Bilingualism: Language and Cognition

cambridge.org/bil

Lexical distributions of Spanish dialects and their processing implications for Chilean Spanish

Julian van Bijnen[1], Roberto A. Ferreira[2,3] and Ton Dijkstra[1]

[1]Donders Centre for Cognition, Radboud Universiteit Nijmegen, Netherlands; [2]Millennium Nucleus for the Science of Learning (MiNSoL), Chile and [3]Facultad de Ciencias de la Educación, Universidad de Talca, Chile

Research Article

Cite this article: van Bijnen, J., Ferreira, R.A. and Dijkstra, T. (2026). Lexical distributions of Spanish dialects and their processing implications for Chilean Spanish. *Bilingualism: Language and Cognition* 1–21. https://doi.org/10.1017/S1366728925100941

Received: 14 March 2025
Revised: 4 December 2025
Accepted: 5 December 2025

Keywords:
dialects; Spanish; Chile; lexical decision; simulation

Corresponding author:
Ton Dijkstra;
Email: t.dijkstra@donders.ru.nl

This research article was awarded Open Data and Open Materials badges for transparent practices. See the Data Availability Statement for details.

Abstract

Word frequency databases like SPALEX and SUBTLEX-ESP treat Spanish as a uniform language, but prior studies and an initial survey (Experiment 1) revealed significant lexical differences between Spanish in Spain and Latin American countries, especially Chile. To establish subjective frequencies of Spanish word usage, an extended survey (Experiment 2) was conducted with Chilean participants, categorizing words by usage area: General, Spain, Chile, and Latin America. Consistent with the initial survey, Chilean participants assigned subjective higher ratings to General and Chilean words. In a lexical decision experiment (Experiment 3), participants responded faster and more accurately to words from these categories. Using survey data, simulations with Multilink+ (Experiment 4) revealed that subjective word ratings better predicted Chilean reaction times than frequencies from existing databases. These findings emphasize the need to address Spanish dialectal differences in research, with word ratings offering a more accurate measure of region-specific lexical nuances than current databases.

Highlights

- Spanish dialects from around the world reveal significant lexical differences.
- In lexical decision, Chileans respond faster to general and Chilean Spanish words.
- Multilink+ simulations support Chilean differences in Spanish lexical usage.
- Word ratings outperform frequency databases for region-specific lexical items.
- Frequency measures for Spanish should be tailored to specific user populations.

1. Introduction

When Google Translate is asked for the Spanish translation of "popcorn," it returns three different translations: *palomitas*, *rosetas*, and *ñaco*. Remarkably, *palomitas* is only used in Spain, Mexico, and parts of Central America, but *rosetas* and *ñaco* are not used in any country! In fact, *rosetas* is an artificial word used in books and movies that are translated to Spanish for the Latin American market. In real life, at least 12 different terms are used for this one concept throughout Hispanoamerica. In Chile, for example, the word for popcorn is *cabritas*, and in Colombia, it is *crispetas*.

The various language variants spread out over the Spanish speaking world are usually referred to by the common term "Spanish," but they are in fact distinct *dialects* (Sorenson, 2021). Dialects are defined as mutually intelligible versions of a language associated with a particular group of speakers, characterized by distinct features in pronunciation, grammar, and vocabulary (Pullum, 2017).

Dialects mainly differ in terms of pronunciation, syntax, and vocabulary. Differences in pronunciation include rhythm, intonation, and the way specific letters or letter combinations are pronounced. Examples are differences in the pronunciation of the letter "s," called *seseo* (Boyd-Bowman, 1956), and the letter combination "ll," called *yeismo* (Rost Bagudanch, 2017). Variation in syntax includes different preferences in word order, use of pronouns, and verb conjugation, for example, the use of *vos* instead of *tú* in Argentina and Uruguay (Penny, 2000; Weinberg & Beatriz, 1977). Vocabulary differences probably contribute most to dialectal differences (Eddington, 2021) arising from colonization and migration (González & Barrios, 2023), linguistic factors (Shea, 2022), and sociolinguistic factors (Jucks & Paus, 2013).

The study of Spanish dialects reveals considerable linguistic variation across syntactic, pragmatic, and lexical domains. Toribio (2000) highlights parametric differences in subject licensing across dialects, illustrating rich syntactic diversity. Gallego (2019) further analyzes variation in phenomena such as word order and differential object marking. Pragmatic variation is also evident: Márquez Reiter (2022) contrasts indirect requests in Peninsular and Uruguayan Spanish, showing how social norms influence indirectness. Orozco (2018), examining

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

Colombian and New York City Spanish, shows how migration shapes linguistic convergence. Eddington (2022) applies computational methods to classify dialect regions based on lexical and grammatical traits, reinforcing the global complexity of Spanish variation. These studies demonstrate the interplay of linguistic and social factors in shaping dialectal diversity.

Complementing these structural approaches, perceptual dialectology reveals how speakers perceive similarities and differences among Spanish varieties. Alfaraz (2022) links perceptual judgments to historical, sociological, and cognitive dimensions, showing how dialect perception reflects broader sociolinguistic contexts. Moreno-Fernandez (2015) found that Spanish speakers perceive Latin American varieties as relatively homogeneous, with Spain seen as more distant; Chile, grouped within the Latin American cluster, showed perceptual variation tied to age and origin. Similarly, Quesada Pacheco (2014) identified perceptual clusters among American Spanish speakers, noting that Chilean Spanish is sometimes viewed as distinct due to isolation and indigenous lexical influence. These studies highlight how subjective beliefs often shape dialect classifications, influencing how lexical variation is understood in daily communication.

Few empirical studies systematically differentiate Spanish dialects, partly due to their number and overlap. Spanish is official in 21 countries (Eddington, 2021), but shared media sources complicate dialect-specific data collection. This lack of clear differentiation affects both public understanding and scholarly analysis. To identify what words people actually use, it is crucial to examine frequency databases.

Lipski (2012, 2017) has contributed extensively to this field, aiming to map dialectal boundaries through lexical data. His research emphasizes the cultural and historical context of vocabulary, for example, African-language influences in Caribbean Spanish (Delgado-Díaz, 2018) and lexical borrowing from indigenous languages in Mexico and the Caribbean. Although Lipski's main focus is phonology and grammar, his work sets the stage for vocabulary-centered studies. Goodrich and Lonigan (2018) show how dialectal exposure affects vocabulary development in bilingual children, underscoring the role of dialects in shaping lexical knowledge. Figure 1, adapted from Lipski (2012), provides a general overview of Spanish dialects, where color coding indicates similarities and distinctions across regions. This map serves as a useful starting point for our investigation and inspires hypotheses regarding lexical divergence.

In the present study, we focus specifically on vocabulary differences between Spanish dialects. Dialects may use different words with the same form but different meanings or use entirely different words for the same concept. An example of the first type is *carro*, which can mean "car" or "cart" depending on the country. An example of the second type includes *frutilla* and *fresa*, both meaning "strawberry" but used in different regions.

Dialects often differ not only in lexical choice but also in the frequency with which particular terms are used. Word frequency databases capture this information by quantifying word usage across language corpora such as newspapers or movie subtitles (New et al., 2007). These databases are constructed through different methodologies, but all rely on counts from written or spoken sources. For example, the SPALEX database (Aguasvivas et al., 2018), developed through collaboration among institutions in Spain, Belgium, and the Netherlands, contains reaction times alongside word frequencies and was the first crowdsourced lexical decision mega-study in Spanish. A close look at SPALEX reveals that most of its source data come from Spain. This regional bias becomes evident in Table 1, where word frequencies favor Spanish forms over their Latin American equivalents, even though most Spanish speakers live in Latin America.

The second database, SUBTLEX-ESP (Cuetos et al., 2011), also aims to represent the Spanish language as a whole. Released in 2011 and comprising 41.5 million words, it was designed to improve upon existing databases. SUBTLEX-ESP used subtitles from 3,523 movies (20,253,754 words) and 257 TV series (21,323,919 words), sourced from three websites: (a) www.argenteam.com, (b) www.subdivx.com, and (c) www.solosubtitulos.com. Although it is hard to determine the exact dialectal representation, user traffic suggests that the primary sources are Argentina, Chile, and Mexico.

These data support the assumption that dialectal origin influences frequency representation. Words common in Argentina or Latin America appear more frequently, yet other dialects remain underrepresented.

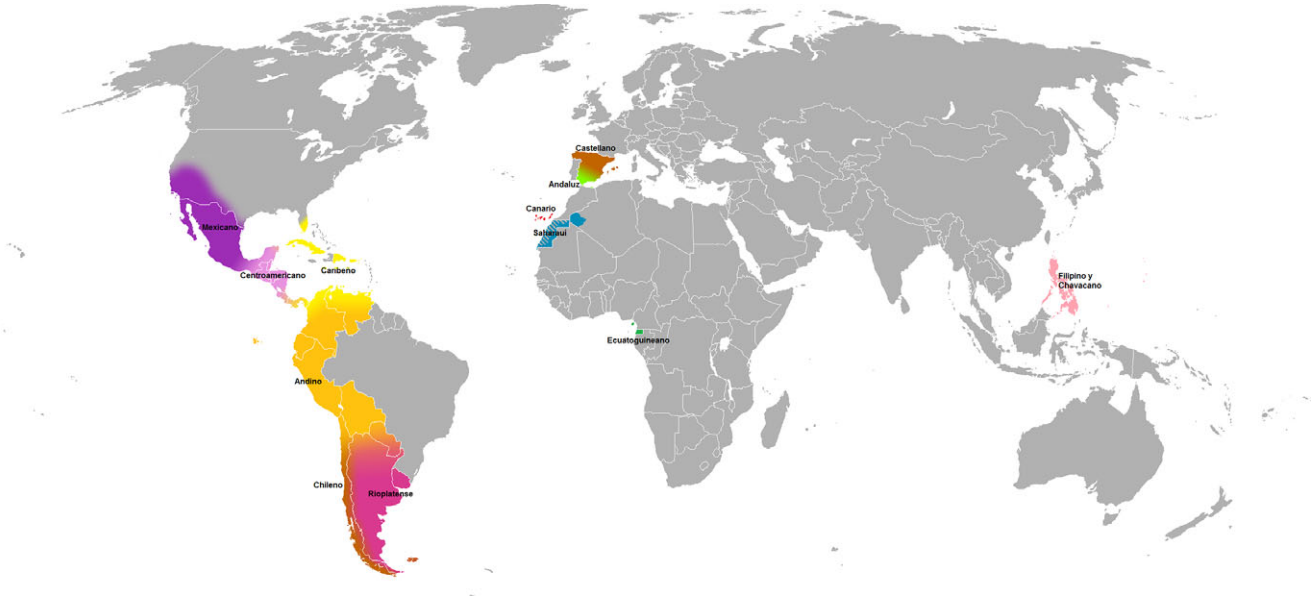


Figure 1. Distribution of Spanish dialects across the world according to Lipski (2012).

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

Table 1. SPALEX and SUBTLEX-ESP occurrences per million (OPM) for typical words from Spain and Latin America

Word Spain	SPALEX	SUBTLEX-ESP	Word Latin America	SPALEX	SUBTLEX-ESP
COCHE	55.5	115	AUTO	26.31	428
ORDENADOR	19.5	6.92	COMPUTADORA	11.3	63.0
ZUMO	0.28	1.95	JUGO	0.06	25.6
CAMARERO	4.27	8.89	MESERO	0.32	9.76
FONTANERO	0.61	2.00	PLOMERO	0.02	4.66

Overall, it is misleading, even incorrect, to treat a diverse language like Spanish as a monolith within a single database. Dialects differ in phonetic realization, prosody, and lexical distribution, and using a single frequency source can hinder accurate dialect classification (Hansen & Liu, 2016). The lack of lexical distinction among Spanish varieties can skew psycholinguistic findings. For instance, Martin et al. (2016) showed that dialect-specific vocabulary affects processing: Listeners adjust to the phonological and lexical cues of dialectal speakers. Ignoring these nuances risks flawed interpretations of comprehension and cognitive function. From both scientific and sociolinguistic perspectives, developing dialect-specific frequency databases appears necessary.

This approach has precedent. English-language research has long distinguished between British and American variants, with separate usage norms and even academic writing guidelines (Biber, 1987; Scott, 1993). Similar divisions exist in French (European, Canadian, and African). Yet Spanish dialects remain comparatively underexamined. This study seeks to address that gap by focusing on vocabulary variation across Spanish dialects, a dimension still insufficiently represented in the literature.

1.1. Present research

Our research is structured into several key stages, each designed to address distinct aspects of lexical variation in Spanish across different regions. We begin by examining lexical distinctions between European Spanish and various Latin American Spanish dialects, with a focus on differences in both word meaning and frequency of use. This initial analysis aims to determine whether a word frequency database constructed from subjective word frequencies offers a more accurate representation of a specific language variant compared with currently available generalized databases.

Previous evidence has shown that the construction of word frequency databases based on subjective word frequencies can indeed offer a more accurate representation of specific language variants compared with generalized databases. For instance, Shade and Altmann (2023) argue that the representation of text through vocabulary and word frequency distributions can reveal significant statistical regularities that are sensitive to the context in which the language is used. Generalized databases often fail to capture the nuances of specific dialects, which can lead to misrepresentations of how frequently certain words are used in particular communities. Additionally, Gimenes and New (2016) argue that subjective database reflecting actual use in everyday communication within a dialect can enhance the validity of psycholinguistic studies by ensuring that the stimuli reflect the actual language experience of the participants.

If such a database proves more representative, it would support the case for developing dialect-specific frequency resources. Conversely, if existing generalized databases are found sufficient, separate dialectal databases may be unnecessary.

To explore this, we will first conduct an international survey to assess the frequency of Spanish word usage across different dialects (Experiment 1). The data gathered from this survey will be analyzed to identify differences among the Spanish dialects of the participating countries. This analysis will highlight cross-dialectal vocabulary differences, particularly at the group level rather than between individual dialect pairs.

Following the survey analysis, we will categorize the words based on regional usage, which will allow us to examine variations in responses, with a special focus on Chilean participants. Building on these initial findings, we will conduct an expanded version of the survey, concentrating exclusively on Chilean respondents (Experiment 2).

Subsequently, the same Chilean participants will engage in a Spanish Lexical Decision task involving all the words from the survey (Experiment 3). This will enable us to convert subjective word ratings from Chile into frequency measures, which we will then apply in computational model simulations using Multilink+ (Dijkstra et al., 2019, 2023; Dijkstra & Peeters, 2023), a model designed for monolingual and multilingual word retrieval. These simulations (Experiment 4) will allow us to compare the predictive accuracy of frequency measures derived from SPALEX and SUBTLEX-ESP with those obtained from subjective word ratings.

In summary, our research aims to achieve two main objectives. The first is to determine whether subjective word frequency measures can better predict Chilean reaction times (RTs) than existing databases such as SUBTLEX-ESP and SPALEX. We hypothesize that word frequencies based on subjective ratings will outperform both databases, as they more accurately reflect the language use of our study participants. The second objective of our study is to investigate how Chilean participants respond to words from different Spanish dialects and assess the extent to which Multilink+ can predict these responses. We expect that Chilean participants will respond faster and more accurately to words aligned with their own dialect and that Multilink+, when using subjective frequency inputs, will effectively capture this pattern.

2. Experiment 1: Survey of Spanish dialects

In Experiment 1, we used a survey to quantify word frequency differences between Spanish dialects. The survey assessed subjective ratings of word use and should be sensitive to geographical dialect differences. Following the earlier mentioned study by Lipski (2012, 2017), we predicted that a north–south division would show up.

2.1. Method

The online survey was conducted using Google Forms and was open to all native Spanish speakers. It was divided into two sections: (1) collected background information, with the exact questions

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

detailed in the Appendix, and (2) word frequency data. The word list used for the experiment consisted of handpicked words, half of which were likely to have different frequencies in different countries and half of which likely to have similar frequencies in different countries. Participants rated these words based on their personal usage frequency, with a scale from 1 ("practically never") to 5 ("very often"). The entire survey, including texts, questions, and answers, was formulated in Spanish.

2.2. Results

Data analysis was conducted using Python (Version 3.10) and the Pandas library (Version 1.4). Initially, 600 participants submitted responses to the survey. To ensure data quality, data from participants exhibiting abnormal response patterns were excluded from the analysis. Abnormal patterns included consistently rating all words with a score of 1 or displaying recurring patterns across responses. Additionally, data from participants who did not provide ratings for 10 or more words were removed. Following these exclusions, the dataset comprised 590 participants.

Outlier detection was performed at the word level. The overall mean rating across all words was 3.48 (standard deviation [SD] = 1.01). No words were identified as outliers, because the mean ratings for all words fell within 3 SDs of the overall mean. Data were collected for the dialects in the following countries (number of participants in parentheses): Spain (185), Uruguay (90), Chile (66), Argentina (48), Mexico (44), Costa Rica (41), Venezuela (33), Colombia (20), and Peru (20).

To compare dialects and potentially group them, we calculated the mean rating for each word in each dialect, excluding zero values. Outlier analysis for each word for each country was performed. As a result, 380 individual entries were removed, as they were more than 3 SD from the mean.

The next step was to divide the array for each dialect into two parts: (a) words expected to have similar usage frequencies across all dialects and control words (LOWDIF array) and (b) words expected to be unique to certain dialects or have different meanings in different dialects (HIGHDIF array). For HIGHDIF words, considering colloquialisms or dialect-specific words, the calculations resulted in Table 2. Note that all regression numbers are absolute, as the direction of the correlation was irrelevant. The table can also be visualized as a heatmap (Figure 2).

We then compared the survey's results to existing collections of word frequency measures, in particular the SPALEX (Aguasvivas et al., 2018) and SUBTLEX-ESP (Cuetos et al., 2011) databases. We calculated correlations between the collected word ratings from the survey and their respective log frequencies from SPALEX and SUBTLEX-ESP to determine which database best reflected word use in each country. These correlations are presented in Table 3.

2.3. Discussion

We hypothesized that our results would reveal a north–south division similar to that observed in Lipski's (2012) dialectal map (see Figure 1). The heatmap in Figure 2 supports this hypothesis and also shows that the Spanish dialect from Spain is significantly different from all others. The highest correlation for Spain was .42 with Colombia, whereas its correlations with Peru (p = .12) and Mexico (p = .37) were not statistically significant at $p < .05$.

Within Latin America, distinct dialectal clusters emerged, with some correlations as low as .38. However, all correlations within Latin America were statistically significant ($p < .005$). The data indicate two major dialect clusters: (1) Uruguay, Chile, and Argentina, and (2) Mexico, Venezuela, Costa Rica, and Colombia, consistent with Lipski's (2008) classification of Spanish dialects into clusters based on phonological features, vocabulary usage, and syntactic structure. Peru demonstrated high correlations with both groups but slightly stronger alignment with the northern cluster (.73 vs. .65), consistent with prior linguistic research (Gallego, 2019).

These findings not only reinforce the concept of regional dialects but also raise questions about how to define dialectal boundaries. Dialects do not conform to national borders, suggesting that regional groupings may be more appropriate than country-based classifications. Given that some Spanish-speaking countries are significantly larger than Spain, substantial internal dialectal variations are to be expected (Orozco, 2018). Future research should explore more fine-grained regional distinctions to improve our understanding of dialect variation.

The primary objective of Experiment 1 was to examine dialectal differences between Spanish from Spain and Latin America and advocate for a more formal classification of these dialects. Our findings reveal substantial discrepancies in vocabulary and word frequency, with the most pronounced divergence between the Spanish spoken in Spain and that spoken in Latin America. These results align with previous corpus-based studies that consistently report systematic lexical and phonetic divergences between Peninsular and Latin American Spanish (Eddington, 2021; Lipski, 2017).

Furthermore, the findings provide strong motivation for expanding surveys to estimate word frequencies in specific dialects in greater detail. Future work should explore spoken language data,

Table 2. Correlations between dialects for HIGHDIF words

	Uruguay	Chile	Argentina	Mexico	Costa Rica	Venezuela	Colombia	Peru
Spain	0.36	0.29	0.41	0.13	0.28	0.29	0.42	0.21
Uruguay		0.82	0.97	0.50	0.47	0.47	0.45	0.62
Chile			0.83	0.54	0.49	0.38	0.48	0.70
Argentina				0.51	0.47	0.48	0.46	0.63
Mexico					0.76	0.66	0.65	0.80
Costa Rica						0.79	0.82	0.74
Venezuela							0.76	0.64
Colombia								0.73

Note: All correlations not involving Spain are significant at $p < .005$.

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

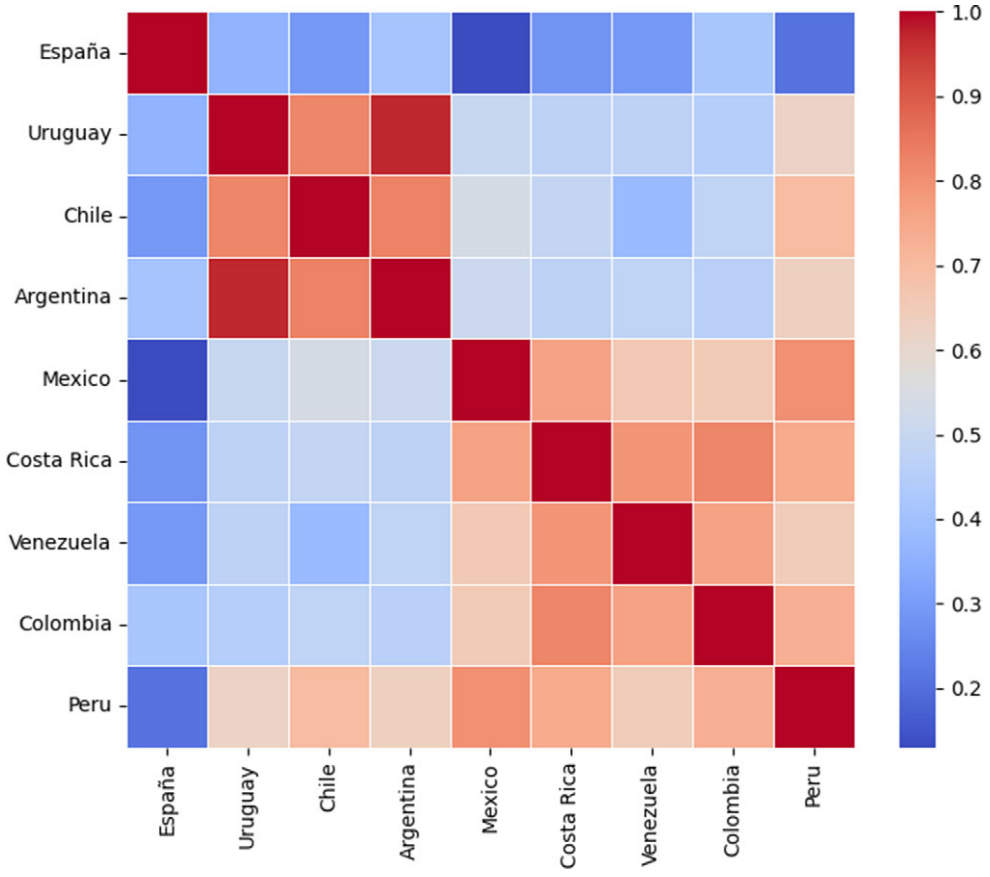


Figure 2. A heatmap of the correlation matrix for all dialect pairs in the survey.

Table 3. Correlations between country ratings and log word frequencies from SPALEX or SUBTLEX-ESP

Country	SPALEX	SUBTLEX-ESP
Spain	0.53	0.38
Uruguay	0.51	0.56
Chile	0.62	0.64
Argentina	0.52	0.58
Mexico	0.66	0.75
Costa Rica	0.58	0.58
Venezuela	0.49	0.55
Colombia	0.52	0.54
Peru	0.63	0.67

Note: All correlations are significant at $p < .001$.

which may further elucidate dialectal divergence beyond written word frequencies (González & Barrios, 2023). The observed correlations and regional divisions suggest that linguistic models and frequency databases should incorporate dialect-specific frequency norms to improve their accuracy and applicability to real-world language use.

The findings lead us to conclude that there are substantial dialectal differences between Spanish spoken in distinct Hispanoamerican countries, in particular between countries in the north versus the south of Latin America. This geographic distinction provides a strong motivation for a more extensive survey to estimate word frequencies in specific dialects in greater detail.

3. Experiment 2: Extended survey of Chilean words

Building on the insights gained from our survey of Spanish dialects (Experiment 1), we turned our attention to the Chilean dialect. Notable for its unique vocabulary and reputation as one of the more challenging dialects for speakers of other Spanish variants to understand (Vivanco Rojas, 2018), Chilean Spanish serves as a compelling focus for further investigation. This phase of the study was conducted in collaboration with the Millennium Nucleus for the Science of Learning in Talca, Chile.

Experiment 1 demonstrated the validity of using surveys as a general measure of word frequency. Accordingly, Experiment 2 employed an updated version of the previous survey, tailored specifically to Chilean participants. Given the presence of a substantial overlap in word materials (discussed later), we anticipated a strong correlation between the newly collected Chilean ratings and those obtained in Experiment 1. Furthermore, Chilean ratings were also compared with ratings from other Spanish-speaking countries, namely Argentina, Mexico, and Spain. Based on lexical and cultural proximities, we predicted the highest correlation to be with ratings from Argentina (.8–.9), followed by Mexico (.7–.8), and finally Spain (below .5). Categorizing the presented word materials into several groups based on their Chilean usage contexts relative to other countries, provided further insight into the size and nature of the differences between Chilean and other Spanish dialects. Although Spanish dialectal variation cannot be reduced to a simple

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

north–south geographical division, the clustering observed here reflects an empirical pattern emerging from our dataset rather than a general claim about Spanish lexical isoglosses.

A major aim of Experiment 2 was to obtain subjective frequency data that can be compared with human lexical decision responses to be collected in Experiment 3 and computational model data simulated in Experiment 4. The obtained survey data allowed us to assess whether word frequencies derived from subjective ratings, focused on the Chilean dialect, can outperform SPALEX and SUBTLEX-ESP databases in experiments and simulations of human word processing. Due to the presence of different types of items in the survey, it provided critical data for evaluating the utility of subjective word ratings in capturing dialect-specific lexical nuances.

3.1. Method

The stimulus list for this survey was an extended version of the previous list, now including more words typically used in Chile. The total number of words increased from 80 to 153, divided into five categories: (a) General – common words in almost all Spanish dialects; (b) Spain – typical words from Spain; (c) Chile – typical words from Chile; (d) Latin America – words used in other countries of Latin America but not in Chile; and (e) Meaning – words with clearly different meanings in different countries. This categorization was designed to create usage-based lexical groupings that would differ in regional association and frequency, rather than representing strict geographical boundaries. It was expected that Chilean participants would react differently to each category, with higher mean word ratings anticipated for categories General and Chile than for Spain, Latin America, and Meaning. Participants were not informed about the word categories to avoid bias.

To avoid suggesting categorical dialect boundaries, we note that Spanish lexical isoglosses form overlapping continua rather than discrete regions. Consequently, some items classified as "Chile," "Spain," or "Latin America" may also occur outside their dominant region. Such cross-regional distributions are inherent to Spanish lexical variation and were therefore not treated as errors, but as part of the empirical phenomenon under study. Because this overlap introduces variability rather than artificial structure, any observed category differences are likely to be conservative rather than inflated.

Stimulus selection was conducted in two stages. First, candidate items were identified using available lexical resources and Chilean/Spanish frequency lists, focusing on words for which there was prior indication of either (a) relative stability across dialects or (b) cross-dialectal variation in usage or association. Second, these candidates were structured so that approximately half of the list consisted of words predicted to vary in regional frequency or usage (Spain, Chile, and Latin America categories), and half consisted of words predicted to be stable across dialects (General category), with an additional set targeting semantic variation (Meaning category). This two-stage selection process ensured that the stimulus set included both frequency-stable and frequency-variable lexical items, allowing direct comparison of their effects within the same participant sample. Items with rare, technical, or highly ambiguous orthography were excluded. Throughout this article, the term "Latin America" is used as a heuristic label referring to non-Peninsular Spanish usage patterns and does not imply homogeneity across Latin American varieties.

A total of 37 participants took part in the survey, each presented with the same 153 words. The survey, conducted digitally through Google Forms, began with background questions followed by the rating of each word in random order. The mean age of the participants was 22.13, ranging from 18 to 31. Of the participants, 67% were female, 30% male, and 3% preferred not to specify. All participants were from Chile and had Spanish as their first language, with 49% also speaking English and 42% not speaking any other language. All participants were either university graduates or current students. These participants also took part in the RT study reported in Experiment 3, which ensured consistency across subjective ratings and behavioral data.

3.2. Results

After all surveys were completed, the participants' subjective ratings were averaged. As in the previous survey, submissions were checked for odd response patterns, such as only using scores of 1 or 5 or recurring patterns. None of these issues were observed among the 37 participants, allowing all data to be included in further analyses.

Next, including only words present in both lists, our new Chilean word ratings were correlated with the previously collected ratings, serving as a validation step to verify consistency. Because the two sets of ratings were based on different participants, some rating differences might be expected. We expected the highest correlation with previous Chilean results. Furthermore, in the original survey, Argentina's dialect was grouped with Chile, while Mexico's was in a different Latin American group, and Spain's was most different from Latin American dialects. We therefore expected lower correlations of the new results from Chile with other Latin American countries, and a minimal correlation with Spain.

The resulting scatterplots are shown in Figure 3. The pattern of correlations supports our hypotheses, as all correlations except for Spain are significant at $p < .0001$. For Spain, no significant correlation was obtained.

We also computed the correlation between the survey ratings and actual word frequencies from SPALEX and SUBTLEX-ESP using log frequencies. This analysis aimed to identify which database our ratings resembled most closely, with an expectation for SUBTLEX-ESP based on literature. This information is crucial for the simulation study (Experiment 4). In Figure 4a, b, SUBTLEX-ESP scores higher with a correlation of .70 compared with SPALEX's .55. SUBTLEX-ESP aligns better with these word ratings, consistent with earlier findings. This is in line with our hypothesis, and SUBTLEX-ESP word frequencies will be used as the basis for the simulation study presented later.

3.3. Discussion

In this experiment, we examined how Chilean participants rated Spanish words across various categories. We hypothesized that the new survey results would show a strong correlation with previous Chilean ratings, a slightly lower correlation with Argentinian ratings, a moderate correlation with Mexican ratings, and a very low correlation with Spanish ratings.

The results aligned with our expectations. Correlation analysis between the new ratings and those from the original survey revealed a very high correlation of .95 ($p < .0001$) with previous Chilean ratings, indicating high consistency despite individual differences. The correlation with Argentinian ratings was also strong (.86, $p < .0001$), reflecting the close lexical relationship between Chilean and Argentinian Spanish (Lipski, 2017). The Mexican ratings showed a moderate correlation of .70, likely due to shared Latin American vocabulary (Gallego, 2019). Finally, Spanish ratings had a negligible correlation of .14, underscoring the substantial differences between Chilean and European Spanish (Orozco, 2018).

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

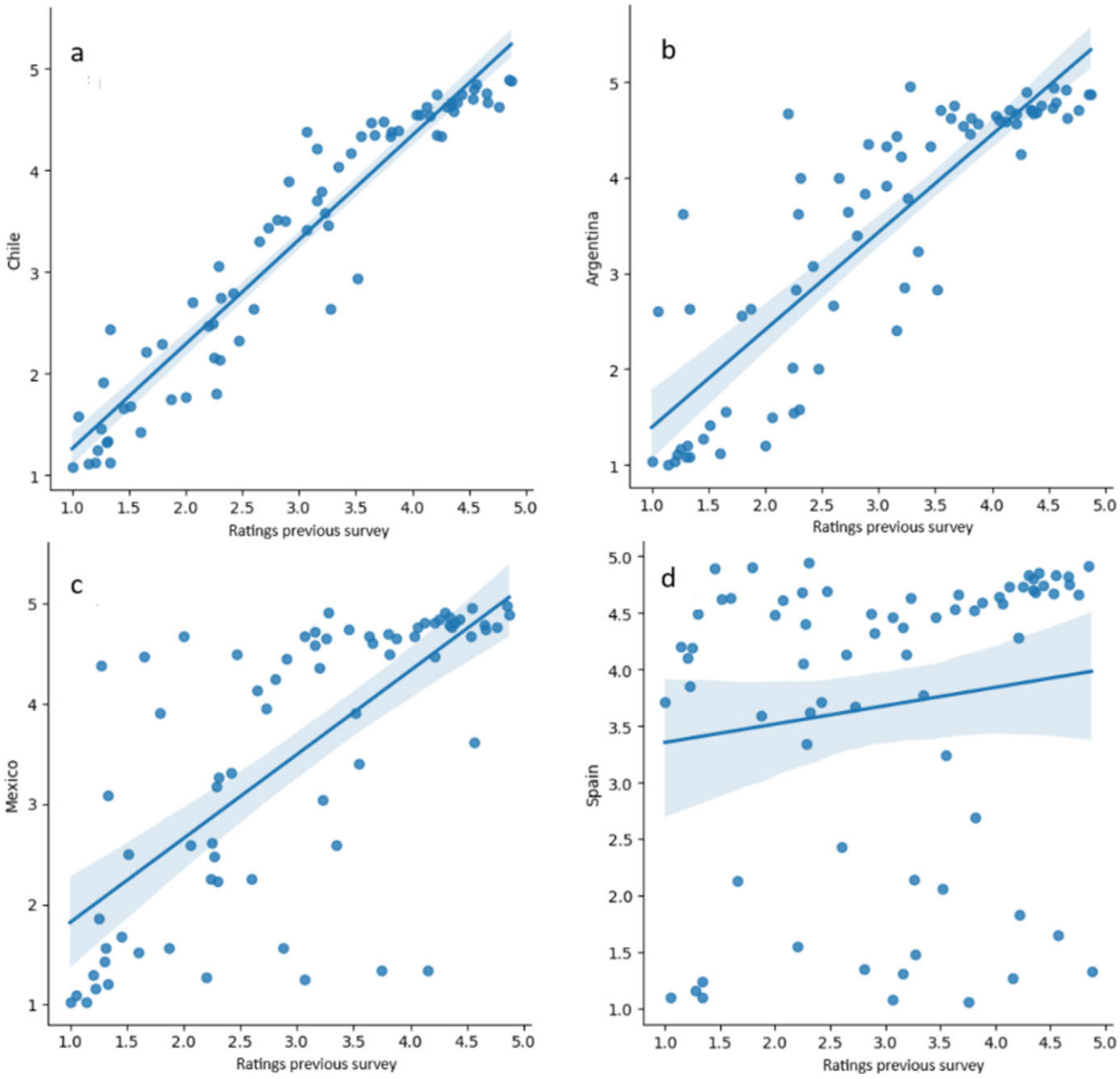


Figure 3. Scatterplots of current Chilean word ratings versus previous ratings (freq_survey) from (a) Chile (r = .95), (b) Argentina (r = .86), (c) Mexico (r = .70), and (d) Spain (r = .14).

The participant sample in this survey primarily consisted of individuals from a similar age group and region in Chile, potentially limiting the generalizability of findings to older populations or other areas of the country. However, comparisons with a previous dataset, which included greater age and geographic diversity, revealed that intracountry variation was minimal compared with intercountry differences (Eddington, 2021). This suggests that regional lexical variation within Chile is substantially smaller than variation between Chile and other Spanish-speaking countries.

The newly added Chilean-specific words were excluded from the correlation calculations but remain integral to the subsequent phases of this study. These words will be particularly relevant in Experiments 3 and 4, which focus on lexical decision tasks and computational modeling.

Mean ratings per category revealed clear patterns. Participants demonstrated greater familiarity with words from the General and Chilean categories, reflecting widespread usage across Spanish-speaking populations and strong regional representation (González & Barrios, 2023). Words from Spain and other Latin American countries received lower familiarity ratings, further validating regional lexical divergence (Sadowsky & Martínez-Gamboa, 2012). Words with different meanings across dialects exhibited high variance, reinforcing the need for further linguistic distinction (Lipski, 2012). These results provide converging evidence that subjective familiarity ratings capture meaningful lexical-distributional differences across dialects.

A key limitation of this classification scheme is that Spanish lexical variation does not form discrete blocks, but rather overlapping continua shaped by historical, sociolinguistic, and media-related influences. Several words classified as "Chile," "Spain," or "Latin America" may therefore also be known or used outside their dominant region. Importantly, such overlap introduces noise rather than artificial structure and thus would be expected to

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

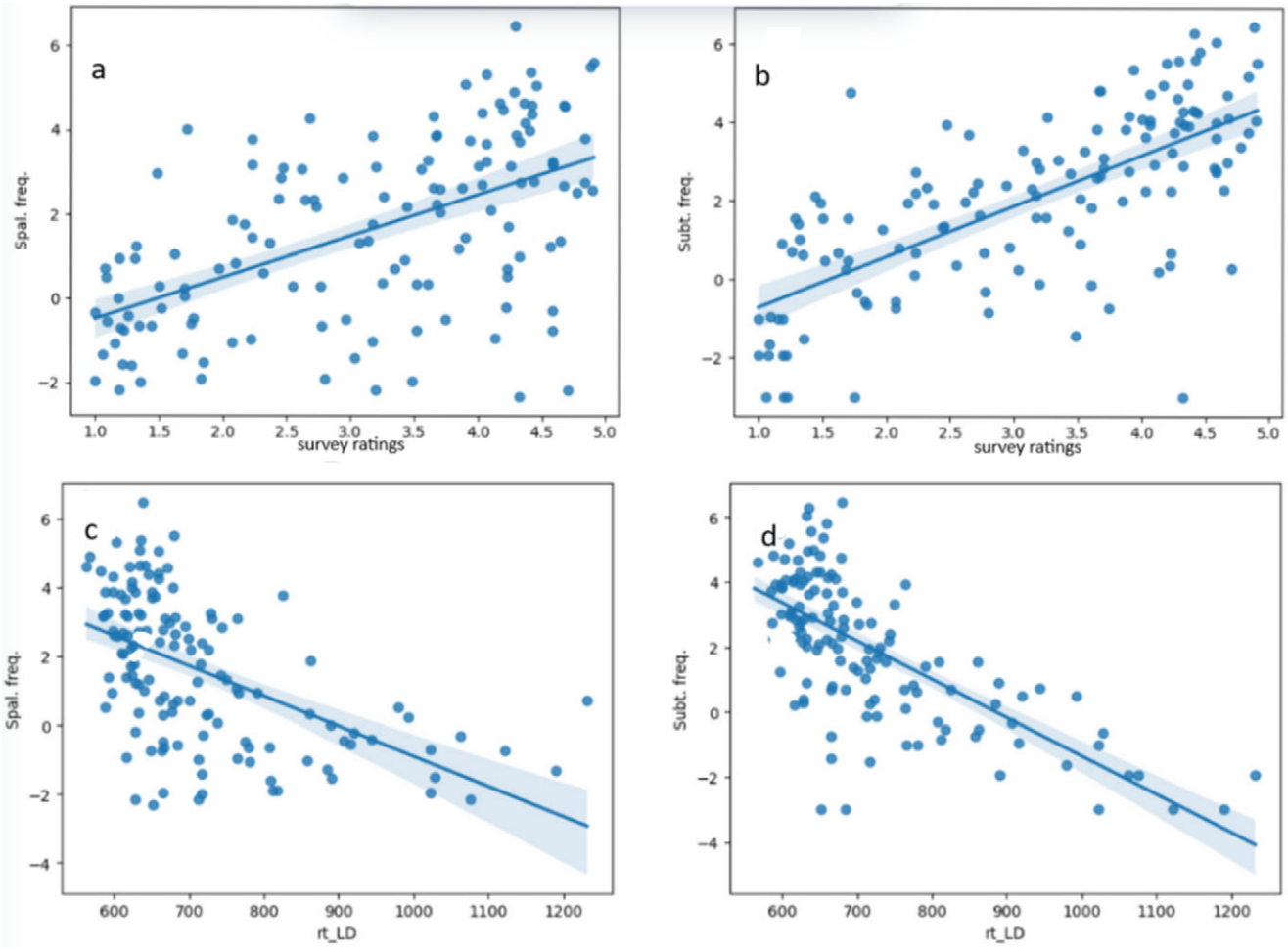


Figure 4. Scatterplots of Chilean survey word ratings versus log word frequencies from (a) SPALEX ($r = .55$) and (b) SUBTLEX-ESP ($r = .70$) and of lexical decision reaction times (rt_LD) versus log word frequencies from (c) SPALEX ($r = -.54$) and (d) SUBTLEX-ESP ($r = -.69$). All significances at $p < .0001$.

attenuate rather than exaggerate category contrasts. Consequently, the differences observed in the present study are likely to be conservative estimates of true regional lexical divergence.

Although the number of words per category varied, this distribution mirrors natural lexical usage patterns across Spanish dialects. This enhances the representativeness of our dataset for future linguistic research. Despite category size differences, the results consistently highlight the distinctiveness of the Chilean dialect and its divergence from other Spanish variants (Rogers & Face, 2024). Future studies should test whether these patterns generalize to spoken language processing and whether continuous measures of dialectal lexical probability may offer an even more precise framework than discrete category labels.

4. Experiment 3: Lexical decision study

In our surveys, we explored how word frequencies and word ratings can measure an individual's familiarity with a word. Reaction times (RTs) provide a more direct and objective measure of word use in a lexical decision experiment. The speed and accuracy of a lexical decision reflect how well a word is represented in the participant's mental lexicon and how fast it can be retrieved from there.

In Experiment 3, we examined whether the subjective word ratings obtained in the surveys align with reaction time (RT) patterns in a lexical decision task. A participant's reaction times for the same word set will differ between Spanish dialects, as word usage frequency varies across dialects. To obtain accurate behavioral measures of word retrieval, we collected RT data for 153 words included in the extended survey from the same group of Chilean participants (see the Appendix for stimulus materials). These RT data will also serve later as a comparison measure to validate predictions from simulations that mimic the human word retrieval process in detail. Such predictions are generated by the Multilink+ model, detailed in Experiment 4.

We hypothesized that Chilean RTs would differ significantly from SPALEX RTs, reflecting dialect-specific variations in word use. This hypothesis was tested by correlating RTs from the two sources. If they are highly correlated, Chilean Spanish may not differ much from other Spanish variants.

Two primary analyses of RTs were conducted. The first examined the entire dataset, comparing RTs to various word metrics collected previously. Additionally, survey word ratings were correlated with RTs to evaluate their predictive value. The second analysis explored the influence of word categories (General, Spain, Chile, and Latin America) on predictive accuracy. We anticipated faster and more accurate responses for words in the General and Chile categories compared with those in the Spain and Latin America categories.

Finally, because words with different meanings across Spanish-speaking countries could confound results, they were excluded from the analysis. Detailed information on word characteristics

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

for the various categories can be found in the Supplementary Material.

4.1. Method

Participants were the same 37 undergraduate students who participated in Experiment 2. Stimuli included 153 words from Experiment 2, categorized based on regional usage: (1) General (n = 58; e.g., *casa*, "house," common across dialects); (2) Spain (n = 18; e.g., *coche*, "car," predominant in Peninsular Spanish); (3) Chile (n = 41; e.g., *cabritas*, "popcorn," Chilean-specific); (4) Latin America (n = 28; e.g., *crispetas*, "popcorn," used elsewhere in Latin America); and (5) Meaning (n = 8; excluded in analyses to avoid semantic confounds, e.g., *carro*, "car" or "cart," varying by dialect). Categories were informed by dialectal corpora (Eddington, 2021; Lipski, 2017) and validated by native Chilean linguists to mitigate biases in perceived exclusivity (Quesada Pacheco, 2014). Words were controlled for length (M = 6.4 letters, SD = 1.7) and OLD20 neighborhood density (M = 1.8, SD = 0.5; Duchon et al., 2013). Nonwords (n = 153) were generated via Wuggy (Keuleers & Brysbaert, 2010), matched for length and Spanish phonotactics, with Levenshtein distances >2 from real words to ensure discriminability. All words and nonwords can be found in the Appendix.

During the task, letter strings were shown in black Courier New 18-point font on a white background, with presentation controlled by PsychoPy 2022.1.0 (Peirce, 2007). Each trial began with a 500-ms fixation cross, followed by a 400-ms blank screen, then a letter string presented for 2 s or until a response was made. Participants were instructed to respond quickly and accurately as they could by pressing a "yes" or a "no" button key if a presented item was an existing Spanish word or a nonword letter string, respectively. They completed 10 practice trials before the experiment, and reaction time was recorded in milliseconds. Including instructions, each session took no longer than 25 min.

4.2. Results

The data underwent a detailed cleaning procedure, beginning with the removal of responses to nonwords and the exclusion of four participants who did not adhere to the instructions. These participants either used the wrong input method or mixed up the keys throughout the experiment. Outlier identification led to the exclusion of one additional participant who had a mean reaction time significantly higher than 2.5 SDs above group average. This left data from 32 participants for further analysis. The dataset was then checked for outliers in both accuracy and reaction time, with problematic items and individual responses being removed.

In the following comparisons, the cleaned data were analyzed overall and per category. The first comparison involved word frequencies from SPALEX and SUBTLEX-ESP to determine which better predicts the RT data. To do this, lexical decision reaction times were correlated with word frequencies from both databases. Based on our previous findings and analysis, SUBTLEX-ESP was expected to show a higher correlation than SPALEX. As shown in Figure 4c, d the findings were in line with this expectation: SUBTLEX-ESP scores a correlation of −.69 compared with SPALEX's −.54. Both correlations are negative, as expected in RT and word frequency comparisons (higher RTs correlate with lower frequencies). Note that in this analysis, log frequencies were used.

Next, correlation scores were calculated between the SPALEX RTs and the RTs obtained in this experiment, as well as between word ratings and the new RTs. These correlations provided further insights into the RT patterns. The correlation between SPALEX RTs and the new RTs was expected to be below .75, given the significant differences between the sources of these two sets. The correlation between word ratings and the new RTs was expected to be high, over .70, indicating that it might outperform word frequencies from SPALEX and SUBTLEX-ESP. Note that all analyses mentioned so far were conducted on the entire cleaned dataset. The following two plots in Figure 5a, b display the correlations between SPALEX RTs and new RTs and between word ratings and new RTs. All mentioned correlations are statistically significant at $p < .005$.

The total error rate over all words, before removing errors, was 6.8 %. The resulting mean RT over all words, after cleaning, was 713 ms with an SD of 133 ms. This leads to a 95% confidence interval for the mean RT of [690.31–735.69]. The mean subjective frequency was 3.12. The nonwords had a mean RT of 919 ms and an error rate of 0.18. Table 4 shows the most important statistics calculated over the identified categories mentioned before. These will be relevant for Experiment 4.

4.3. Discussion

Our lexical decision study tested two primary hypotheses. First, we expected SUBTLEX-ESP word frequencies to better predict the newly obtained reaction times than those of SPALEX. Second, we expected the new Chilean RTs to differ significantly from SPALEX RTs.

Our results were in line with the first hypothesis: Scatter plots indicated that Chilean RTs correlate more strongly with SUBTLEX-ESP than with SPALEX log word frequencies. Although both datasets contained outliers, SPALEX displayed more variance and a larger number of extreme values, reinforcing the conclusion that SUBTLEX-ESP is a more reliable predictor of Chilean RTs (Cuetos et al., 2011).

Regarding the second hypothesis, Chilean RTs showed a correlation of only .62 with SPALEX RTs. This relatively modest correlation for "the same words in the same language" underscores the necessity of region-specific Spanish lexical decision RTs, further supporting the distinctiveness of Spanish dialects (Aguasvivas et al., 2018).

Interestingly, subjective word ratings demonstrated an even stronger correlation with RTs than either SPALEX or SUBTLEX-ESP word frequencies. This suggests that subjective ratings provide a highly reliable estimate of lexical accessibility, likely because both measures were obtained from the same participant group, thus capturing properties of their mental lexicon (Balota et al., 2001).

Category-specific analyses further supported these conclusions. General and Chilean words exhibited higher accuracy (0.99 and 0.97), lower mean RTs (651 and 688 ms), and higher mean ratings (3.80 and 3.59). In contrast, Spanish and Latin American words excluding Chile showed lower accuracy (0.79 and 0.87), higher mean RTs (837 and 816 ms), and lower mean ratings (1.55 and 1.99), consistent with participants' relative unfamiliarity with these words (Eddington, 2021).

These differences were statistically significant ($p < .005$), despite variations in word count across categories. The observed disparities in word distribution reflect natural lexical variations across Spanish dialects. The robustness of these category effects underscores the importance of incorporating region-specific frequency measures in lexical decision studies (Lipski, 2012).

In conclusion, this study highlights the necessity of subjective word ratings and region-specific RT norms in psycholinguistic research. Experiment 4 will further evaluate the predictive value of these measures with respect to human language processing through a computational simulation using the Multilink+ model.

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

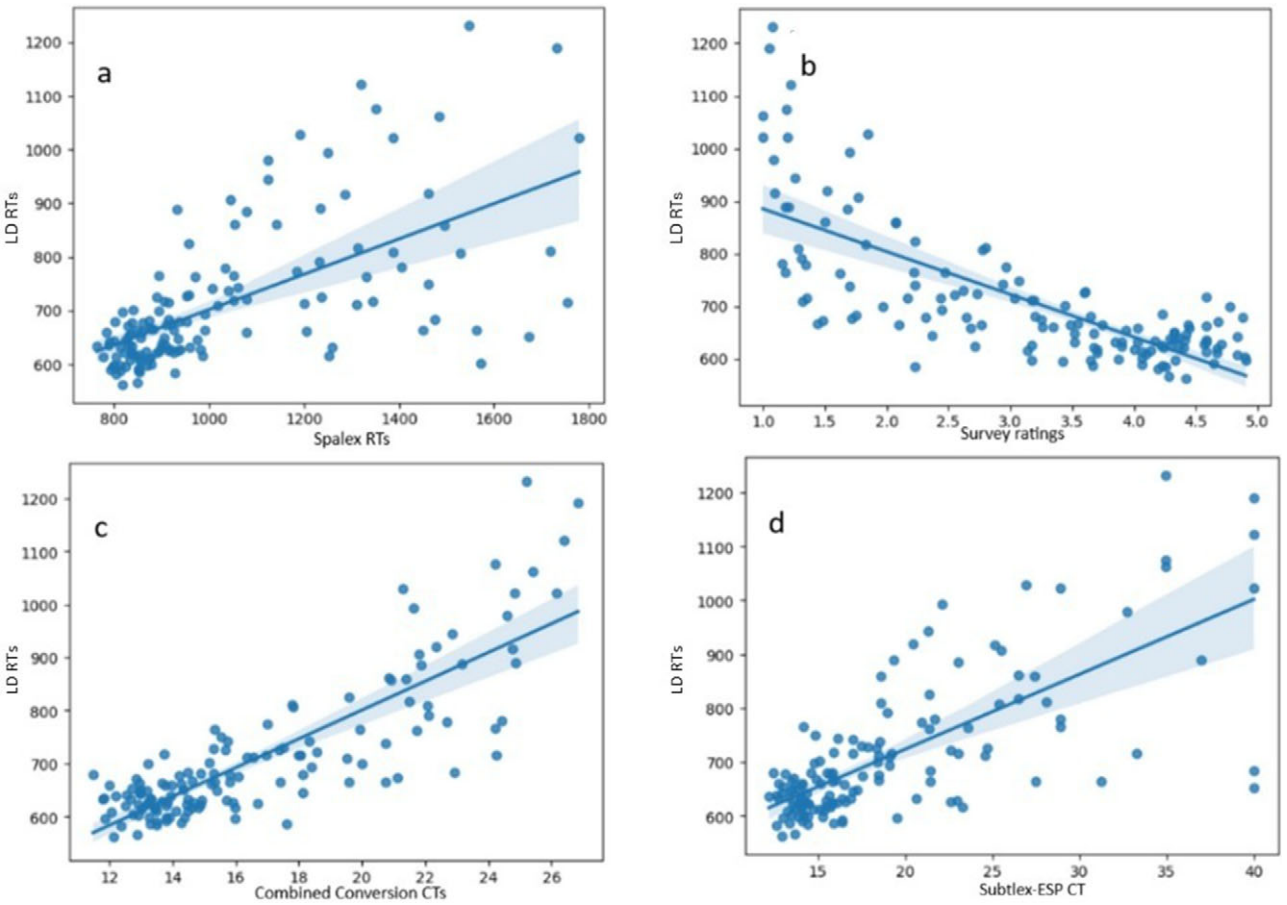


Figure 5. Scatterplots of lexical decision reaction times (LD RTs) versus (a) SPALEX RTs (r = .62), (b) survey word ratings (r = −.75), (c) cycle times from combined converted frequencies (r = .84), and (d) SUBTLEX-ESP frequencies (r = .73). All significances at p < .0001.

Table 4. Summary statistics for all word categories in Experiment 3 (lexical decision)

Category	Number of words	Mean percentage correct (SD)	Mean RT in ms (SD)	Mean rating (SD)
General	74	0.99 (0.02)	651 (58)	3.80 (0.77)
Spain	25	0.79 (0.26)	837 (169)	1.55 (0.47)
Chile	18	0.97 (0.04)	688 (70)	3.59 (0.94)
Latin America	19	0.87 (0.14)	816 (160)	1.99 (0.93)

Note: Standard deviations (SDs) are given in parentheses.

5. Experiment 4: Simulation study

The frequency counts in databases like SUBTLEX-ESP and the subjective ratings collected in Experiments 1 and 2 reflect the frequency of word use in a larger population and a specific group of language users, respectively. A reaction time, such as that collected in Experiment 3, can be seen as the frequency-sensitive duration of the psychological processes of word retrieval and identification (Thul et al., 2024). These processes are themselves characterized and implemented by the computational Multilink+ model. Going beyond a verbal account, its computations not only allow a qualitative but also a quantitative test of the assumptions expressed in the underlying psycholinguistic theory. This theory defines various lexical representations and processes in multilinguals, how processing is adapted to the task at hand, and what temporal patterns will result for different conditions. The cycle times that Multilink+ produces for different types of words reflect the constraints proposed by the underlying theory. The extent to which the complex interwoven series of constraints assumed in the model is similar to those in human word recognition is reflected in the correlations between model cycle times and empirical RTs. Correlations will be high when the model follows correctly and precisely, step-by-step, the human multistage word recognition process.

However, every model, including Multilink+, by necessity represents a simplified version of the subdomain it simulates. Non-implemented processes will lead to differences between simulation and empirical results. In human lexical decision, reaction times are influenced not only by lexical and sublexical processing but also by a range of additional factors such as attentional fluctuations (e.g., due to fatigue), individual response criteria, and meta-judgments. As a consequence, perfect correlations between simulation and empirical results cannot be reached. In light of this observation, the high correlations that Multilink+ generally produces (see Dijkstra et al., 2019, 2023) are impressive. They indicate that Multilink+ does an excellent job in capturing key aspects of human word recognition. As such, it is a useful tool to test predictions about how different variants of Spanish are processed. In the present paper, the model tests the hypothesis that various types of Spanish words are used more frequently or less frequently depending on item type and

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

the variant of Spanish that is considered, with RT-patterns for item types reflecting such variant-dependent processing.

To test our hypotheses about how Spanish word recognition would be sensitive to dialectal differences, we conducted simulation Experiment 4. Various simulations were performed, involving the use of lexicon files with SPALEX frequencies, SUBTLEX-ESP frequencies, and word frequencies derived from word ratings. To assess whether word ratings outperform existing database frequencies, we compared Multilink+'s cycle times in all cases to the reaction times, expecting the rating-based frequencies to significantly outperform both SUBTLEX-ESP and SPALEX.

As specified below, the word ratings from the extended survey were converted into frequencies to simulate and predict RT patterns across different item conditions in line with the research questions from Experiment 1. The conversion is necessary, because the Multilink+ model, just like the corpus data, represents (log) frequency of word use in terms of occurrence counts. However, ratings reflect participants' perceptions of frequency in a different, nonlinear fashion that is also sensitive to the rating scale used. As a reference basis for the conversion of the subjective ratings collected in the surveys, SUBTLEX-ESP frequencies were used, because they were proven to be more appropriate for Chilean Spanish than SPALEX frequencies.

This analysis was done across all cleaned RT data and within specific item categories. We anticipated that the converted frequencies will be the best predictor overall, especially for words from the categories Chile and Latin America, where differences in reaction times between Chileans and other Latin Americans are likely to be most evident.

5.1. Method

In order to apply subjective word ratings in Multilink+, they must first be converted into word frequencies to match database scales. Three methods to do so were tested.

(1) SD Conversion: Ratings were converted based on their distance from their mean in SD, mapped onto the corresponding SD range of SUBTLEX-ESP frequencies. Simply said, this involved mapping the rating distribution on the frequency distribution by turning the mean rating into a mean frequency, and adapting the SD of ratings to the SD of frequencies. However, this linear mapping approach limited the distribution of the frequencies based on ratings (0.91–258 vs. the usual 0.01–500+) and led to unwanted results for specific words. For instance, *argolla* had a below-average rating of 2.6, but after conversion received a disproportionately high frequency (8.70 vs. 0.22 in SUBTLEX-ESP).

(2) Factor Method: Ratings were converted by calculating their factor relative to the mean and applying it to the SUBTLEX-ESP log frequency. For example, *billete* has a mean rating of 4.25 and a SUBTLEX-ESP log frequency of 1.028. With the rating mean at 3.07, the rating of 4.25 is 1.384 times the mean. This factor was multiplied by the log frequency of the word, resulting in a log frequency of 1.423. This nonlinear mapping approach better aligned conversions with original word frequencies but restricted expressiveness: The value multiplied with log frequency could never be higher than 5 (max rating)/3.07 (average rating) = 1.629 times the original rating.

(3) Combined Conversion Method: Frequencies from SD Conversion and Factor Method were averaged, balancing distribution and expressiveness. The stimulus list with newly generated frequencies was added to the existing lexicon file and fed into Multilink+, which generated a cycle time (CT) for each word. Any CT value of 0, indicating that Multilink+ did not recognize the word within the limit of 40 CTs, was replaced with a value of 40. This adjustment is reasonable as such words are likely to have higher RTs, corresponding to higher CTs. Finally, analogous simulations with the same words were performed using their SPALEX and SUBTLEX-ESP frequencies.

5.2. Results

The three resulting lists of cycle times were correlated with the corresponding RTs from the Chilean lexical decision study (Experiment 3). Separate correlations per category were calculated, with Table 5 as a result.

5.3. Discussion

In Experiment 4, we used Multilink+ to test predictions for item reaction times based on subjective ratings. We hypothesized that using subjective ratings would outperform frequency measures from SUBTLEX-ESP and SPALEX, in correlations between cycle times (CTs) and RTs, both overall and within specific item categories.

Our results were in line with these hypotheses. Consistent with prior findings, SUBTLEX-ESP outperformed SPALEX as a predictor of RTs for Chilean participants, with correlations of .73 versus .52. When subjective ratings were integrated into Multilink+, the model achieved an even higher correlation of .84, considerably surpassing SUBTLEX-ESP. This pattern confirms that subjective familiarity ratings provide a more dialect-sensitive estimate of lexical accessibility than frequency norms derived primarily from Peninsular Spanish sources (Brysbaert & Cortese, 2011; Gimenes & New, 2016; Kuperman & Van Dyke, 2013).

As emphasized in Experiment 2, Spanish lexical variation does not consist of discrete dialectal blocks but of overlapping continua shaped by historical, sociolinguistic, and media-driven processes. Consequently, some words classified as "Chile," "Spain," or "Latin America" may also be known or occasionally used outside their dominant region. This overlap does not undermine the categorization logic, which is based on relative dominance rather than lexical exclusivity. Instead, such cross-regional overlap introduces natural variability that would attenuate, rather than inflate, observed differences. The strong category effects and model–human correlations observed here therefore likely represent conservative estimates of true regional lexical divergence.

We further examined how different frequency measures predicted RTs of Chilean speakers and Multilink+ CTs across item categories. CTs using SPALEX consistently showed weak correlations with RTs, reinforcing that SPALEX is not well suited to predicting Chilean lexical access. This pattern was expected, given that SPALEX primarily reflects Peninsular Spanish input and therefore lacks sensitivity to Latin American lexical distributions (Cuetos et al., 2011; Meneses et al., 2018)

Table 5. Correlations between lexical decision reaction times (RTs) and cycle times (CTs) based on different word frequency sources

Word category	CT SPALEX	CT SUBTLEX-ESP	Combined conversion
General	0.29	0.67	0.69
Spain	0.56	0.80	0.77
Chile	0.24	0.19	0.83
Latin America	0.39	0.56	0.75

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

Considering CTs based on SUBTLEX-ESP and converted subjective frequencies, both performed better than SPALEX. However, for item categories General and Spain, the differences between the two predictors were minimal, whereas Chile and Latin America showed significant discrepancies. This likely reflects the fact that SUBTLEX-ESP aggregates lexical input from multiple Latin American regions, which makes it a reasonably good estimator for words common across dialects, but insufficiently precise for dialect-specific lexical patterns characteristic of Chilean Spanish (Eddington, 2021; Meneses et al., 2018).

This distinction became evident when analyzing category-specific correlations. Words common across all Spanish dialects (General) and words from Spain yielded similar correlations between SUBTLEX-ESP and subjective ratings (.67 vs. .69 and .80 vs. .77, respectively). However, significant differences arose between words from Chile and from other countries in Latin America. The correlation for Chilean words was significantly higher when using subjective ratings (.83 vs. .19), while Latin American words not used in Chile correlated better with SUBTLEX-ESP than subjective ratings (0.75 vs. 0.56). These differences were statistically significant at $p < .01$. This divergence demonstrates that the predictive accuracy of frequency measures depends heavily on whether their underlying corpora capture the relevant dialectal distribution for the target population.

Overall, simulated word frequencies based on subjective ratings captured human lexical retrieval dynamics, as expressed by Multilink+, more effectively than existing databases. These findings underscore the limitations of using a single, pan-dialectal frequency measure to represent a language spoken across diverse regions. Unlike SPALEX and SUBTLEX-ESP, subjective ratings provide a more accurate representation of lexical access in dialect-specific speakers frequencies (Brysbaert & Cortese, 2011; Gimenes & New, 2016; Kuperman & Van Dyke, 2013), reinforcing the importance of regionalized linguistic resources for psycholinguistic research (González & Barrios, 2023). Future studies should investigate how additional linguistic variables, such as phonological neighborhood density, semantic richness, contextual usage patterns, and continuous measures of dialectal lexical probability, can be integrated into computational models to further refine predictions of region-specific lexical access.

6. General discussion

This study examined the variation among Spanish dialects and evaluated the efficacy of subjective word ratings as an alternative to conventional word frequency databases. Across four experiments, our findings underscore that Spanish is far from uniform; instead, dialectal variation exerts systematic and measurable effects on lexical processing and computational modeling.

In Experiment 1, we documented substantial vocabulary differences among Spanish dialects, with interdialect correlations as low as .13. These differences were strongly mediated by geographic location. While some dialects showed a closer alignment with SPALEX frequency data, others were better predicted by SUBTLEX-ESP. This dissociation likely reflects differences in corpus composition and participant demographics, highlighting that frequency norms inherit the regional biases of their underlying data sources (Aguasvivas et al., 2018; Cuetos et al., 2011).

Experiment 2 expanded our focus on Chile by augmenting the stimulus list and categorizing words into four groups: General, Spain, Chile, and Latin America. The new ratings exhibited an exceptionally high correlation with the Chilean ratings from Experiment 1 ($r = .95$), underscoring their reliability. Moreover, the Chilean ratings corresponded more closely with SUBTLEX-ESP than with SPALEX, affirming SUBTLEX-ESP's greater relevance for Latin American dialects (Eddington, 2021). The highest familiarity ratings for General and Chilean words further indicate that lexical accessibility is shaped by both pan-dialectal familiarity and local usage, supporting the need for dialect-sensitive frequency norms (Brysbaert & Cortese, 2011; Gimenes & New, 2016; Kuperman & Van Dyke, 2013).

Experiment 3 provided further insights by demonstrating that subjective ratings were superior predictors of Chilean lexical decision times compared with both SPALEX and SUBTLEX-ESP. Words in the General and Chile categories were recognized more quickly and accurately, highlighting the pivotal role of dialect familiarity and contextualized lexical experience in word recognition (Balota et al., 2001).

Building on these findings, Experiment 4 employed the Multilink+ model to simulate lexical retrieval using subjective ratings. The model achieved a strong correlation with observed RTs ($r = .84$), outperforming simulations based on SUBTLEX-ESP frequencies ($r = .73$). This convergence between behavioral and computational evidence demonstrates that subjective ratings capture dialectal nuances that objective corpora often underrepresent, strengthening the case for integrating dialect-specific information into computational models of word recognition (Gimenes & New, 2016).

6.1. Implications for linguistic research

Our findings challenge the prevailing assumption that Spanish can be treated as a monolithic language in linguistic research. Even modest lexical differences between dialects produce measurable shifts in familiarity ratings, reaction times, and computational cycle times, revealing a level of variation with major implications for psycholinguistic modeling. For instance, although Argentina and Uruguay exhibit high lexical similarity, considerable variation is evident when considering the broader Spanish-speaking world (Lipski, 2017). Given the magnitude of Spanish's global population, treating it as a uniform linguistic system risks masking genuine cross-regional differences that are relevant to both theory and practice.

The superior performance of Chilean-tuned frequencies in our simulations highlights the importance of developing region-specific linguistic resources. Future studies should replicate Experiments 3 and 4 in other Latin American regions to test whether subjective ratings outperform corpus-based frequencies across dialects and whether dialect regions differ in the degree to which corpora represent their lexical environments (Orozco, 2018). Similarly, applying this approach to other languages with significant dialectal variation, such as French, could yield valuable insights.

6.2. Practical applications

Tailored word frequency databases that reflect specific dialects and regions have wide-ranging practical applications. They can enhance the validity of psycholinguistic stimulus selection, improve natural language processing tools such as speech recognition, predictive text, and machine translation, and support lexicographic and pedagogical efforts. Moreover, subjective ratings offer a viable method for capturing the nuances of real-world language use, especially in scenarios where traditional frequency databases fall short in accounting for regional variation (Thul et al., 2024). This is particularly relevant in languages where corpora are composed

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

disproportionately of media or formal registers that do not reflect everyday lexical experience.

From an educational standpoint, acknowledging and incorporating dialectal diversity has profound implications. Language instruction tailored to local dialects may improve comprehension, reduce cognitive load, and support more realistic expectations of learner vocabulary. This dialect-sensitive approach might also facilitate the teaching of foreign languages, such as English, by drawing informative parallels between dialect-specific structures and second-language grammar (cf. Wheeler & Swords, 2006).

Furthermore, developing dedicated dialect databases can illuminate cultural distinctions within Latin America. Regional terms – for example, *crispetas* in Colombia and *pochoclo* in Argentina for "popcorn" – reflect unique cultural identities. Similarly, the tendency among Chileans to refer to their language as *Chileno* rather than *Español* underscores the deep connection between dialect and national identity (Sadowsky & Martínez-Gamboa, 2012).

7. Conclusion

In conclusion, our study demonstrates that treating Spanish as a homogeneous entity obscures significant dialectal differences with both theoretical and practical ramifications. Subjective word ratings provide a powerful, cost-effective tool for developing dialect-specific frequency norms and yield more accurate predictions of lexical processing than existing databases for millions of speakers. For over 500 million native Spanish speakers, and millions of language learners, embracing dialectal diversity is not merely descriptive but also essential. Integrating regional variation into linguistic research, computational models, and educational practice enriches scientific understanding and reflects the lived linguistic reality of Spanish-speaking communities worldwide.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/S1366728925100941.

Data availability statement. Data and materials for the project are openly available at http://doi.org/10.1017/S1366728925100941.

Funding statement. This work was supported by Agencia Nacional de Investigación y Desarrollo (ANID), Chile (Millennium Science Initiative Program – NCS2022_026). It did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Open access funding provided by Radboud University Nijmegen.

Ethical standard. This study received ethical approval from the Ethics Committee at Universidad Católica del Maule, Chile. All procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Competing interests. The authors declare none.

References

Aguasvivas, J. A., **Carreiras, M.**, **Brysbaert, M.**, **Mandera, P.**, **Keuleers, E.**, & **Duñabeitia, J. A.** (2018). SPALEX: A Spanish lexical decision database from a massive online data collection. *Frontiers in Psychology*, **9**, Article 2156. https://doi.org/10.3389/fpsyg.2018.02156

Alfaraz, G. G. (2022). Dialectología perceptiva del español. In F. Moreno-Fernández & R. Caravedo, *Dialectología hispánica / The Routledge Handbook of Spanish Dialectology* (1st Edition, pp. 105–116). Routledge.

Balota, D. A., **Pilotti, M.**, & **Cortese, M. J.** (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, **29**(4), 639–647. https://doi.org/10.3758/BF03200465

Biber, D. (1987). A textual comparison of British and American writing. *American Speech*, **62**(2), 99–119. https://doi.org/10.2307/455273

Boyd-Bowman, P. (1956). The regional origins of the earliest Spanish colonists of America. *PMLA/Publications of the Modern Language Association of America*, **71**(5), 1152–1172. https://doi.org/10.2307/460533

Brysbaert, M., & **Cortese, M. J.** (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *Quarterly Journal of Experimental Psychology*, **64**(3), 545–559. https://doi.org/10.1080/17470218.2010.503374

Cuetos, F., **González-Nosti, M**, **Barbón A.**, & **Brysbaert, M.** (2011). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicológica Journal* **32** (2), 133–143.

Delgado-Díaz, G. (2018). Dialectal variation of the preterit and imperfect: A closer look to Puerto Rican and Buenos Aires Spanish. *Revista Española de Lingüística Aplicada/Spanish Journal of Applied Linguistics*, **31**(1), 64–93. https://doi.org/10.1075/resla.15048.del

Dijkstra, T., & **Peeters, D.** (2023). *The new psychology of language: From body to mental model and back*. Routledge/Taylor & Francis. https://doi.org/10.4324/9781003326274

Dijkstra, T., **Peeters, D.**, **Hieselaar, W.**, & **van Geffen, A.** (2023). Orthographic and semantic priming effects in neighbour cognates: Experiments and simulations. *Bilingualism: Language and Cognition*, **26**(2), 371–383. https://doi.org/10.1017/S1366728922000591

Dijkstra, T., **Wahl, A.**, **Buytenhuijs, F.**, **Halem, N.**, **Al-Jibouri, Z.**, **de Korte, M.**, & **Rekké, S.** (2019). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition* **22**, 4, 657–679. https://doi:10.1017/S1366728918000287

Duchon, A., **Perea, M.**, **Sebastián-Gallés, N.**, **Martí, A.**, & **Carreiras, M.** (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, **45**(4), 1246–1258. https://doi.org/10.3758/s13428-013-0326-1

Eddington, D. E. (2021). A corpus analysis of some usage differences among Spanish-speaking countries. *Dialectologia*, **27**, 71–95. https://doi.org/10.1344/Dialectologia2021.27.4

Eddington, D. E. (2022). Grouping Spanish-speaking countries by dialect: A corpus dialectometric approach. *Isogloss. Open Journal of Romance Linguistics*, **8**(1), 1–30. https://doi.org/10.5565/rev/isogloss.207

Gallego, Á. J. (Ed.). (2019). *The syntactic variation of Spanish dialects*. Oxford University Press.

Gimenes, M., & **New, B.** (2016). Worldlex: Twitter and blog word frequencies for 66 languages. *Behavior Research Methods*, **48**(3), 963–972. https://doi.org/10.3758/s13428-015-0621-0

González, R. G., & **Barrios, J. F. C.** (2023). Object person marking in two under-represented Spanish dialects of Mexico. *Isogloss. Open Journal of Romance Linguistics*, **9**(2), Article 2. https://doi.org/10.5565/rev/isogloss.203

Goodrich, J. M., & **Lonigan, C. J.** (2018). Development of first- and second-language vocabulary knowledge among language-minority children: Evidence from single language and conceptual scores. *Journal of Child Language*, **45**(4), 1006–1017. https://doi.org/10.1017/S0305000917000538

Hansen, J. H. L., & **Liu, G.** (2016). Unsupervised accent classification for deep data fusion of accent and language information. *Speech Communication*, **78**, 19–33. https://doi.org/10.1016/j.specom.2015.12.004

Jucks, R., & **Paus, E.** (2013). Different words for the same concept: Learning collaboratively from multiple documents. *Cognition and Instruction*, **31**(2), 227–254. https://doi.org/10.1080/07370008.2013.769993

Keuleers, E., & **Brysbaert, M.** (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods* **42**, 627–633. https://doi.org/10.3758/BRM.42.3.627

Kuperman, V., & **Van Dyke, J. A.** (2013). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology: Human Perception and Performance*, **39**(3), 802–823. https://doi.org/10.1037/a0030859

Lipski, J. M. (2008). *Varieties of Spanish in the United States*. Georgetown University Press.

Lipski, J. M. (2012). Geographical and Social Varieties of Spanish: An Overview. In J. I. Hualde, A. Olarrea, & E. O'Rourke, *The Handbook of Hispanic Linguistics* (pp. 1–26). John Wiley & Sons. https://doi.org/10.1002/9781118228098.ch1

Lipski, J. M. (2017). Dialects of Spanish and Portuguese. In C. Boberg, J. Nerbonne, & D. Watt, *The Handbook of Dialectology* (pp. 498–509). John Wiley & Sons. https://doi.org/10.1002/9781118827628.ch30

Márquez Reiter, R. (2022). A contrastive study of conventional indirectness in Spanish: Evidence from Peninsular and Uruguayan Spanish. *Pragmatics. Quarterly Publication of the International Pragmatics Association (IPrA)*, **12**(2), 135–151. https://doi.org/10.1075/prag.12.2.02mar

Martin, C. D., Garcia, X., Potter, D., Melinger, A., & Costa, A. (2016). Holiday or vacation? The processing of variation in vocabulary across dialects. *Language, Cognition and Neuroscience*, **31**(3), 375–390. https://doi.org/10.1080/23273798.2015.1100750

Meneses, A., Uccelli, P., Santelices, M. V., Ruiz, M., Acevedo, D., & Figueroa, J. (2018). Academic language as a predictor of reading comprehension in monolingual Spanish-speaking readers: Evidence from Chilean early adolescents. *Reading Research Quarterly*, **53**(2), 223–247. https://doi.org/10.1002/rrq.192

Moreno-Fernandez, F. (2015). La percepción global de la similitud entre variedades de la lengua española. In K. Jeppesen & J. Lindschouw (Eds.), *Les variations diasystématiques et leurs interdépendances dans les langues romanes*. (pp. 217–238). Travaux de Linguistique Romane. Editions de Linguistique et Philologie. https://doi.org/10.13140/RG.2.1.2362.9528

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, **28**(4), 661–677. https://doi.org/10.1017/S014271640707035X

Orozco, R. (2018). *Spanish in Colombia and New York City: Language contact meets dialectal convergence*. John Benjamins. https://doi.org/10.1075/impact.46

Peirce, J. W. (2007). PsychoPy – Psychophysics software in Python. *Journal of Neuroscience Methods*, **162** (1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Penny, R. (2000). *Variation and change in Spanish*. Cambridge University Press. https://doi.org/10.1017/CBO9781139164566

Pullum, G. K. (2017). Theory, data, and the epistemology of syntax. In M. Konopka & A. Wöllstein (Eds.), *Grammatische Variation: Empirische Zugänge und theoretische Modellierung* (pp. 283–298). De Gruyter. https://doi.org/10.1515/9783110518214-016

Quesada Pacheco, M. Á. (2014). División dialectal del español de América según sus hablantes Análisis dialectológico perceptua. *Boletín de Filología*, **49** (2), 257–309.

Rogers, B. M. A., & Face, T. L. (2024). Chilean Spanish intonational plateaus and their pragmatic functions: A case of contact with Mapudungun. *Languages*, **9**(7), Article 7. https://doi.org/10.3390/languages9070246

Rost Bagudanch, A. (2017). Variation and phonological change: The case of yeísmo in Spanish. *Folia Linguistica*, **51**(1), 169–206. https://doi.org/10.1515/flin-2017-0005

Sadowsky, S. & Martínez-Gamboa, R. (2012). *LIFCACH 2.0: Word frequency list of Chilean Spanish (Lista de Frecuencias de Palabras del Castellano de Chile), version 2.0*. Zenodo. http://doi.org/10.5281/zenodo.268043

Scott, J. C. (1993). Preparing business correspondence the British Way. *The Bulletin of the Association for Business Communication*, **56**(2), 10–17. https://doi.org/10.1177/108056999305600203

Shade, B., & Altmann, E. G. (2023). Quantifying the dissimilarity of texts. *Information*, **14**(5), Article 5. https://doi.org/10.3390/info14050271

Shea, C. (2022). Mapping different L1 dialects to L2 words. *Linguistic Approaches to Bilingualism*, **12**(5), 598–627. https://doi.org/10.1075/lab.18020.she

Sorenson, T. D. (2021). *The dialects of Spanish: A lexical introduction*. Cambridge University Press. https://doi.org/10.1017/9781108924580

Thul, R., March, J., Dijkstra, T., & Conklin, K. (2024). Stratified distributional analysis: A novel perspective on RT distributions. *Quarterly Journal of Experimental Psychology*, **78**(8), 1740–1756. https://doi.org/10.1177/17470218241288516

Toribio, A. J. (2000). Setting parametric limits on dialectal variation in Spanish. *Lingua*, **110**(5), 315–341. https://doi.org/10.1016/S0024-3841(99)00044-3

Vivanco Rojas, R. A. (2018). La apasionante forma de hablar de los chilenos. El español de Chile: generalidades. Contextos (UMCE).

Weinberg, F. D., & Beatriz, M. (1977). *La constitución del paradigma pronominal de voseo.*

Wheeler, R. S., & Swords, R. (2006). *Code-switching: Teaching standard English in urban classrooms*. National Council of Teachers of English (NCTE).

Appendix: Stimulus materials for Experiment 3 (lexical decision)

Words

Word	Region	Frequency (SPALEX OPM)
accidente	General	48.81527
agua	General	277.367169
alarma	General	12.428009
alcancía	General	0.471127
anillo	General	23.692821
arte	General	163.874265
avión	General	48.100456
bañera	General	1.806529
barba	General	13.942114
bigote	General	3.924976
billetera	General	0.42239
bohemia	General	6.436571
bolsa	General	54.780714
bote	General	10.264073
botella	General	15.618677
botón	General	7.827209
bóveda	General	10.426531

(*Continued*)

(*Continued*)

Word	Region	Frequency (SPALEX OPM)
brazo	General	48.779529
cama	General	45.514131
cámara	General	134.498676
camisa	General	13.724421
carcasa	General	1.358146
casco	General	22.786308
cazador	General	10.27707
cepillo	General	1.65057
cereza	General	1.410132
cielo	General	103.582988
cigarro	General	3.249153
ciudad	General	657.553775
culpa	General	42.365702
distancia	General	82.716929
escultura	General	17.668892
estatua	General	17.353725
grasa	General	8.944917
horario	General	14.513965
horno	General	9.367307
hueso	General	13.604202
humo	General	21.850552
impresora	General	2.007976
jabón	General	3.931475
madrugada	General	26.386369
mañana	General	247.169544
manta	General	5.786741
manzana	General	8.126131
marinero	General	5.841977
melodía	General	8.873436
mentira	General	16.018323
mesa	General	100.246108
metal	General	47.912006
micrófono	General	4.25639
miedo	General	89.159999
miseria	General	21.81806
niño	General	100.070654
nube	General	15.043577
olor	General	24.219184
patrimonio	General	43.785582
pierna	General	15.72265
piña	General	2.488851
playa	General	39.084058
poeta	General	71.647066
precio	General	96.350374

(*Continued*)

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

(*Continued*)

Word	Region	Frequency (SPALEX OPM)
prisa	General	22.331426
problema	General	220.286054
puerta	General	158.977792
pureza	General	10.767692
reloj	General	26.145932
rizo	General	0.922759
salida	General	80.111109
silla	General	24.804032
sonido	General	64.443695
velocidad	General	76.033422
ventana	General	41.56641
vergüenza	General	23.380903
viaje	General	105.165325
viático	General	0.256683
voz	General	205.573891
zapato	General	3.418109
bayeta	Spain	0.445134
biberón	Spain	0.477625
bocadillo	Spain	1.020234
bolígrafo	Spain	0.477625
bombilla	Spain	1.289914
braga	Spain	2.518093
camarero	Spain	4.275885
chaval	Spain	2.521342
ciervo	Spain	3.726778
coche	Spain	55.531269
cotilla	Spain	0.162458
empollón	Spain	0.090976
escayola	Spain	1.988481
felpudo	Spain	0.224192
follón	Spain	0.295673
fontanero	Spain	0.610841
fregona	Spain	0.526363
frigorífico	Spain	1.218432
grifo	Spain	2.293902
hucha	Spain	0.428888
móvil	Spain	24.261423
ordenador	Spain	19.543653
palangana	Spain	0.675824
pincho	Spain	0.588097
zumo	Spain	2.83651
afiche	Chile	0.194949
ampolleta	Chile	0.090976

(*Continued*)

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

(*Continued*)

Word	Region	Frequency (SPALEX OPM)
audífono	Chile	0.064983
auto	Chile	26.314888
boleto	Chile	2.001478
calzón	Chile	0.773298
celular	Chile	13.051846
choclo	Chile	0.571851
durazno	Chile	1.387388
florería	Chile	0.064983
frutilla	Chile	0.34441
galón	Chile	0.565353
lavatorio	Chile	0.302171
mamadera	Chile	0.100724
pololo	Chile	0.048737
purpurina	Chile	0.068232
refrigerador	Chile	0.698568
sostén	Chile	2.713042
tina	Chile	3.684539
auricular	Latin America	1.289914
bochinche	Latin America	0.217693
bombillo	Latin America	0.087727
bombona	Latin America	0.506868
canilla	Latin America	0.334663
chincheta	Latin America	0.064983
computadora	Latin America	11.251816
cremallera	Latin America	0.740807
floristería	Latin America	0.172205
fresa	Latin America	1.99498
jugo	Latin America	5.5593
labial	Latin America	0.425639
lavabo	Latin America	1.315907
melocotón	Latin America	0.961749
mesero	Latin America	0.318417
patata	Latin America	3.457098
pintalabios	Latin America	0.097475
plomero	Latin America	0.15271
viaducto	Latin America	1.624576
argolla	Meaning	3.154927
baúl	Meaning	2.131444
billete	Meaning	8.707729
buzo	Meaning	0.630336
cacho	Meaning	2.144441
canguro	Meaning	0.981244
carátula	Meaning	1.166446

(*Continued*)

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

(*Continued*)

Word	Region	Frequency (SPALEX OPM)
carro	Meaning	16.75588
cartel	Meaning	12.80816
churro	Meaning	0.172205
piso	Meaning	35.237061
saco	Meaning	10.93015
torta	Meaning	2.505097
Nonwords		
absarino		
agués		
ahuima		
allojar		
alómila		
amiseita		
anedior		
arnaca		
arza		
aunáfero		
auváfono		
avar		
avenjo		
bayira		
beche		
becinura		
becio		
bimita		
bimpa		
bimulo		
binucato		
boledora		
bolontrón		
boneador		
borle		
botacidad		
botrio		
brauga		
brima		
briño		
brogón		
buca		
budán		
bulgue		
burol		
busión		

(*Continued*)

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

(*Continued*)

Word	Region	Frequency (SPALEX OPM)
cadel		
cafar		
cafo		
callaja		
callunta		
camovo		
canvina		
caper		
capuján		
carlante		
casficero		
castulleo		
ceceña		
celinte		
cergo		
cernén		
cesca		
chapán		
chespura		
chirno		
chueno		
cifapa		
colilós		
comija		
cosedio		
crevamiento		
crisa		
cíctel		
cífrea		
cínara		
desceño		
duseta		
empuejo		
enségeno		
escruedo		
eserpe		
flauda		
focatega		
fojiela		
follezo		
fopustero		
frestón		
frono		

(*Continued*)

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

(*Continued*)

Word	Region	Frequency (SPALEX OPM)
frullo		
fucedal		
gepio		
griba		
gruda		
horsidor		
huda		
huemo		
huror		
huste		
incanio		
juemo		
jueto		
lapera		
lapitina		
lasadio		
maidora		
maigonal		
malamio		
malmiz		
mancotre		
miserpa		
miula		
mivero		
moce		
mocémica		
modriz		
momén		
mándola		
mírtica		
mísame		
móleo		
naral		
nedo		
nupe		
opeza		
opiantador		
palaneo		
palgo		
palquenada		
pamba		
pazel		
pecheza		

(*Continued*)

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press

(*Continued*)

Word	Region	Frequency (SPALEX OPM)
peicha		
pemar		
pemina		
pentol		
peplijo		
perceo		
perda		
pernón		
pervegía		
pálsar		
pérgilo		
péspano		
rechensa		
recho		
relgo		
sarlo		
sata		
sochero		
socote		
sáscope		
taveo		
ténel		
veligre		
veluria		
vempazo		
vencora		
veocia		
vercola		
verol		
zapla		
zuedo		

https://doi.org/10.1017/S1366728925100941 Published online by Cambridge University Press