "to spy on," albergar "to shelter"), verbs expressing growth (e.g., madurar "to mature," nacer "to be born"), and verbs indicating a change of state (e.g., florecer "to bloom") do not take the -ble suffix. To identify such verbs, we relied on Levin's (1993) classification for English verbs, which provided a framework for recognizing verbs falling under the categories of change of state, existence, and appearance. These categories were then translated into Spanish by a native speaker. Additionally, we consulted Mendikoetxea's (1999) classification to ensure a comprehensive and accurate selection of verbs that align with the identified semantic groups of verbs.

7.1.1.2 The suffix -dizo

In Spanish, the suffix -dizo can be used to form adjectives that denote a tendency or disposition to perform a certain action or to have a certain quality. Compared to -ble, the suffix -dizo is considered a low-productivity suffix in Spanish. Fábregas (2020) argues that the properties of -dizo are similar to those of -ble.

b. pegar (to stick)	→	*pegadizo (sticky)*
<Event <Cause (External Causer) <<Patient >		*<State <Patient (the entity that is sticky)>*

This thematic scheme represents the argument structure of the verb "pegar" where the internal argument (the object being stuck or hit) is acted upon by the originator (the person or entity performing the action). However, -dizo exhibits different behavior than -ble, in that -dizo does not impose an external causation constraint. The addition of the suffix -ble to a verb typically implies that the action can be performed by an external agent like a person, tool, animal or force. Although with -dizo it is often convey suddenness or intensity, they do not necessarily indicate that an external force caused the action.

4.

a. resbalar (to slip) → *resbaladizo (slippery)*

<Event <Cause (External Causer) <<Patient > *<State <Patient (the entity that is slippery)>*

For example, "resbaladizo" is a Spanish adjective derived from the verb "resbalar" (to slip or slide) with the addition of "-dizo." The resulting adjective means slippery or slick, but the addition of -dizo does not imply external causation. Slippery surfaces can occur naturally or can be intentionally created without an external force, so the suffix -dizo in "resbaladizo" does not impose an external causation constraint on the adjective.

Good candidates, for verbs whose -dizo can attach to, are the antiaccusative verbs denoting a psychological state. Antiaccusative verbs are intransitive verbs that show an event affecting its subject, while giving no semantic or syntactic indication of the cause of the event. The single argument of the anticausative verb (its subject) is a patient, that is, what undergoes an action:

<Event <<Patient >

Therefore, the use of the suffix -dizo is commonly preferred when forming deverbal adjectives from verbs that describe events caused internally, due to the properties of the internal argument that are sufficient to initiate a change of state. Reflexive verbs in Spanish are ideal examples of such verbs.

5.

a. *enojarse (get angry)→ enoja-dizo (short-tempered)*
b. *olvidarse (forget about) → olvida-dizo (forgetful)*
c. *enamorarse (fall in love) → enamora-dizo (inclined to fall in love)*

There are few examples, found in Fábregas (2020), where the base verb seems to denote something other than a psychological state: *huir (run away), alzarse (raise up), apartarse (distance oneself), encontrarse (to meet), arrimar-se (to come closer) and soltarse (to slip away).*

In summary, argument structure constraints provide a choice between two affixes: -ble and -dizo. The -ble suffix mandates that the external and internal arguments in the verbal base must be distinct, whereas -dizo does not impose this limitation. Thus, internally-caused events can only make dispositional adjectives with -dizo, while -ble specialises in the externally-caused events.

Importantly for the current study, same as -ble , because of the overlap in the meaning and distribution, -dizo doesn't attach to certain verbs based on semantic restrictions. Namely, it doesn't attach to verbs of existence and appearance, growing, and change of state, as described in the previous section.

7.1.2 Suffixes added to nouns

7.1.2.1 The suffix -udo

In Spanish, the suffix -udo is commonly added to nouns (N-udo) that denote a body part (human or animal, or both), a material object, or, rarely, an abstract concept, and imparts an intensive nuance to the resulting word. Carriazo (2014) shows three main types of meanings conveyed by the suffix, namely "full of / endowed with / composed of," with variations depending on the category of the noun.

[Noun] + [suffix] > [qualifying adjective]"

The most abundant group is that of derived adjectives whose base is a noun that designates a body part. Here -udo has an augmentative meaning but considered excessive due to their size or number, such as :

barba (beard) → bearded

cabello (hear) → hairy

diente (tooth) → buck-toothed

panza (belly) → panzudo (paunchy)

The other categories are composed of derivatives whose base designates an object or a quality, including some cases of plant part names. We can distinguish those that serve to express a relationship or similarity and those that can be paraphrased as "having, suffering, or showing X", where the base (X) can be a concrete or abstract noun, such as:

capricho (caprice) → caprichudo (capicious)

cáscara (shell) → cascarudo (tick shelled)

coraje (courage) → corajudo (courageous)

7.1.2.2 The suffix -ano

The suffix -ano is added to nouns to form relational adjectives that express an origin relationship between a noun denoting a place and a kind. This suffix is commonly used to form gentilic adjectives, which describe people or things that come from a particular place.

[Noun] + [suffix] > [relational adjective]

The suffix -ano can be applied to a large number of toponyms, including cities, regions, provinces, countries, continents, and other proper nouns of place such as: *Mexicano (Mexican), Cubano (Cuban), Venezolano (Venezuelan), Colombiano (Colombian), etc*. It is also used to form a good number of adjectives derived from proper names of people, whether they are first names such as: *dominicano (*Dominican)*, franciscano (*Franciscan)*, gregoriano* (Gregorian) or surnames such as: copernicano (Copernican) ,and galileano,(Galilean).

These adjectives can have qualifying or relational uses (such as *mundano* meaning "worldly" and *urbano* meaning "urban"), and some of them are used as nouns as well as adjectives (such as *aldeano* meaning "villager," *ciudadano* meaning "citizen," etc)

THE EXPERIMENT

8. Methodology

8.1 Materials

Four suffixes were used for the experiment, resulting in a 2 (Productivity: High, Low) x “ (Base: Verb, Noun) study design. Below we show the suffixes used in the current study, with the base category they select on the left side of the arrow, and the category they create on the right.

High Productivity	***-ble***	V	→	Adj
Low Productivity	***-dizo***	V	→	Adj
High Productivity	***-ano***	N	→	Adj
Low Productivity	***-udo***	N	→	Adj

Based on these suffixes pseudowords violating categorial and semantic rules were created .

8.1.2 Pseudo-Words with Semantic and Categorial Violations

-ble

Section 7.1.1.1 elucidates that the Spanish suffix -ble is primarily affixed to transitive verbs featuring an agent subject, necessitating the presence of an internal argument and implying an originator. Of particular significance is the observation that the suffix -ble in Spanish does not attach to verbs that describe appearance, existence, or change of state.

- For the Semantic Violations (SemViol), the restriction on the attachment of the -ble suffix to certain verbs was primarily based on semantics, rather than argument structure (e.g. Manouilidou & Stockall, 2014). Our approach was influenced by previous work on South Slavic languages, where a similar restriction was applied (Ristić et al., 2022). To establish this semantic restriction, we included verbs of

change of state, verbs of existence, and verbs of appearance as the bases, as they appeared to be incompatible with the suffix.

- ❖ Categorial Violations (CatViol) were created by adding -ble to nouns (e.g. piña 'pinapple'- *piñable)

Fig. 2: Categorial vs. Semantic Violations -ble

In our study, we utilized a set of 28 intransitive verbs, specifically change of state and existence verbs, as the basis for the SemViol items. The verbs used in our study were sourced primarily from Mendikoetxea (1999) and Levin (1993). Hernandez (2006) notes that these semantic groups of verbs are predominantly unaccusative, aligning with the argument structure restrictions that were previously described. Mendikoetxea (1999) identifies unaccusative verbs falling into two main categories. The first category encompasses verbs indicating a change of state or location, such as *abrir(se)* 'open', *hundir(se)* 'sink', *caer* 'fall', *florecer* 'bloom', and more. The second category consists of verbs expressing appearance or existence, such as *aparecer* 'appear', *llegar* 'arrive', *existir* 'exist', *venir* 'come', *suceder* 'happen', and others. Additionally, we included verbs of existence from Levin (1993) and several verbs of existence or growth selected by a native speaker, ensuring a comprehensive and accurate selection of verbs that align with the semantic groups of verbs identified.

Additionally, we employed a set of 29 concrete nouns representing various inanimate objects, fruits, and other similar entities as the basis for the CatViol conditions. Our aim was to construct a morphologically homogeneous sample of bases for our study, avoiding polysemy. For SemViol items, verbs that had prefixes or other additional morphemes attached to them were deleted. Furthermore, we omitted words which might be potentially ambiguous in terms of whether the base form is a verb or a noun, as in the case of *peinable* ('comb-able') which could interpreted as derived from *peinar* ('to comb') or *peine* ('comb'). This applied for every suffix.

-dizo

Section 7.1.1.2 provides an explanation regarding the Spanish suffix -dizo, which is appended to transitive verbs that possess both an agent argument and an internal argument, in a manner similar to the suffix -ble. However, it is crucial to note that this suffix is not applied to verbs denoting appearance, existence, or change of state.

- To create semantic violations (SemViol), we utilized a set of 28 intransitive verbs (the same as for -ble) which are mostly verbs of change of state, appearance and existence.
- To create categorial violations (CatViol) -dizo was added to nouns (e.g. *apio* 'celery' **apiodizo*). It is important to note that the nouns used for creating CatViol were the same employed with the suffix -ble, as previously described.

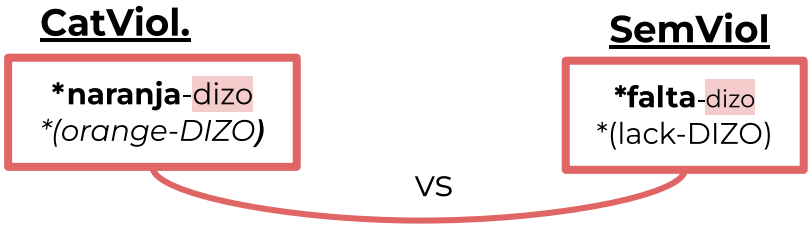


Fig. 3: Categorial vs. Semantic Violations -dizo

-udo

Sections 7.1.2.1 explains that the suffix -udo is added to nouns in Spanish to intensify their meaning. It is commonly used with body parts, objects, and abstract concepts. It can indicate an excessive size or number of body parts, and express relationships, similarities, or possession of objects and qualities.

- To create semantic violations (SemViol), -udo was added to concrete nouns referring to objects after removing the inflectional suffix (e.g. *mesa* 'table' – **mesudo*).
- To create categorial violations (CatViol), -udo was added to verbs after removing the infinitival suffix (e.g. *correr* 'run' – **corrudo*)

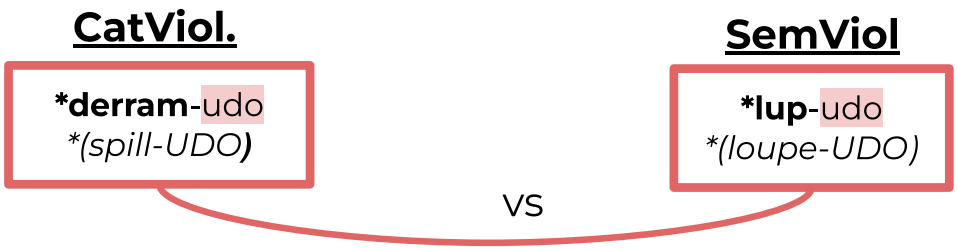


Fig. 4: Categorial vs. Semantic Violations -udo

-ano

Section 7.1.2.2 explains that the suffix -ano is commonly used to form gentilic adjectives, indicating the people or things from a specific place. It can also create adjectives from common nouns, with qualifying or relational meanings. Due to these rules, nouns referring to objects were intentionally used to create semantic violations.

- To create semantic violations (SemViol), -ano was added to a category of nouns referring to concrete objects (e.g. *mesa* 'table' -**mesano*).
- To create categorial violations (CatViol), -ano was added to verbs (e.g. *comprar* 'buy'- **comprano*)

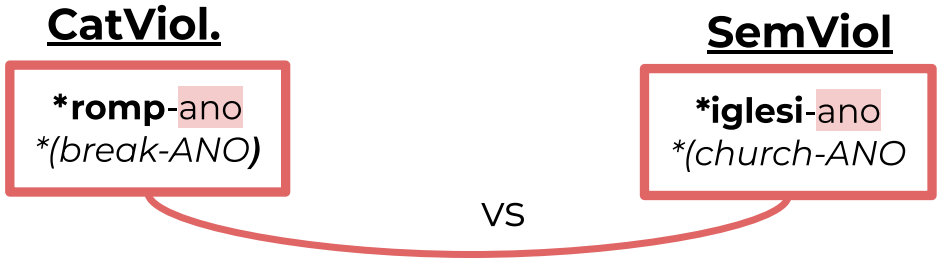


Fig. 5: Categorial vs. Semantic Violations -ano

We employed a uniform collection of 29 specific concrete nouns as the foundation for creating semantic violation (SemViol) conditions for both -udo and -ble. Similarly, we utilized an identical set of 28 action verbs as the basis for generating categorial violation (CatViol) conditions. The details of these noun and verb sets were previously explained in the -ble section.

Other remarks

To ensure consistency in our study, we took specific steps regarding the materials used. Firstly, we made sure that the verb bases were carefully matched in terms of both frequency and length when compared to the bases of other pseudowords employed in the research. Similar attention was given to the selection of nouns used to create violations. We aimed to form a set of noun bases and verb bases that were morphologically homogeneous, excluding those that contained prefixes or were ambiguous in terms of word category. Additionally, all bases were group-matched on frequency. To achieve these objectives, we utilized the Colombian sub-corpus within the Sketch engine as a reference (which will be described later). This corpus served as a valuable resource for the selection process and enabled us to control the variables effectively. These measures were implemented to ensure a controlled and balanced experimental design.

8.1.3 Grammatical items

We selected existing real words, more specifically derived adjectives in Spanish. They were selected utilizing the Sketch Engine Corpora. These adjectives were carefully chosen to ensure an equal base frequency distribution across four different suffixes: "-ble," "-dizo," "-udo," and "-ano." By matching the frequency of these derived adjectives, we aimed to control for any potential bias or influence of word frequency on participants' lexical decision task performance. In total, 118 grammatical items were used.

8.1.4 Non-words

Following Manouilidou & Stockall (2014), we also included non-words. These serve as a useful “sanity check”, as there should be no processes of evaluation of separate phonemes with these stimuli, and they should be quickly discarded as non-existent (quicker than pseudowords). They were created by modifying existing words from the Grammatical condition. These modifications involved changing a single letter while following phonotactic rules (e.g. **cadelludo* from cabelludo (hairy)). In total, we created 40 such non-words.

8.1.5 Fillers

In order to balance the number of existing and non-existing words in the experiment, we also included fillers, which were existing complex words in Spanish but with different types of suffixes (e.g. jardinero 'gardener'). Those fillers were selected based on 4 conditions: low base frequency - low affix productivity, low affix productivity - high affix productivity, high base frequency - high affix productivity , high base frequency - low affix productivity . In total, 40 fillers were used.

Organization of materials into lists

In our study, we organized the materials into two lists, each containing a total of 310 items. These items were distributed as follows: 118 grammatically correct words (gramm), 40 non-words (NWs), 40 filler items, and 112 pseudowords. To ensure controlled distribution and avoid repetition, we employed a Latin Square design specifically for the pseudowords. A Latin Square design is a method that ensures each level of a factor appears exactly once with every level of another factor. In our case, it allowed us to divide the pseudowords, which shared the same bases with different suffixes, into two lists, and avoid base repetition within the experiment. This design helped us achieve a balanced distribution of the pseudowords, ensuring that each stem occurred only once in each list as well as minimizing potential biases and enhance the validity and reliability of our experimental results.

8.2 The corpora used: Sketch Engine

Kilgarriff et al. (2014) introduced The Sketch Engine, a highly regarded corpus management and text analysis tool that has gained widespread adoption. It is designed to assist linguists, lexicographers, and language researchers explore large corpora to extract valuable linguistic information. The Sketch Engine can handle diverse language corpora, supports multiple languages, and offers flexible query functionalities through its Corpus Query Language (CQL). It utilizes various data sources to provide its corpus management and text analysis capabilities. The data primarily comes from multiple origins. Firstly, web corpora are collected by crawling the internet, ensuring coverage of various languages and domains. Additionally, Lexical Computing Limited has obtained licenses for high-quality

corpora such as the British National Corpus (BNC) and the Corpus of Contemporary American English (COCA), as well as domain-specific corpora. These licensed corpora enrich the linguistic data available for analysis.

The Spanish Corpora

The Spanish Web corpus (esTenTen) is a comprehensive text corpus created from internet texts belonging to the TenTen corpus family with a target size of 10+ billion words in over 40 languages. It encompasses a subcorpora of European and American Spanish, sourced from web domains in their respective continents and a portion of the Spanish Wikipedia. The corpus includes texts from various top-level domains corresponding to different Spanish-speaking countries, such as Argentina, Bolivia, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, Guatemala, Honduras, Mexico, Nicaragua, Paraguay, Peru, El Salvador, Spain, Uruguay, Venezuela, and Wikipedia.

For linguistic investigations, the Wordlist tool in Sketch Engine enables users to generate frequency lists of various types in the Spanish Web corpus. This includes frequency lists of nouns, verbs, adjectives, and other parts of speech, as well as words with specific character patterns (beginning, ending, or containing certain characters) and words based on attributes such as word forms, tags, lemmas, and more. The Wordlist tool also provides three different frequency measures: frequency, frequency per million, and ARF (Average Relative Frequency).

In addition to its diverse sources and linguistic investigation tools, the Spanish Web corpus offers a specialized subset known as the Colombian Subcorpus, explicitly sourced from web domains in Colombia. This subset provides a focused collection of texts that accurately represents the Spanish language as used in Colombia.

We chose to utilize these Corpora based on the aforementioned information, as it demonstrated sufficient robustness for our study. The entire process encompassed multiple stages. Firstly, using Sketch Engine, we filtered the American Spanish Web 2011 corpus (esamTenTen11). Subsequently, Wordlist feature was employed to obtain word frequencies. Specifically, we focused on adjectives ending with the suffixes (-udo, -ano, -ble, -dizo),

which were relevant to our study. The minimum frequency threshold was set at 5. Furthermore, we narrowed down our analysis by filtering the subcorpus, specifically targeting the *Colombian_domain_.co*. After performing these filtering steps, we obtained a list of words along with their absolute frequency, indicating how many times each word was found in the corpus. Additionally, it included the frequency per million tokens, representing the relative frequency of each word in relation to the entire corpus or subcorpus. After downloading the generated list of words, an additional crucial step was undertaken. The list was carefully reviewed and verified by a native speaker to ensure that all the words included were indeed authentic and valid. By following this comprehensive procedure, it was aimed to maintain the accuracy and reliability of the data obtained for our study.

8.3 Pre-experiment norming task: offline acceptability judgment

As part of our research, we first conducted an acceptability judgment task before the lexical decision task, using the same methodology as Manouilidou (2007) and Manouilidou & Stockall (2014). Our goal was to explore how participants' conscious metalinguistic knowledge is related to their implicit knowledge. This task also helped us eliminate any items that resulted in unclear judgments, which improved the reliability of the online reaction time results.

We used an online survey platform called LimeSurvey to collect acceptability judgments. Participants rated items on a 1-5 Likert scale, and the task included three conditions: SemViol, CatViol, and grammatical items (Gramm). Based on previous research by Manouilidou (2007) and Manouilidou & Stockall (2014), and in line with our overall predictions for the lexical decision task, we anticipated an increase in acceptability ratings for the SemViol in comparison to CatViol.

Based on previous research by Lázaro et al. in 2015, we predicted that pseudo-words with low-productivity suffixes in the SemViol condition would have lower acceptance rates than those with high-productivity suffixes. By doing so, we were able to examine how the productivity of those suffixes affects acceptability judgments in Spanish.

Furthermore, we expanded the study by incorporating nouns as a new category. Our focus was on analyzing the acceptability ratings of affixes that attach to nouns and those that attach to verbs. We hypothesized that the acceptability patterns for both nouns and verbs would be similar.

We fit linear mixed effects models with Condition (CatViol, SemViol, Gramm) x Suffix (-ble, -dizo, -ano, -udo) as fixed effects and by-item and by-subject random intercepts. In our results, we observed the expected pattern of higher acceptability ratings for SemViol items compared to CatViol items, but this pattern was only evident for certain suffixes, -ble ($z=-8.72$, $p<.0001$) and -dizo ($z=-4.05$, $p=.0002$). Surprisingly, for the suffixes -ano and -udo, SemViol items were perceived as equally unacceptable as CatViol items, indicating that violating a semantic rule was just as unacceptable as violating a category. This finding is intriguing because -ano and -udo have different levels of productivity, as well as -ble and -dizo, suggesting that productivity alone cannot explain the observed pattern. However, both -ano and -udo are noun-attaching suffixes, which raised the possibility of differences between noun-attaching and verb-attaching affixes when it comes to violating post-decomposition rules. As these results did not provide conclusive evidence of a productivity effect, further investigation was warranted through an online lexical decision task to gain a better understanding of these findings.

It is important to note that the acceptability judgment task served as a preliminary step to our main experiment, which involved a lexical decision task. The acceptability judgment task acted as a "norming" task, allowing us to test and select a set of items for use in the subsequent lexical decision task.

8.4 Main Experiment: Lexical decision task

We conducted one online lexical decision task in Spanish to analyze the acceptability ratings of words and pseudowords. In addition to CatViol, SemViol, and Gramm items, non-words (NWs) and fillers were included in the study to compare the processing differences between pseudowords and NWs, as observed in previous behavioral studies (Manouilidou, 2007; Manouilidou & Stockall, 2014). NWs are expected to be easier to reject due to the absence of existing stems, allowing for a useful comparison with pseudowords that require the decomposition and post-decomposition evaluation of both the

stem and the affix.

Drawing from previous lexical decision tasks in Greek, English, and Slovenian (Manouilidou, 2007; Manouilidou et al., 2015; Manouilidou & Stockall, 2014; Ristić et al., 2022), as well as the aforementioned acceptability ratings, we anticipate higher accuracy for CatViol items compared to SemViol items. Moreover, we predict shorter response times (RTs) for CatViol items compared to SemViol items across the four suffixes (-ble, -dizo,-udo, -ano). Additionally, based on the results of the acceptability judgment task, where the effect described in the study by Lazaro et al. (2015) was not observed, we aim to confirm these findings. We expect shorter RTs and lower error rates for pseudo-words with SemViol using low-productivity suffixes (-dizo, -udo) compared to those with high-productivity suffixes (-ble, -ano), which in turn might diminish the relevant contrast. However, we hypothesize that there will be no significant differences in RTs and error rates between noun-attaching and verb-attaching suffixes.

8.5 Participants

This study included 60 Native Colombian Spanish speakers between the ages of 18 and 66 years old, who volunteered to participate. Of the 60 participants, 36 were male and 24 were female. The mean age of the participants was 31.5 years old (SD = 11.9). Participants had diverse backgrounds and levels of education. Recruitment was conducted online through social media platforms and online forums, and all participants provided informed consent prior to taking part in the study. Prior to the experiment, participants were provided with an explanation of the study's purpose and detailed instructions on how to participate.

8.6 Procedure

Before the experiment, participants received a brief explanation of the experiment's aims. They also completed a language questionnaire, providing information about their native language(s), country of birth, and age. The experiment was created using Psychopy and was conducted online through Pavlovia. Participants were randomly assigned to one of the two lists (see Materials). The task involved presenting letter strings in the middle of the screen. Participants had to determine whether each string represented a real word in their

native language by pressing the appropriate arrow key on the keyboard (left arrow for "no" and right arrow for "yes"). The experiment lasted approximately 15-20 minutes.

Before the main task, participants were presented with a consent form, instructions, and 10 practice trials. The experiment consisted of 310 trials per list. Each trial began with a 300 ms fixation cross, followed by the stimulus displayed for 2 seconds or until the participant responded. A blank screen appeared between stimuli, with a random duration ranging from 300 to 500 ms. Following the completion of the experiment, participants were asked additional questions regarding their understanding of the task and word building rules, for us to have some insights into whether there were any potential strategies employed in word processing or not.

9. Analysis

The initial steps involved removing skipped trials where participants did not respond using the designated "left" or "right" keys. Next, we converted response times from seconds to milliseconds for easier analysis. Fillers were excluded from the analysis. We then focused on grammatical conditions and calculated item-level accuracy across the participants. Any items with an accuracy below 50% were removed from the dataset. Given the relatively low frequency of some of the existing items, especially the one with low productivity suffixes, we believed that this threshold provided a sufficiently good filter. There were 28 grammatical items removed. Further scrutiny was applied to participants' performance on grammatical items. Participants with an accuracy on grammatical items below 70% were excluded from subsequent analyses. 13 participants were removed in total. Response times below 200 milliseconds were considered too short and were therefore trimmed from the dataset, resulting in 0.1% of data being removed. All analyses were performed using R.

Linear mixed-effects models analysis (for RTs) and generalized linear mixed-effects models (for accuracy) in the *lme4* package in R were fit (R Core Team, 2017). The model used had RTs or accuracy as the dependent variable, and Condition (SemViol, CatViol, Gramm) X Prefix (ble, dizo, udo, ano) as fixed factors. By-subject and by-item random slopes and intercepts were included by default. We progressively increased model complexity following Barr, Levy, Scheepers, & Tily (2013) to choose the best random

effects structure. We started with the basic model, coded as *model_acc_basic<- glmer (response.corr ~ suffix * condition + (1|participant)+ (1|stimuli), family=binomial)* for accuracy, and *model_RT_basic<- lmer (RT ~ suffix * condition + (1|participant)+ (1|stimuli), REML=FALSE) for RTs.*

We performed model comparisons using the ANOVA function (e.g. *anova (model_acc_1, model_acc_2)*). If the less complex model showed better goodness of fit, we proceeded to make the model more complex further. The resulting most complex converging model was *model_acc_3a<- glmer (response.corr ~ suffix * condition + (1+condition|participant)+ (1|stimuli), data = d3, family=binomial, control = glmerControl(optimizer = "bobyqa"))* for accuracy, and the basic model (above) for RTs.

Paired comparisons were performed using the *emmeans* function in R, and following the formula: *pairs (emmeans (model_acc1, "Condition", by = "Suffix"))*. The default Tukey adjustment method for the p-value was applied. The effects across the analyses were considered significant when t/z values were above 2.

10. Results

Accuracy

Condition	Mean	SD
CatViol	**0.86**	**0.35**
Gramm	0.86	0.35
NonWords	0.89	0.31
SemViol	**0.69**	**0.46**

Table 1: Averaged Response Accuracy by Condition

The results of the analysis reveal interesting patterns in the accuracy of responses to pseudowords under different conditions. The average accuracy rates for the CatViol and Gramm conditions were found to be comparable, suggesting that participants performed similarly when faced with these two conditions. In contrast, the SemViol condition showed a lower average accuracy rate, indicating that participants encountered more difficulties in

accurately responding to pseudowords in this condition. On the other hand, the NonWords condition exhibited the highest average accuracy rate, suggesting that participants had the highest level of success in accurately responding to pseudowords in this particular condition

Condition	Suffix	Mean	SD
CatViol	ano	0.85	0.35
	ble	0.86	0.34
	dizo	0.89	0.32
	udo	0.84	0.37
Gramm	ano	0.85	0.36
	ble	0.9	0.3
	dizo	0.87	0.34
	udo	0.81	0.4
NonWords	ano	0.88	0.33
	ble	0.92	0.28
	dizo	0.87	0.33
	udo	0.9	0.3
SemViol	ano	0.86	0.35
	ble	0.38	0.49
	dizo	0.74	0.44
	udo	0.81	0.39

Table 2 : Averaged Response Accuracy by Condition and suffix

In the CatViol condition, participants demonstrated consistently high mean accuracy across all suffixes. The -ano suffix achieved a mean accuracy of 0.85, while the -ble, -dizo, and -udo suffixes had mean accuracy of 0.86, 0.89, and 0.84, respectively. Within the Gramm condition, most suffixes exhibited similar mean accuracy levels. The -ano and -dizo suffixes achieved mean accuracies of 0.85 and 0.87, respectively. The -ble suffix showed a slightly lower mean accuracy of 0.81. The -udo suffix had a mean accuracy of 0.9.In the NonWords condition, participants demonstrated consistently high mean accuracy across all suffixes. The -ano suffix had a mean accuracy of 0.88, while the -ble -dizo and -udo suffixes achieved mean accuracies of 0.92, 0.87, and 0.9, respectively. Interestingly, the SemViol condition exhibited variations in mean accuracy across different suffixes. The -ano suffix had a mean accuracy of 0.86, while the -ble, -dizo, and -udo suffixes yielded

lower mean accuracies of 0.38, 0.74, and 0.81, respectively. Notably, the -ble suffix showed the lowest mean accuracy compared to the other suffixes. These findings suggest that participants faced challenges in processing pseudowords with semantic violations, with different suffixes eliciting varying degrees of accuracy and variability in individual processing and response patterns.

	Estimate Std.	Error	z value	Pr(>|z|)
(Intercept)	2.25323	0.25507	8.834	**< 2e-16 *****
suffixble	0.08952	0.29324	0.305	0.76014
suffixdizo	0.3263	0.29744	1.097	0.27262
suffixudo	-0.13162	0.28778	-0.457	0.6474
conditionGramm	-0.08622	0.37241	-0.232	0.81691
conditionNonWords	0.25765	0.39716	0.649	0.5165
conditionSemViol	-0.06599	0.30459	-0.217	0.82848
suffixble:conditionGramm	0.37773	0.41589	0.908	0.36375
suffixdizo:conditionGramm	-0.14365	0.44791	-0.321	0.74843
suffixudo:conditionGramm	-0.28566	0.41865	-0.682	0.49502
suffixble:conditionNonWords	0.11128	0.54939	0.203	0.83949
suffixdizo:conditionNonWords	-0.37708	0.55089	-0.684	0.49367
suffixudo:conditionNonWords	0.25326	0.54639	0.464	0.64299
suffixble:conditionSemViol	-2.90233	0.41199	-7.045	**1.86e-12 *****
suffixdizo:conditionSemViol	-1.19459	0.41593	-2.872	**0.00408 ****
suffixudo:conditionSemViol	-0.28449	0.41248	-0.69	0.49038

Table 3: The output of the best Fitting Model for Accuracy Analysis

The table 3 provides the estimated coefficients, standard errors, z-values, and p-values for the best-fitting model for the accuracy analysis. Significance codes indicate the level of statistical significance: " (<0.001), " (<0.01), " (<0.05), '.' (<0.1), and ' ' (≥0.1). The interaction term for the suffix -ble and semantic violation condition is highly significant (p = 1.86e-12. z = -7.045), as well as between the suffix -dizo and the semantic violation condition (p = 0.00408, z = -2.872), suggesting a strong interaction effect between Suffix and Condition factors. To get a better look at the results, it is beneficial to perform paired comparisons and *emmeans* to get a better look at the results. The outputs of the paired comparisons between condition are presented below.

Contrast	Estimate	SE	df	z.ratio	p.value
CatViol - Gramm	-0.2915	0.357	Inf	-0.817	0.8464
CatViol - NonWords	-0.3689	0.403	Inf	-0.916	0.7962
CatViol - SemViol	**2.9683**	**0.295**	**Inf**	**10.050**	**<.0001**
Gramm - NonWords	-0.0774	0.423	Inf	-0.183	0.9978
Gramm - SemViol	3.2598	0.331	Inf	9.840	<.0001
NonWords - SemViol	3.3373	0.393	Inf	8.489	<.0001

Table 4: Paired comparisons for the accuracy analysis for suffix -ble

The analysis of the -ble suffix shows significant accuracy differences. No significant differences were observed between CatViol and Gramm, CatViol and NonWords, and Gramm and NonWords. However, significant differences were found between CatViol and SemViol ($p < .0001$, $z = 10.050$), Gramm and SemViol ($p < .0001$, $z = 9.840$), and NonWords and SemViol ($p < .0001$, $z = 8.489$).

Contrast	Estimate	SE	df	z.ratio	p.value
CatViol - Gramm	0.2299	0.394	Inf	0.584	0.9369
CatViol - NonWords	0.1194	0.403	Inf	0.296	0.9910
CatViol - SemViol	**1.2606**	**0.303**	**Inf**	**4.159**	**0.0002**
Gramm - NonWords	-0.1104	0.450	Inf	-0.245	0.9948
Gramm - SemViol	1.0307	0.370	Inf	2.785	0.0274
NonWords - SemViol	1.1412	0.394	Inf	2.899	0.0196

Table 5: Paired comparisons for the accuracy analysis for suffix -dizo

The analysis of the -dizo suffix reveals significant differences in processing across the examined conditions. No significant differences were found between CatViol and Gramm, CatViol and NonWords, Gramm and NonWords, and Gramm and SemViol. However, significant differences were observed between CatViol and SemViol ($p = 0.0002$, $z= 4.159$) and between NonWords and SemViol ($p = 0.0196$, $z = 2.899$).

Contrast	Estimate	SE	df	z.ratio	p.value
CatViol - Gramm	0.0862	0.372	Inf	0.232	0.9956
CatViol - NonWords	-0.2577	0.397	Inf	-0.649	0.9160
CatViol - SemViol	**0.0660**	**0.305**	**Inf**	**0.217**	**0.9964**
Gramm - NonWords	-0.3439	0.438	Inf	-0.785	0.8614
Gramm - SemViol	-0.0202	0.364	Inf	-0.056	0.9999
NonWords - SemViol	0.3236	0.402	Inf	0.805	0.8521

Table 6: Paired comparisons for the accuracy analysis for suffix -ano

Contrast	Estimate	SE	df	z.ratio	p.value
CatViol - Gramm	0.3719	0.359	Inf	1.036	0.7284
CatViol - NonWords	-0.5109	0.398	Inf	-1.285	0.5728
CatViol - SemViol	**0.3505**	**0.298**	**Inf**	**1.176**	**0.6422**
Gramm - NonWords	-0.8828	0.429	Inf	-2.058	0.1671
Gramm - SemViol	-0.0214	0.346	Inf	-0.062	0.9999
NonWords - SemViol	0.8614	0.399	Inf	2.157	0.1355

Table 7: Paired comparisons for the accuracy analysis for suffix -udo

For both -udo and -ano suffixes, the analysis indicates that there were no significant differences in the accuracy across the conditions.

Reaction Times (RTs)

Condition	Mean	SD
CatViol	**1135.08**	**310.36**
Gramm	998.19	288.14
NonWords	1096.82	312.46
SemViol	**1154.54**	**320.23**

Table 8: Averaged RTS by Condition

Table 8 uncovers striking patterns in response times (RTs) across different conditions. The CatViol and Gramm conditions exhibited similar average RTs, indicating comparable performance by participants in these conditions. In contrast, the SemViol condition displayed a slightly higher average RT, suggesting that participants faced increased

challenges in responding to stimuli in this condition. The NonWords condition stood out with the lowest average RT.

Condition	Suffix	Mean	SD
CatViol	ano	1117.74	306.59
	ble	1123.91	306.60
	dizo	1174.52	309.07
	udo	1125.56	316.36
Gramm	ano	980.95	276.75
	ble	975.47	288.43
	dizo	1029.38	280.52
	udo	1019.83	299.42
NonWords	ano	1088.08	307.26
	ble	1098.68	316.53
	dizo	1115.12	315.93
	udo	1085.62	310.23
SemViol	ano	1112.24	301.77
	ble	1133.34	325.44
	dizo	1220.76	319.90
	udo	1153.38	323.59

Table 9: Averaged RTS (in milliseconds) by Condition and suffix

Table 9 shows that participants in the CatViol condition exhibited varying mean reaction times (RTs) across different suffixes. The -dizo suffix had the highest mean RT (1174.52 ms), while the -ano suffix had the lowest (1117.74 ms). Similarly, in the SemViol condition, mean RTs showed differences across suffixes, with the -dizo suffix having the highest mean RT (1220.76 ms) and the -ble suffix having the lowest (1133.34 ms). Overall, participants took longer to process SemViol compared to CatViol.

Variable	Estimate	Std. Error	t value
(Intercept)	1135.4785	27.3977	41.444
suffixble	-9.2940	28.4581	-0.327
suffixdizo	41.3522	28.4049	1.456
suffixudo	8.4631	28.2579	0.299
conditionGramm	-159.0809	29.8261	-5.334
conditionNonWords	-35.9887	37.1772	-0.968
conditionSemViol	-10.3208	28.6950	-0.360
suffixble:conditionGramm	-4.0680	40.3224	-0.101
suffixdizo:conditionGramm	-0.6366	43.3077	-0.015
suffixudo:conditionGramm	25.5560	41.4906	0.616
suffixble:conditionNonWords	13.9036	52.5919	0.264
suffixdizo:conditionNonWords	-9.4313	52.7415	-0.179
suffixudo:conditionNonWords	-19.3271	52.5014	-0.368
suffixble:conditionSemViol	82.4951	43.1725	1.911
suffixdizo:conditionSemViol	70.2176	41.0040	1.712
suffixudo:conditionSemViol	31.9515	40.7447	0.784

Table 10: The output of the best Fitting Model for RTs Analysis

Table 10 shows the estimates, standard errors, and t-values for the best fitting model in the RT analysis. The results of the analysis revealed a significant effect of the Condition variable (t = -5.334) on the dependent variable, when compared to the reference level CatViol. As for accuracy, the coefficients for suffix -ble, suffix -dizo, and -suffix -udo were not statistically significant (ts < 1.46), indicating no significant main effects of the Suffix factor when the reference level is -ano.

However, interactions between certain variables demonstrated significant effects. The interaction between suffix -ble and the semantic violation condition showed a statistically significant effect (t = 1.911). Similarly, the interaction between suffix -dizo and the semantic violation condition displayed a significant effect (t = 1.712), as well as the interaction suffix -udo and semantic violation condition (t = 0.784). To get a better look at the results and understand which contrasts drive the interactions, it is beneficial to perform paired comparisons and *emmeans* to get a better look at the results. The outputs of the paired comparisons per condition are presented below.

contrast	estimate	SE	df	z.ratio	p.value
CatViol - Gramm	163.1	27.1	Inf	6.011	<.0001
CatViol - NonWords	22.1	37.2	Inf	0.594	0.9340
CatViol - SemViol	**-72.2**	**32.3**	**Inf**	**-2.237**	**0.1133**
Gramm - NonWords	-141.1	36.0	Inf	-3.919	0.0005
Gramm - SemViol	-235.3	30.9	Inf	-7.620	<.0001
NonWords - SemViol	-94.3	40.0	Inf	-2.356	0.0857

Table 11: Paired comparisons RTs analysis for suffix -ble

Table 11 presents the statistical analysis results for reaction times (RTs) associated with the -ble suffix. Significant differences were observed between CatViol and Gramm (163.1, $p < .0001$, $z = 6.011$), Gramm and NonWords (-141.1, $p = 0.0005$, $z = -3.919$), and Gramm and SemViol (-235.3, $p < .0001$, $z = -7.620$). No significant differences were found between CatViol and NonWords, and NonWords and SemViol. In the CatViol vs SemViol, the z ratio for these observed differences was above 2 (t= -2.237), indicating a strong difference. Additionally, there was an obvious numeric trend in the reaction times, further supporting the contrast.

contrast	estimate	SE	df	z.ratio	p.value
CatViol - Gramm	159.7	31.4	Inf	5.086	<.0001
CatViol - NonWords	45.4	37.4	Inf	1.214	0.6179
CatViol - SemViol	**-59.9**	**29.3**	**Inf**	**-2.045**	**0.1716**
Gramm - NonWords	-114.3	39.6	Inf	-2.887	0.0203
Gramm - SemViol	-219.6	32.0	Inf	-6.858	<.0001
NonWords - SemViol	-105.3	37.9	Inf	-2.777	0.0281

Table 12: Paired comparisons RTs analysis for suffix -dizo

Table 12 presents the results of the statistical analysis for reaction times (RTs) associated with the -dizo suffix. Statistically significant differences were observed between CatViol and Gramm (159.7, $p < .0001$, $z = 5.086$), Gramm and NonWords (-114.3, $p = 0.0203$, $z = -2.887$), and Gramm and SemViol (-219.6, $p < .0001$, $z = -6.858$). No significant differences were found between CatViol and NonWords , and NonWords and SemViol. As

for -ble, the z-ratio was above 2 for the SemViol vs. CatViol contrast (t= -2.045), but the p-value did not not survive correction and was above the significance threshold.

contrast	estimate	SE	df	z.ratio	p.value
CatViol - Gramm	159.1	29.8	Inf	5.334	<.0001
CatViol - NonWords	36.0	37.2	Inf	0.968	0.7677
CatViol - SemViol	**10.3**	**28.7**	**Inf**	**0.360**	**0.9841**
Gramm - NonWords	-123.1	38.4	Inf	-3.203	0.0074
Gramm - SemViol	-148.8	30.3	Inf	-4.910	<.0001
NonWords - SemViol	-25.7	37.6	Inf	-0.683	0.9034

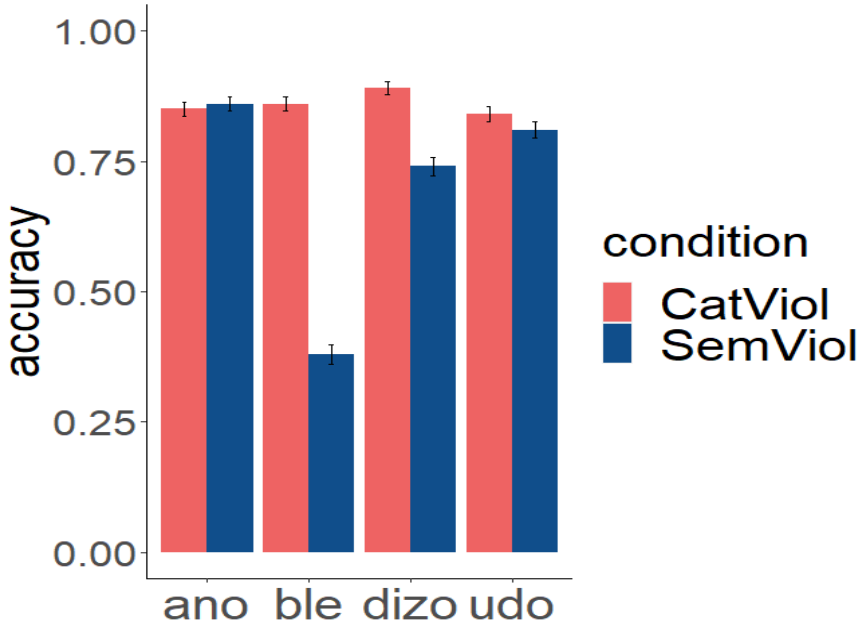
Table 13: Paired comparisons RTs analysis for suffix -ano

Table 13 presents the results of the statistical analysis for reaction times (RTs) associated with the -ano suffix. Significant differences were observed between CatViol and Gramm (159.1, $p < .0001$, $z = 5.334$), Gramm and NonWords (-123.1, $p = 0.0074$, $z = -3.203$), and Gramm and SemViol (-148.8, $p < .0001$, $z = -4.910$). No significant differences were found between the other conditions.

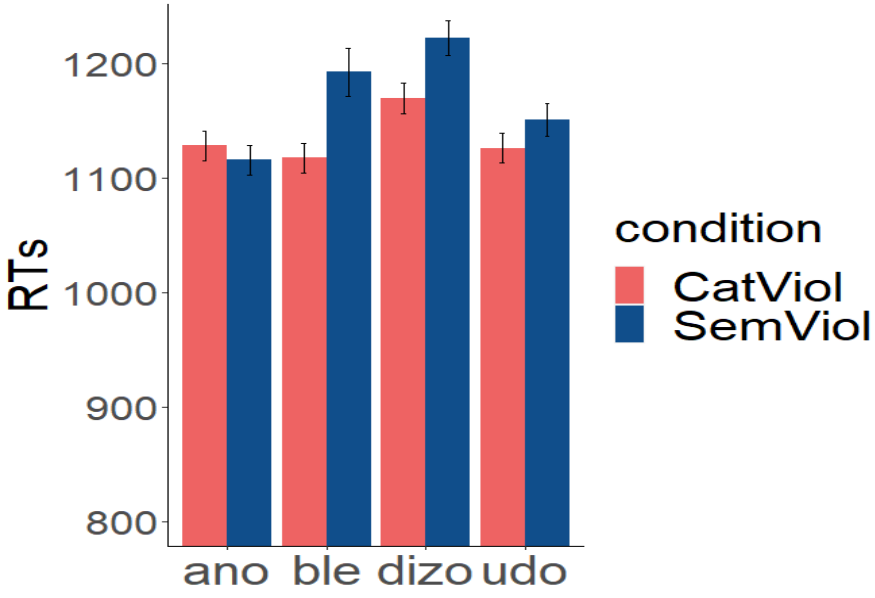
contrast	estimate	SE	df	z.ratio	p.value
CatViol - Gramm	133.5	28.8	Inf	4.628	<.0001
CatViol - NonWords	55.3	37.1	Inf	1.492	0.4422
CatViol - SemViol	**-21.6**	**28.9**	**Inf**	**-0.748**	**0.8776**
Gramm - NonWords	-78.2	37.5	Inf	-2.087	0.1574
Gramm - SemViol	-155.2	29.5	Inf	-5.268	<.0001
NonWords - SemViol	-76.9	37.5	Inf	-2.050	0.1700

Table 14: Paired comparisons RTs analysis for suffix -udo

Table 14 presents the results of the statistical analysis for reaction times (RTs) associated with the -udo suffix. Statistically significant differences were observed between CatViol and Gramm (133.5, $p < .0001$, $z = 4.628$) and Gramm and SemViol (-155.2, $p < .0001$, $z = -5.268$). No significant differences were found between the other interactions.



Plot 1: Results accuracy



Plot 2: Results RTs

11. Discussion

In the present study, we investigated the evaluation of semantic well-formedness in comparison to categorial well-formedness, as well as the effects of affix productivity and base category on morphological processing, specifically focusing on accuracy and reaction times (RTs) associated with different suffixes. Our objective was to contribute to the existing literature and enhance our understanding of the complex processes involved in morphological processing. Additionally, we sought to expand the scope of research by investigating a yet unattested language within this research line, Spanish, and a generally under-studied dialect of Spanish, namely Colombian Spanish. To achieve this, we investigated the dissociation of effects related to the two distinct post-decomposition processes in visual word recognition, namely licensing and semantic composition (Schreuder & Baayen, 1995). Drawing from previous behavioral studies (Manouilidou, 2007; Manouilidou et al., 2015; Manouilidou & Stockall, 2014), we adopted a paradigm that compared pseudowords violating Categorial rules and Semantic rules of affix attachment.

Comparison to Previous Studies in this Research Line

The study's findings support previous research and demonstrate a cross-linguistic pattern. Participants faced greater difficulties in providing accurate responses to pseudowords in the SemViol condition compared to the CatViol condition for -ble and -dizo. This disparity was evident in the lower average accuracy rate observed in the SemViol condition. However, the most intriguing discovery is the absence of this effect in pseudowords with suffixes attached to nouns, such as -udo and -ano. This finding aligns with the results obtained in Bangla (Moitra et al., 2022), where the expected effect was not observed for denominal complex pseudowords.

When comparing the response times (RTs) of participants in the SemViol and CatViol conditions, it was found that, on average, participants took longer to respond in the SemViol condition. A closer examination of the mean for the CatViol vs. SemViol interaction condition revealed significant differences, particularly for the -ble suffix. Processing times were notably longer in the SemViol condition compared to the CatViol condition. Similar results were observed for the -dizo suffix, with significant differences in

RTs detected. It is important to note that although the z ratio for the SemViol vs. CatViol contrast for the -ble suffix indicated a potential effect by exceeding the usual threshold of 2, the corresponding p-value in the post-hoc paired comparisons did not withstand correction and did not reach the significance level. Also, the model output suggested significant interactions involving these suffixes, suggesting that the effect is present prior to adjustment for multiple comparisons. Nevertheless, these findings are supported by numerical differences in RTs across the conditions, suggesting that semantic violations lead to prolonged processing times compared to category violations. Importantly, similar to the findings for accuracy, the analysis revealed no significant differences in response times between the SemViol and CatViol conditions for the -udo and -ano suffixes. These suffixes, which are added to nouns, did not exhibit any notable variations in processing times between these two conditions.

The aforementioned findings, intriguingly, unveiled a notable reliance on the base category, with noun-attaching affixes not displaying the expected CatViol vs. SemViol contrast. Specifically, violations of CatViol and SemViol were found to be equally detrimental in terms of acceptability, response times, and accuracy for -udo and -ano, suggesting that the specific suffix used does not significantly influence the processing of derived words. Instead, the primary factors in understanding and processing derived words are the underlying base category and the affix restrictions, both categorial and semantic. This perspective is further supported by a study conducted by Ristić et al. (2022) on South Slavic languages, where violations of purely semantic (rather than argument structure) restrictions resulted in longer rejection times compared to category violations in both Slovenian and BCS. This emphasizes the crucial role of evaluating semantic well-formedness in the processing of complex words, as a separate sub-process from evaluating category well-formedness.

Contrasting Affix Productivity Studies

In the context of our study, another important factor we sought to investigate was affix productivity in the processing of morphological structures. This investigation was motivated by previous studies (Burani and Thornton, 2003; Plag and Baayen, 2009; Lazaro et al., 2016) that had demonstrated the potential influence of affix productivity on morphological processing. Contrary to the findings of previous studies and specifically the

one in Spanish by Lazaro et al. (2015), our study did not reveal a significant effect of affix productivity on morphological processing within the dialect of Colombian Spanish. Despite our deliberate selection of suffixes with substantial disparities in terms of high and low productivity for both verbs and nouns, we found no compelling evidence of a pronounced effect.

Models of morphological processing

Our findings have important implications for models of morphological processing, particularly concerning the analysis of complex words derived from verbs and nouns. The model proposed by Schrauder and Bayen in 1995 receives support when studying complex words derived from verbs. However, we did not observe the same effect for complex words derived from nouns. This suggests that the evaluation of category and semantic well-formedness for noun+suffix combinations may involve either two distinct processes leading to the same outcome or a single process that considers both aspects simultaneously. To arrive at more definitive conclusions, it is imperative to conduct further research utilizing techniques that enable precise temporal dissociation of the two subprocesses.

Besides, in our research, our primary focus was on emphasizing the significance of semantics when processing pseudowords, with less emphasis on argument structure considerations. Additionally, the contribution of our research lies in the fact that by incorporating semantic restrictions, we gain a more direct understanding of the semantic composition subprocess. Previous studies primarily relied on argument structure considerations to draw conclusions about the semantic composition process. However, argument structure can be seen as residing somewhere between syntax and semantics (Bresnan, 2015). By using only semantic rules, we can confidently and clearly state that the evaluation focuses on the way meanings are combined.

Finally, in relation to the subprocesses described in Schrauder and Baayen's model, our study indicates that suffix productivity did not have a notable effect. This suggests that regardless of productivity, the differentiation between CatViol and SemViol remains consistent, indicating no significant variation.

NWs vs other stimuli

Comparing non-words (NWs) with other types of stimuli, such as grammatical words, categorial violations (CatViol), and semantic violations (SemViol), yielded interesting findings. In terms of accuracy, there were no significant differences between NWs and grammatical words, regardless of the suffix (-udo, -ano, -dizo, -ble). This suggests that participants performed equally well in distinguishing between these two types of stimuli, despite NWs being easily identifiable as non-words. However, when considering reaction times, grammatical words were processed faster than NWs for most suffixes, except for the -udo suffix, This exception could be attributed to the relatively low productivity of -udo in the Spanish language. The limited exposure, frequency, and familiarity of words containing this particular suffix may have contributed to an overall increase in reaction times for grammatical items, thereby making them comparable to the other conditions. The processing speed advantage for grammatical words aligns with expectations, as participants were able to differentiate between plausible words and uninterpretable NWs.

Comparing CatViol and NWs, no significant differences were observed in terms of accuracy or reaction times across the examined suffixes. This indicates a consistent processing pattern for both types of stimuli, supporting previous research (Manouilidou & Stockall, 2014)

In the comparison between SemViol and NWs, accuracy and reaction times varied across different suffixes. For the -ano and -udo suffixes, NWs performed as accurately as SemViol, while for the -ble and -dizo suffixes, NWs showed lower accuracy. In terms of reaction times, NWs exhibited similar speed to SemViol for most suffixes, except for the -dizo suffix, where NWs had significantly slower reaction times.

Overall, the results suggest that participants' ability to differentiate between NWs and other stimuli is influenced by various factors, including the visual appearance of the stimuli and their adherence to phonotactic rules. The adherence to phonotactic rules of NWs may have led to their recognition as novel words, potentially contributing to comparable accuracy and reaction times for the pseudowords. Additionally, the challenge in decomposing NWs might explain the slower reaction times observed for certain suffixes.

Gramm vs CatViol

We compared the Gramm and CatViol conditions. As Gramm words are existing words, they should be processed quickly and more accurately. Surprisingly, we didn't find any significant differences in accuracy between the two conditions for any of the examined suffixes. However, by looking at reaction times, we consistently found that the Gramm condition had significantly faster RTs compared to the CatViol condition for all the examined suffixes. In summary, the CatViol condition consistently took longer to be rejected for each of the suffix conditions we studied. Therefore, if the participants take longer to process categorial violations, it indicates that they are indeed decomposing and actively evaluating these stimuli to determine their status as words or non-words.

Does the effect found in -ble depend purely on argument structure?

We additionally investigated the behavior of the suffix -ble, which drew our attention given the amount of research dedicated to it in Spanish, and the complexity of the attachment rules. Given that it is classified as a high-productivity suffix, one would expect participants to provide accurate answers when they are presented with pseudowords. Surprisingly, we discovered that its accuracy was unexpectedly low, even in comparison to the other verb-attaching suffix, -dizo. Intrigued by this inconsistency, we decided to investigate if the observed effects could be attributed to the argument structure after all.

To delve deeper into this phenomenon, we took into account two relevant studies. The first study, conducted by Oltra-Massuet (2014), proposes two crucial conditions for a verb to be eligible for the addition of the -ble suffix: the presence of an internal argument (which is not always explicitly stated) and an implication of an originator. Consequently, intransitive verbs generally do not meet these conditions. The second by Albano et al. (2021) focused on the compatibility of the -ble suffix with unaccusative and unergative verbs. As stated before, these types of verbs were not expected to generate forms with the -ble suffix. By analyzing data from electronic sources, real-life situations, and native speakers' judgments, the researchers argued that the participles of unaccusative verbs, akin to those of transitive verbs, could function as noun modifiers resembling adjectives. Notably, they could form constructions with the neuter article "lo" (l*o vivido* 'what was lived' → lo **vivible* 'what can be lived').

Additionally, the study conducted by Oltra-Massuet (2014) revealed that the function of "lo" within an internal argument acts as the head of a noun phrase, meeting the requirements of the -ble suffix. Moreover, this research demonstrated that certain derived adjectives ending in -ble (which are pseudowords) such as **ocurrible* (‘occur-able’) and **estornudable* (‘sneeze-able’) exhibit a semantic interpretation akin to other adjectives sharing the same suffix. These findings indicate a common understanding or meaning among these adjectives, suggesting that the conditions for -ble derivation are influenced by external factors, particularly the contextual structure or syntax. Oltra-Massuet (2014) say that one possible explanation for this phenomenon is that these verbs inherently possess a "degree argument," as referred to by the researcher. An illustration of this is found in the verb *dormir* (‘to sleep’), where we can create expressions such as *Durmió todo lo dormible* (‘He slept all that is sleepable’), despite the fact that **dormible* (‘sleep-able’) does not actually exist as a word. In this example, the degree argument could be interpreted as "the amount which can be slept," although it remains implicit.

Building upon the aforementioned studies, we proceeded with an analysis of our selected verbs, which formed the basis for generating semantic violations for -ble (and -dizo). We categorized these verbs based on their argument structure, including the “degree argument”, encompassing verbs (9) capable of functioning as both transitive and intransitive, unaccusative verbs (10) featuring a degree internal argument ("argumentyes"), and unaccusative verbs (10) devoid of any detectable internal argument ("no").

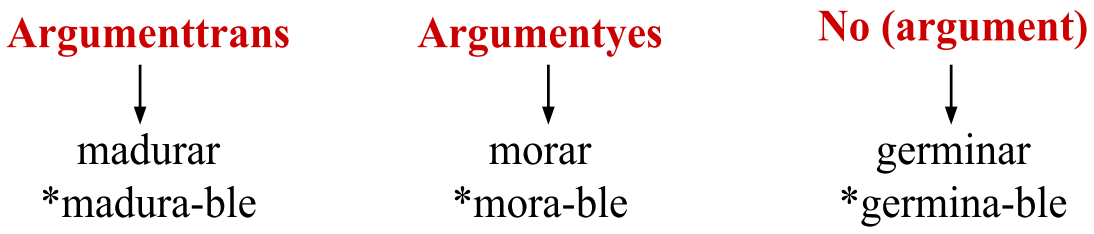


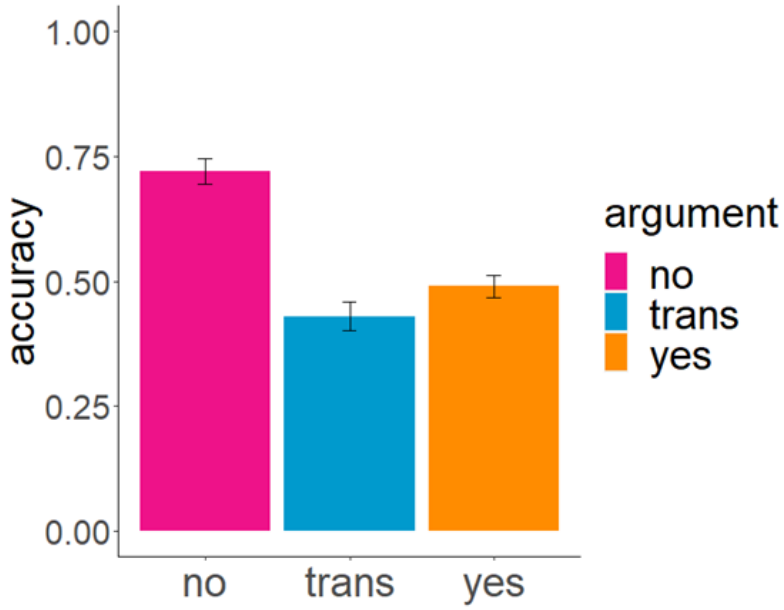
Fig. 6: Categorization arguments structure -ble

Statistical analysis was performed both on RTs and accuracy, with the verb type (argumenttrans, argumentyes, no argument) as factor.

The analysis revealed significant findings regarding the influence of the "argument" variable on both accuracy and reaction times (RTs). In terms of accuracy, the "argumenttrans" and "argumentyes" levels demonstrated negative coefficients ($p < 0.001$, $z = -3.904$ and $p < 0.001$, $z = -3.28$, respectively), indicating a significant decrease compared to the reference level ("no argument"). These results highlight the importance of considering the "argument" variable, as it strongly influences the outcome with lower accuracy values associated with the "trans" and "yes" levels. Further examination of paired comparisons revealed a significant difference between the "no" level of the "argument" variable and both the "trans" and "yes" levels ($p < 0.05$, $z = 0.09$, $z = 0.923$, respectively). However, no significant difference is observed between the "trans" and "yes" levels.

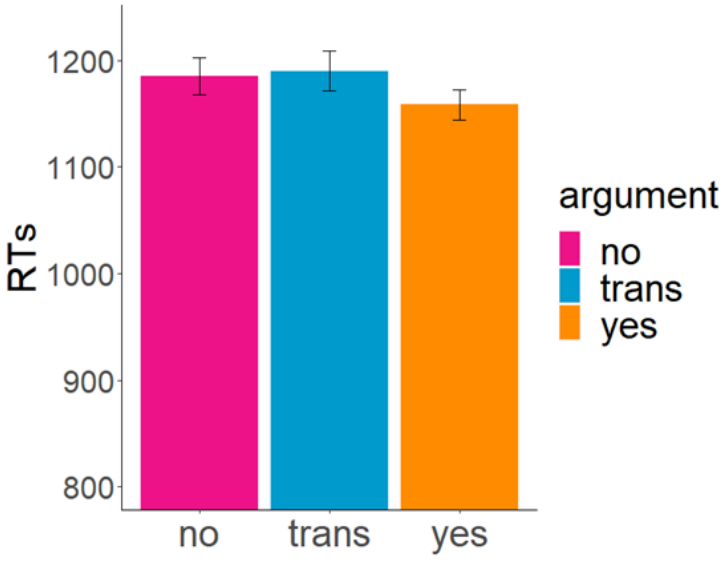
On the other hand, when considering RTs, the impact of the "argumenttrans" and "argumentyes" levels is not statistically significant compared to the reference level ("no argument"). Further analysis of paired comparisons confirms the lack of statistical significance in the observed differences in RTs among the compared levels ("no argument", "argumenttrans" and "argumentyes").

The findings of this analysis suggest that the accuracy in identifying pseudowords containing unaccusative verbs without any internal argument was significantly higher compared to the other two categories. This indicates that these pseudowords with argumentless verbs were more easily recognized as non existing, when combined with -ble, suggesting that the argument structure information plays an important role in the well-formedness evaluation. On the other hand, there were no noticeable differences in response times across the three categories, suggesting that it took a similar amount of time to reject all types of pseudowords. This implies that there might be another factor at play that made the decision equally difficult for all the pseudowords, and this factor could be related to semantics, i.e. the semantic restriction that the SemViol items were primarily based on in the current study.



.

Plot 3: Results accuracy argument analysis -ble



Plot 4: Results RTs argument analysis -ble

Therefore, it has become evident that both argument structure and semantics play influential roles, particularly in the context of verbs (the argument restrictions don't exist for nouns). Finally, it remains an open question whether these observed effects can be solely attributed to the base category or if they are more closely linked to the specific semantic restrictions imposed by these particular suffixes. In our study we have delved into a specific range of semantic restrictions. In the case of -ble and -dizo, for both we primarily focused on verbs of appearance and existence. In the case of -udo and -ano we introduce additional semantic restrictions of violating noun meanings (nouns not referring to places or people, nouns not referring to body parts). However, we acknowledge that our exploration of semantic restrictions has been limited in scope. It is crucial to recognize that there may exist other relevant ones that we did not investigate, which could have yielded different results. Therefore, it is essential to acknowledge the potential impact of unexplored semantic restrictions on our findings. By considering a broader array of semantic restrictions, future studies can provide a more comprehensive understanding.

Conclusion

In conclusion, our study investigated how Spanish speakers process complex words with different suffixes when rule violations occur. We discovered that violations of semantic rules for verb-attaching affixes posed greater difficulty in understanding the meaning of pseudowords, leading to lower accuracy and longer response times compared to violations of categorial rules. This emphasizes that this separate evaluation subprocess (separate from category evaluation) is related to semantics in the processing of complex words. Additionally, we observed that argument structure, primarily present in verb-attaching affix -ble, also might have influenced processing by determining constituent compatibility. Surprisingly, when analyzing affixes attached to nouns, both types of violations had similar effects, indicating that the base category of the word had a more significant impact on processing than the specific suffix used. It appears that both noun category and semantic violations are equally detrimental. This novel finding contributes to the existing research, which previously concentrated on verbs, and raises intriguing questions about word category processing. Additionally, it is important to note that our study differs from previous research by not finding a significant impact of suffix productivity on morphological processing. These findings deepen our understanding of how Spanish speakers analyze and comprehend complex words, emphasizing the importance of considering semantic well-formedness and base category in morphological processing.

References

Albano, H., Vaninetti, A., Crespo, C., & Belart, N. G. (2021, June 15). 2021-06-15. Un problema de comportamiento verbal: Verbos inergativos que se comportan como inacusativos e inacusativos como inergativos. Ideas. https://p3.entendiste.ar/index.php/ideas/article/view/5605

Baayen, Harald & Schreuder, Robert. (2006). Morphological Processing. doi: 10.1002/0470018860.s00254.

Baayen, R. H., & Lieber, R. (1991). Productivity and English derivation: A corpus-based study. Linguistics, 29(5), 801-844.

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. Journal of Memory and Language, 55(2), 290-313.

Bauer, L. (1983). English word-formation. Cambridge University Press.

Bauer, L. (2001). Morphological Productivity. Cambridge: Cambridge. University Press.

Bauer, L. (2015). The Oxford handbook of morphology. Oxford University Press.

Bojana Ristić, Karin Kavčič, Rok Žaucer, Linnaea Stockall, Christina Manouilidou. Processing morphologically complex words in South Slavic. Poster presentation at SinFonIJA 15 (Conference on Syntax, Phonology and Language Analysis), September 22–24 2022, University of Udine, Italy.

Booij, G. (2010). Construction morphology. Oxford University Press.

Bresnan, J. (2015). Lexicality and Argument Structure. *Lexical-Functional Syntax*, 21–38. https://doi.org/10.1002/9781119105664.ch3

Butterworth, B. (1983). Lexical representation. In Butterworth B.L. (ed.). Language Production, Vol. 2: Development, Writing and Other Language. Processes. London: Academic.

Bybee, J., & Moder, C. (2019). Morphological derivation. In Oxford Research Encyclopedia of Linguistics. Oxford University Press. doi: 10.1093/acrefore/9780199384655.013.248

Bybee, J. (2001). Phonology and language use. Cambridge: Cambridge University Press.

Burani, C., & Thornton, A. M. (2003). Effects of suffixation on sensitivity to inflectional morphology in reading: Evidence from Italian. Language and Cognitive Processes, 18(5-6), 625-666.

Carriazo Ruiz, J. R. (2014). Lexicografía, Morfología Y Etimología: El Sufijo -udo y sus derivados en el e-DECH. UVaDOC Principal. http://uvadoc.uva.es/handle/10324/21999

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. Psychological Review, 108(1), 204–256. https://doi.org/10.1037/0033-295x.108.1.204

Fábregas, A. (2020). Dispositional adjectives: Description and analysis. In Morphologically derived adjectives in Spanish. essay, John Benjamins Publishing Company.

Hernández, R. (2006). A typological approach to the ordering of adverbials: Weight, argumenthood and EPP. International Journal of Basque Linguistics, 39, 141-159

Joanisse, M. F., & Seidenberg, M. S. (1999). Impairments in verb morphology after brain injury: A connectionist model. Proceedings of the National Academy of Sciences, 96(13), 7592-7597.

Kilgarriff, A., Baisa, V., Bušta, J. et al. The Sketch Engine: ten years on. Lexicography ASIALEX 1, 7–36 (2014). https://doi.org/10.1007/s40607-014-0009-9

Kiparsky, P. (1982). Word-formation and the lexicon. In I. Roca (Ed.), Logical issues in language acquisition (pp. 161-193). Foris Publications.

Lázaro, M. (2012). The Effects of Base Frequency and Affix Productivity in Spanish. The Spanish Journal of Psychology, 15(2), 505–512. https://doi.org/10.5209/rev

Lázaro, M., Sainz, J., Illera, V., (2015). The role of derivative suffix productivity in the visual word recognition of complex words. Psicológica. 36.

Lázaro M, Acha J, Illera V, Sainz JS. (2016). Written Type and Token Frequency Measures of Fifty Spanish Derivational Morphemes. Span J Psychol. 2016 Nov 8;19:E75. doi: 10.1017/sjp.2016.75. PMID: 27821213.

Levin, B. (1993). English verb classes and alternations: A preliminary investigation. University of Chicago Press.

Manouilidou, C., & Stockall, L. (2014). Teasing apart syntactic category vs. argument structure information in deverbal word formation: A comparative psycholinguistic study. Italian Journal of Linguistics, 26(2), 71-98.

Manouilidou, C. (2007). Thematic Constraints in Deverbal word formation: psycholinguistic evidence from pseudo-words.

Manouilidou, C., Dolenc, B., Marvin, T., & Pirtošek, Z. (2015). Processing Complex pseudo-words in mild cognitive impairment: The interaction of preserved morphological rule knowledge with compromised cognitive ability. Clinical Linguistics & Phonetics, 30(1), 49–67. https://doi.org/10.3109/02699206.2015.1102970

Marjanovič, K., Manouilidou, C., & Marvin, T. (2013). Word-formation rules in Slovenian agentive deverbal nominalization: A psycholinguistic study based on pseudo-words.

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. Cognition, 8(1), 1-71.

Mendikoetxea, A. (1999). Las construcciones sintácticas fundamentales. Relaciones temporales, aspectuales y modales. In Gramática descriptiva de la lengua española (Vol. 2, pp. 1575–1630).

Neophytou, K., Manouilidou, C., Stockall, L., & Marantz, A. (2018). Syntactic and semantic restrictions on morphological recomposition: MEG evidence from Greek. Brain and Language, 183(May), 11–20. https://doi.org/10.1016/j.bandl.2018.05.003

Oltra Massuet , I. M. (2014). Deverbal adjectives at the interface: A crosslinguistic investigation into the morphology, syntax and semantics of -ble. De Gruyter Mouton.

Ortega, S. V., & Fabregat, S. (2005). In Morfología Léxica: La Formación de las palabras. story, Gredos.

Pinker, S. (1999). Words and rules: The ingredients of language. New York: Basic Books.

Plag, I. (2003). Word-Formation in English. https://doi.org/10.1017/cbo9780511841323

Plag, I., & Baayen, R. H. (2009). Morphological productivity across speech and writing. Linguistics, 47(2), 353-396.

Pustejovsky, J. (1991). The Generative Lexicon. Computational Linguistics, 17(4):409–441.

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. Language and Cognitive Processes, 23(7-8), 942-971.

Real Academia Española (2014). Sufijo. In Diccionario de la lengua española (23rd ed.). Retrieved from https://dle.rae.es/sufijo.

Ristić, B., Kavčič, K., Žaucer,R., Stockall, L., Manouilidou, C. (September 22–24 2022) Processing morphologically complex words in South Slavic. Poster presentation at SinFonIJA 15 (Conference on Syntax, Phonology and Language Analysis), University of Udine, Italy.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), Morphological aspects of language processing (pp. 131–154). Lawrence Erlbaum Associates, Inc.

Stockall, L., Manouilidou, C., Gwilliams, L., Neophytou, K., & Marantz, A. (2019). Prefix Stripping Re-Re-Revisited: MEG Investigations of Morphological Decomposition and Recomposition. Frontiers in Psychology, 10(September), 1–17. https://doi.org/10.3389/fpsyg.2019.01964

Moitra, S., Chacón, D., Stockall, L. (September 7-9 2022) Morphological recomposition and the concrete/abstract distinction in Bangla. Poster presented at AMLaP 28 (Architectures and Mechanisms for Language Processing), University of York, UK.

Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. Journal of Verbal Learning and Verbal Behavior, 14(6), 638-647. https://doi.org/10.1016/s0022-5371(75)80051-x

Appendix A - Stimuli for List 1

Condition	Suffix	Stimuli	Stem	Stem frequency	Stem frequency per million	Length stem without suffix	Word length	Word frequency	Word frequency per million
Gramm	udo	cabelludo	cabello	7595	15.30666	6	9	595	1.19914
Gramm	udo	peludo	pelo	8258	16.64284	3	6	297	0.59856
Gramm	udo	sesudo	seso	166	0.33455	3	6	214	0.43129
Gramm	udo	concienzudo	conciencia	33721	67.95994	8	11	435	0.88
Gramm	udo	barbudo	barba	1589	3.20241	4	7	163	0.3285
Gramm	udo	picudo	pico	9210	18.56146	3	6	146	0.29424
Gramm	udo	sañudo	saña	183	0.36881	3	6	111	0.2237
Gramm	udo	testarudo	testa	226	0.45547	6	9	92	0.18541
Gramm	udo	corajudo	coraje	1765	3.55711	5	8	77	0.15518
Gramm	udo	conchudo	concha	2261	4.55673	5	8	51	0.10278
Gramm	udo	huesudo	hueso	7396	14.9056	4	7	39	0.0786
Gramm	udo	lanudo	lana	1416	2.85375	3	6	38	0.07658
Gramm	udo	cascudo	casco	6924	13.95435	4	7	15	0.03
Gramm	udo	velludo	vello	815	1.64252	4	7	36	0.07255
Gramm	udo	cornudo	cuerno	962	1.93878	4	7	34	0.06852
Gramm	udo	melenudo	melena	311	0.62678	5	8	29	0.05845
Gramm	udo	membrudo	miembro	113107	227.9513	5	8	17	0.03426
Gramm	udo	colmilludo	colmillo	294	0.59252	7	10	4	0.00806
Gramm	udo	ceñudo	ceño	220	0.44338	3	6	16	0.03225
Gramm	udo	faldudo	falda	2008	4.04684	4	7	14	0.02822
Gramm	udo	bigotudo	bigote	589	1.18705	5	8	13	0.0262
Gramm	udo	carnudo	carne	19732	39.76708	4	7	12	0.02418
Gramm	udo	narizudo	nariz	4677	9.42584	5	8	12	0.02418
Gramm	udo	linajudo	linaje	1204	2.42649	5	8	11	0.02217
Gramm	udo	ganchudo	gancho	1487	2.99684	5	8	11	0.02217
Gramm	udo	barrigudo	barriga	1223	2.46478	6	9	11	0.02217
Gramm	udo	patudo	pata	3897	7.85386	3	6	10	0.02015
Gramm	udo	cejudo	ceja	3,136	6.32	3	6	5	0.01
Gramm	udo	cabezudo	cabeza	49985	100.7378	5	8	9	0.01814
Gramm	udo	forzudo	fuerza	107606	216.8648	4	7	7	0.01411
Gramm	udo	nervudo	nervio	4468	9.00463	4	7	7	0.01411
Gramm	udo	felpudo	felpa	107	0.21564	4	7	6	0.01209
Gramm	ble	portable	portar	10187	20.53047	5	8	581	1.17092
Gramm	ble	sociable	social	490889	989.3179	5	8	401	0.80816
Gramm	ble	bailable	bailar	7755	15.62911	5	8	370	0.74568
Gramm	ble	esperable	esperar	138607	279.343	6	9	225	0.45346
Gramm	ble	vencible	vencer	22357	45.0574	5	8	39	0.0786
Gramm	ble	agotable	agotar	12521	25.23432	5	8	110	0.22169
Gramm	ble	pensable	pensar	120420	242.6896	5	8	89	0.17937
Gramm	ble	vendible	vender	48645	98.03717	5	8	118	0.23781
Gramm	ble	atacable	atacar	19315	38.92667	5	8	48	0.09674
Gramm	ble	mutable	mutar	811	1.63446	4	7	100	0.20154
Gramm	ble	cobrable	cobrar	29934	60.32778	5	8	38	0.07658
Gramm	ble	anulable	anular	7918	15.95762	5	8	61	0.12294
Gramm	ble	oxidable	oxidar	1689	3.40394	5	8	60	0.12092
Gramm	ble	mudable	mudar	1467	2.95653	4	7	56	0.11286
Gramm	ble	pasable	pasar	296235	597.0201	4	7	56	0.11286
Gramm	ble	servible	servir	84532	170.3624	5	8	51	0.10278
Gramm	ble	curable	curar	5617	11.32028	4	7	340	0.68522
Gramm	ble	imitable	imitar	3774	7.60597	5	8	45	0.09069
Gramm	ble	asumible	asumir	73355	147.8367	5	8	41	0.08263
Gramm	ble	saturable	saturar	2746	5.53418	6	9	38	0.07658
Gramm	ble	pagable	pagar	105745	213.1142	4	7	37	0.07457
Gramm	ble	bebible	beber	8570	17.27163	4	7	34	0.06852

Gramm	ble	decible	decir	829670	1672.083	4	7	32	0.06449
Gramm	ble	ubicable	ubicar	108680	219.0293	5	8	26	0.0524
Gramm	ble	resoluble	resolver	76432	154.038	6	9	18	0.03628
Gramm	ble	vinculable	vincular	62707	126.3772	7	10	15	0.03023
Gramm	ble	salvable	salvar	15878	31.99988	5	8	13	0.0262
Gramm	ble	referible	referir	113351	228.443	6	9	12	0.02418
Gramm	ble	cantable	cantar	15796	31.83462	5	8	11	0.02217
Gramm	ble	flotable	flotar	1587	3.19838	5	8	11	0.02217
Gramm	ble	licuable	licuar	1126	2.2693	5	8	8	0.01612
Gramm	ble	penable	penar	423	0.8525	4	7	4	0.00806
Gramm	dizo	corredizo	correr	31532	63.54832	5	9	289	0.58244
Gramm	dizo	escurridizo	escurrir	913	1.84002	7	11	192	0.38695
Gramm	dizo	quebradizo	quebrar	7189	14.48842	6	10	192	0.38695
Gramm	dizo	movedizo	mover	26304	53.01202	4	8	167	0.33657
Gramm	dizo	resbaladizo	resbalar	589	1.18705	7	11	167	0.33657
Gramm	dizo	plegadizo	plegar	1025	2.06574	5	9	132	0.26603
Gramm	dizo	huidizo	huir	7191	14.49245	3	7	88	0.17735
Gramm	dizo	voladizo	volar	6871	13.84754	4	8	87	0.17534
Gramm	dizo	olvidadizo	olvidar	31464	63.41128	6	10	85	0.17131
Gramm	dizo	antojadizo	antojar	825	1.66267	6	10	56	0.11286
Gramm	dizo	advenedizo	advenir	148	0.29827	6	10	55	0.11084
Gramm	dizo	pegadizo	pegar	9331	18.80532	4	8	54	0.10883
Gramm	dizo	levadizo	levantar	25353	51.09541	4	8	51	0.10278
Gramm	dizo	asustadizo	asustar	3563	7.18073	6	10	47	0.09472
Gramm	dizo	enamoradizo	enamorar	5289	10.65924	7	11	38	0.07658
Gramm	dizo	arrojadizo	arrojar	14810	29.84748	6	10	16	0.03225
Gramm	dizo	caedizo	caer	43599	87.86767	3	7	13	0.0262
Gramm	dizo	anegadizo	anegar	562	1.13263	5	9	12	0.02418
Gramm	dizo	bebedizo	beber	8570	17.27163	4	8	11	0.02217
Gramm	dizo	tornadizo	tornar	7228	14.56702	5	9	11	0.02217
Gramm	dizo	encontradizo	encontrar	435587	877.8645	8	12	9	0.01814
Gramm	dizo	elevadizo	elevar	36260	73.07694	5	9	6	0.01209
Gramm	ano	quindiano	Quindio	1148	2.31363	6	9	263	0.53004
Gramm	ano	caucano	Cauca	62535	126.0305	4	7	535	1.07822
Gramm	ano	persiano	Persia	266	0.53609	5	8	96	0.19347
Gramm	ano	indiano	India	7752	15.62307	4	7	472	0.95125
Gramm	ano	paisano	Paisa	5225	10.53025	4	7	292	0.58849
Gramm	ano	jordano	Jordán	957	1.9287	4	7	154	0.31037
Gramm	ano	texano	Texas	2735	5.51201	3	6	29	0.05845
Gramm	ano	hawaiano	Hawai	336	0.67716	5	8	191	0.38493
Gramm	ano	troyano	troya	785	1.58206	4	7	115	0.23177
Gramm	ano	keniano	Kenia	619	1.24751	4	7	113	0.22774
Gramm	ano	neivano	Neiva	20657	41.63129	4	7	98	0.19751
Gramm	ano	tunjano	Tunja	8516	17.1628	4	7	76	0.15317
Gramm	ano	pisano	Pisa	862	1.73724	3	6	34	0.06852
Gramm	ano	siberiano	Siberia	696	1.40269	6	9	68	0.13704
Gramm	ano	jamaicano	Jamaica	1754	3.53494	6	9	64	0.12898
Gramm	ano	espartano	Esparta	207	0.41718	6	9	59	0.11891
Gramm	ano	villano	villa	17308	34.88185	4	7	101	0.20355
Gramm	ano	toledano	Toledo	1823	3.674	5	8	74	0.14914
Gramm	ano	liberiano	Liberia	301	0.60662	6	9	46	0.09271
Gramm	ano	riojano	Rioja	345	0.6953	4	7	40	0.08061
Gramm	ano	segoviano	Segovia	1598	3.22054	6	9	31	0.06248
Gramm	ano	octaviano	Octavio	2062	4.15567	6	9	26	0.0524
Gramm	ano	luliano	Lulio	48	0.09674	4	7	11	0.02217
Gramm	ano	zuliano	Zulia	1470	2.96258	4	7	15	0.03023
Gramm	ano	guineano	Guinea	568	1.14472	5	8	19	0.03829
Gramm	ano	bejarano	Béjar	29	0.05845	5	8	19	0.03829
Gramm	ano	gambiano	Gambia	116	0.23378	5	8	8	0.01612
Gramm	ano	koreano	Koreano	11	0.02217	4	7	7	0.01411
Gramm	ano	soriano	Sor	1517	3.0573	4	7	208	0.41919
Gramm	ano	siciliano	Sicilia	327	0.65902	6	9	303	0.61065
Gramm	ano	aldeano	aldea	2854	5.75184	4	7	204	0.41113

Gramm	ano	samoano	Samoa	87	0.17534	4	7	20	0.04031
CatViol	ano	abatano	abatir	2299	4.63	4	7	NA	NA
CatViol	ano	abolano	abolir	1412	2.85	4	7	NA	NA
CatViol	ano	absorbano	absorber	6337	12.77	6	9	NA	NA
CatViol	ano	andano	andar	33541	67.6	3	6	NA	NA
CatViol	ano	camuflano	camuflar	2333	4.7	6	9	NA	NA
CatViol	ano	comprimano	comprimir	2525	5.09	7	10	NA	NA
CatViol	ano	constreñano	constreñir	672	1.35	8	11	NA	NA
CatViol	ano	cosechano	cosechar	2997	6.04	6	9	NA	NA
CatViol	ano	encontrano	encontrar	435587	877.86	7	10	NA	NA
CatViol	ano	derivano	derivar	42931	86.52	5	8	NA	NA
CatViol	ano	derramano	derramar	2332	4.7	6	9	NA	NA
CatViol	ano	difuminano	difuminar	383	0.77	7	10	NA	NA
CatViol	ano	dividano	dividir	22563	45.47	5	8	NA	NA
CatViol	ano	dormano	dormir	13204	26.61	4	7	NA	NA
CatViol	udo	elududo	eludir	2710	5.46	4	7	NA	NA
CatViol	udo	escondudo	esconder	10111	20.38	6	9	NA	NA
CatViol	udo	estropeudo	estropear	512	1.03	7	10	NA	NA
CatViol	udo	leudo	leer	74830	150.81	2	5	NA	NA
CatViol	udo	llorudo	llorar	8368	16.86	4	7	NA	NA
CatViol	udo	maquilludo	maquillar	1234	2.49	7	10	NA	NA
CatViol	udo	masticudo	masticar	755	1.52	6	9	NA	NA
CatViol	udo	mordudo	morder	2245	4.52	4	7	NA	NA
CatViol	udo	notificudo	notificar	17441	35.15	7	10	NA	NA
CatViol	udo	persuadudo	persuadir	1581	3.19	7	10	NA	NA
CatViol	udo	quebrudo	quebrar	7189	14.49	5	8	NA	NA
CatViol	udo	repeludo	repeler	726	1.46	5	8	NA	NA
CatViol	udo	rompudo	romper	20296	40.9	4	7	NA	NA
CatViol	udo	subudo	subir	32415	65.33	3	6	NA	NA
CatViol	udo	vertudo	verter	3225	6.5	4	7	NA	NA
SemViol	ano	alcobano	alcoba	2783	5.61	5	8	NA	NA
SemViol	ano	astuciano	astucia	830	1.67	6	9	NA	NA
SemViol	ano	atributano	atributo	7638	15.39	7	10	NA	NA
SemViol	ano	blusano	blusa	788	1.59	4	7	NA	NA
SemViol	ano	bronzano	bronca	240	0.48	5	8	NA	NA
SemViol	ano	bufandano	bufanda	258	0.52	6	9	NA	NA
SemViol	ano	cacerolano	cacerola	284	0.57	7	10	NA	NA
SemViol	ano	canoano	canoa	1319	2.66	4	7	NA	NA
SemViol	ano	casetano	caseta	1691	3.41	5	8	NA	NA
SemViol	ano	cebrano	cebra	470	0.95	4	7	NA	NA
SemViol	ano	camarano	cámara	73947	149.03	5	8	NA	NA
SemViol	ano	escenano	escena	15326	30.89	5	8	NA	NA
SemViol	ano	estrellano	estrella	18540	37.36	7	10	NA	NA
SemViol	udo	guitarrudo	guitarra	5617	11.32	7	10	NA	NA
SemViol	udo	lupudo	lupa	1101	2.22	3	6	NA	NA
SemViol	udo	macetudo	maceta	405	0.82	5	8	NA	NA
SemViol	udo	maletudo	maleta	2852	5.75	5	8	NA	NA
SemViol	udo	mascarudo	máscara	2506	5.05	6	9	NA	NA
SemViol	udo	iglesiudo	iglesia	39760	80.13	6	9	NA	NA
SemViol	udo	muelludo	muelle	2952	5.95	5	8	NA	NA
SemViol	udo	paelludo	paella	249	0.5	5	8	NA	NA
SemViol	udo	botelludo	botella	6643	13.39	6	9	NA	NA
SemViol	udo	perfumudo	perfume	1979	3.99	6	9	NA	NA
SemViol	udo	podiudo	podio	1374	2.77	4	7	NA	NA
SemViol	udo	cumbrudo	cumbre	9,466	19.08	5	8	NA	NA
SemViol	udo	toalludo	toalla	2272	4.58	5	8	NA	NA
SemViol	udo	tuercudo	tuerca	665	1.34	5	8	NA	NA
CatViol	dizo	aguacatedizo	aguacate	1469	2.96	8	12	NA	NA
CatViol	dizo	ajodizo	ajo	1910	3.85	3	7	NA	NA
CatViol	dizo	algadizo	alga	1669	3.36	4	8	NA	NA
CatViol	dizo	sandiadizo	sandia	239	0.48	6	10	NA	NA
CatViol	dizo	chiledizo	chile	31234	62.95	5	9	NA	NA
CatViol	dizo	cerezadizo	cereza	535	1.08	6	10	NA	NA

CatViol	dizo	patatadizo	patata	421	0.85	6	10	NA	NA
CatViol	dizo	cirueladizo	ciruela	364	0.73	7	11	NA	NA
CatViol	dizo	lechugadizo	lechuga	950	1.91	7	11	NA	NA
CatViol	dizo	peradizo	pera	910	1.83	4	8	NA	NA
CatViol	dizo	granadadizo	granada	12960	26.12	7	11	NA	NA
CatViol	dizo	fuegodizo	fuego	20026	40.36	5	9	NA	NA
CatViol	dizo	hongodizo	hongo	6238	12.57	5	9	NA	NA
CatViol	dizo	cebolladizo	cebolla	3203	6.46	7	11	NA	NA
CatViol	ble	naranjable	naranja	5154	10.39	7	10	NA	NA
CatViol	ble	platanible	plátano	3969	8	7	10	NA	NA
CatViol	ble	apiable	apio	379	0.76	4	7	NA	NA
CatViol	ble	pimientible	pimienta	2138	4.31	8	11	NA	NA
CatViol	ble	piñable	piña	1810	3.65	4	7	NA	NA
CatViol	ble	rabanible	rábano	230	0.46	6	9	NA	NA
CatViol	ble	romerible	romero	9966	20.09	6	9	NA	NA
CatViol	ble	setable	seta	384	0.77	4	7	NA	NA
CatViol	ble	teible	té	2783	5.61	2	6	NA	NA
CatViol	ble	tomatible	tomate	4206	8.48	6	9	NA	NA
CatViol	ble	trigible	trigo	3326	6.7	5	8	NA	NA
CatViol	ble	uvable	uva	1866	3.76	3	6	NA	NA
CatViol	ble	zanahoriable	zanahoria	1627	3.28	9	12	NA	NA
CatViol	ble	tierrable	tierra	87571	176.49	6	9	NA	NA
SemViol	dizo	acaecedizo	acaecer	2113	4.26	6	10	NA	NA
SemViol	dizo	acechadizo	acechar	668	1.35	6	10	NA	NA
SemViol	dizo	albergadizo	albergar	7084	14.28	7	11	NA	NA
SemViol	dizo	brotadizo	brotar	2138	4.31	5	9	NA	NA
SemViol	dizo	constadizo	constar	21653	43.64	6	10	NA	NA
SemViol	dizo	crecedizo	crecer	49484	99.73	5	9	NA	NA
SemViol	dizo	cubridizo	cubrir	44661	90.01	5	9	NA	NA
SemViol	dizo	enfermadizo	enfermar	2022	4.08	7	11	NA	NA
SemViol	dizo	expiradizo	expirar	1504	3.03	6	10	NA	NA
SemViol	dizo	fallecedizo	fallecer	14472	29.17	7	11	NA	NA
SemViol	dizo	faltadizo	faltar	35804	72.16	5	9	NA	NA
SemViol	dizo	lactadizo	lactar	258	0.51996	5	9	NA	NA
SemViol	dizo	fenecedizo	fenecer	566	1.14	6	10	NA	NA
SemViol	dizo	florecedizo	florecer	2076	4.18	7	11	NA	NA
SemViol	ble	germinable	germinar	997	2.01	7	10	NA	NA
SemViol	ble	gestable	gestar	3631	7.32	5	8	NA	NA
SemViol	ble	languidecible	languidecer	191	0.38	10	13	NA	NA
SemViol	ble	levitable	levitar	23	0.04635	6	9	NA	NA
SemViol	ble	madurable	madurar	2524	5.09	6	9	NA	NA
SemViol	ble	medrable	medrar	131	0.26	5	8	NA	NA
SemViol	ble	morible	morir	53192	107.2	4	7	NA	NA
SemViol	ble	morable	morar	3384	6.82	4	7	NA	NA
SemViol	ble	nacible	nacer	58605	118.11	4	7	NA	NA
SemViol	ble	ocurrible	ocurrir	80137	161.5	6	9	NA	NA
SemViol	ble	originable	originar	22483	45.31	7	10	NA	NA
SemViol	ble	residible	residir	11172	22.52	6	9	NA	NA
SemViol	ble	sobrable	sobrar	5460	11	5	8	NA	NA
SemViol	ble	surgible	surgir	47031	94.78	5	8	NA	NA
NonWords	udo	cadelludo	cabelludo	NA	NA	9	9	NA	NA
NonWords	udo	pemudo	peludo	NA	NA	6	6	NA	NA
NonWords	udo	barludo	barbudo	NA	NA	7	7	NA	NA
NonWords	udo	pimudo	picudo	NA	NA	6	6	NA	NA
NonWords	udo	mebenudo	melenudo	NA	NA	8	8	NA	NA
NonWords	udo	bilotudo	bigotudo	NA	NA	8	8	NA	NA
NonWords	udo	nabigudo	narigudo	NA	NA	8	8	NA	NA
NonWords	udo	pamudo	patudo	NA	NA	6	6	NA	NA
NonWords	udo	cadezudo	cabezudo	NA	NA	8	8	NA	NA
NonWords	udo	nerludo	nervudo	NA	NA	7	7	NA	NA
NonWords	ble	fuvible	fusible	NA	NA	7	7	NA	NA
NonWords	ble	fanible	falible	NA	NA	7	7	NA	NA
NonWords	ble	muvable	mutable	NA	NA	7	7	NA	NA

NonWords	ble	atulable	anulable	NA	NA	8	8	NA	NA
NonWords	ble	paxable	pasable	NA	NA	7	7	NA	NA
NonWords	ble	cuvable	curable	NA	NA	7	7	NA	NA
NonWords	ble	salurable	saturable	NA	NA	9	9	NA	NA
NonWords	ble	padable	pagable	NA	NA	7	7	NA	NA
NonWords	ble	uticable	ubicable	NA	NA	8	8	NA	NA
NonWords	ble	reterible	referible	NA	NA	9	9	NA	NA
NonWords	dizo	quedradizo	quebradizo	NA	NA	10	10	NA	NA
NonWords	dizo	mopedizo	movedizo	NA	NA	8	8	NA	NA
NonWords	dizo	resnaladizo	resbaladizo	NA	NA	11	11	NA	NA
NonWords	dizo	pletadizo	plegadizo	NA	NA	9	9	NA	NA
NonWords	dizo	huibizo	huidizo	NA	NA	7	7	NA	NA
NonWords	dizo	vonadizo	voladizo	NA	NA	8	8	NA	NA
NonWords	dizo	pemadizo	pegadizo	NA	NA	8	8	NA	NA
NonWords	dizo	lezadizo	levadizo	NA	NA	8	8	NA	NA
NonWords	dizo	asuftadizo	asustadizo	NA	NA	10	10	NA	NA
NonWords	dizo	becedizo	bebedizo	NA	NA	8	8	NA	NA
NonWords	ano	paizano	paisano	NA	NA	7	7	NA	NA
NonWords	ano	jorlano	jordano	NA	NA	7	7	NA	NA
NonWords	ano	trorano	troyano	NA	NA	7	7	NA	NA
NonWords	ano	keliano	keniano	NA	NA	7	7	NA	NA
NonWords	ano	siteriano	siberiano	NA	NA	9	9	NA	NA
NonWords	ano	jalaicano	jamaicano	NA	NA	9	9	NA	NA
NonWords	ano	esfartano	espartano	NA	NA	9	9	NA	NA
NonWords	ano	tovedano	toledano	NA	NA	8	8	NA	NA
NonWords	ano	sevoviano	segoviano	NA	NA	9	9	NA	NA
NonWords	ano	guiceano	guineano	NA	NA	8	8	NA	NA
Fillers	low bf high ap	frutero	NA	NA	NA	7	7	NA	NA
Fillers	low bf high ap	pescador	NA	NA	NA	8	8	NA	NA
Fillers	low bf high ap	duradero	NA	NA	NA	8	8	NA	NA
Fillers	low bf high ap	aceitoso	NA	NA	NA	8	8	NA	NA
Fillers	low bf high ap	tintero	NA	NA	NA	7	7	NA	NA
Fillers	low bf high ap	tramposo	NA	NA	NA	8	8	NA	NA
Fillers	low bf high ap	quesero	NA	NA	NA	7	7	NA	NA
Fillers	low bf high ap	armonioso	NA	NA	NA	9	9	NA	NA
Fillers	low bf high ap	montañero	NA	NA	NA	9	9	NA	NA
Fillers	low bf high ap	pagajoso	NA	NA	NA	8	8	NA	NA
Fillers	low bf low ap	pelotazo	NA	NA	NA	8	8	NA	NA
Fillers	low bf low ap	temible	NA	NA	NA	7	7	NA	NA
Fillers	low bf low ap	vagancia	NA	NA	NA	8	8	NA	NA
Fillers	low bf low ap	fianza	NA	NA	NA	6	6	NA	NA
Fillers	low bf low ap	vigilancia	NA	NA	NA	10	10	NA	NA
Fillers	low bf low ap	grasiento	NA	NA	NA	9	9	NA	NA
Fillers	low bf low ap	montaje	NA	NA	NA	7	7	NA	NA
Fillers	low bf low ap	labranza	NA	NA	NA	8	8	NA	NA
Fillers	low bf low ap	patinaje	NA	NA	NA	8	8	NA	NA
Fillers	low bf low ap	carruaje	NA	NA	NA	8	8	NA	NA
Fillers	high bf high ap	cristalero	NA	NA	NA	10	10	NA	NA
Fillers	high bf high ap	ayudante	NA	NA	NA	8	8	NA	NA
Fillers	high bf high ap	papelera	NA	NA	NA	8	8	NA	NA
Fillers	high bf high ap	jardinero	NA	NA	NA	9	9	NA	NA
Fillers	high bf high ap	pensador	NA	NA	NA	8	8	NA	NA
Fillers	high bf high ap	librero	NA	NA	NA	7	7	NA	NA
Fillers	high bf high ap	limpiador	NA	NA	NA	9	9	NA	NA
Fillers	high bf high ap	letrista	NA	NA	NA	8	8	NA	NA
Fillers	high bf high ap	mentiroso	NA	NA	NA	9	9	NA	NA
Fillers	high bf high ap	coronilla	NA	NA	NA	9	9	NA	NA
Fillers	high bf low ap	arboleda	NA	NA	NA	8	8	NA	NA
Fillers	high bf low ap	extrañeza	NA	NA	NA	9	9	NA	NA
Fillers	high bf low ap	pobreza	NA	NA	NA	7	7	NA	NA
Fillers	high bf low ap	veraniego	NA	NA	NA	9	9	NA	NA
Fillers	high bf low ap	alteza	NA	NA	NA	6	6	NA	NA
Fillers	high bf low ap	bajeza	NA	NA	NA	6	6	NA	NA

Fillers	high bf low ap	golpazo	NA	NA	NA	7	7	NA	NA
Fillers	high bf low ap	perruno	NA	NA	NA	7	7	NA	NA
Fillers	high bf low ap	simpleza	NA	NA	NA	8	8	NA	NA
Fillers	high bf low ap	patriota	NA	NA	NA	8	8	NA	NA

Appendix B - Stimuli for List 2

Condition	Suffix	Stimuli	Stem	Stem frequency	Stem frequency per million	Length stem without suffix	Word length	Word frequency	Word frequency per million
Gramm	udo	cabelludo	cabello	7595	15.30666	6	9	595	1.19914
Gramm	udo	peludo	pelo	8258	16.64284	3	6	297	0.59856
Gramm	udo	sesudo	seso	166	0.33455	3	6	214	0.43129
Gramm	udo	concienzudo	conciencia	33721	67.95994	8	11	435	0.88
Gramm	udo	barbudo	barba	1589	3.20241	4	7	163	0.3285
Gramm	udo	picudo	pico	9210	18.56146	3	6	146	0.29424
Gramm	udo	sañudo	saña	183	0.36881	3	6	111	0.2237
Gramm	udo	testarudo	testa	226	0.45547	6	9	92	0.18541
Gramm	udo	corajudo	coraje	1765	3.55711	5	8	77	0.15518
Gramm	udo	conchudo	concha	2261	4.55673	5	8	51	0.10278
Gramm	udo	huesudo	hueso	7396	14.9056	4	7	39	0.0786
Gramm	udo	lanudo	lana	1416	2.85375	3	6	38	0.07658
Gramm	udo	cascudo	casco	6924	13.95435	4	7	15	0.03
Gramm	udo	velludo	vello	815	1.64252	4	7	36	0.07255
Gramm	udo	cornudo	cuerno	962	1.93878	4	7	34	0.06852
Gramm	udo	melenudo	melena	311	0.62678	5	8	29	0.05845
Gramm	udo	membrudo	miembro	113107	227.9513	5	8	17	0.03426
Gramm	udo	colmilludo	colmillo	294	0.59252	7	10	4	0.00806
Gramm	udo	ceñudo	ceño	220	0.44338	3	6	16	0.03225
Gramm	udo	faldudo	falda	2008	4.04684	4	7	14	0.02822
Gramm	udo	bigotudo	bigote	589	1.18705	5	8	13	0.0262
Gramm	udo	carnudo	carne	19732	39.76708	4	7	12	0.02418
Gramm	udo	narizudo	nariz	4677	9.42584	5	8	12	0.02418
Gramm	udo	linajudo	linaje	1204	2.42649	5	8	11	0.02217
Gramm	udo	ganchudo	gancho	1487	2.99684	5	8	11	0.02217
Gramm	udo	barrigudo	barriga	1223	2.46478	6	9	11	0.02217
Gramm	udo	patudo	pata	3897	7.85386	3	6	10	0.02015
Gramm	udo	cejudo	ceja	3136	6.32	3	6	5	0.01
Gramm	udo	cabezudo	cabeza	49985	100.7378	5	8	9	0.01814
Gramm	udo	forzudo	fuerza	107606	216.8648	4	7	7	0.01411
Gramm	udo	nervudo	nervio	4468	9.00463	4	7	7	0.01411
Gramm	udo	felpudo	felpa	107	0.21564	4	7	6	0.01209
Gramm	ble	portable	portar	10187	20.53047	5	8	581	1.17092
Gramm	ble	sociable	social	490889	989.3179	5	8	401	0.80816
Gramm	ble	bailable	bailar	7755	15.62911	5	8	370	0.74568
Gramm	ble	esperable	esperar	138607	279.343	6	9	225	0.45346
Gramm	ble	vencible	vencer	22357	45.0574	5	8	39	0.0786
Gramm	ble	agotable	agotar	12521	25.23432	5	8	110	0.22169
Gramm	ble	pensable	pensar	120420	242.6896	5	8	89	0.17937
Gramm	ble	vendible	vender	48645	98.03717	5	8	118	0.23781
Gramm	ble	atacable	atacar	19315	38.92667	5	8	48	0.09674
Gramm	ble	mutable	mutar	811	1.63446	4	7	100	0.20154
Gramm	ble	cobrable	cobrar	29934	60.32778	5	8	38	0.07658
Gramm	ble	anulable	anular	7918	15.95762	5	8	61	0.12294
Gramm	ble	oxidable	oxidar	1689	3.40394	5	8	60	0.12092
Gramm	ble	mudable	mudar	1467	2.95653	4	7	56	0.11286
Gramm	ble	pasable	pasar	296235	597.0201	4	7	56	0.11286
Gramm	ble	servible	servir	84532	170.3624	5	8	51	0.10278

Gramm	ble	curable	curar	5617	11.32028	4	7	340	0.68522
Gramm	ble	imitable	imitar	3774	7.60597	5	8	45	0.09069
Gramm	ble	asumible	asumir	73355	147.8367	5	8	41	0.08263
Gramm	ble	saturable	saturar	2746	5.53418	6	9	38	0.07658
Gramm	ble	pagable	pagar	105745	213.1142	4	7	37	0.07457
Gramm	ble	bebible	beber	8570	17.27163	4	7	34	0.06852
Gramm	ble	decible	decir	829670	1672.083	4	7	32	0.06449
Gramm	ble	ubicable	ubicar	108680	219.0293	5	8	26	0.0524
Gramm	ble	resoluble	resolver	76432	154.038	6	9	18	0.03628
Gramm	ble	vinculable	vincular	62707	126.3772	7	10	15	0.03023
Gramm	ble	salvable	salvar	15878	31.99988	5	8	13	0.0262
Gramm	ble	referible	referir	113351	228.443	6	9	12	0.02418
Gramm	ble	cantable	cantar	15796	31.83462	5	8	11	0.02217
Gramm	ble	flotable	flotar	1587	3.19838	5	8	11	0.02217
Gramm	ble	licuable	licuar	1126	2.2693	5	8	8	0.01612
Gramm	ble	penable	penar	423	0.8525	4	7	4	0.00806
Gramm	dizo	corredizo	correr	31532	63.54832	5	9	289	0.58244
Gramm	dizo	escurridizo	escurrir	913	1.84002	7	11	192	0.38695
Gramm	dizo	quebradizo	quebrar	7189	14.48842	6	10	192	0.38695
Gramm	dizo	movedizo	mover	26304	53.01202	4	8	167	0.33657
Gramm	dizo	resbaladizo	resbalar	589	1.18705	7	11	167	0.33657
Gramm	dizo	plegadizo	plegar	1025	2.06574	5	9	132	0.26603
Gramm	dizo	huidizo	huir	7191	14.49245	3	7	88	0.17735
Gramm	dizo	voladizo	volar	6871	13.84754	4	8	87	0.17534
Gramm	dizo	olvidadizo	olvidar	31464	63.41128	6	10	85	0.17131
Gramm	dizo	antojadizo	antojar	825	1.66267	6	10	56	0.11286
Gramm	dizo	advenedizo	advenir	148	0.29827	6	10	55	0.11084
Gramm	dizo	pegadizo	pegar	9331	18.80532	4	8	54	0.10883
Fillers	dizo	levadizo	levantar	25353	51.09541	4	8	51	0.10278
Gramm	dizo	asustadizo	asustar	3563	7.18073	6	10	47	0.09472
Gramm	dizo	enamoradizo	enamorar	5289	10.65924	7	11	38	0.07658
Gramm	dizo	arrojadizo	arrojar	14810	29.84748	6	10	16	0.03225
Gramm	dizo	caedizo	caer	43599	87.86767	3	7	13	0.0262
Gramm	dizo	anegadizo	anegar	562	1.13263	5	9	12	0.02418
Gramm	dizo	bebedizo	beber	8570	17.27163	4	8	11	0.02217
Gramm	dizo	tornadizo	tornar	7228	14.56702	5	9	11	0.02217
Gramm	dizo	encontradizo	encontrar	435587	877.8645	8	12	9	0.01814
Gramm	dizo	elevadizo	elevar	36260	73.07694	5	9	6	0.01209
Gramm	ano	quindiano	Quindio	1148	2.31363	6	9	263	0.53004
Gramm	ano	caucano	Cauca	62535	126.0305	4	7	535	1.07822
Gramm	ano	persiano	Persia	266	0.53609	5	8	96	0.19347
Gramm	ano	indiano	India	7752	15.62307	4	7	472	0.95125
Gramm	ano	paisano	Paisa	5225	10.53025	4	7	292	0.58849
Gramm	ano	jordano	Jordán	957	1.9287	4	7	154	0.31037
Gramm	ano	texano	Texas	2735	5.51201	3	6	29	0.05845
Gramm	ano	hawaiano	Hawai	336	0.67716	5	8	191	0.38493
Gramm	ano	troyano	troya	785	1.58206	4	7	115	0.23177
Gramm	ano	keniano	Kenia	619	1.24751	4	7	113	0.22774
Gramm	ano	neivano	Neiva	20657	41.63129	4	7	98	0.19751
Gramm	ano	tunjano	Tunja	8516	17.1628	4	7	76	0.15317
Gramm	ano	pisano	Pisa	862	1.73724	3	6	34	0.06852
Gramm	ano	siberiano	Siberia	696	1.40269	6	9	68	0.13704
Gramm	ano	jamaicano	Jamaica	1754	3.53494	6	9	64	0.12898
Gramm	ano	espartano	Esparta	207	0.41718	6	9	59	0.11891
Gramm	ano	villano	villa	17308	34.88185	4	7	101	0.20355

Gramm	ano	toledano	Toledo	1823	3.674	5	8	74	0.14914
Gramm	ano	liberiano	Liberia	301	0.60662	6	9	46	0.09271
Gramm	ano	riojano	Rioja	345	0.6953	4	7	40	0.08061
Gramm	ano	segoviano	Segovia	1598	3.22054	6	9	31	0.06248
Gramm	ano	octaviano	Octavio	2062	4.15567	6	9	26	0.0524
Gramm	ano	luliano	Lulio	48	0.09674	4	7	11	0.02217
Gramm	ano	zuliano	Zulia	1470	2.96258	4	7	15	0.03023
Gramm	ano	guineano	Guinea	568	1.14472	5	8	19	0.03829
Gramm	ano	bejarano	Béjar	29	0.05845	5	8	19	0.03829
Gramm	ano	gambiano	Gambia	116	0.23378	5	8	8	0.01612
Gramm	ano	koreano	Koreano	11	0.02217	4	7	7	0.01411
Gramm	ano	soriano	Sor	1517	3.0573	4	7	208	0.41919
Gramm	ano	siciliano	Sicilia	327	0.65902	6	9	303	0.61065
Gramm	ano	aldeano	aldea	2854	5.75184	4	7	204	0.41113
Gramm	ano	samoano	Samoa	87	0.17534	4	7	20	0.04031
CatViol	ano	eludano	eludir	2710	5.46	4	7	NA	NA
CatViol	ano	escondano	esconder	10111	20.38	6	9	NA	NA
CatViol	ano	estropeano	estropear	512	1.03	7	10	NA	NA
CatViol	ano	leano	leer	74830	150.81	2	5	NA	NA
CatViol	ano	llorano	llorar	8368	16.86	4	7	NA	NA
CatViol	ano	maquillano	maquillar	1234	2.49	7	10	NA	NA
CatViol	ano	masticano	masticar	755	1.52	6	9	NA	NA
CatViol	ano	mordano	morder	2245	4.52	4	7	NA	NA
CatViol	ano	notificano	notificar	17441	35.15	7	10	NA	NA
CatViol	ano	persuadano	persuadir	1581	3.19	7	10	NA	NA
CatViol	ano	quebrano	quebrar	7189	14.49	5	8	NA	NA
CatViol	ano	repelano	repeler	726	1.46	5	8	NA	NA
CatViol	ano	rompano	romper	20296	40.9	4	7	NA	NA
CatViol	ano	subano	subir	32415	65.33	3	6	NA	NA
CatViol	ano	vertano	verter	3225	6.5	4	7	NA	NA
CatViol	udo	abatudo	abatir	2299	4.63	4	7	NA	NA
CatViol	udo	aboludo	abolir	1412	2.85	4	7	NA	NA
CatViol	udo	absorbudo	absorber	6337	12.77	6	9	NA	NA
CatViol	udo	andudo	andar	33541	67.6	3	6	NA	NA
CatViol	udo	camufludo	camuflar	2333	4.7	6	9	NA	NA
CatViol	udo	comprimudo	comprimir	2525	5.09	7	10	NA	NA
CatViol	udo	constreñudo	constreñir	672	1.35	8	11	NA	NA
CatViol	udo	cosechudo	cosechar	2997	6.04	6	9	NA	NA
CatViol	udo	encontrudo	encontrar	435587	877.86	7	10	NA	NA
CatViol	udo	derivudo	derivar	42931	86.52	5	8	NA	NA
CatViol	udo	derramudo	derramar	2332	4.7	6	9	NA	NA
CatViol	udo	difuminudo	difuminar	383	0.77	7	10	NA	NA
CatViol	udo	dividudo	dividir	22563	45.47	5	8	NA	NA
CatViol	udo	dormudo	dormir	13204	26.61	4	7	NA	NA
SemViol	ano	guitarrano	guitarra	5617	11.32	7	10	NA	NA
SemViol	ano	lupano	lupa	1101	2.22	3	6	NA	NA
SemViol	ano	macetano	maceta	405	0.82	5	8	NA	NA
SemViol	ano	maletano	maleta	2852	5.75	5	8	NA	NA
SemViol	ano	mascarano	máscara	2506	5.05	6	9	NA	NA
SemViol	ano	iglesiano	iglesia	39760	80.13	6	9	NA	NA
SemViol	ano	muellano	muelle	2952	5.95	5	8	NA	NA
SemViol	ano	paellano	paella	249	0.5	5	8	NA	NA
SemViol	ano	botellano	botella	6643	13.39	6	9	NA	NA
SemViol	ano	perfumano	perfume	1979	3.99	6	9	NA	NA
SemViol	ano	podiano	podio	1374	2.77	4	7	NA	NA

SemViol	ano	cumbreano	cumbre	9466	19.08	5	9	NA	NA
SemViol	ano	toallano	toalla	2272	4.58	5	8	NA	NA
SemViol	ano	tuercano	tuerca	665	1.34	5	8	NA	NA
SemViol	udo	alcobudo	alcoba	2783	5.61	5	8	NA	NA
SemViol	udo	astuciudo	astucia	830	1.67	6	9	NA	NA
SemViol	udo	atributudo	atributo	7638	15.39	7	10	NA	NA
SemViol	udo	blusudo	blusa	788	1.59	4	7	NA	NA
SemViol	udo	bronzudo	bronce	4621	9.31	5	8	NA	NA
SemViol	udo	bufandudo	bufanda	258	0.52	6	9	NA	NA
SemViol	udo	caceroludo	cacerola	284	0.57	7	10	NA	NA
SemViol	udo	canoudo	canoa	1319	2.66	4	7	NA	NA
SemViol	udo	casetudo	caseta	1691	3.41	5	8	NA	NA
SemViol	udo	cebrudo	cebra	470	0.95	4	7	NA	NA
SemViol	udo	camarudo	cámara	73947	149.03	5	8	NA	NA
SemViol	udo	escenudo	escena	15326	30.89	5	8	NA	NA
SemViol	udo	estrelludo	estrella	18540	37.36	7	10	NA	NA
CatViol	dizo	naranjadizo	naranja	5154	10.39	7	11	NA	NA
CatViol	dizo	platanodizo	plátano	3969	8	7	11	NA	NA
CatViol	dizo	apiodizo	apio	379	0.76	4	8	NA	NA
CatViol	dizo	pimientodizo	pimienta	2138	4.31	8	12	NA	NA
CatViol	dizo	piñadizo	piña	1810	3.65	4	8	NA	NA
CatViol	dizo	rabanodizo	rábano	230	0.46	6	10	NA	NA
CatViol	dizo	romerodizo	romero	9966	20.09	6	10	NA	NA
CatViol	dizo	setadizo	seta	384	0.77	4	8	NA	NA
CatViol	dizo	tedizo	té	2783	5.61	2	6	NA	NA
CatViol	dizo	tomatedizo	tomate	4206	8.48	6	10	NA	NA
CatViol	dizo	trigodizo	trigo	3326	6.7	5	9	NA	NA
CatViol	dizo	uvadizo	uva	1866	3.76	3	7	NA	NA
CatViol	dizo	zanahoriadizo	zanahoria	1627	3.28	9	13	NA	NA
CatViol	dizo	tierradizo	tierra	87571	176.49	6	10	NA	NA
CatViol	ble	aguacatible	aguacate	1469	2.96	8	11	NA	NA
CatViol	ble	ajible	ajo	1910	3.85	3	6	NA	NA
CatViol	ble	algable	alga	1669	3.36	4	7	NA	NA
CatViol	ble	sandiable	sandia	239	0.48	6	9	NA	NA
CatViol	ble	chilible	chile	31234	62.95	5	8	NA	NA
CatViol	ble	cerezable	cereza	535	1.08	6	9	NA	NA
CatViol	ble	patatable	patata	421	0.85	6	9	NA	NA
CatViol	ble	ciruelable	ciruela	364	0.73	7	10	NA	NA
CatViol	ble	perable	pera	910	1.83	4	7	NA	NA
CatViol	ble	granadable	granada	12960	26.12	7	10	NA	NA
CatViol	ble	fuegible	fuego	20026	40.36	5	8	NA	NA
CatViol	ble	hongible	hongo	6238	12.572	5	8	NA	NA
CatViol	ble	lechugable	lechuga	950	1.91459	7	10	NA	NA
CatViol	ble	cebollable	cebolla	3203	6.46	7	10	NA	NA
SemViol	dizo	germinadizo	germinar	997	2.01	7	11	NA	NA
SemViol	dizo	gestadizo	gestar	3631	7.32	5	9	NA	NA
SemViol	dizo	languidecedizo	languidecer	191	0.38	10	14	NA	NA
SemViol	dizo	levitadizo	levitar	23	0.04635	6	10	NA	NA
SemViol	dizo	maduradizo	madurar	2524	5.09	6	10	NA	NA
SemViol	dizo	medradizo	medrar	131	0.26	5	9	NA	NA
SemViol	dizo	moridizo	morir	53192	107.2	4	8	NA	NA
SemViol	dizo	moradizo	morar	3384	6.82	4	8	NA	NA
SemViol	dizo	nacedizo	nacer	58605	118.11	4	8	NA	NA
SemViol	dizo	ocurridizo	ocurrir	80137	161.5	6	10	NA	NA
SemViol	dizo	originadizo	originar	22483	45.31	7	11	NA	NA

SemViol	dizo	resididizo	residir	11172	22.52	6	10	NA	NA
SemViol	dizo	sobradizo	sobrar	5460	11	5	9	NA	NA
SemViol	dizo	surgidizo	surgir	47031	94.78	5	9	NA	NA
SemViol	ble	acaecible	acaecer	2113	4.26	6	9	NA	NA
SemViol	ble	acechable	acechar	668	1.35	6	9	NA	NA
SemViol	ble	albergable	albergar	7084	14.28	7	10	NA	NA
SemViol	ble	brotable	brotar	2138	4.31	5	8	NA	NA
SemViol	ble	constable	constar	21653	43.64	6	9	NA	NA
SemViol	ble	crecible	crecer	49484	99.73	5	8	NA	NA
SemViol	ble	cubrible	cubrir	44661	90.01	5	8	NA	NA
SemViol	ble	enfermable	enfermar	2022	4.08	7	10	NA	NA
SemViol	ble	expirable	expirar	1504	3.03	6	9	NA	NA
SemViol	ble	fallecible	fallecer	14472	29.17	7	10	NA	NA
SemViol	ble	faltable	faltar	35804	72.16	5	8	NA	NA
SemViol	ble	lactable	lactar	258	0.51996	5	8	NA	NA
SemViol	ble	fenecible	fenecer	566	1.14	6	9	NA	NA
SemViol	ble	florecible	florecer	2076	4.18	7	10	NA	NA
NonWords	udo	cadelludo	cabelludo	NA	NA	9	9	NA	NA
NonWords	udo	pemudo	peludo	NA	NA	6	6	NA	NA
NonWords	udo	barludo	barbudo	NA	NA	7	7	NA	NA
NonWords	udo	pimudo	picudo	NA	NA	6	6	NA	NA
NonWords	udo	mebenudo	melenudo	NA	NA	8	8	NA	NA
NonWords	udo	bilotudo	bigotudo	NA	NA	8	8	NA	NA
NonWords	udo	nabigudo	narigudo	NA	NA	8	8	NA	NA
NonWords	udo	pamudo	patudo	NA	NA	6	6	NA	NA
NonWords	udo	cadezudo	cabezudo	NA	NA	8	8	NA	NA
NonWords	udo	nerludo	nervudo	NA	NA	7	7	NA	NA
NonWords	ble	fuvible	fusible	NA	NA	7	7	NA	NA
NonWords	ble	fanible	falible	NA	NA	7	7	NA	NA
NonWords	ble	muvable	mutable	NA	NA	7	7	NA	NA
NonWords	ble	atulable	anulable	NA	NA	8	8	NA	NA
NonWords	ble	paxable	pasable	NA	NA	7	7	NA	NA
NonWords	ble	cuvable	curable	NA	NA	7	7	NA	NA
NonWords	ble	salurable	saturable	NA	NA	9	9	NA	NA
NonWords	ble	padable	pagable	NA	NA	7	7	NA	NA
NonWords	ble	uticable	ubicable	NA	NA	8	8	NA	NA
NonWords	ble	reterible	referible	NA	NA	9	9	NA	NA
NonWords	dizo	quedradizo	quebradizo	NA	NA	10	10	NA	NA
NonWords	dizo	mopedizo	movedizo	NA	NA	8	8	NA	NA
NonWords	dizo	resnaladizo	resbaladizo	NA	NA	11	11	NA	NA
NonWords	dizo	pletadizo	plegadizo	NA	NA	9	9	NA	NA
NonWords	dizo	huibizo	huidizo	NA	NA	7	7	NA	NA
NonWords	dizo	vonadizo	voladizo	NA	NA	8	8	NA	NA
NonWords	dizo	pemadizo	pegadizo	NA	NA	8	8	NA	NA
NonWords	dizo	lezadizo	levadizo	NA	NA	8	8	NA	NA
NonWords	dizo	asuftadizo	asustadizo	NA	NA	10	10	NA	NA
NonWords	dizo	becedizo	bebedizo	NA	NA	8	8	NA	NA
NonWords	ano	paizano	paisano	NA	NA	7	7	NA	NA
NonWords	ano	jorlano	jordano	NA	NA	7	7	NA	NA
NonWords	ano	trorano	troyano	NA	NA	7	7	NA	NA
NonWords	ano	keliano	keniano	NA	NA	7	7	NA	NA
NonWords	ano	siteriano	siberiano	NA	NA	9	9	NA	NA
NonWords	ano	jalaicano	jamaicano	NA	NA	9	9	NA	NA
NonWords	ano	esfartano	espartano	NA	NA	9	9	NA	NA
NonWords	ano	tovedano	toledano	NA	NA	8	8	NA	NA

NonWords	ano	sevoviano	segoviano	NA	NA	9	9	NA	NA
NonWords	ano	guiceano	guineano	NA	NA	8	8	NA	NA
Fillers	low bf high ap	frutero	NA	NA	NA	7	7	NA	NA
Fillers	low bf high ap	pescador	NA	NA	NA	8	8	NA	NA
Fillers	low bf high ap	duradero	NA	NA	NA	8	8	NA	NA
Fillers	low bf high ap	aceitoso	NA	NA	NA	8	8	NA	NA
Fillers	low bf high ap	tintero	NA	NA	NA	7	7	NA	NA
Fillers	low bf high ap	tramposo	NA	NA	NA	8	8	NA	NA
Fillers	low bf high ap	quesero	NA	NA	NA	7	7	NA	NA
Fillers	low bf high ap	armonioso	NA	NA	NA	9	9	NA	NA
Fillers	low bf high ap	montañero	NA	NA	NA	9	9	NA	NA
Fillers	low bf high ap	pagajoso	NA	NA	NA	8	8	NA	NA
Fillers	low bf low ap	pelotazo	NA	NA	NA	8	8	NA	NA
Fillers	low bf low ap	temible	NA	NA	NA	7	7	NA	NA
Fillers	low bf low ap	vagancia	NA	NA	NA	8	8	NA	NA
Fillers	low bf low ap	fianza	NA	NA	NA	6	6	NA	NA
Fillers	low bf low ap	vigilancia	NA	NA	NA	10	10	NA	NA
Fillers	low bf low ap	grasiento	NA	NA	NA	9	9	NA	NA
Fillers	low bf low ap	montaje	NA	NA	NA	7	7	NA	NA
Fillers	low bf low ap	labranza	NA	NA	NA	8	8	NA	NA
Fillers	low bf low ap	patinaje	NA	NA	NA	8	8	NA	NA
Fillers	low bf low ap	carruaje	NA	NA	NA	8	8	NA	NA
Fillers	high bf high ap	cristalero	NA	NA	NA	10	10	NA	NA
Fillers	high bf high ap	ayudante	NA	NA	NA	8	8	NA	NA
Fillers	high bf high ap	papelera	NA	NA	NA	8	8	NA	NA
Fillers	high bf high ap	jardinero	NA	NA	NA	9	9	NA	NA
Fillers	high bf high ap	pensador	NA	NA	NA	8	8	NA	NA
Fillers	high bf high ap	librero	NA	NA	NA	7	7	NA	NA
Fillers	high bf high ap	limpiador	NA	NA	NA	9	9	NA	NA
Fillers	high bf high ap	letrista	NA	NA	NA	8	8	NA	NA
Fillers	high bf high ap	mentiroso	NA	NA	NA	9	9	NA	NA
Fillers	high bf high ap	coronilla	NA	NA	NA	9	9	NA	NA
Fillers	high bf low ap	arboleda	NA	NA	NA	8	8	NA	NA
Fillers	high bf low ap	extrañeza	NA	NA	NA	9	9	NA	NA
Fillers	high bf low ap	pobreza	NA	NA	NA	7	7	NA	NA
Fillers	high bf low ap	veraniego	NA	NA	NA	9	9	NA	NA
Fillers	high bf low ap	alteza	NA	NA	NA	6	6	NA	NA
Fillers	high bf low ap	bajeza	NA	NA	NA	6	6	NA	NA
Fillers	high bf low ap	golpazo	NA	NA	NA	7	7	NA	NA
Fillers	high bf low ap	perruno	NA	NA	NA	7	7	NA	NA
Fillers	high bf low ap	simpleza	NA	NA	NA	8	8	NA	NA
Fillers	high bf low ap	patriota	NA	NA	NA	8	8	NA	NA