Short Title: SPOKEN VERB PROCESSING IN SPANISH

Spoken Verb Processing in Spanish: An Analysis Using a New Online Resource

Semilla M. Rivera[1, 2]

Elizabeth A. Bates[3, 4, *]

Araceli Orozco-Figueroa[3]

Nicole Y. Y. Wicha[1, 2]

[1] University of Texas at San Antonio, Department of Biology

[2] University of Texas Health Science Center at San Antonio, Research Imaging Center

[3] University of California at San Diego, Center for Research in Language

[4] University of California at San Diego, Department of Cognitive Science

[*]Deceased

Address correspondence to:

Nicole Y. Y. Wicha

Department of Biology

University of Texas at San Antonio

One UTSA Circle

San Antonio, Texas 78249

Phone: (210) 458-7013 Fax: (210) 458-5658

Email: Nicole.Wicha@UTSA.edu

Short Title: SPOKEN VERB PROCESSING IN SPANISH

Spoken Verb Processing in Spanish: An Analysis Using a New Online Resource

Abstract

Verbs are one of the basic building blocks of grammar, yet few studies have examined the grammatical, morphological, and phonological factors contributing to lexical access and production of Spanish verb inflection. This report describes an online dataset that incorporates psycholinguistic dimensions for 50 of the most common early-acquired Spanish verbs (accessible at http://crl.ucsd.edu/experiments/svi/). Using this dataset, predictors of response time from stimulus onset and mean differences at offset are examined. Native Spanish speakers, randomly assigned to one of two tasks, listened to pre-recorded verbs and either repeated the verb (single word shadowing) or produced its corresponding pronoun. Factors such as stimulus duration, number of syllables, syllable stress position, and specific levels of initial phoneme facilitated both shadowing of a verb and production of its pronoun. Higher frequency verbs facilitated faster verb repetition while verbs with alternative pronouns increased response time to pronoun production. Mean differences at offset (stimulus duration is removed) indicated that listeners begin speaking earlier when the verb is longer and multisyllabic compared to shorter, monosyllabic words. These results highlight the association between psycholinguistic factors and response time measures of verb processing, in particular features unique to languages like Spanish, such as alternative pronoun and tense.

Introduction

Like nouns, verbs represent basic units of grammar and are essential components in language use. Generally, nouns are defined as the class of words referring to entities and verbs as the class referring to processes (Laudanna and Voghera, 2002). Much of what is known about word recognition and production is based on the study of nouns. But recently, research on verb recognition and production has burgeoned in studies of language development (Labelle et al., 2002; Negro et al., 2005), aging (Bird et al., 2001; Mackay et al., 2002; Morrison et al., 2003; Marini et al., 2005; Persson et al., 2004), aphasia (Barde et al., 2006; Berndt and Mitchum 1991; Caramazza and Hillis, 1991; Hillis et al., 2002; Plunkett and Bandelow, 2006), bilingualism (Dopke, 1998; Pillai et al., 2003), and speech errors (Arnaud, 1999; Poulisse, 1999). Still, less is known about the extent to which the grammatical, morphological, and phonological characteristics of verbs contribute to differences in word recognition and production.

Depending on the language, a verb may vary in form according to factors such as tense, gender, person, and number (singular or plural). Unlike English verbs, which undergo little inflection, many Romance languages such as Spanish and Italian have a rich inflectional morphology. Psycholinguistic studies show that adults and children who speak Italian rely heavily on morphological cues for the recognition and sentence interpretation of nouns as well as verbs (e.g., Bates et al., 1995; Bates et al., 1999; Devescovi et al., 1998; MacWhinney et al., 1984). The present report takes advantage of the inflectional morphology of Spanish verbs to determine which linguistic and psycholinguistic factors influence verb shadowing and pronoun production.

To assist in the study of Spanish inflection we implement a dataset that incorporates grammatical, morphological, and phonological dimensions of Spanish verbs, which we refer to as the Spanish Verb Inventory (SVI). Only a handful of online databases have been developed to obtain indices specific to Spanish words. Theses include LEXESP[1] (Sebastian-Galles et al., 2000), Corpus del Español (Davies, 2001), BuscaPalabras (Davis and Perea, 2005), EuroWordNet (Vossen et al., 1998), the International Picture naming project (Szekely et al., 2004), and C-ORAL-ROM (Crestie and Moneglia, 2000). There is currently, however, no online lexical dataset free of charge for computing linguistic/psycholinguistic dimensions specific to Spanish inflected verbs. Such a dataset would facilitate the study of lexical processing in a morphologically rich language. The first part of this report describes a lexical Excel spreadsheet designed specifically for calculating summary statistics across several psycholinguistic features characteristic of Spanish verbs[2]. Using these data, we then characterize the factors contributing to spoken word recognition and production across two levels of processing. These analyses shed light on psycholinguistic factors shared by other languages but also call attention to verb features unique to the Spanish language, such as verb tense and inflected verb forms that are compatible with multiple subject pronouns.

Part I: Spanish Verb Inventory

Spanish Language Databases. Currently, there are few databases that provide psycholinguistic measures of Spanish inflected verbs. Two excellent datasets, LEXESP (Sebastián et al., 2000) and BuscaPalabras (B-Pal; http://www.uv.es/mperea/) (Davis and Perea, 2005) include many psycholinguistic and linguistic variables. The LEXESP query system is an extensive lexical database for Spanish words. It contains a total of 5,020,930 word tokens (including some inflected verbs) and incorporates measures of word frequency (for 166,494 words), number of syllables, stress location, and word pronunciations. As noted by Davis and Perea (2005), many of these words are proper nouns, words containing non-alphabetic characters, pseudowords, or non-Spanish words. Norms for imageability, concreteness, and familiarity are also included for 6,500 of the most frequent words. Of these, however, imageability and concreteness values are available for only 21 (2.3%) of the 920 inflected verbs in the present study. In addition, LEXESP does not contain other variables such as age of acquisition, orthographic or phonological neighborhood measures, or syllable frequencies (Davies and Perea, 2005).

In 2005, Davis and Perea created BuscaPalabras (B-Pal; http://www.uv.es/mperea/), a lexical program based on 31,491 Spanish word types found in LEXESP. B-Pal includes several indices not found in LEXESP such as age of acquisition, syllable-based measures (such as token and type syllable frequency), orthographic neighborhood measures, phonological statistics such as the word's pronunciation, initial and number of phonemes, stress pattern, syllable count, and occurrences of homophones, valence, and arousal. The program also includes information about orthographic similarity such as transposed letter neighbors and embedded-word similarity, and enables researchers to integrate up to three user-defined indices not previously included in the original database. However, B-Pal is limited in that it includes few inflected verb forms. Although 49 of the 50 infinitive verb forms are included, B-Pal contains only 51 (5.5%) of the 920 inflected verb forms found in the present study. In addition, indices such as tense, class, stress type, regularity, and the object pronoun associated with each verb are not part of B-Pal. Furthermore, B-Pal does not include phonemic variables such as vocal location of articulation (e.g., lateral or voiced palatal), manner and place of articulation, root and stem spoken durations, and audio sound recordings for each verb. Note that although B-Pal and LEXESP include frequency measures, the most comprehensive and updated frequency database for written

Spanish words is the Davies corpus (2001), which is freely available online (http://www.corpusdelespanol.org/).

Spanish Verb Inventory. When it comes to the study of inflected verbs, the availability of online psycholinguistic databases is thus limited. To address this shortcoming, we present the Spanish Verb Inventory (SVI), a dataset constructed to include psycholinguistic measures particular to Spanish inflected verbs. The verb inventory consists of 50 of the earliest acquired common Spanish verbs (Table 1), conjugated across person, number, and four verb tenses, for a total of 920 unique inflected verb forms (see method section for details). For each word, queries can be made across grammar, morphology, and phonology (see Appendix for a complete list of variables). Like LEXESP and B-Pal, SVI includes measures of word length, word and syllable frequency, and subjective ratings of concreteness (infinitive verb forms only). Similar to B-Pal, SVI includes phonological indices such as initial phoneme, stress pattern, number of syllables, vowel-consonant structure of the lexeme (e.g., *hacemos* [we do] has a CVCVCVC structure), and phonemic variables, such as root and stem. However, we provide these variables for words that are not within the B-Pal database.

INSERT TABLE 1 ABOUT HERE

In presenting this dataset, we also acknowledge its shortcomings. Some factors such as phonological and orthographic neighborhood, which have proved important for lexical access, are not available for verbs and were not included herein. However, another feature of SVI is that it is amenable to the addition of indices as they become available in future studies.

In brief, SVI is the first dataset that provides indices that are relevant for psycholinguistic studies of inflected Spanish verbs that are unavailable from other databases. SVI is unique in that it is specific for the study of Spanish verbs and includes variables that have never been included in other lexical databases. These variables include the following: 1) alternative pronouns across verb forms, 2) the duration of each word and word part (stem, root, word ending) in milliseconds, 3) mean response times for two production tasks, 4) audio sounds files for the 920 Spanish inflected verbs. The dataset is presented as a Microsoft Excel spreadsheet, along with a codebook describing each of the variables. These tools can be accessed free of charge from the Center for Research in Language website at http://crl.ucsd.edu/experiments/svi/.

PART II: Factors Related to Spoken Word Recognition, Lexical Access, and Verb Production

In developing this online dataset, we also aimed to provide researchers interested in verb processing with analyses characterizing how different linguistic and psycholinguistic factors relate to word recognition and production. In particular, evidence suggests that factors such as word frequency, measures of word length (Cuetos et al., 1999; Prado and Ullman, under review; Spieler and Balota, 1997), phonetic characteristics (Treiman et al., 1995) and verb regularity/irregularity (Marslen-Wilson and Tyler, 1997; Sonnenstuhl et al., 1999; Stanners et al., 1979) affect lexical access. Moreover, prior studies suggest that performance can change when participants are asked to focus explicitly on grammatical dimensions (Bates et al., 1995; Bates et al., 1996). Verb tense has also been found to influence verb processing in inflected verb forms (Carreiras et al., 1997; Kostic and Havelka, 2002). Not only are these important factors for word recognition and production, these are factors that can potentially act as confounds in studies of language processing more generally.

In the present study, we tested the effect of many of these psycholinguistic dimensions on response time using two tasks: a word repetition task and a pronoun production task. In both tasks, native speakers of Spanish listened to spoken verbs. In the first task (called single word shadowing, see Bates & Liu, 1996), participants were asked to repeat each target word as quickly as possible without making a mistake. This task does not necessarily require specific attention to verb conjugation or conscious decision about morphological markers, though these factors have been shown to influence even simple word repetition (Bates et al., 1995). Such attention and reflection were more important in the second task, in which participants were asked to generate a subject pronoun that agreed with each verb. To perform the latter task, participants had to monitor the person and number inflection on the target verb and to make a deliberate decision about a suitable subject pronoun (e.g., given the second person singular imperfect form *corrías* [you were running], participants would have to generate the second person singular pronoun *tu)*. This is a metalinguistic task, and it could be argued that it is artificial in Spanish, for two reasons: (a) Spanish is a pro drop language in which subject pronouns are frequently omitted and (b) subject pronouns, when they do occur, are more likely in preverbal position; yet in this task, speakers generated the pronoun after the verb. On the other hand, subject pronouns are still a high frequency phenomena, and because overt subject pronouns tend to be used for emphasis, they are more common in post verbal position than many other subject types (e.g., the sentence *Corría yo*, literally “Was running I,” can be translated more fully as “I was the one who was running”). The main advantage of the pronoun generation task, as a complement to word repetition, is that it permits us to investigate the factors that influence performance when speakers are forced to attend to and reflect upon grammatical information. Although it is obvious that the latter task will be more taxing and thus lead to longer response times than the repetition task, we have made available the mean differences in RT between two tasks, as these are informative for knowing how long on average it takes listeners to process and produce each stimulus dependent on the task demands.

Method

Participants

Data were obtained from 60 native speakers of Spanish (39 females and 21 males), ages 17 through 25 years (M=20.9) who were students of the college of Humanities at the Universidad Autónoma de Baja California living in Tijuana, Mexico[3]. Prior to testing, each participant completed a language history questionnaire to assess biographical information regarding all contact with their native and any other languages and to identify them as native Spanish speakers (i.e., contact since birth and dominant language used at time of testing). Information concerning handedness and past auditory or linguistic disability was also collected. All participants were right-handed with no prior history of disabilities that could hinder their performance on the experimental tasks. Participants were compensated for their participation.

Materials

The verbs used in this dataset are 50 of the first 100 verbs acquired by Spanish speaking children, obtained from the Spanish version of the MacArthur-Bates Communicative Development Inventory (CDI) (Fenson and Dale, 1993) (Table 1). The same 50 verbs were used in a study of Italian inflected verbs (Devescovi et al., unpublished). These 50 verbs were chosen because they were among the first 100 verbs learned in both languages. Each of the 50 verbs (i.e., lexemes) appeared as isolated inflected verbs (i.e., morphosyntactic word form) in four indicative tenses (imperfect, preterite, future, and present), three persons (1st, 2nd and 3rd person) and number (singular and plural) combinations. Although four tenses, three persons, and two numbers should lead to 24 inflected morphosyntactic forms per verb, there are several repeated lexical forms across verb conjugations in Spanish, resulting in only 18 or 19 unique word forms depending on the verb (Table 2). First, the 2nd and 3rd person plural forms across all tenses are always identical in Spanish (e.g., *ellos/ellas van, ustedes van* – "they [masc/fem] go", "you [plural] go"). Second, the 1st and 3rd person singular forms of the imperfect tense for all verbs are identical (e.g., *yo abría, él/ella abría*– "I was opening, he/she was opening"). Finally, the 1st person plural forms in the present and preterite tenses for 30 of the verbs used are identical (e.g., *nosotros abrimos/abrimos* – "we open/opened", but not *nosotros vemos/vimos* – "we see/saw"). Each of the 920 inflected verbs were recorded by a female native speaker of Spanish in a sound-attenuating booth and converted from digital audiotape to individual digital sound files. The sound files were normalized and cleaned, with the blank space before word onset and after word offset removed. The average duration of the sound files was 624.8 ms with a range between 274.0 and 965.0 ms. These sound files are Windows PCM files and can be played online using any Windows media software at http://crl.ucsd.edu/experiments/svi/.

INSERT TABLE 2 ABOUT HERE

Procedure

Participants were randomly assigned to one of two tasks: verb repetition (single word shadowing) (N=30) or pronoun production (N=30). Note that our sample size is smaller than other larger norming studies but is still comparable to several other norming studies (for example, those listed at http://crl.ucsd.edu/~aszekely/ipnp/ under "Studies"). In both tasks participants listened to pre-recorded verbs in random order and were asked to either repeat the word as soon as they knew what it was (single word shadowing) or to produce the first pronoun that came to mind upon listening to the verb[4]. In each case, participants were instructed to respond as quickly as possible upon hearing the stimulus, without making errors. They were asked to avoid making any sounds that could interfere with the answer (given that the voice key recorded the first perceived sound as the onset of a response) and to avoid correcting themselves after a response was given.

A brief practice session consisting of 20 verbs that were not used in the experimental session was given to each individual prior to the actual session. Participants were asked to fixate their attention on a cross (+) in the middle of a blank computer screen and were informed that a series of inflected verbs were going to be presented through the headphones. In the pronoun production task, participants were not prompted on which pronoun to say. Thus, for some verbs, more than one correct pronoun could potentially be produced. For instance, the verb *corría* [was running] could elicit either of the following correct pronoun response: *él* [he], *ella* [she], *yo* [I], or *usted* [you formal]. The actual response was noted, with any of these responses considered correct.

Individuals were tested one at a time in a quiet cubicle with an experimenter present. Each participant sat in front of a computer monitor and wore headphones with a sensitive built-in microphone with adjustable volume. The headphones were connected to the Voice Operated Relay (VOR) of a Carnegie Mellon University button box, a voice activated key that gives results in milliseconds[5] (Cohen et al., 1993). A tie microphone connected to a magnetic tape recorder was used to record the actual voice responses for offline verification. Before testing, participants read a list of words into the microphone to adjust the sensitivity of the VOR for each participant individually. The experimenter also wore headphones (connected to the computer via a two-prong connector), and hand recorded each correct response and all naming errors on a score sheet during testing.

The stimuli were presented directly from a Macintosh computer using PsyScope software. The experiment began when the participant pressed a button on the keyboard. Three lists of all 920 verbs were created, each with a different random order. Participants were randomly assigned to one list. The verbs were presented in a continuous sequence with a 1500 ms inter-stimulus interval between each word. Participants were given 3000 ms to respond from the onset of each stimulus, but as soon as a voice response was detected the inter-stimulus interval or ISI began. A black dot appeared at the bottom of the computer screen when a response was received. An “NR” for “no response” was automatically marked in the data file if the 3000 ms period ended prior to the participant’s response. There were eight rest breaks throughout the experiment (approximately 102 verbs presented per block). Participants were allowed to continue when ready, by pressing any button on the keyboard. The entire experimental session lasted approximately 45 minutes.

Transitivity and Concreteness Norming. Depending on their contextual use, verbs can belong to either transitive (requiring a direct object) or intransitive categories. Concreteness is usually employed when referring to nouns (the word “though” is less concrete than “dog”), but verbs may also be described by varying degrees of concreteness (e.g., action verbs like “kick” versus verbs of being like “am”). Currently, there are no resources available to obtain normed ratings of concreteness or transitivity for all of the Spanish inflected verbs used herein. Normed values of transitivity (and concreteness) for Spanish inflected verbs would entail undertaking a more extensive study. This is not the primary goal of this study, but is a worthy endeavor for a future study. We did, however, want to provide an estimate of concreteness and transitivity for the 50 infinitive verb forms used in the present study to examine whether these variables were associated with response time. Ratings were therefore collected using a short questionnaire from a separate group of 30 native speakers of Spanish living in San Antonio, Texas. Subjects rated each verb lexeme on a 5-point Likert scale that ranged from high to low concreteness and transitivity. The mean values for all ratings across subjects were calculated and while they are not normative are made available to readers as an estimate of concreteness and transitivity for the 50 infinitive verb forms used in the study. They can be accessed online along with the SVI dataset at http://crl.ucsd.edu/experiments/svi/.

Main Analyses and Results

Mean response time (RT) in milliseconds (ms) was the primary dependent variable across all analyses. Two operationally defined measures of RT were examined: (1) RT measured from word onset - from the onset of the stimulus to the onset of the verbal response, (2) RT measured from word offset (stimulus onset minus stimulus duration ms). The duration of each stimulus sound file was determined by audio review and visual inspection of the speech waveforms. Onset RT provides a standard RT measure, while offset RT eliminates the duration of the word as a factor in response delay. Accuracy was also measured as the percentage of responses that were correctly repeated in the single word-shadowing task, or for which a pronoun was correctly generated in the pronoun production task. Given the high accuracy, (only 3.9% errors for pronoun production and 1.9 % for repetition) errors were not analyzed further. Only onset RT was used as a dependent measure in multiple regression analysis, while the other measures were used for descriptive statistics of the means only.

Data were examined by item (N=920 items by 30 participants for each task). Only correct responses were used for the analyses. Mechanical error (e.g., no voice detected when response was given or voice detected early) and participant error (e.g., incorrect pronoun or misheard word) were coded offline in each participant's data file, and removed from the dataset prior to averaging. Trials removed due to error amounted to 2,275 or 8.2% of the total data in the pronoun production task and 1,357 or 4.9% of the total data in the repetition task. Mean RTs were measured from stimulus offset and onset after removing univariate outliers exceeding three standard deviations from the mean (an additional 988 trials or 3.6% of the total data for repetition and 1,743 trials or 6.3% of the total data for pronoun production). The remaining 23,582 pronoun trials (85.4% of the data) and 25,255 repetition trials (91.5% of the data) were used for further analyses.

Verb frequency counts were logarithmically transformed to improve pairwise linearity and to reduce skewness. A variable called "alternative pronoun" consisting of four groups was created to categorize verbs that take more than one alternative pronoun (e.g., the pronoun for the imperfect form of *correr* [to run], *corría* can be either *yo* [I], *ella* [she], *él* [he], *or used* [you formal]. Categorical variables were also created for verb class, syllable stress, and first phoneme sound articulation. Finally, multivariate outliers were assessed across the ten independent variables (Table 3) through Mahalanobis distance[6], indicating that at least one variable was skewed (i.e. the independent variable, "number of tenses"). This variable contributed to multivariate outliers with a Mahalanobis distance greater than X^2=29.59, p<0.001. Removal of this factor reduced the number of outliers to one factor. This factor was not extreme and therefore, was retained. Residual plots showed no large departures from a linear association.

INSERT TABLE 3 ABOUT HERE

Mean Response Time Differences Across Tasks

A two-sample t-test was used to compare mean onset RTs for the repetition and pronoun production tasks. Producing a pronoun corresponding to a spoken verb took on average 350 ms longer (M=1242.1, SD=119.5) than repeating the verb (M=885.5, SD=67.3) (t [1468.5] = 65.3, p<0.001; equal variance not assumed)[7].

Mean Response Time Difference Across Independent Variables

Analysis of variance (ANOVA) with Dunnett T3 correction was used to examine mean differences in RT across the following categorical factors: “verb class” (e.g., *-er,- ir,- ar*), “number of syllables”, “tense”[8], “alternative pronoun”[9], “syllable stress position”, “canonical stress”, “verb regularity”, “voiced/voiceless pronunciation”, and “first phoneme” sound pattern (e.g., fricative, stop, nasal). Partial eta squared (PES)[10] was used to determine effect size. All remaining means analyses were conducted on each task separately.

Means Comparisons from Onset. Univariate analysis showed significant mean differences across verb class, number of syllables, tense, stress position, verb regularity, and first phoneme articulation in both tasks, as well as alternative pronoun (relevant to the pronoun task only) and canonical stress in the pronoun production task (Table 4). In particular, *-er* verbs, verbs containing one to two syllables, stress occurring on the first syllable, irregular verbs, verbs beginning with a stop sound, and canonical stress were associated with faster RTs than comparable conditions. In the pronoun production task, RTs were slower for verbs with multiple alternative pronouns compared to verbs with fewer or no alternative pronouns.

INSERT TABLE 4 ABOUT HERE

Means Comparisons from Offset. There were significant mean differences measured from stimulus offset (with duration of the stimulus removed) and a reversal in pattern as compared to stimulus onset, for number of syllables and stress position in both tasks (Table 5). In particular, words containing four or five syllables elicited faster offset RTs than one, two, or three syllable verbs. Furthermore, stress occurring on later syllables was associated with faster RTs compared to earlier stressed syllables.

INSERT TABLE 5 ABOUT HERE

Predictors of Response Time at Onset

Correlation coefficients (Pearson's correlation) were used to investigate linearity between the dependent (RT) and the ten independent variables (Table 3) and to examine multicollinearity between each independent variable. Independent variables with a correlation of 0.8 or greater were considered multicollinear (Katz, 2006). Pearson's correlations were classified as weak for r=0.20, moderate for r=0.50, and strong for r=0.80 (Cohen, 1988). Given that a significant p-value does not always mean the presence of a strong relationship with large sample sizes (Odberg et al., 2001), a predictor variable that had a correlation of 0.20 or greater with RT (the dependent variable) was considered for inclusion in a standard linear regression analyses. Using this method, all of the independent variables enter the regression equation at one time. Each independent variable is evaluated in terms of what it adds to the prediction that is different from the predictability afforded by all of the other variables (Tabachnick and Fidell, 2001). Separate item analyses were performed for the repetition and pronoun production tasks. The percent of unique and shared variation explained by each predictor variable was estimated in each model. A $P<0.05$ was used to denote statistical significance.

In univariate analyses, five factors were associated with the time required to repeat a verb measured from stimulus onset: stimulus duration, number of syllables, word frequency, and certain levels of stress position and first phoneme sound articulation. Of these, stimulus duration, word frequency, second syllable stress, and the fricative and vowel phonemic sounds contributed to predicting RT for repeating a verb in multivariate regression analyses (Table 6). Together, these variables accounted for 62 percent of the variance in RT ($P < 0.001$). Stimulus duration contributed the largest to the effect size, accounting for 21.5 percent of the variance in the model.

INSERT TABLE 6 ABOUT HERE

Verb characteristics bivariately associated with pronoun production measured from stimulus onset included alternative pronoun, stress position, number of syllables, stimulus duration, and specific levels of first phoneme sound articulation. In multivariate regression analyses, verbs with a greater number of alternative pronouns, stress occurring at a later position in the verb, stimulus duration, and verbs beginning with a nasal or vowel phonemic sounds significantly contributed to predicting pronoun production (Table 6). Together, these variables accounted for 32.5 percent of the variation in onset RT ($P < 0.001$), with the largest effect size attributed to verbs having multiple pronouns.

Concreteness and Transitivity

While we did not have RTs for each of the 50 verb lexemes, we took the averaged response times of the root (the portion of each verb that is common across all verb conjugations) and used this value as a dependent measure for correlation with ratings of transitivity and concreteness. No linear association was found between the mean root response time when compared to mean values for concreteness (Pearson r= 0.01) and transitivity (Pearson r=0.03). Although, it is worth noting that the trend was such that mean root response times were faster for verbs judged to be more concrete and transitive than those judged to be less concrete and intransitive.

Summary

In summary, factors such as number of syllables, stress position, stimulus duration, and certain levels of a verb's initial phoneme facilitate both verb shadowing as well as pronoun production at onset. Of particular relevance to Spanish verbs, was the finding that verbs with more than one possible pronoun lead to slower RTs compared to verbs having only one possible subject pronoun. This finding suggests that at least some grammatical aspects unique to the Spanish language influence response time. These findings also suggest that a listener can prepare a response prior to the end of auditory stimulus presentation. This is evident from the reversal of RT effects when comparing onset to offset RTs. When measured from stimulus onset, listeners were able to respond faster to verbs with fewer syllables and verbs with stress on the first syllable than to verbs with stress appearing later. However, when the stimulus duration was subtracted from the RT (e.g., RT measured from stimulus offset), the direct inverse relationship was observed.

Discussion

As research into the processing and production of Spanish language continues to grow, so too will be the greater utility for quick, user friendly online lexical resources that provide psycholinguistic frequency and summary statistics suitable for use with standard spreadsheets. Only recently have lexical databases such as LEXESP and BuscaPalabras incorporated psycholinguistic measures of Spanish words. We present the first dataset available free of charge dedicated to psycholinguistic measures specific to Spanish inflected verbs including measures of grammar, morphology, word length, word frequency, phonology, and response time. This is also the first study of inflected Spanish verbs to include concreteness ratings. The Spanish Verb Inventory is a lexical dataset developed to facilitate the study of Spanish verbs, as described in Part I.

In Part II of this report, we examined several characteristics found in the Spanish Verb Inventory to determine which ones best predicted response time to repeat a verb and produce its pronoun. Our results highlight four main factors that contribute to RT: word length, stress position, phonetic patterns, and word frequency. Though each of these topics merits an extensive review, herein we briefly discuss these factors as per our findings and highlight their relation to each body of literature independently.

Word Length

The relationship between word length and the time it takes to respond could have three potential outcomes. First, if the time it takes to produce a verbal response is independent of word length then, by removing the duration of the stimulus from the response time, there would be no difference between short and long words. Second, if response time strictly depends on the length of the word, such that longer words elicit longer preparation times then, when removing the duration of the stimulus from response time, longer words should lead to longer response times than shorter words. The third option is that preparation time begins at some point after stimulus onset and not stimulus offset. In this case, removing the stimulus duration from the response time could elicit shorter response times for longer than shorter words. If this relationship were such that it takes a fixed amount of acoustic information before one can begin to prepare a response, then there would be no difference in response times from the onset of shorter or longer words. Our comparison of the number of syllables and stress position measured from stimulus onset suggest that words containing more syllables elicit longer response times. However, when the effect of stimulus length was removed by looking at the offset latencies, an inverse result was observed: verbs containing more syllables resulted in shorter offset RTs in both tasks. This indicates that listeners begin to prepare a response at some point from stimulus onset, but that this time is not a fixed preparation time. This notion was supported in post hoc analysis by the fact that when the stimulus duration was removed, many RTs yielded negative values, indicating that individuals were able to respond before the end of the word (these were negative responses for words that had positive responses measured from stimulus onset. Of 26,070 responses in the pronoun production task, 1135 (4.4 percent) produced negative offset times. Similarly, of 26,444 individuals responses in the repetition task, 1745 (6.6 percent) produced negative offset times. Only responses that had positive onset RT values were used for these analyses, hence these negative offset values were necessarily a result of voice onset times occurring between the onset and offset of the stimulus. In the repetition task, onset RT increased with increasing stimulus duration (r=0.74). Conversely, a strong significant negative correlation was observed for offset RT and stimulus duration (-0.84), with faster RTs resulting with increasing stimulus duration. A similar but weaker relationship was observed in the pronoun task.

This is supported by previous research that suggest that when it comes to two word pairs, the point of articulation depends on both the time required to prepare a word, as well as the length of a preceding word (Griffin et al., 2003). Griffin et al. (2003) proposed that articulation of shorter words reduced the amount of "last second preparation" time available for preparing the next word, such that speakers delayed pronouncing the first word while preparing the second. Speakers had more time during speech to prepare longer words, so pronunciation of the first word was initiated earlier (see also Meyer et al., 2003; Schriefers and Teruel, 1999).

Studies suggest that this early recognition process may be facilitated by early acoustic information (e.g., O’Rourke & Holcomb, 2002; Wheeldon & Levelt, 1995). O’Rourke and Holcomb (2002) tested the impact of acoustic input on word recognition using event related potentials (ERP) for words ranging in duration from 600 to 900 ms. The authors found that N400 peak latency and RT measured from word onset were faster for stimuli when the acoustic uniqueness point occurred earlier in the word than points occurring later. To quantify the time course of processing, the authors evaluated the onset and offset of the N400 using consecutive t-tests at each electrode site contrasting words with early and late uniqueness points. Significant positive t-test (i.e., early uniqueness point had a more negative peak than late uniqueness point) were found early in the time course (at about 400 ms). Significant negative t-tests (i.e., later uniqueness point had a more negative peak than earlier uniqueness points) were found later in the time course (at about 550 ms), with the earliest difference between early and late uniqueness points occurring almost 200 ms before the offset of the shortest stimuli. Based on these studies, it is evident that early acoustic input of units, smaller than the complete word, facilitates word recognition.

Although determining the true uniqueness point of each of our 920 verbs was beyond the scope of this study, we approximated the uniqueness point by subtracting the duration of the root from the stimulus duration. The root represents the portion of each verb that is common across all verb conjugations, especially in non-stem changing regular verbs. Hence, the following syllable often carries the unique conjugation information for each word. Our data showed a weak positive correlation between post root verb duration and onset RT (r=0.40) in the repetition task but not in the pronoun production task (r=0.063). While the uniqueness point did not contribute to producing the pronoun of the verb in the present study, our findings support those of previous research suggesting that articulation can begin at some point prior to hearing the completion of a word.

Stress Position

Another feature of words that may determine the ability to recognize them prior to hearing them to completion is lexical stress. A word's metric shape includes information about both the number of syllables in the word and the position of stress (i.e., the syllable that is stressed). According to the WEAVER++ model (Levelt et al., 1999), accessing word forms entails activation of the word's morphological makeup, its metric shape and its segmental makeup. In Spanish, the syllable is considered to be a basic sublexical processing unit (Alvarez et al., 2001; Barber et al., 2004; Carreiras et al., 1993; Perea and Carreiras, 1998). The position of stress varies systematically across verb conjugations in Spanish, with stress most commonly occurring on the second to the last syllable (canonical stress for Spanish words). Several studies have shown an effect of stress position on word recognition and production. For instance, Jansma and Schiller (2004) instructed participants to press a button if the stress of a bi-syllabic Dutch noun represented by a picture occurred on the first syllable and avoid pressing the button if the stress occurred on the second syllable. The authors found that mean decision latencies were significantly faster for words whose stress occurred on the first syllable compared to the second syllable. In another study, Schiller et al. (2006) used a picture naming and self-monitoring task to investigate the effects of stress on onset latency time. Participants were asked to name a picture consisting of a bi-syllabic or tri-syllabic word (having initial, pre-final, or final stress) or to suppress overt naming of the pictures and instead press a button to determine if a picture had initial or final stress. There was a significant subject advantage in onset latency for picture names with final stress compared to picture names with initial stress in bi-syllabic words, which disappeared in trisyllabic words. Mean RT in the self monitoring task showed a significant advantage of the initial stress condition over the final stress in bi-syllabic words (i.e., the canonical stress position), and initial stress facilitated responses to trisyllabic words. The authors concluded that encoding of stress follows a rightward incremental pattern. These findings suggest that stress position is an important factor in recognizing and processing verbs

Phonetic Patterns

Evidence suggests that initial phonemic sounds affect the articulatory motor components of naming performance (Spieler & Balota, 1997). In this study initial phoneme as a factor did not strongly correlate with response time in either task. However, when examining the individual levels of initial phoneme our results showed some significant relationships between certain initial phonemes and RT. A comparison of mean differences (Table 4) suggested that verbs containing a stop or nasal phoneme were repeated faster than other initial sounds. When the variance of other factors was considered, fricative and liquid phonemes were significant predictors of response time when repeating a verb, increasing response time compared to other initial phonemes (Table 6). These findings are partially consistent with previously reported findings.

For example, in Treiman et al. (1995) participants were instructed to read a word as soon as it appeared on a screen. In a standard regression analysis that controlled for consistency, neighborhood size, word length, word frequency and familiarity, the authors found that faster onset RTs were produced for words having liquids/semivowels or nasal initial sounds than fricatives and affricates.

One could argue that this effect is simply due to mechanical error, given the evidence that voice keys are sensitive to the acoustic properties of the initial phoneme (see Kessler et al., 2002; Pechmann et al., 1989; Rastle & Davis, 2002; Tyler et al., 2005). In particular, voiceless initial phonemes can fail to trigger the device, eliciting delayed RTs compared to voiced-initial phonemes (Trieman et al., 1995). However, a means analysis (Table 4) shows the opposite pattern – verbs with voiceless onset were repeated *faster* than voiced-initial phonemes, indicating that these differences are not solely due to a bias in the trigger device. Moreover, similar effects were found for pronoun production, where the words triggering the voice key are all voiced (vowel onset) and unrelated physically to the verb heard. In particular, producing an appropriate pronoun for verbs starting with a stop or nasal phoneme was faster than for other initial sounds, and nasal phonemes significantly predicted faster response times, while vowel sounds predicted slower response times. It is not clear from these data what the effect of initial phoneme on pronoun production means, but it is likely due to salience differences of the initial sound during perception of the word. Nevertheless, these data indicate that initial phoneme can be a significant contributing factor to performance in recognizing and processing verbs.

Word Frequency

The effect of word frequency on recognition and production has been examined extensively across a variety of psycholinguistic dimensions. Lexical frequency effects have been observed across a number of behavioral tasks including auditory word recognition (Connie et al., 1990), visual word recognition (Alegre & Gordon, 1999; Spieler & Balota, 2000; Treiman et al., 1995), picture naming tasks (Alario et al., 2002; Bachoud et al., 1998; Bates et al., 2003; Cuetos et al., 1999; Cuetos et al., 2006; Cuetos, submitted; Navarrete et al., 2006), and eye movement tasks (Pynte & Kennedy, 2006). These studies find that higher frequency words facilitate lexical access and production compared to low frequency words. In the present study, we found that word frequency contributed to predicting response time for verb repetition when other covariates were controlled for, but was not a significant factor in producing a pronoun.

Word frequency has also been examined in relation to word length. For instance, Trieman et al. (1995) found an interaction between word frequency and word length such that increases in word length produced lower RTs for low frequency words as compared to high frequency words. Similarly, Spieler and Balota (1997) found that frequency, as a sole predictor, accounted for 7.3 percent of the variance in naming latency for 2,870 monosyllabic words. When word length and neighborhood frequency were entered into the model all three variables accounted for 21.7 percent of the variance.

Magnetoencephalography (MEG) (e.g., Assodollahi & Pulvermüller, 2003) and ERP (e.g., Hauk & Pulvermüller, 2004) studies with visually presented word stimuli have also investigated the relationship between word frequency and word length. In particular, low frequency words led to stronger amplitude responses primarily in the left occipitotemporal cortex compared to high frequency words (Assodollahi and Pulvermüller, 2003). This occurred at early time intervals for short words (between 120-170 ms post stimulus onset) while longer words showed a frequency effect later in the brain response (225-250 ms post stimulus onset). Interestingly, syllable frequency has the opposite effect on RT compared to whole word frequency (Alvarez et al., 2001; Barber et al., 2004; Carreiras et al., 1993; Conrad & Jacobs, 2004). The assumption is that words with high frequency syllables trigger a larger number of lexical candidates because they are shared by more words (Alvarez et al., 2001; Barber et al., 2004) and thus lead to longer RTs until the uncertainty is resolved. For instance, Carreiras et al. (1993) used a lexical decision task and a naming task to examine the effect of lexical (or word) frequency and the positional frequency of each syllable in 144 bi- and tri-syllabic words. Results showed that RT was significantly faster for high frequency words than for low frequency words. However, syllable frequency produced an inhibitory effect: low frequency syllables produced faster RTs compared to high frequency syllables. Similar results were found for bi- and tri-syllabic words across both tasks, and have been observed in several recent behavioral (Alvarez et al., 2001, Experiment 1 and 2; Conrad & Jacobs, 2004, Experiment 1; Perea & Carreiras, 1998) and electrophysiology studies (e.g., Barber et al., 2004).

In the present study, we attempted to measure first and second syllable frequency by using BPAL, the Spanish standard database used in previous studies. The online database only includes first syllable frequencies for 244 verbs (26.5 percent of total). Of these, second syllable frequencies were available for 91 verbs (9.9 percent). Of the 224 verbs with first syllable frequencies, we did find a weak correlation between first syllable frequency measured from stimulus onset (r= -0.131) and offset (r=0.123) in the repetition task, but not in the pronoun production task. Although this correlation is weak, it provides support for the inhibitory effect discussed above, in that, words that share their first syllable with many words (i.e., more frequent first syllable) take longer to repeat than less frequent first syllable words.

In brief, our results show that word frequency, and syllable frequency, can also significantly contribute to verb recognition and processing.

Conclusion

This report describes a lexical dataset that incorporates grammatical, morphological, phonological, and phonemic psycholinguistic dimensions including concreteness ratings particular to Spanish verbs. Using this dataset, we demonstrated significant differences in auditory verb recognition and processing across two psycholinguistic measures that are supported by past research, while accounting for other factors that could potentially affect lexical access. In particular, our findings suggest that measures of word length, word stress, word frequency, phonetic composition, and verbs with alternative pronouns contribute to differences in response times for both repeating a verb and producing an appropriate pronoun. Clearly, these data do not account for the total range of potentially explanatory or confounding factors that may contribute to predicting RT (e.g., measures of phonological neighborhoods, word familiarity, imageability, concreteness, and transitivity). Furthermore, while grammatical factors unique to Spanish (such as multiple pronouns) were associated with RT, analysis of other factors such as verb tense was limited due to the fact that several of the lexical forms repeat across verb conjugation. However these results provide psycholinguists with an overview of some of the important factors currently available to study Spanish inflected verbs. The online dataset presented in this manuscript offers researchers a tool for addressing future questions related to verb processing in Spanish.

Acknowledgments

This project was started in 1995 by Elizabeth Bates, in collaboration with Antonella Devescovi and Nicole Wicha, as a starting point for creating a foundation for verb processing in languages with rich inflectional morphology, in particular Italian and Spanish. The work has been supported over the years by a grant to Elizabeth Bates, "Cross-linguistic studies of aphasia," NIH/NIDCD R01 DC00216, and faculty startup funds from UTSA to Nicole Wicha. We would like to thank the many individuals who have provided valuable support and advice, in particular, Iliana Reyes (the voice for the verbs) and Robert A. Buffington for technical assistance, Mark Davies and Manuel Carreiras for assistance with their databases and edification in linguistic terminology, Marta Ortega-Llebaria, Mary Ellen Garcia, and Fred F. Jehle for providing assistance with linguistic rules of Spanish verbs, and Fred Dick, Vic Ferreira and 2 anonymous reviewers who have provided valuable criticism and comments on previous versions of this manuscript. In no way are these individuals responsible for any remaining errors in this manuscript. We would also like to acknowledge our collaboration with the College of Humanities at the Autonomous University of Baja California (UABC) in Tijuana as our valued participant population, and in particular Lourdes Gavaldón de Barreto, for her invaluable assistance on this and many other projects.

Appendix

Percentages in (%) are out of the total of 920 verbs.

1. Grammatical Dimensions.

    - Verb tense – imperfect (21.7%), preterite (23.9%), future (27.2%), and present indicative (27.2%).
    - Person – 1st (40.2%), 2nd (21.7%), and 3rd person (38.1%).
    - Alternative pronoun – (43.5%) verbs that can take more than one pronoun (e.g., *caía* can represent more than one person --*yo caía* [I was falling], *él caía* [he was falling]) (see also footnote 9).
    - Number – singular (59.8%) and plural (40.2%).
    - Verb class – *-ar* (52.9%), *-er* (33.0%), *-ir* (14.0%).

2. Morphological Characteristics (word form):

    - Regularity: regular (84.0%) and irregular verbs (16.0%).
    - Stem Changing – (4.7%) changes in vowel stem of the word when inflected. For instance, the vowel *o* changes to *ue* or *u,* and the vowel *e* in the stem changes to *ie* or *i* (e.g., the *o* in *dormir* [to sleep] changes to *ue* in *yo duermo* [I sleep]).
    - Root, Stem and Suffix

3. Measures of Word Length:

    - Number of characters – ranging from 2 to 11 (mean=6.73).
    - Number of syllables – ranging from 1 to 5 (mean=2.86)
    - Stimulus duration – mean duration 624.8 ms; range 274 – 965 ms
    - Length of the root – the root was defined as the simplest form of the lexical morpheme, after all affixes (all bound and free forms of the morpheme) are removed;

the root was measured from word onset to the offset of the root in milliseconds (ms) (e.g., the root *corr-* in the verb *corremos* [we run]).

- Length of the suffix after the root – measured from the offset of the root to the end of the word in ms (e.g., the suffix *-emos* in the verb *corremos* [we run]).
- Length of the stem – As with root length, stem length was measured from the onset of the word to the offset of the stem (in ms). The stem was defined as the root plus the thematic vowel immediately following the root of the infinitive form of a regular verb (e.g., the stem *corre-* in the verb *corremos* [we run]) (Linares et al., 2006). In the case of an inflected irregular verb not containing the thematic vowel of the infinitive verb form, the stem is equal to the root of the word. In the verb *corrimos* [we ran], for instance, both the root and stem are denoted as *corr-* since the thematic vowel *e* is replaced with *i*. Similarly, the root *d* in the verb *doy* [I give] is the same as its stem, since the infinitive of the verb *dar* [to give] contains the thematic vowel *a* instead of *o*. The same rule was applied to stem changing verbs (e.g., in the verb *dormir* [to sleep], the root and stem for the inflected verb *duermo* is denoted as *duerm-*).
- Length of the suffix after the stem – measured from the offset of the stem to the end of the word (in ms).
- Structure – consonant and vowel structure of the word

4. <u>Measures of Frequency:</u>

- Lemma Frequency – frequency counts were obtained from Davies' (2001) Corpus del Español, the largest online word frequency database for written Spanish words. Only counts that occurred during the 1900s (20 million words) were included and consisted of words obtained from written literature, oral texts, newspapers, and encyclopedias. All 50 verb-lemmas (100 percent) were retrieved from the corpus. Additional frequency corpora, such as Juilland, A., & Chang-Rodriguez, A. (1964) and Alameda, J.R. & Cuetos, F., (1995) are included in the dataset, but were not used for analysis in part II given that not all verb forms were available.

- Lexeme Frequency – lexeme frequency counts were also obtained from Davies' (2001) Corpus del Español for the 1900s data. A total of 54 out of a total of 920 verbs (5.8 percent) were either not found in the corpus or contained a frequency of zero during that century.
- First syllable Frequency – first syllable frequency counts for a total of 244 inflected verbs out of 920 (26.5 percent) were available from the BuscaPalabras program (Davis & Perea, 2005).

5. <u>Phonetic Dimensions:</u>

- Syllable stress – the syllable position carrying the stress of the word on the first (22.4%), second (47.5%), third (25.2%), and fourth (4.9%) syllable.
- Canonical Stress – stress that falls on the penultimate syllable. This is the most common and default stress position for Spanish words, unless marked by a stress accent (e.g., *camino* [I walk] versus *caminó* [he/she/it/you walked]). A verb is defined as canonical (62.4%) or noncanonical (37.6%).
- Stress Type – in traditional Spanish grammar, stress is defined based on four categories. Although stress patterns in Spanish are assigned from right to left, they are presented from left to right as follows: oxytone [*aguda*] (stress on the final syllable); paroxytone [*llana* or *grave*] (penultimate stress), proparoxytone [*esdrújula*]

(antepenultimate stress or stress on the third to last syllable), sobresdrújula [*sobresdrújula*] (preantepenultimate stress on the fourth to last syllable). All proparoxtyones and *sobresdrújulas* have written accent marks. The frequency of stress in our verb list includes: 62.4% llana, 32.2% aguda, and 5.4% esdrújula.

- First phoneme sound articulation – provides sub-classification of obstruent and sonorant constants for the first phoneme of each inflected verb. Verbs are categorized as beginning with a fricative [f, s, v, z] (18.3%), stop [t, k, b, d, g, p] (52.8%), liquid [l, r] (8.0%), vowel [a, e, i, o, u] (14.9%), or nasal [m, n] (6.0%) sound.
- Voiced/Non-voiced articulation - characterizes sounds that are produced with vibration of the vocal cords (in English [b] and [d] are voiced as opposed to [p] and [t] which are voiceless). Of the verb list, 53.5% are voiced, while 46.5% are voiceless.
- Consonant Phonetics – classifies each word according to whether it is bilabial (22.7%), labiodental (0.5%), dental (14.1%), alveolar (23.8%), palatal (1.8%), velar (22.1%), or central or middle (7.4%).

6. <u>Additional linguistic dimensions:</u>

- Transitivity* – whether or not a verb can take a direct object (based on a subjective rating on a 5-point scale, Likert-type questionnaire, mean of N=30): mean transitivity rating = 2.94, range = 1.97 to 4.0
- Concreteness* – how concrete a verb is (based on a subjective rating on a 5-point scale, Likert-type questionnaire, mean of N=30): mean concreteness rating = 3.92, range = 1.73 to 4.77

* See methods section.

7. <u>RT Data:</u>

- Verb Shadowing – Mean voice onset times from the onset of the stimulus (mean = 885.48, range = 392.04 [repetition] and offset of the stimulus (mean = 260.74, range = 554.02) for each verb entry from 30 participants on an auditory shadowing task (see methods)
- Pronoun Production – Mean voice onset times from the onset of the stimulus (mean = 1242.09, range = 939.08) and offset of the stimulus (mean = 617.46, range = 1044.88) for each verb entry from 30 participants on a pronoun production task (see methods).

Footnotes

1. The LEXESP database is available on CD-ROM. It can be purchased from the website of the Universitat de Barcelona: www.ub.es/edicions/libros/v14.htm.

2. The lexical dataset is available as an Excel file along with the sound files at http://crl.ucsd.edu/experiments/svi/.

3. Active data collection occurred in 1995.

4. Out of the 920 verbs used, 42 (4.6%) can also be classified as nouns (e.g. *camino* can be used to mean “I walk” or to refer to a “road”), one (0.1%) can be used as an adjective, two (0.2%) can be used as adverbs, two (0.2%) can be used as interjections, and two are inflected verbs for an alternative verb (*sentar*) (Cassell, 1968; Garcia-Pelayo and Gross, 1976). Given that the instructions to participants indicated that a series of inflected verbs were to be auditorily presented to them, and that the majority of the stimuli can only be verbs, we assume that participants were accessing the verbal form of these words. However, without a secondary measure, such as a semantic association measure, we cannot know this for sure. This is something to pursue in future studies.

5. A complete description of this device is available on the on the Center for Research in Language (CRL) website at http://crl.ucsd.edu/experiments/svi/.

6. The Mahalanobis distance identifies a factor as a potential outlier when the contributions of each factor are considered together and if the factor produces a value greater than a given cutoff value (in this case, 29.59).

7. Since the assumption of equal variance was not met, 1448.5 degrees of freedom were used instead of 1838 (N=1840). This places greater restrictions on the number of values in a sample that are free to vary.

8. Analyzing differences in processing speed across tense is less straightforward in Spanish because, as mentioned in the text, while there are four tenses, several of the lexical forms repeat across verb conjugations. For instance, the 1st, 2nd (formal), and 3rd person singular forms of the imperfect tense for all verbs are identical (e.g., *yo abría*, *él/ella abría*, *usted abría – I was opening*, *he/she was opening*, *you (formal) were opening*) (see Table 2). Thus, analysis of these values was limited to the present, past, and future tense singular forms.

9. We created a variable (called alternative pronoun) for verbs with alternative object pronouns to determine whether verbs with a greater number of alternative pronouns contributed to predicting response time in the pronoun production task only. Four categories were used to classify verbs that: (1) have no alternative pronouns, (2) can take either él/ella/usted, (3) can take either ellos/ustedes, and (4) can take either yo/él/ella.

10. This value describes the proportion of variability in the dependent variable that is attributable by the independent variable. PES ranges from 0 to 1 and is considered a weak effect if between 0.00 and 0.04, a moderate effect if between 0.05 and 0.14, and a strong effect if greater than 0.14 (Tabachnick & Fidell, 2001). Eta square, an estimate of systematic variance in the population was not reported because its value for a particular independent variable is influenced by the number and significance of other independent variables in the design.

References

Alameda, J. R., & Fernando, C. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano* (Vol. II). Oviedo: University of Oviedo Press.

Alario, F. X., Costa, A., & Caramazza, A. (2002). Frequency effects in noun phrase production: Implication for models of lexical access. *Language and Cognitive Processes, 17*(3), 299-320.

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language, 40*(1), 41-61.

Alvarez, C. J., Carreiras, M., & Taft, M. (2001). Syllables and morphemes: Contrasting frequency effects in Spanish. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(2), 545-555.

Andonova, E., D'Amico, S., Devescovi, A., & Bates, E. (2004). Gender and lexical access in Bulgarian. *Percept Psychophys, 66*(3), 496-507.

Arciuli, J., & Cupples, L. (2003). Effects of stress typicality during speeded grammatical classification. *Lang Speech, 46*(Pt 4), 353-374.

Arnaud, P. J. L. (1999). Target-error resemblance in French word substitution speech errors and the mental lexicon. *Applied Psycholinguistics, 20*(2), 269-287.

Assadollahi, R., & Pulvermuller, F. (2003). Early influences of word length and frequency: A group study using MEG. *Neuroreport: For Rapid Communication of Neuroscience Research, 14*(8), 1183-1187.

Aydelott, J., & Bates, E. (2004). Effects of acoustic distortion and semantic context on lexical access. *Language and Cognitive Processes, 19*(1), 29-56.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (Version 2)

[CD Rom]. Philadelphia.

Bachoud-Levi, A.-C., Dupoux, E., Cohen, L., & Mehler, J. (1998). Where is the length effect? A cross-linguistic study of speech production. *Journal of Memory and Language, 39*(3), 331-346.

Ball C. N. (2004). TheWeb Frequency Indexer. Department of Linguistics, Georgetown University. http://www.georgetown.edu/faculty/ballc/webtools/web_freqs.html.

Balota, D. A., Watson, J. M., Duchek, J. M., & Ferraro, F. R. (1999). Cross-modal semantic and homograph priming in healthy young, healthy old, and in Alzheimer's disease individuals. *Journal of the International Neuropsychological Society, 5*(7), 626-640.

Barber, H., Vergara, M., & Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: evidence from ERPs. *Neuroreport, 15*(3), 545-548.

Barde, L. H., Schwartz, M. F., & Boronat, C. B. (2006). Semantic weight and verb retrieval in aphasia. *Brain Lang, 97*(3), 266-278.

Bates, E., D'Amico, S., Jacobsen, T., Szekely, A., Andonova, E., Devescovi, A., et al. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review, 10*(2), 344-380.

Bates, E., Devescovi, A., & D'Amico, S. (1999). Processing complex sentences: A cross-linguistic study. *Language and Cognitive Processes, 14*(1), 69-123.

Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Percept Psychophys, 58*(7), 992-1004.

Bates, E., Devescovi, A., Pizzamiglio, L., D'Amico, S., & Hernandez, A. (1995). Gender and lexical access in Italian. *Percept Psychophys, 57*(6), 847-862.

Bates, E., Devescovi, A., & Wulfeck, B. (2001). Psycholinguistics: a cross-language perspective.

Annu Rev Psychol, 52, 369-396.

Bates, E., & Liu, H. (1996). Cued shadowing. *Language and Cognitive Processes, 11*(6), 577-581.

Bentrovato, S., Devescovi, A., D'Amico, S., & Bates, E. (1999). Effect of grammatical gender and semantic context on lexical access in Italian. *J Psycholinguist Res, 28*(6), 677-693.

Bentrovato, S., Devescovi, A., D'Amico, S., Wicha, N., & Bates, E. (2003). The effect of grammatical gender and semantic context on lexical access in Italian using a timed word-naming paradigm. *J Psycholinguist Res, 32*(4), 417-430.

Berndt, R. S., Mitchum, C. C., & Price, T. R. (1991). Short-term memory and sentence comprehension. An investigation of a patient with crossed aphasia. *Brain, 114 (Pt 1A),* 263-280.

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behav Res Methods Instrum Comput, 33*(1), 73-79.

Bird, H., Howard, D., & Franklin, S. (2000). Why is a verb like an inanimate object? Grammatical category and semantic category deficits. *Brain Lang, 72*(3), 246-309.

Buchanan, l., & Westbury, C. (2000). Wordmine database: Probablistic values for all four to seven letter words in English Language. http://www.wordmine.org/.

Caramazza, A., & Hillis, A. E. (1991). Lexical organization of nouns and verbs in the brain. *Nature, 349*(6312), 788-790.

Carreiras, M., Alvarez, C. J., & de Vega, M. (1993). Syllable frequency and visual word recognition in Spanish. *Journal of Memory and Language, 32*(6), 766-780.

Carreiras, M., & Perea, M. (2004). Naming pseudowords in Spanish: Effects of syllable frequency. *Brain and Language, 90*(1-3), 393-400.

Carreiras, M., & Perea, M. (2004). Naming pseudowords in Spanish: Effects of syllable frequency. *Brain and Language, 90*(1-3), 393-400.

Chen, S., & Bates, E. (1998). The dissociation between nouns and verbs in Broca's and Wernicke's aphasia: Findings from chinese. *Aphasiology, 12*(1), 5-36.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum.

Cohen J.D., MacWhinney B., Flatt M., and Provost J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. Behavioral Research Methods, Instruments, and Computers, 25(2), 257-271.

Coltheart, V., & Evans, J. S. (1981). An investigation of semantic memory in individuals *Memory and Cognition, 9*(5), 524-532.

Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(6), 1084-1096.

Conrad, M., & Jacobs, A. M. (2004). Replicating syllable frequency effects in Spanish in German: One more challenge to computational models of visual word recognition. *Language and Cognitive Processes, 19*(3), 369-390.

Content, A., Mousty, P., & Radeau, M. (1990). BRULEX: Une base de données lexicales informatisée pour le français écrit et parlé. *L'Année psychologique, 90*, 551-556.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of Lexical Stress on Lexical Access in English: Evidence from Native and Non-native Listeners. *Language and Speech, 45*(3), 207-228.

Cresti, E., & Moneglia, M. (Eds.). (2005). *C-ORAL-ROM: Integreated reference corpora for* *spoken romance languages*. Florence: John Benjamins Publishing Company.

Cuetos, F., Alvarez, B., Gonzalez-Nosti, M., Meot, A., & Bonin, P. (2006). Determinants of lexical access in speech production: role of word frequency and age of acquisition. *Mem Cognit, 34*(5), 999-1010.

Cuetos, F., Bonin, P., Ramón Alameda, J., Chalard, M., & Caramazza, A. (submitted). The specific-word frequency effect in speech production: Evidence from Spanish and French. *Language and Cognitive Processes*.

Cuetos, F., Ellis, A. W., & Alvarez, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments & Computers, 31*(4), 650-658.

Davies, M. (2001). Corpus of historical Spanish prose 1200-1900/Corpus del Español. http://www.corpusdelespanol.org.

Davis, C. J. (2005). N-watch: a program for deriving neighborhood size and other psycholinguistic statistics. *Behav Res Methods, 37*(1), 65-70.

Davis, C. J., & Perea, M. (2005). BuscaPalabras: a program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behav Res Methods, 37*(4), 665-671.

de Vega Rodríguez, M., Alvarez, C., & Carreiras Valiña, M. F. (1992). Estudio estadístico de la ortografía castellana (1): La frecuencia silábica. *Cognitiva, 4*(1), 75-106.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychol Rev, 93*(3), 283-321.

Devescovi, A., D'Amico, S., Bentrovato, S., & Bates, E. Processing inflected verbs in Italian. Unpublished Manuscript. Università degli Studi di Roma 'La Sapienza' & University of

California, San Diego.

Devescovi, A., D'Amico, S., & Gentille, P. (1999). The development of sentence comprehension in Italian: a reaction time study. *First Language, 19*, 129-163.

Dollaghan, C., Biber, M., & Campbell, T. (1993). Constituent syllable effects in a nonsense-word repetition task. *J Speech Hear Res, 36*(5), 1051-1054.

Dopke, S. (1998). Competing language structures: the acquisition of verb placement by bilingual German-English children. *J Child Lang, 25*(3), 555-584.

ELRA/ELDA. (1995). The European Language Resources Association. from http://www.tc-forum.org/topictr/tr05thee.htm.

Eriksen, C. W., Pollack, M. D., & Montague, W. E. (1970). Implicit speech: Mechanism in perceptual encoding? *Journal of Experimental Psychology, 84*(3), 502-507.

Federmeier, K. D., Segal, J. B., Lombrozo, T., & Kutas, M. (2000). Brain responses to nouns, verbs and class-ambiguous words in context. *Brain,* 123(Pt 12), 2552-2566.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., et al. (1993). *MacArthur communicative development inventories: User's guide and technical manual*. San Diego: Singular Publishing Group, Inc.

Fischler, I. (1977). Semantic facilitation without association in a lexical decision task. *Memory and Cognition*, *5*, 335–339.

Fry, D. B. (1979). *The physics of speech*. Cambridge: Cambridge University Press.

García-Pelayo y Gross, R., Durand, M. (1976). Diccionario moderno: Español-Inglés, English-Spanish. New York: Ediciones Larousse.

Gentner, D. (1981). Some interesting differences between verbs and nouns. *Cognition and Brain Theory, 4*(2), 161-178.

Gooch, A., & Garcia de Pareded, A.(1968). Cassell's Spanish dictionary, Spanish-English, English-Spanish. New York: Funk & Wagnalls.

Goodman, G. O., McClelland, J. L., & Gibbs, R. W., Jr. (1981). The role of syntactic context in word recognition. *Mem Cognit, 9*(6), 580-586.

Griffin, Z. M. (2003). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin & Review, 10*(3), 603-609.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics, 28*(4), 267-283.

Halle, P. A., & de Boysson-Bardies, B. (1992). *Word recognition before production of first words?* Paper presented at the International Conference on Spoken Language Processing.

Hauk, O., & Pulvermuller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clin Neurophysiol, 115*(5), 1090-1103.

Hillis, A. E., Tuffiash, E., & Caramazza, A. (2002). Modality-specific deterioration in naming verbs in nonfluent primary progressive aphasia. *J Cogn Neurosci, 14*(7), 1099-1108.

Holcomb, P. J., & Anderson, J. E. (1993). Cross-modal semantic priming: A time-course analysis using event-related brain potentials. *Language and Cognitive Processes, 8*(4), 379-411.

Jansma, B. M., & Schiller, N. O. (2004). Monitoring syllable boundaries during speech production. *Brain and Language, 90*(1-3), 311-317.

Juilland, A., & Chang-Rodríguez, E. (1964). *Frequency dictionary of spanish words*. La Haya: Mouton and Co.

Katz, M. H. (2006). *Multivariable analysis : A practical guide for clinicians*. Cambridge: Cambridge University Press.

Kessler, B., Treiman, R., & Mullennix, J. (2002). Phonetic biases in voice key response time measurements. *Journal of Memory and Language, 47*(1), 145-171.

Klapp, S. T., Anderson, W. G., & Berrian, R. W. (1973). Implicit speech in reading: Reconsidered. *Journal of Experimental Psychology, 100*(2), 368-374.

Kolinsky, R. (1998). Spoken word recognition: A stage-processing approach to language differences. *European Journal of Cognitive Psychology, 10*(1), 1-40.

Labelle, M., Godard, L., & Longtin, C. M. (2002). Grammatical and situational aspect in French: a developmental study. *J Child Lang, 29*(2), 301-326.

Laudanna, A., & Voghera, M. (2002). Nouns and verbs as grammatical classes in the lexicon. *Rivista di Linguistica, 14*(1), 9-26.

Levelt, W. J. (2001). Spoken word production: a theory of lexical access. *Proc Natl Acad Sci U S A, 98*(23), 13464-13471.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behav Brain Sci, 22*(1), 1-38; discussion 38-75.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*(1), 1-36.

Linares, R. E., Rodriguez-Fornells, A., & Clahsen, H. (2006). Stem allomorphy in the Spanish mental lexicon: evidence from behavioral and ERP experiments. *Brain Lang, 97*(1), 110-120.

Lindfield, K. C., Wingfield, A., & Goodglass, H. (1999). The contribution of prosody to spoken word recognition. *Applied Psycholinguistics, 20*(3), 395-405.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: the neighborhood activation model. *Ear Hear, 19*(1), 1-36.

Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1991). Similarity neighborhoods of spoken words. In *Cognitive models of speech processing* (pp. 122 - 147). Cambridge: MIT Press.

Lukatela, G., Kostic, A., Feldman, L. B., & Turvey, M. T. (1983). Grammatical priming of inflected nouns. *Mem Cognit, 11*(1), 59-63.

Mackay, A. I., Connor, L. T., Albert, M. L., & Obler, L. K. (2002). Noun and verb retrieval in healthy aging. *J Int Neuropsychol Soc*, *8*(6), 764-770.

MacKay, D. G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer-Verlag New York, Inc.

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior, 23,* 127-150.

Marini, A., Boewe, A., Caltagirone, C., & Carlomagno, S. (2005). Age-related differences in the production of textual descriptions. *J Psycholinguist Res*, *34*(5), 439-463.

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, *244*(5417), 522-523.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition, 25*(1-2), 71-102.

Marslen-Wilson, W. D., & Tyler, L. K. (1997). Dissociating types of mental computation. *Nature, 387*(6633), 592-594.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18*(1), 1-86.

Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language, 48*(1), 131-147.

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*(6), 1146-1160.

Miller, G. A. (1998). WordNet: an electronic lexical database. Cambridge, MA: MIT Press.

Morrison, C. M., Hirsh, K. W., & Duggan, G. B. (2003). Age of acquisition, ageing, and verb production: normative and experimental data. *Q J Exp Psychol A, 56*(4), 705-730.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76*(2), 165-178.

Navarrete, E., Basagni, B., Alario, F. X., & Costa, A. (2006). Does word frequency affect lexical selection in speech production? *The Quarterly Journal of Experimental Psychology, 59*(10), 1681-1690.

Negro, I., Chanquoy, L., Fayol, M., & Louis-Sidney, M. (2005). Subject-verb agreement in children and adults: serial or hierarchical processing? *J Psycholinguist Res, 34*(3), 233-258.

Norris, D. (1994). A connectionist model of continuous speech recognition. *Cognition, 52*(3), 189-234.

Odberg, T., Jakobsen, J. E., Hultgren, S. J., & Halseide, R. (2001). The impact of glaucoma on the quality of life of patients in Norway. II. Patient response correlated to objective data. *Acta Ophthalmol Scand, 79*(2), 121-124.

O'Rourke, T. B., & Holcomb, P. J. (2002). Electrophysiological evidence for the efficiency of spoken word processing. *Biological Psychology, 60*(2-3), 121-150.

Pechmann, T., Reetz, H., & Zerbst, D. (1989). The unreliability of voice key measurements.

Sprache & Kognition, 8(2), 65-71.

Perani, D., Cappa, S. F., Schnur, T., Tettamanti, M., Collina, S., Rosa, M. M., et al. (1999). The neural correlates of verb and noun processing. A PET study. *Brain, 122 (Pt 12)*, 2337-2344.

Perea, M., & Carreiras, M. (1998). Effects of syllable frequency and syllable neighborhood frequency in visual word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 24*(1), 134-144.

Persson, J., Sylvester, C. Y., Nelson, J. K., Welsh, K. M., Jonides, J., & Reuter-Lorenz, P. A. (2004). Selection requirements during verb generation: differential recruitment in older and younger adults. *Neuroimage, 23*(4), 1382-1390.

Phelps, E. A., Hyder, F., Blamire, A. M., & Shulman, R. G. (1997). FMRI of the prefrontal cortex during overt verbal fluency. *Neuroreport, 8*(2), 561-565.

Pillai, J. J., Araque, J. M., Allison, J. D., Sethuraman, S., Loring, D. W., Thiruvaiyaru, D., et al. (2003). Functional MRI study of semantic and phonological language processing in bilingual subjects: preliminary findings. *Neuroimage, 19*(3), 565-576.

Plunkett, K., & Bandelow, S. (2006). Stochastic approaches to understanding dissociations in inflectional morphology. *Brain Lang, 98*(2), 194-209.

Poulisse, N. (1999). *Slips fo the tongue: Speech errors in first and second production*. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Prado, E., & Ullman, M. (under review). Can imageability help us draw the line between storage and composition? *Journal of Experimental Psychology*.

Pynte, J., & Kennedy, A. (2006). An influence over eye movements in reading exerted from beyond the level of the word: Evidence from reading English and French. *Vision*

Research, 46(22), 3786-3791.

Rastle, K., & Davis, M. H. (2002). On the complexities of measuring naming. *Journal of Experimental Psychology: Human Perception and Performance, 28*(2), 307-314.

Roelofs, A. P. A. (2003). Modeling the relation between the production and recognition of spoken word forms. In A. S. Meyer, & Schiller, N.O. (Ed.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 115-158). Berlin: Mouton de Gruyter.

Roelofs, A. (2004). Error biases in spoken word planning and monitoring by aphasic and nonaphasic speakers: comment on Rapp and Goldrick (2000). Psychol Rev, 111(2), 561-572; discussion 573-580.

Scarborough, D., Cortese, C., & Scarborough, H. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology*, *3*, 1-17.

Schiller, N. O., Jansma, B. M., Peters, J., & Levelt, W. J. M. (2006). Monitoring metrical stress in polysyllabic words. *Language and Cognitive Processes, 21*(1-3), 112-140.

Schriefers, H., & Teruel, E. (1999). Phonological facilitation in the production of two-word utterances. *European Journal of Cognitive Psychology, 11*(1), 17-50.

Sebastián, N., Martí, M. A., Carreiras, M., & Cuetos, F. (2000). LEXESP: Léxico. Informatizado del Español. Barcelona: Ediciones de la Universitat de Barcelona, http://clic.fil.ub.es/.

Serra, M., Serrat, E., Sole, R., Bel, A., & Aparici, M. (2000). *La Adquisición del Lenguaje (Language Acquisition)*. Barcelona: Editorial Ariel.

Shapiro, K. A., Moo, L. R., & Caramazza, A. (2006). Cortical signatures of noun and verb production. *Proc Natl Acad Sci U S A, 103*(5), 1644-1649.

Sonnenstuhl, I., Eisenbeiss, S., & Clahsen, H. (1999). Morphological priming in the German mental lexicon. *Cognition*, *72*(3), 203-236.

Spieler, D. H., & Balota, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science, 8*(6), 411-416.

Spieler, D. H., & Balota, D. A. (2000). Factors influencing word naming in younger and older adults. *Psychol Aging, 15*(2), 225-231.

Stanners, R. F., Neiser, J. J., Hernon, W. P., & Hall, R. (1979). Memory representation for morphologically related words. *Journal of Verbal Learning & Verbal Behavior, 18*(4), 399-412.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., et al. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language, 51*(2), 247-250.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics* (4th ed.). Boston: Allyn and Bacon.

Treiman, R., Mullennix, J., Bijeljac-Babic, R., & Richmond-Welty, E. D. (1995). The special role of rimes in the description, use, and acquisition of English orthography. *J Exp Psychol Gen, 124*(2), 107-136.

Turner, J. E., Valentine, T., & Ellis, A. W. (1998). Contrasting effects of age of acquisition and word frequency on auditory and visual lexical decision. *Mem Cognit, 26*(6), 1282-1291.

Tyler, L. K., & Wessels, J. (1983). Quantifying contextual contributions to word-recognition processes. *Perception & Psychophysics, 34*(5), 409-420.

Tyler, M. D., Tyler, L., & Burnham, D. K. (2005). The delayed trigger voice key: An improved analogue voice key for psycholinguistic research. *Behavior Research Methods, 37*(1), 139-147.

Vossen, O. (1998). *Eurowordnet: A multilingual database with lexical semantic networks*. Amsterdam: Kluwer Academic Publishers.

Wayland, S. C., Wingfield, A., & Goodglass, H. (1989). Recognition of isolated words: The dynamics of cohort reduction. *Applied Psycholinguistics, 10*(4), 475-487.

Weiller, C., Isensee, C., Rijntjes, M., Huber, W., Muller, S., Bier, D., et al. (1995). Recovery from Wernicke's aphasia: a positron emission tomographic study. *Ann Neurol, 37*(6), 723-732.

Wheeldon, L. R., & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language, 34*(3), 311-334.

Table 1

List of Experimental Verbs

Spanish Infinitive	English Infinitive	Spanish Infinitive	English Infinitive
abrir	to open	ir	to go
acabar	to finish	jugar	to play
ayudar	to help	lavar	to wash
bailar	to dance	leer	to read
besar	to kiss	llorar	to cry
buscar	to look for	mirar	to look at
caer	to fall	morder	to bite
caminar	to walk	nadar	to swim
cantar	to sing	peinar	to comb
cerrar	to close	poder	to be able
cocinar	to cook	poner	to put
comer	to eat	prender	to turn on
comprar	to buy	querer	to want
correr	to run	regalar	to give
cortar	to cut	romper	to tear
dar	to give	saber	to know
deber	to owe	salir	to go out
decir	to say	saltar	to jump
dibujar	to draw	saludar	to greet

table continues

Spanish Infinitive	English Infinitive	Spanish Infinitive	English Infinitive
dormir	to sleep	sentir	to feel
entrar	to enter	soplar	to blow
esconder	to hide	tocar	to touch
esperar	to expect	traer	to bring
gritar	to shout	venir	to come
hacer	to do	ver	to see

Table 2

Simple Tense of the Indicative Mode in Spanish Verbs

Person/Number		Present	Imperfect	Future	Preterite
		Conjugation I (caminar - to walk)			
1st	sing	camino	caminaba	caminaré	caminé
2nd	sing	caminas	caminabas	caminarás	caminaste
3rd	sing	camina	caminaba	caminará	caminó
1st	plur	caminamos	caminábamos	caminaremos	caminamos
2nd	plur	caminan	caminaban	caminarán	caminaron
3rd	plur	caminan	caminaban	caminarán	caminaron
		Conjugation II (correr - to run)			
1st	sing	corro	corría	correré	corrí
2nd	sing	corres	corrías	correrás	corriste
3rd	sing	corre	corría	correrá	corrió
1st	plur	corremos	corríamos	correremos	corrimos
2nd	plur	corren	corrían	correrán	corrieron
3rd	plur	corren	corrían	correrán	corrieron
		Conjugation III (dormir - to sleep)			
1st	sing	duermo	dormía	dormiré	dormí
2nd	sing	duermes	dormías	dormirás	dormiste
3rd	sing	duerme	dormía	dormirá	durmió
1st	plur	dormimos	dormíamos	dormiremos	dormimos

table continues

Person/Number		Present	Imperfect	Future	Preterite
2nd	plur	duermen	dormían	dormirán	durmieron
3rd	plur	duermen	dormían	dormirán	durmieron

Note. 2nd person plural can express 3rd person forms; 3rd person singular forms can express 2nd person forms; 1st, 2nd, 3rd persons singular forms and 1st person plural forms of *dormir* denote irregular verbs across all verb tenses; 1st person imperfect is expressed by 3rd person imperfect; 1st person plural present tense is expressed by 1st person plural preterite tense.

Table 3

Bivariate Correlations Between Dependent and Independent Variables (N = 1840)

	Variables	1	2	3	4	5	6	7	8	9	10	11	12	13
	Repetition (n=920)													
1	Onset RT	--												
2	Offset RT	-0.26	--											
3	Verb Class	0.16	-0.11	--										
4	Tense	0.032	-0.22	0.14	--									
5	Syllable	0.53	-0.67	0.19	0.11	--								
6	Stress	-0.36	0.45	-0.20	-0.05	-0.75	--							
7	Canonical	0.06	-0.11	-0.05	0.14	-0.03	0.43	--						
8	Regularity	-0.16	0.27	-0.29	-0.08	-0.38	0.37	0.06	--					
9	Phoneme	-0.03	-0.07	-0.04	0.00	0.18	-0.16	0.00	-0.12	--				
10	Voicing	0.04	0.12	0.10	0.00	0.04	-0.03	-0.03	-0.08	0.60	--			
11	Duration	0.74	-0.84	0.17	0.17	0.76	-0.51	0.11	-0.27	0.03	-0.05	--		
12	Frequency	-0.35	0.58	-0.26	-0.05	-0.53	0.60	0.22	0.36	0.00	0.03	-0.60	--	
13	Alternative Pronoun	-0.03	0.27	0.02	-0.14	-0.04	0.02	0.14	-0.05	0.00	0.00	-0.20	0.31	--
	Production (n=920)													
1	Onset RT	--												
2	Offset RT	0.60	--											
3	Verb Class	0.12	-0.04	--										

table continues

	Variables	1	2	3	4	5	6	7	8	9	10	11	12	13
4	Tense	-0.12	-0.25	0.14	--									
5	Syllable	0.28	-0.40	0.19	0.11	--								
6	Stress	-0.31	0.17	-0.20	-0.05	-0.75	--							
7	Canonical	-0.07	-0.14	-0.05	0.14	-0.03	0.43	--						
8	Regularity	-0.18	0.08	-0.29	-0.08	-0.38	0.37	0.06	--					
9	Phoneme	0.07	0.03	-0.04	0.00	0.18	-0.16	0.00	-0.12	--				
10	Voicing	0.06	0.09	0.10	0.00	0.04	-0.03	-0.03	-0.08	0.60	--			
11	Duration	0.28	-0.60	0.17	0.17	0.76	-0.51	0.11	-0.27	0.03	-0.05	--		
12	Frequency	-0.09	0.42	-0.26	-0.05	-0.53	0.60	0.22	0.36	0.00	0.03	-0.60		
13	Alternative Pronoun	0.34	0.46	0.02	-0.14	-0.04	0.02	0.14	-0.05	0.00	0.00	-0.21	0.31	--

Table 4

Mean Differences in Onset Response Time Across Grammatical, Phonological, and Morphological Characteristics (N = 1840)

Variable/Task	n	M	SD	Partial Eta^2
Class				
Repetition				
ER	304	869.26***	72.81	0.029
IR	129	893.36[a]	65.13	
AR	487	893.53	62.36	
Pronoun				
ER	304	1222.99**	122.35	0.015
IR	129	1237.54[b]	119.35	
AR	487	1255.22	116.18	
Tense[c]				
Repetition				
Present	200	865.69	64.94	0.022
Past	170	876.90	69.83	
Future	200	889.50	66.26	
Pronoun	Pronoun			
Present	200	1193.72	133.14	0.036
Past	170	1213.00	107.44	

table continues

Variable/Task	n	M	SD	Partial Eta^2
Future	200	1246.64	103.45	
Number of syllables				
Repetition				
One or two	308	844.67***	56.94	0.285
Three	434	889.40	55.83	
Four or five	178	946.58	59.62	
Pronoun				
One or two	308	1190.32***	106.63	0.096
Three	434	1264.70[d]	117.33	
Four or five	178	1276.57	116.60	
Syllable stress position				
Repetition				
First syllable stress	206	848.16	56.30	
Second syllable stress	437	884.28	63.48	
Fourth or third syllable stress	277	915.15***	66.49	0.128
Pronoun				
First syllable stress	206	1184.23	108.23	
Second syllable stress	437	1239.74	120.17	
Fourth or third syllable stress	277	1288.84***	106.35	0.099

table continues

Variable/Task	n	M	SD	Partial Eta^2
Canonical stress				
Repetition				0.003
Canonical	574	888.53	64.95	
Noncanonical	346	880.43	70.75	
Pronoun				
Canonical	574	1236.09*	121.73	0.004
Noncanonical	346	1252.05	115.11	
Verb regularity				
Repetition				
Regular	773	890.06***	65.98	0.024
Irregular	147	861.45	69.09	
Pronoun				
Regular	773	1251.31***	118.75	0.031
Irregular	147	1193.64	111.62	
Initial phoneme				
Repetition				
Fricative	168	926.05***	64.92	0.120
Stop	486	867.30	57.47	
Liquid	74	902.53[e]	62.95	
Nasal	55	867.67[f]	47.90	

table continues

Variable/Task	n	M	SD	Partial Eta^2
Vowel	137	898.19	83.21	
Pronoun				
Fricative	168	1266.43[g]***	103..72	0.040
Stop	486	1224.47	115.60	
Liquid	74	1253.47[e]	112.00	
Nasal	55	1210.91[f]	95.97	
Vowel	137	1281.13	146.13	
Voicing				
Repetition				
Voiced	429	888.44	65.84	0.002
Voiceless	491	882.11	68.79	
Pronoun				
Voiced	491	1248.53	128.75	0.003
Voiceless	429	1234.72	107.54	
Alternative pronoun				
Pronoun				
No alternative	520	1209.51***	109.34	0.131
Él/ella/usted	150	1267.76[h]	129.52	
Ellos/Ustedes	200	1274.97	99.00	
Yo/él/ella	50	1372.45***	123.26	

table continues

Note. [a]Verbs that end with *-ir* are not significantly different from *-ar* verbs. [b]Verbs that end with *-ir* are not significantly different from *-er* and *-ar* verbs. [c]Due to repeating lexical forms in the singular imperative tense and across all plural forms these values are limited to the singular present, past, and tense forms. [d]Three syllable verbs are not significantly different from the four or five syllable group. [e]Liquids are not significantly different from vowels. [e]Nasals are not significantly different from stops. [f]Fricatives are not significantly different from liquids or vowels. [h]Él/ella/usted are not significantly different from ustedes/ellos.

*$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

Table 5

Mean Differences in Response Time Measured from Stimulus Offset for Number of Syllables an Stress Position (N = 1840)

Variable/Task	n	M	SD	Partial Eta^2
Repetition				
Number of syllables				
One or two	308	330.39***	62.85	0.447
Three	434	245.23	60.46	
Four or five	178	178.07	62.47	
Syllable stress position				
First syllable stress	206	338.31	68.49	
Second syllable stress	437	244.18	72.51	
Third or fourth syllable stress	277	229.20***	71.53	0.259
Pronoun				
Number of syllables				
One or two	308	676.04***	111.19	0.168
Three	434	620.76	139.40	
Four or five	178	508.06	142.28	
Syllable stress				
First syllable stress	206	674.38	126.47	

table continues

Variable/Task	n	M	SD	Partial Eta^2
Second syllable stress	437	599.87[a]	146.03	
Third or fourth syllable stress	277	602.89***	141.76	0.045

Note. [a]Second syllable stress position is not significantly different from third or fourth stress position.

*p < 0.05. **p < 0.01. ***p < 0.001.

Table 6

Summary of Standard Regression Analysis for Onset Response Time Across Repetition and Pronoun Tasks (N = 920).

Predictor by Task	B	SE B	Beta	Uniqueness (sr^2) (%)
Onset Repetition				
Stimulus duration	0.49	0.02	0.87***	21.5
Number of syllables	8.32	4.07	0.09*	0.2
Syllable stress position				
First syllable	reference			
Second syllable	11.85	4.41	0.09*	0.3
Third or fourth syllable	12.22	6.53	0.08	
Word frequency	-5.23	0.91	-0.18***	1.3
First phoneme sound				
Fricative	29.44	3.93	0.17***	2.3
Liquid	18.83	5.26	0.08***	0.5
Nasal	5.87	6.04	0.02	
Vowel	6.77	4.25	0.04	
Stop	reference			
Onset Pronoun				
Alternative pronoun				

table continues

Predictor by Task	B	SE B	Beta	Uniqueness (sr^2) (%)
No alternatives	reference			
Ustedes/ellos	82.31	8.29	0.28***	7.2
Usted/él/ella	100.76	10.03	0.31***	7.4
Yo/él/ella	201.86	15.45	0.38***	12.5
Syllable stress position				
First syllable	reference			
Second syllable	23.11	10.19	0.10*	0.4
Third or fourth syllable	76.95	13.98	0.30***	2.2
Number of syllables	35.51	9.80	0.21***	1.0
Stimulus duration	0.46	0.05	0.46***	6.4
First phoneme sound				
Fricative	12.20	9.27	0.04	0.1
Liquid	6.14	12.34	0.01	0.02
Nasal	-29.02	14.04	-0.06*	0.3
Vowel	34.34	9.72	0.11***	0.9
Stop	reference			

Note. Repetition: R^2 = 62.1% (unique = 24.5%; shared = 37.6%). Pronoun: R^2 = 32.5% (unique = 38.0%; shared = 00%).

*$p < 0.05$ **$p < 0.01$ ***$p < 0.001$