Completion norms for 112 Spanish sentences

SCOTT A. McDONALD and MONICA TAMARIZ
University of Edinburgh, Edinburgh, Scotland

Sentence completion norms are a valuable resource for researchers interested in studying the effects of context on word recognition processes. Norms for 112 Spanish sentences were compiled with the use of experimental software accessed over the World-Wide Web. Several measures summarizing the distribution of responses for each sentence are reported, including Schwanenflugel's (1986) multiple-production measure of sentence constraint strength, the type–token ratio, and the information-theoretic measure of redundancy. The complete set of completion norms is available at http://www.ling.ed.ac.uk/~monica/spanish_completion_norms.html.

Investigations of the interaction between sentence context and word recognition processes have benefited from the availability of sentence completion norms. Published normative data for sentences typically consist of the most frequent words produced as a completion for each sentence with the use of the Cloze procedure (Taylor, 1953) or a closely related method (Schwanenflugel, 1986). The distribution of responses—the set of completions generated for each sentence frame, together with their production probabilities—allows a number of useful measures of contextual constraint to be computed in order to quantify the degree to which the preceding sentence context constrains its final word. Such measures position sentence contexts along a scale ranging from highly constraining (in which perhaps only a single completion word is possible) to completely unconstraining (in which the set of possible completions is extremely large). Sets of experimental materials can then be selected from the normative data and matched on such characteristics as constraint strength. Normative sentence completion data are currently available for English (e.g., Bloom & Fischler, 1980; Schwanenflugel, 1986) and French (Robichon, Besson, & Faita, 1996) only; no comparable norms exist for Spanish. The purpose of the present paper was to make a large set of sentence completion norms available to researchers studying language processing in Spanish speakers.

Completion norms are a valuable resource for diverse areas of research in experimental psycholinguistics. For example, word recognition processes have been extensively investigated with the normative completion data. The basic finding has been facilitation of a sentence-final word when it is highly constrained by the preceding context, compared with presentation in a less constraining, or "neutral," sentence context (e.g., Schwanenflugel & Shoben, 1985). Normed completion data have also proved useful for study-

ing language production; for instance, the degree to which a sentence context constrains its final word has been shown to interact with word frequency in the task of naming a picture following a sentence context (Griffin & Bock, 1998).

In addition to sentence constraint strength, the predictability of the sentence-final word has been a variable of interest. Kleiman (1980) found no benefit of a constraining context on lexical decision time for unexpected (low cloze probability) completion words, though Stanovich and West (1983) reported a context effect on naming latency for unexpected completions. Schwanenflugel and LaCount (1988) demonstrated an interaction between sentence constraint strength and predictability, arguing that high-constraint sentence contexts permit the facilitation of a smaller set of completion words than do low-constraint contexts.

Neuropsychological research employs sentence completion data in order to delimit the areas of the brain involved in various aspects of sentence processing. For example, in a study using *event-related brain potentials* (ERPs), Federmeier and Kutas (1999) showed that an unexpected sentence-final word elicited a larger N400 component (a component of the ERP waveform associated with semantic processing) than did an expected final word when it was presented to either the left or right visual field. However, when the unexpected completion word was a member of the expected word's semantic category, lateralization differences were observed, with smaller effects for the right visual field (i.e., left hemisphere).

Finally, normative sentence completion data have been used to study language processing in impaired populations. Berndt, Mitchum, Haendiges, and Sandson (1997) reported differential processing of sentence-final nouns and verbs by aphasic patients. Nebes and Brady (1991) conducted a sentence comprehension study with Alzheimer's patients and found that the time to judge whether a word could sensibly complete a sentence fragment varied with the amount of constraint strength.

Although completion norms have proven useful for controlling contextual constraint in a number of language pro-

cessing studies, less attention has been devoted to the clarification of the properties of the preceding sentence context that influence the set of responses produced. Both semantic and syntactic properties of the context words have been demonstrated to affect the corresponding properties of subsequently produced items. For instance, Kohn and Cragnolino (1998) reported that when subjects were asked to devise a sentence with a given verb, they often supplied a word that was semantically or associatively related to the verb as its direct object. There is also a growing literature on the related phenomenon of syntactic priming: Subjects are more likely to produce a syntactic structure identical to the one that they have just processed than they are to produce viable alternative structures (e.g., Bock, 1986; Pickering & Branigan, 1999). For example, Cleland and Pickering (2001) demonstrated that the priming of noun phrase structures was enhanced when the prime and target contained semantically related nouns. Thus, both semantic and syntactic characteristics of the previous context can affect the choice of subsequent syntactic categories or structures.

In the present study, we compiled a set of 112 Spanish sentences that intuitively varied in the degree to which they constrained their final word and elicited responses from native Spanish speakers using Schwanenflugel's (1986) multiple-production paradigm. Unlike the standard cloze procedure, for which only one response is requested from each subject, subjects supply up to three completion words for each sentence by using the multiple-production method. This method overcomes the Cloze procedure's implicit assumption that a single completion is made prominent when in fact more than one may readily come to mind.

We used an interactive software package specifically designed for administering psychological experiments over the World-Wide Web. This package, WebExp 2.1 (Keller, Corley, Corley, Konieczny, & Todirascu, 1998), uses a program written in Java to display stimuli and record responses. The primary advantage in using the Web for collecting response data of this kind is that subjects can participate without physically being in the same location. Demographic data (e.g., name, age, and language background) are easily collected by having the subjects fill in a short questionnaire. WebExp ensures the authenticity of the subjects taking part by using several methods. For instance, subjects must supply their e-mail address. WebExp automatically sends a "thank-you" e-mail to this address at the end of the experiment; bounced e-mails indicate an invalid address.[1] The IP (Internet Protocol) address of the computer used to connect to the Internet is logged in order to prevent the submission of more than one set of response data from the same subject.

Evidence that response data collected via the Web is as valid as data collected with traditional methods (e.g., questionnaires filled out in the laboratory or classroom) comes from Corley and Scheepers (2000) and Keller and Alexopoulou (2001). (See Corley, Keller, & Scheepers, 2000, for a critical overview of the validity issue.) Using WebExp, Corley and Scheepers (2000) replicated syntactic (grammatical structure) priming effects originally obtained with paper questionnaire-based methods. Keller and Alexopoulou found that sentence acceptability judgments collected over the Web corroborated with those elicited in a laboratory setting. Further evidence for the validity of the Web-based procedure is supplied by the present study, since the results are consistent with norming studies in which English materials and conventional methodology have been used.

METHOD

Subjects

A call for volunteers was posted to several Spanish language Usenet groups and was distributed by colleagues in Spain. Some of the subjects were undergraduate students at Madrid Complutense University and obtained course credit for taking part. Forty-five subjects participated in the study. The data for 1 subject were removed because he did not indicate Spanish as his first language, and 2 subjects failed to respond to any items after proceeding beyond the instruction phase. The data from 42 subjects remained for analysis. Ages ranged from 17 to 41 ($M = 26$). Eight (19%) of the subjects indicated that they were speakers of Latin American Spanish (3 from Mexico, 2 from Colombia, and 1 each from Chile, Peru, and Venezuela).[2]

Materials

One hundred and twelve sentences were constructed in order to represent a broad range of sentence constraint. Sentence fragments ranged in length from 4 to 15 words ($M = 8.3$, $SD = 2.0$). The final word of 88 of the fragments was either a determiner (e.g., *el*, *la*, *un*) or an adjective (e.g., *último*), which constrained the grammatical gender of the supplied completion word. All sentences were similar in syntactic structure, containing on average one and a half clauses ($SD = 0.5$, range: 1–3). Fifty-four percent of the final clauses (i.e., the clause containing the missing item) were headed by transitive verbs, 25% were headed by instransitive verbs, and 21% by copulative verbs (e.g., *estar*, *ser*). The number and percentage of content words per sentence fragment ranged from 2 (29%) to 7 (71%), with an mean of 4.1 (50%) and standard deviation of 1.2 (9%).

In order to make the task manageable, the materials were divided into two versions of 56 items each; each subject saw only one version. All stimuli are listed in Appendix B. The following is an example sentence: *El médico le aconsejó mudarse al ___.* "*The doctor advised him to move to the ___.*"

Procedure

The data were collected over the Internet. The subjects accessed the experiment's Web page with their Web browser, which established a connection to the experimental server running WebExp 2.1 (Keller et al., 1998). The subjects were assigned to one of two versions; the version number alternated every time the experimental server was accessed. The subjects were first presented with a page of instructions, which were modified from those used by Schwanenflugel (1986). An example sentence fragment and three possible completion words were included in the instructions page (see Appendix A). The subjects were instructed to supply three possible completion words for each item and to do the task rapidly, entering the first words that came to mind.

Next, the subjects were required to fill out a demographic questionnaire, providing personal details such as name, age, sex, and the region where they grew up. The experimental phase consisted of the sequential presentation of 56 incomplete sentences. Order of presentation was randomized individually for each subject. Responses were entered via the terminal keyboard into a blank text window located immediately below the displayed sentence fragment.

RESULTS AND DISCUSSION

Three issues were encountered with the response data that needed to be resolved before further analysis: (1) Spelling errors were corrected, and missing diacritics were replaced (not all Web browser/platform configurations support the Spanish character set). (2) If a subject had provided more than three completions for a fragment, only the first three were retained. (3) A small proportion of responses (2.67%) consisted of two or more words. These multiword responses were classified as either compound words or phrases by the second author, a professional lexicographer. Only compounds (e.g., *guarda forestal*, "forestry officer") were deemed to be valid and thus were retained for further analysis. Forty percent of the multiword responses were classified as compounds.

Measures of Constraint

As is detailed in Schwanenflugel (1986), the production probability of each response to its associated sentence frame was computed as the number of subjects who provided that response divided by the total number of subjects who contributed at least one valid response for that item. Following Schwanenflugel, sentence constraint strength was defined as the probability of the most frequently produced completion word. Constraint strength (CS) ranged from 0.125 to 1.00 ($M = 0.493$, $SD = 0.245$). Consistent with Schanenflugel's results for English materials, the more strongly constraining the sentence frame, the smaller the set of unique completion words produced (Pearson $r = -.786$, $p < .001$).

Before continuing with our analysis of constraint strength, we briefly summarize several descriptive properties of the dataset produced. Overall, 6 (0.14%) responses were grammatically incorrect, and 5 (0.12%) were semantically implausible completions. Although the sentence frames were written with the aim of eliciting nouns as completion words, the most frequent completions for two items were adjectives (*sola* for item 69 and *complicada* for item 111). Considering the entire set of responses produced, 41 (0.95%) were adjectives, 10 (0.23%) were verbs, and 7 (0.16%) were pronouns or numerals.

We computed several other measures from the distribution of responses besides constraint strength. First, we computed two measures that are closely related to CS, but which additionally summarize the shape of the response distribution: the type–token ratio (TTR), and the estimated redundancy of the sentence context. Although the type–token ratio (for a given sentence, the number of different words, or *types*, divided by the total number of completions, or *tokens*, generated) also reflects the constraint strength of the sentence context, it is sensitive to the variety of completion words supplied. For example, if the sentence fragment was so strongly constraining that only one final word was possible, CS would be 1.0 and TTR would be very small ($1/n$ subjects). If more than one word was possible, yet every subject's set of responses had a particular word in common, CS would still be 1.0, but TTR would be larger than in the previous case.

Following Tulving and Gold (1963), we assume that the redundancy of a sentence context can be estimated from the production probabilities of the words elicited to complete that sentence. Redundancy is a measure of the amount of information that the sentence context provides about its final word; it is 1 if the final word is completely predictable from the preceding context, and it is 0 if each member of the set of generated completion words was produced exactly once. Like TTR, redundancy captures differences in the shape of the response distribution between sentences with the same CS value, because it takes the production probabilities of all responses into account. Redundancy (R) is defined by subtracting the entropy of (or alternatively, the amount of uncertainty in) the response distribution (H) over the maximum possible entropy of that distribution (H_{max}) from 1:

$$R = 1 - \frac{H}{H_{max}}$$

$$H = \sum_{i=1}^{n} p_i \frac{1}{\log_2 p_i}$$

$$H_{max} = \log_2 n,$$

where n is the number of distinct responses, or types, and p_i is the production probability of response$_i$.

Table 1 displays the correlations between response measures. The type–token ratio ($M = .543$, $SD = .165$; range, .222–.944) was significantly negatively correlated with sentence constraint strength ($r = -.740$, $p < .001$). The more strongly a sentence constrained its final word, the narrower the range of completion words supplied. Redundancy ($M = .122$, $SD = .080$; range, .006–.332) was also highly correlated with CS ($r = .922$, $p < .001$), indicating that the more information that was contained in the preceding context about the sentence-final word (or the more predictable the final word was from the context), the larger the proportion of subjects providing the same response. Tulving and Gold (1963) reported a lower correlation of .439 between the cloze probability of the final word and the redundancy of the sentence context; the discrepancy between their correlation coefficient and ours is likely due the difference between the standard Cloze procedure and the multiple-production paradigm. As pointed out by Schwanenflugel (1986), the multiple-production procedure tends to shift CS values upward compared with the Cloze procedure.

The fourth measure we computed for each item was the mean number of responses per subject (MR). MR should

Table 1
Correlation Matrix for Five Measures Computed
from the Distribution of Responses

Variable	1	2	3	4	5
CS	1.00				
TTR	−0.740**	1.00			
R	0.922**	−0.827**	1.00		
MR	−0.104	0.360**	−0.356**	1.00	
*P*1	0.672**	−0.547**	0.662**	−0.272*	1.00

Note—CS, sentence constraint strength; TTR, type–token ratio; *R*, redundancy; MR, mean number of responses; *P*1, probability of the most expected completion being the first response.   *$p < .01$.   **$p < .001$.

also capture aspects of the degree to which a sentence constrains its final word; the more constraining the context, the fewer possibilities exist that fit the context, and so the harder it should be to think of potential completion words. MR ranged from 1.696 to 2.467 ($M = 2.074, SD = .200$); over the entire dataset, only one completion word was provided 46% of the time, compared with 4% for two and 50% for three completions. Interestingly, there was no reliable correlation between CS and MR ($r = -.104, p = .274$). However, both redundancy and the type–token ratio were significantly correlated with MR ($r = -.356, p < .001$, and $r = .360, p < .001$, respectively), indicating that these two measures may capture the difficulty involved in producing alternative responses better than CS does.

Finally, we computed the probability that the most expected completion word was the first word produced in the set of (up to three) responses provided by each subject. This measure ($P1$) intuitively captures how quickly the response comes to mind. The multiple production method does not distinguish the case in which the most expected completion word is typically given as a primary response from the case in which the word is produced as a subject's second or third choice. Of most interest here is the correlation between $P1$ and CS. Is it the case that the stronger the sentence constraint, the more likely that the most expected response will be produced before any other? There was a highly significant correlation of $r = .672$ ($p < .001$) between the two measures, indicating that the more constraining the context, the greater the likelihood that the most expected completion was produced as a first response.

Correlates of Contextual Constraint

What properties of the preceding sentence context are responsible for the varying degrees of constraint observed? One view considers constraint to be a reflection of expectations generated by the context about the meaningful properties of upcoming words. In this view, context is assumed to activate the set of semantic features most consistent with the sentence thus far (e.g., Federmeier & Kutas, 1999), or alternatively, to determine the number and makeup of semantic feature restrictions on upcoming words (see, e.g., Schwanenflugel & LaCount, 1988). On this account, contextual constraint reflects some combination of world knowledge retrieved from long-term memory associated with the words in the context and the cognitive model of the discourse (or situation) currently being built.

Although we cannot directly measure the impact of world knowledge and discourse factors on the degree of constraint, there are several observable characteristics of the sentence fragments that could conceivably predict sentence constraint strength. For instance, constraint generally increases with sentence length (Van Petten & Kutas, 1990). Other potential predictors of constraint are the number of content words in the fragment and the proportion of these words that are strong lexical associates—roughly, words that tend to co-occur more often than expected by chance—of the most expected completion. Lexical associates are often strong cues to the meaning of an upcoming word, and their role in explanations of constraint-related phenomenon such as sentence priming is controversial (e.g., Moss & Marslen-Wilson, 1993; Simpson, Peterson, Casteel, & Burgess, 1989).

Table 2 displays the correlations between the four response measures and sentence length and the number and percentage of content words in the sentence fragment. A low, though significant, correlation was found between CS and the number of words in the sentence ($r = .203, p < .05$). There was also a reliable correlation between CS and the number of content words in the fragment when the class of content words was defined to include nouns, verbs, adjectives, and adverbs ($r = .206, p < .05$), and when only the noun and verb content words were considered ($r = .225, p < .05$). The comparable correlation coefficients for these two definitions of *content word* indicate that the adjectival and adverbial context words did not contribute much to the degree that our stimuli sentences constrained their final words. The lack of a corresponding correlation between CS and the percentage of content words (a measure that controls for sentence length) suggests, however, that it is simply the length of the sentence fragment that is important. TTR was significantly correlated with both sentence length ($r = -.275, p < .01$), the number of noun/verb content words ($r = -.246, p < .05$), and the number of content words ($r = -.241, p < .05$). Again, the longer the sentence fragment, the more constrained the distribution of responses (as indexed by the type–token ratio). Redundancy was the only other constraint measure to reliably correlate with any of the stimulus variables; the correlation coefficient between R and the number of noun/verb content words was $.230$ ($p < .05$).

We can only provide a rough analysis of the relationship between the number of words in the sentence lexically associated to the most expected completion word and CS. The prediction is that constraining contexts will contain more lexical associates of the completion than will less constraining contexts. As a crude test of this prediction, we looked up the most expected completion word for each item in the Clave (1997) Spanish dictionary and calculated the overlap (the number of shared content words) between its corresponding sentence frame and its definition in the Clave dictionary. Sentence constraint strength was significantly correlated with the number of lexical associates in

Table 2
Correlations Between Four Stimulus Characteristics and Four Measures Computed from the Distribution of Responses

Stimulus Variable	Response Measure			
	CS	TTR	R	MR
Sentence length	0.203*	−0.275**	0.175	0.088
Number of nouns/verbs	0.225*	−0.246*	0.230*	−0.019
Number of content words	0.206*	−0.241*	0.181	−0.003
Percent content words	.059	−.028	.057	−.150

Note—CS, sentence constraint strength; TTR, type–token ratio; R, redundancy; MR, mean number of responses.   *$p < .05$.   **$p < .01$.

the sentence ($r = .354$, $p < .001$), indicating that the presence of such words might cue the production of the same completion word across subjects.

In summary, in the present paper, we analyzed normative sentence completion data elicited from native Spanish speakers over the World-Wide Web. These norms (presented in truncated form in Appendix B) consist of the set of responses generated for each sentence, in descending order of the production probability of each response.[3] We reported four summary measures of the response distribution for each sentence. Following Schwanenflugel (1986), sentence constraint strength was defined as the probability of the most frequently produced response. We also investigated three other measures that can be easily calculated from the distribution of completion words. Redundancy and the type–token ratio were highly correlated with constraint strength, yet we additionally took the shape of the response distribution into account. The mean number of responses measure appears to capture something different than constraint strength. The availability of these four measures should be of interest to researchers using these norms to study the effects of contextual constraint in Spanish speakers.

REFERENCES

BERNDT, R. S., MITCHUM, C. C., HAENDIGES, A. N., & SANDSON, J. (1997). Verb retrieval in aphasia: Characterizing single word impairments. *Brain & Language*, **56**, 68-106.

BLOOM, P. A., & FISCHLER, I. (1980). Completion norms for 329 sentence contexts. *Memory & Cognition*, **8**, 631-642.

BOCK, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, **18**, 355-387.

Clave: Diccionario de uso del español actual (1997). Madrid: Ediciones SM.

CLELAND, A., & PICKERING, M. (2001, September). Priming of noun phrase structure. *Seventh Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP-2001)*. Saarbrücken, Germany.

CORLEY, M., KELLER, F., & SCHEEPERS, C. (2000). *Psycholinguistic research via the World Wide Web: A methodological evaluation.* Unpublished manuscript, University of Edinburgh and University of Glasgow.

CORLEY, M., & SCHEEPERS, C. (2000). *Syntactic priming in English: Evidence from response latencies.* Manuscript submitted for publication.

FEDERMEIER, K. D., & KUTAS, M. R. (1999). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research*, **8**, 373-392.

GRIFFIN, Z. M., & BOCK, K. (1998). Constraint, word frequency and the relationship between lexical processing levels in spoken word production. *Journal of Memory & Language*, **38**, 313-338.

KELLER, F., & ALEXOPOULOU, T. (2001). Phonology competes with syntax: Experimental evidence for the interaction of word order and accent placement in the realization of information structure. *Cognition*, **79**, 301-372.

KELLER, F., CORLEY, M., CORLEY, S., KONIECZNY, L., & TODIRASCU, A. (1998). *WebExp: A Java toolbox for Web-based psychological experiments* (Tech. Rep. No. HCRC/TR-99). Edinburgh, U.K.: University of Edinburgh, Human Communication Research Centre.

KLEIMAN, G. M. (1980). Sentence frame contexts and lexical decision: Sentence-acceptability and word-relatedness effects. *Memory & Cognition*, **8**, 336-344.

KOHN, S. E., & CRAGNOLINO, A. (1998). The role of lexical co-occurrence in aphasic sentence production. *Applied Psycholinguistics*, **19**, 631-646.

MOSS, H. E., & MARSLEN-WILSON, W. D. (1993). Access to word meanings during spoken language comprehension: Effects of sentential semantic context. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 1254-1276.

NEBES, R. D., & BRADY, C. B. (1991). The effect of contextual constraint on semantic judgments by Alzheimer patients. *Cortex*, **27**, 237-246.

PICKERING, M., & BRANIGAN, H. (1999). Syntactic priming in language production. *Trends in Cognitive Sciences*, **3**, 136-141.

ROBICHON, F., BESSON, M., & FAITA, F. (1996). Completion norms for 744 French linguistic contexts of differing formats. *Canadian Journal of Experimental Psychology*, **50**, 205-233.

SCHWANENFLUGEL, P. J. (1986). Completion norms for final words of sentences using a multiple production measure. *Behavior Research Methods, Instruments, & Computers*, **18**, 363-371.

SCHWANENFLUGEL, P. J., & LACOUNT, K. L. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 344-354.

SCHWANENFLUGEL, P. J., & SHOBEN, E. J. (1985). The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory & Language*, **24**, 232-252.

SIMPSON, G. B., PETERSON, R. R., CASTEEL, M. A., & BURGESS, C. (1989). Lexical and sentence context effects in word recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 88-97.

STANOVICH, K. E., & WEST, R. F. (1983). On priming by a sentence context. *Journal of Experimental Psychology: General*, **112**, 1-36.

TAYLOR, W. L. (1953). "Cloze Procedure": A new tool for measuring readability. *Journalism Quarterly*, **30**, 415-433.

TULVING, E., & GOLD, C. (1963). Stimulus information and contextual information as determinants of tachistoscopic recognition of words. *Journal of Experimental Psychology*, **66**, 319-327.

VAN PETTEN, C., & KUTAS, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition*, **18**, 380-393.

NOTES

1. In the present study, three *thank you* messages were bounced owing to problems with the e-mail address supplied.

2. No systematic differences were observed between the responses supplied by the Latin American subjects and those elicited from the Spanish subjects. None of the most expected sentence completion words were particular to the vocabulary of Latin American Spanish. Note that the Web-Exp methodology is ideal for investigating hypotheses about dialectal differences, since speakers of the dialects of interest can easily be reached via the Internet.

3. The full set of response data is available for downloading from the World-Wide Web at http://www.ling.ed.ac.uk/~monica/spanish_completion_norms.html or can be obtained in electronic form from the authors.

APPENDIX A
Sentence Completion Instructions

Este experimento es un estudio de cómo hablantes nativos de castellano usan su lengua. Tendrá que completar varias frases; es una tarea sencilla y durará aproximadamente 20 minutos.

En el experimento verá varios fragmentos de frases; una línea como _____ representa una palabra que falta. Tendrá que pensar en tres palabras posibles que puedan terminar cada frase. Cada palabra debe completar la frase en una manera aceptable (incluya el artículo si es necesario).

APPENDIX A (Continued)

Por ejemplo, si el fragmento es:

Cuando hace mucho calor, nos gusta ir a la _____.

Tres terminaciones apropiadas podrían ser:

playa, piscina, terraza

Haga la tarea bastante rápido. Estamos interesados en las primeras palabras que se le ocurran y que terminen bien la frase. Dé por lo menos una respuesta (palabra o artículo + palabra) que complete la frase.

Se presentará cada fragmento en una ventana. Habrá una ventana vacía abajo donde debe escribir sus respuestas. Separe las diferentes respuestas con comas. Después de escribir sus respuestas, pulse «ENTER» y el próximo fragmento aparecerá. (Después de pulsar «ENTER», no será posible cambiar sus respuestas).

APPENDIX B

Sentences are ordered by decreasing constraint strength (CS). The production probability of the three most expected responses for each sentence immediately follows the response, and the final four values are the four measures calculated from the distribution of responses: CS, redundancy (R), type–token ratio (TTR), and mean number of responses (MR). (The complete dataset, which includes the entire response distribution, can be downloaded from http://www.ling.ed.ac.uk/~monica/spanish_completion_norms.html.)

Completion	Prob	Completion	Prob	Completion	Prob	CS	TTR	MR	R
1. Mientras comía pescado, se atragantó con una ___.									
espina	1.0	escama	.27	raspa	.13	1.0	0.406	2.133	0.245
2. Los parados opinan que deberían tener los mismos ___.									
derechos	1.0	trabajos	.19	sueldos	.19	1.0	0.3	1.905	0.298
3. El panadero sacó los pasteles del ___.									
horno	1.0	mostrador	.13	frigorífico	.13	1.0	0.5	2.133	0.239
4. Una de las modelos tropezó en la ___.									
pasarela	.94	escalera	.31	entrada	.19	0.938	0.364	2.063	0.244
5. Para abrir esta puerta necesitas una ___.									
llave	.93	tarjeta	.27	clave	.27	0.933	0.424	2.2	0.227
6. El cura de repente dejó de hablar y sufrió un colapso en mitad de la ___.									
misa	.93	homilía	.2	comunión	.13	0.933	0.455	2.2	0.206
7. La duración de su sentencia será determinada por el ___.									
juez	.93	jurado	.5	fiscal	.21	0.929	0.333	2.357	0.212
8. En la plaza mayor la gente suele dar de comer a las ___.									
palomas	.91	aves	.3	ratas	.09	0.913	0.268	1.783	0.297
9. Delante del cine había una larga ___.									
cola	.91	fila	.26	carretera	.09	0.913	0.333	1.696	0.332
10. Nuestros resultados estuvieron por encima de la ___.									
media	.91	mayoría	.14	clase	.14	0.909	0.368	1.727	0.299
11. Todos los muebles en la casa de mi abuela estaban cubiertos de ___.									
polvo	.87	sábanas	.33	telarañas	.13	0.867	0.5	2.267	0.2
12. Cuando nuestro hijo aprenda a decir la hora, le regalaremos un ___.									
reloj	.87	cronómetro	.27	cronógrafo	.13	0.867	0.515	2.2	0.197
13. Siempre se pone seria cuando hablamos de ese ___.									
tema	.86	asunto	.27	problema	.14	0.864	0.302	1.955	0.241
14. Después de cambiar la rueda nos dejamos el ___.									
gato	.86	tornillo	.14	coche	.14	0.864	0.452	1.909	0.247
15. Estaba jugando en los columpios y se cayó al ___.									
suelo	.86	charco	.19	barro	.14	0.857	0.41	1.857	0.254
16. En cuanto nos sentamos, el camarero nos trajo la ___.									
carta	.86	comida	.29	bebida	.21	0.857	0.419	2.214	0.182
17. No pudo encontrar una corbata que haga juego con el ___.									
traje	.81	pantalón	.38	vestido	.06	0.813	0.5	2.125	0.213
18. Mi tío el millonario guarda su revólver en el ___.									
cajón	.8	armario	.5	escritorio	.15	0.8	0.333	2.1	0.246

APPENDIX B (Continued)

Completion	Prob	Completion	Prob	Completion	Prob	CS	TTR	MR	R
19. La partida de caza salió bastante temprano en busca de la ___.									
presa	.8	zorra	.2	comida	.13	0.8	0.576	2.2	0.168
20. El bombero se mojó porque había un agujero en la ___.									
manguera	.8	pared	.27	acera	.2	0.8	0.387	2.067	0.174
21. No pudimos jugar al tenis porque había agua en la ___.									
pista	.78	cancha	.52	superficie	.09	0.783	0.275	1.739	0.329
22. Empezarán a filmar en cuanto el escritor complete el ___.									
guión	.77	libro	.18	trabajo	.09	0.773	0.415	1.864	0.211
23. Por fin el pintor completó el último ___.									
cuadro	.76	retrato	.14	boceto	.14	0.762	0.439	1.952	0.192
24. Mis padres viven en un edificio de seis ___.									
pisos	.73	plantas	.67	vecinos	.13	0.733	0.353	2.267	0.215
25. El médico le aconsejó mudarse al ___.									
campo	.73	mar	.4	sur	.27	0.733	0.3	2	0.184
26. Todo me parece bastante limpio, salvo la ___.									
cocina	.7	habitación	.13	piscina	.09	0.696	0.523	1.913	0.179
27. El único hombre que faltaba en la boda era su ___.									
padre	.7	hermano	.39	novio	.3	0.696	0.244	1.957	0.197
28. El observador de aves salió de casa sin los ___.									
prismáticos	.7	binoculares	.26	anteojos	.17	0.696	0.422	1.957	0.2
29. Estaba jugando al tenis y se rompió la ___.									
raqueta	.69	pelota	.31	red	.25	0.688	0.333	2.25	0.144
30. A mi madre le sorprendió mucho recibir la ___.									
carta	.67	noticia	.6	postal	.13	0.667	0.457	2.333	0.187
31. Un gran grupo de niños vino para ver al ___.									
payaso	.65	mago	.13	circo	.13	0.652	0.5	1.913	0.164
32. La maestra se puso delante de la ___.									
pizarra	.65	mesa	.3	clase	.26	0.652	0.222	1.957	0.14
33. Es divertido comer comida china con unos ___.									
palillos	.65	amigos	.57	chinos	.17	0.652	0.293	1.783	0.286
34. Por causa de la niebla, no pudimos encontrar su ___.									
casa	.61	coche	.35	calle	.17	0.609	0.4	1.957	0.18
35. Por la mañana va a la ciudad a comprar la ___.									
comida	.6	leche	.4	prensa	.33	0.6	0.406	2.133	0.159
36. La adivina predijo la llegada de una plaga de ___.									
langostas	.6	mosquitos	.27	hormigas	.2	0.6	0.457	2.333	0.098
37. Con esa armadura y montado en un caballo parece un ___.									
caballero	.59	rey	.18	guerrero	.18	0.591	0.409	2	0.135
38. Se molestó en venir hasta aquí para darnos la ___.									
noticia	.57	carta	.5	enhorabuena	.29	0.571	0.406	2.286	0.153
39. Para que le sirva, esta blusa necesita un ___.									
arreglo	.57	botón	.33	lavado	.24	0.571	0.366	1.952	0.167
40. Cuando nació nuestra hija, mi hermano se convirtió en un ___.									
tío	.57	celoso	.21	padrino	.14	0.571	0.688	2.286	0.094
41. El perro dio la vuelta a la ___.									
manzana	.57	esquina	.39	casa	.26	0.565	0.364	1.913	0.193
42. La bruja sabe utilizar bien sus ___.									
poderes	.55	hechizos	.23	trucos	.18	0.545	0.444	2.045	0.131
43. A su esposa no le gusta descubrir pintalabios en el ___.									
cuello	.54	saco	.15	espejo	.15	0.538	0.655	2.231	0.081
44. Preferiríamos que devolviera el ___.									
dinero	.53	bolso	.13	uniforme	.07	0.533	0.765	2.267	0.08
45. Fue una decisión errónea poner al peor futbolista de ___.									
delantero	.53	portero	.47	defensa	.27	0.533	0.429	2.333	0.123

APPENDIX B (Continued)

Completion	Prob	Completion	Prob	Completion	Prob	CS	TTR	MR	*R*
46. Todos los meses pagamos el alquiler al ___.									
casero	.52	dueño	.3	propietario	.22	0.522	0.293	1.783	0.14
47. Los estudiantes tenían que escribir las respuestas en distintas ___.									
hojas	.5	casillas	.45	partes	.14	0.5	0.35	1.818	0.167
48. Cuando se estropeó el motor del barco, tuvimos que usar la ___.									
vela	.5	lancha	.33	radio	.17	0.5	0.422	1.875	0.162
49. Cuando paseaba por el parque se inventó una nueva ___.									
canción	.5	historia	.23	poesía	.14	0.5	0.561	1.864	0.131
50. Al cabo de cinco minutos, volvieron a la ___.									
casa	.5	carga	.25	reunión	.19	0.5	0.541	2.313	0.087
51. Después de mirar el eclipse, no notó ningún cambio en la ___.									
vista	.48	visión	.24	luna	.24	0.476	0.39	1.952	0.119
52. Les debemos siete mil pesetas de acuerdo con sus ___.									
cuentas	.47	cálculos	.33	honorarios	.2	0.467	0.6	2.333	0.093
53. Es más pequeña que una ___.									
hormiga	.47	pulga	.4	moneda	.2	0.467	0.529	2.267	0.104
54. El mes que viene, el gobierno tendrá que elegir un nuevo ___.									
presidente	.45	ministro	.36	representante	.18	0.455	0.452	1.909	0.141
55. No sabe nadie por qué hay sangre en el ___.									
suelo	.44	piso	.31	lavabo	.19	0.438	0.556	2.25	0.087
56. El alcalde del pueblo convocó a todos los ___.									
ciudadanos	.44	vecinos	.31	habitantes	.25	0.438	0.486	2.313	0.101
57. Habló con el mecánico que llevaba el ___.									
mono	.43	taller	.33	asunto	.14	0.429	0.5	1.81	0.13
58. Una semana antes de la sesión, el director me pidió un ___.									
favor	.41	informe	.27	consejo	.18	0.409	0.488	1.864	0.115
59. En primer lugar nos trajo la ___.									
comida	.41	carta	.18	sopa	.09	0.409	0.667	1.909	0.089
60. Por causa de la lluvia, las campistas pasaron toda la tarde en su ___.									
tienda	.4	casa	.27	refugio	.2	0.4	0.486	2.467	0.07
61. La lingüista pasó seis meses con los sabios de la isla estudiando su ___.									
idioma	.4	lengua	.33	lenguaje	.2	0.4	0.529	2.267	0.081
62. Ella fue la última persona que vio al nuevo ___.									
profesor	.4	inquilino	.2	alumno	.2	0.4	0.647	2.267	0.064
63. Después de cenar, me llevó mucho tiempo fregar la ___.									
vajilla	.4	cocina	.33	loza	.27	0.4	0.5	2.133	0.099
64. Aunque buscaron por todas partes, todavía faltaba un ___.									
niño	.4	rincón	.13	pasaje	.13	0.4	0.758	2.2	0.054
65. Al menos una docena de clientes estaba esperando para pagar en la ___.									
caja	.4	cola	.27	tienda	.2	0.4	0.5	2.133	0.076
66. Creo que esto se lleva en el ___.									
bolsillo	.38	bolso	.24	pelo	.19	0.381	0.523	2.095	0.086
67. Mi hermana no soporta la idea de comer ___.									
sola	.36	verdura	.09	tomate(s)	.09	0.364	0.682	2	0.063
68. El técnico dijo que tendría que reparar el ___.									
televisor	.36	ordenador	.28	vídeo	.12	0.36	0.489	1.88	0.107
69. El incendio del lunes pasado destruyó completamente la ___.									
casa	.35	oficina	.1	montaña	.1	0.35	0.7	2	0.057
70. Se sospecha que las ratas del casco antiguo son responsables de la ___.									
epidemia	.35	peste	.3	enfermedad	.26	0.348	0.386	1.913	0.116
71. Están a la espera de la ___.									
respuesta	.35	noticia	.3	carta	.13	0.348	0.614	1.913	0.102
72. Es lógico que esté dentro del ___.									
cajón	.33	grupo	.13	baúl	.13	0.333	0.788	2.2	0.04

APPENDIX B (Continued)

Completion	Prob	Completion	Prob	Completion	Prob	CS	TTR	MR	*R*
73. En la ladera de la montaña está la cabaña del ___.									
ermitaño	.33	pastor	.27	leñador	.27	0.333	0.636	2.2	0.07
74. Para cuidar el ganado en los grandes ranchos, necesitan ___.									
vaqueros	.32	perros	.32	gente	.14	0.318	0.581	1.955	0.096
75. No soporta las rutas con ___.									
curvas	.32	guía	.14	calor	.14	0.318	0.6	1.818	0.068
76. La prueba tendrá lugar mañana en el ___.									
aula	.32	salón	.18	auditorio	.18	0.318	0.595	1.909	0.073
77. Todo el mundo se congregó para mirar el ___.									
espectáculo	.31	eclipse	.25	acontecimiento	.19	0.313	0.639	2.25	0.056
78. Nuestro hijo acaba de hacer un dibujo de un ___.									
árbol	.3	coche	.3	perro	.26	0.304	0.489	2.043	0.109
79. Está claro que hay un problema con sus ___.									
padres	.3	amigos	.22	estudios	.13	0.304	0.667	1.957	0.071
80. No es difícil olvidarse del ___.									
nombre	.29	pasado	.21	reloj	.14	0.286	0.75	2.286	0.037
81. Mis amigas están hartas del ___.									
profesor	.29	trabajo	.21	novio	.14	0.286	0.742	2.214	0.039
82. Lo más importante es dar la bienvenida a su ___.									
invitado	.29	amigo	.21	tío	.14	0.286	0.719	2.286	0.038
83. Desde lejos me parece una larga ___.									
caminata	.29	cola	.21	avenida	.21	0.286	0.71	2.214	0.045
84. Resulta que la bolsa no contenía nada excepto ___.									
papeles	.27	dinero	.23	basura	.23	0.273	0.643	1.909	0.08
85. Nadie ha visto todavía su ___.									
obra	.27	casa	.27	cara	.18	0.273	0.605	1.955	0.082
86. Ésto es simplemente un asunto de ___.									
familia	.27	negocios	.2	honor	.13	0.267	0.743	2.333	0.035
87. Tocaron bien incluso sin la ___.									
flauta	.27	batería	.27	partitura	.2	0.267	0.625	2.133	0.056
88. Nos evitamos la molestia de recoger las ___.									
sobras	.27	maletas	.27	sillas	.13	0.267	0.697	2.2	0.046
89. No es posible volar de Madrid a Australia sin ___.									
escalas	.27	dinero	.27	billete	.27	0.267	0.559	2.267	0.057
90. Mil pesetas es demasiado caro para una caja de ___.									
cerillas	.27	bombones	.27	cigarrillos	.2	0.267	0.618	2.267	0.058
91. En el mundo moderno es difícil encontrar a un auténtico ___.									
amigo	.27	caballero	.2	profesional	.13	0.267	0.824	2.267	0.032
92. En el maletero sólo hay un trozo de una ___.									
rueda	.27	manta	.27	cuerda	.2	0.267	0.686	2.333	0.048
93. Creo que la mejor foto es ésta de las ___.									
niñas	.27	vacaciones	.2	montañas	.2	0.267	0.765	2.267	0.04
94. Todo el mundo sabe usar una ___.									
cuchara	.26	pluma	.17	computadora	.17	0.261	0.605	1.87	0.068
95. Me dijo que costaría ochocientas pesetas hacer un ___.									
regalo	.25	cambio	.25	marco	.17	0.25	0.759	2.417	0.031
96. Pensaron que tenía dificultades con la ___.									
lectura	.24	tarea	.19	escritura	.19	0.238	0.674	2.048	0.054
97. Lo que más le gusta a mi padre es observar al ___.									
vecino	.24	perro	.24	cielo	.19	0.238	0.615	1.857	0.071
98. Me alegra que no tengas ningún problema con la ___.									
vecina	.23	niña	.15	literatura	.15	0.231	0.806	2.385	0.022
99. Si hubiera hecho caso, habría visto al ___.									
ladrón	.23	doctor	.14	cantante	.14	0.227	0.667	1.909	0.046

APPENDIX B (Continued)

Completion	Prob	Completion	Prob	Completion	Prob	CS	TTR	MR	*R*
100. Lo único que se rompió fue la ___.									
ventana	.23	pierna	.23	taza	.14	0.227	0.548	1.909	0.062
101. Tiene el mismo tamaño que una ___.									
pelota	.22	pulga	.17	naranja	.17	0.217	0.558	1.87	0.066
102. Mis padres pensaban que había seis ___.									
personas	.22	niños	.13	ratones	.09	0.217	0.825	1.739	0.036
103. Me di cuenta que nada estaría listo para los ___.									
invitados	.22	exámenes	.17	juegos	.13	0.217	0.615	1.696	0.055
104. Todo lo que ves está incluido excepto los ___.									
zapatos	.2	impuestos	.2	muebles	.13	0.2	0.794	2.267	0.026
105. Nunca tuve la menor duda de que eran ___.									
novios	.2	pareja	.13	hermanos	.13	0.2	0.886	2.333	0.015
106. Las guitarristas y la cantante se fueron del estudio sin la ___.									
grabación	.2	maqueta	.16	guitarra	.16	0.2	0.5	1.84	0.061
107. El tema de conversación son sus ___.									
problemas	.2	preocupaciones	.13	padres	.13	0.2	0.829	2.333	0.019
108. Uno de los fotógrafos saltó de la ___.									
terraza	.19	silla	.19	ventana	.14	0.19	0.61	1.952	0.051
109. Con este tubo podemos hacer una ___.									
tubería	.18	flauta	.18	conexión	.14	0.182	0.75	1.818	0.045
110. No le apetece enseñarnos el ___.									
resultado	.15	cuadro	.15	camino	.15	0.154	0.906	2.462	0.009
111. La palabra que no pude recordar antes es ___.									
complicada	.13	ésta	.09	rara	.09	0.13	0.878	1.783	0.014
112. Otra cosa a tener en cuenta es la ___.									
edad	.13	economía	.13	voluntariedad	.06	0.125	0.944	2.25	0.006