

Native Spanish Speaker Intuition in Noun Gender Assignment

JENS H. CLEGG

Indiana University-Purdue University Fort Wayne

Abstract

When an English word is borrowed into Spanish it has no specific gender and in order for the word to be used in the language it must be assigned a gender, either masculine or feminine. There are many different factors that may play a role in the assignment of gender to new English-origin words. This paper concentrates on English word final phonemes that do not correspond with the phonemes usually found in word final position in Spanish. For this study, words that are phonotactically possible in English but do not, in fact, exist were given to native Spanish speakers. The native Spanish speakers were then asked to assign the new words a gender. The words used in the study were designed to appear to be one gender according to their terminal phoneme. (ie. words ending in /o/ are generally masculine and words ending in /a/ are feminine) The definitions given to the words were designed to contradict the predictable gender associated with the terminal phoneme. In 90% of the cases the participants overlooked the definition and assigned the word to the gender associated with the terminal phoneme. This indicates that the meaning of the word and its function does not directly influence gender assignment. This study does indicate that the single most important factor in assigning gender to words borrowed from English into Spanish is the terminal phoneme.

1. Introduction and Review of the Literature

In Spanish all nouns have a specific grammatical gender that is an integral part of the word. The gender of nouns is important since Spanish syntactic rules mandate that all descriptors of the noun agree in gender with that noun. In English, however, nouns have no specific grammatical gender. This creates a challenge when an English-origin word is used in Spanish. In order for the word to be used by Spanish speakers, it must be assigned a specific gender, either masculine or feminine so that Spanish agreement rules can be followed.

This leads to the question of how the Spanish noun gender system works. Bull (1965) studied the patterns for gender of all of the Spanish nouns in a

Spanish/English dictionary. Bull's motivation was to discover generalized noun gender rules that could be taught to students instead of using rote memorization to learn the gender of each individual noun. He focused his study on the terminal, or final, grapheme and in some cases groups of graphemes of the words. Bull found that words that end in *-a*, *-d*, *-ción*, *sis*, and *-itis* will be feminine in gender 98% of the time, whereas words ending in any other grapheme will be assigned a masculine gender 96% of the time. Bull claims that students who learn these rules and follow them will select the correct noun gender 97% of the time.

Bull (1965) also indicated that 52% of the total nouns in the dictionary are masculine and 45% feminine. There is a three percent discrepancy in the numbers presented by Bull which is accounted for by ambivalent words which may use both masculine and feminine semantic gender. These findings would indicate that there is a slight predominance of the masculine gender in the Spanish language, but that, overall, the two genders are nearly equally represented.

Bergen (1978) further studied and confirmed the findings of Bull (1965) refining them by adding four additional generalizations. Bergen found that the endings *-umbre* (*la servidumbre* 'servitude'), *-ie* (*la especie* 'the species'), and *-z* (*la luz* 'the light') should be associated with feminine gender. Bergen also suggested that nouns of Greek origin ending in *-ma* (*el problema* 'the problem') should be associated with masculine gender. Bergen also coined the term 'LONERS' as an acronym for endings associated with masculine gender.

Teschner et al. (1984) looked in depth at the findings of Bull (1965) and the refinements of Bergen (1978) and found that there were some inaccuracies. Teschner et al. used a much larger dictionary (89,000 words) and carefully excluded words that were ambivalent for gender. They first point out that students are unaware of words of Greek origin and that further, there are many words ending in *-ma* that have feminine gender. They also comment that the new refinements *-umbre* and *-ie* are low productivity forms and are not frequent enough to justify the creation of separate rules. Based on the larger dictionary Teschner et al. conclude that the endings *-z*, *-n*, and *-s* are in fact indeterminate for gender since they only slightly favor one gender or the other. The *do*, however, confirm that *-a*, *-d*, and *-ión* are feminine endings and that *-l*, *-o*, *-r*, and *-e* are masculine endings.

Based on the findings of Bull (1965), Bergen (1978) and Teschner and Russell (1984) it appears that there is a clear correlation between the terminal

grapheme, or group of graphemes, and the gender of Spanish nouns. This leads to two questions: 1. Are native speakers aware of this terminal grapheme system? 1. When an English-origin noun is borrowed into Spanish do speakers use this system to give the new noun a gender?

To answer these questions Poplack et al. (1982) analyzed a series of English-origin words encountered in Puerto Rican Spanish. They identified several factors that may play a role in the gender assignment process. The factor they determined to be the most significant in predicting gender assignment was the biological sex that was associated with the referent, either male or female. An example of this is the word *el wachiman* which refers to a male security guard and thus has been assigned the Spanish masculine article *el* 'the'. Words in this category were assigned that specific biological sex regardless of any other factor. The biological factor, though usually an absolute, is not highly useful in the process of gender assignment since it only applies to a small group of words that refer specifically to male or female animate entities.

Other factors that their study found to be important in determining gender were: 1. Analogical gender where the loanword is assigned the same gender as its existing Spanish equivalent as in *la milque* 'milk' which has the same gender as *la leche*. Analogical gender, however, only functions when there is a clear existing language synonym. 2. Phonological gender where the word is assigned a gender based on its terminal phoneme. In their analysis Poplack et al. change from terminal grapheme (the last letter or group of letters) to terminal phoneme (the last sound or group of sounds), hereafter TP. This change allows the researchers to focus more on the sounds instead of orthographic irregularities and condense some sounds into one group, ie. *-s* and *-z* are in fact */s/*. Poplack et al. claim that all the informants in their study regardless of sex, age, or social status responded within similar parameters. They also found that by the age of five, children had already developed a system for assigning gender, and that adults and children assigned the same gender to the same words.

Natalicio (1983) evaluated the accuracy of the rules developed in Bull (1965) in an attempt to determine if native Spanish speakers possessed instincts or intuitions that they use when assigning gender. Uncommon Spanish words, half of which conformed to Bull's rules, and the other half which contradicted them, were selected from a dictionary. The words that were chosen were words that were extant in the Spanish language but unknown or unfamiliar to most Spanish speakers. The participants in the study were fifty students from El Paso,

Texas, and fifty from Juarez, Mexico. They were asked to assign the words a gender, using the articles *el* or *la*. The study found that in most instances the assigned genders agreed with the rules set forth by Bull. Natalicio (1983: 52-53) also found that "...not only do native Spanish speakers favor *el* over *la* when assigning articles to unfamiliar nouns, but they also appear to have stronger convictions about assigning *el* to a noun than they do in assigning *la*." She also claims that native Spanish speakers do have an instinctive sense of gender and that teaching gender assignment to nouns on a case by case approach is inefficient. Natalicio also suggested that her experiment should be repeated using nonsense words instead of unfamiliar Spanish words to make the results more reliable.

Banfield (1994) used existing English words that were unfamiliar to native Spanish speakers accompanied by pictures of the items. He showed the pictures to the participants, told them the word and then asked them to provide the correct definite article for each word. He found that there is a preference for the masculine definite article *el*. Banfield also concluded that two factors were important in assigning gender the terminal phoneme of the word and the gender of an already existing Spanish word that has the same meaning ie. analogical gender. However, in cases where the new word had no semantic equivalent in Spanish, analogy does not help in the assignment since there is nothing with which to compare the new unique term.

Smead (2000) confirms the findings of Poplack et al. (1982) that the two most important factors in the assignment of gender to English-origin nouns are the biological sex of the referent and the terminal phoneme, hereafter TP, of the noun. Smead also finds that variable assignment with TPs is largely biased toward the masculine with the exception of /a/. Smead (2000) also posits a new factor called hyperonymic gender wherein words are assigned a gender based on their presence in a super-ordinate class of words. For example, a new sport, ie. Ultimate Frisbee, would be assigned masculine gender since all existing Spanish sports are masculine.

Otheguy and Lapidus (2001) echo the same findings of previous researchers for bilingual Spanish speakers in New York. Otheguy and Lapidus claim that /a/ is the only productive feminine ending. They propose that the assignment of masculine gender, in most other cases, is an adaptive strategy used by bilinguals. Spanish noun gender has exceptions for each of Bull's rules, for example, *el día* 'day', *el problema* 'problem', and *la mano* 'hand', all of which,

in standard Spanish, have the opposite gender expected based on the gender associated with the TP. Assuming gender assignment is rule based, native Spanish speakers must memorize these exceptions to the rule and, in general, do not deviate from standard gender assignment with these exceptions (Garcia 1998). Otheguy and Lapidus (2001) claim that the memorization-intensive gender system of Spanish is too much of a cognitive load for bilinguals. According to Otheguy and Lapidus (2001: 216-217), these bilinguals have too much going on in their minds and as an “evolutionary adaptation” have simplified the Spanish gender system such that /a/ is feminine and all other endings are masculine.

Clegg et al. (2009) studied borrowings found in spontaneous speech from sociolinguistic interviews of bilingual speakers living in Mora, New Mexico. The results indicated that there was a clear preference for masculine gender with 89% of the tokens. Clegg et al. also found that words ending in the TPs corresponding to the “loners” group were assigned masculine gender in 89% of cases.

Eddington (2002) tested Spanish gender assignment patterns using a technique known as the analogical modeling of language. In his study Eddington used a computer analysis program to determine whether the gender assignment process follows predictable patterns or is random. Eddington used the 2,416 most frequent nouns in the Spanish language as found in Juilland and Chang Rodriguez’s (1964) frequency dictionary for Spanish. Eddington compared the assignment of gender based on the TP to gender assignment based on the final or penultimate syllable. He found that the penultimate rhyme combined with the final syllable correlated with the gender of nouns accurately in 95% of cases and that the gender associated with the TP correlated with the gender of the same nouns in 93.6% of cases. As Eddington (2002) points out this is not a significant difference.

The research which has been done on gender assignment in Spanish indicates that native Spanish speakers do have an internalized system for determining noun gender and that this system is in place at a young age and is uniform to most Spanish speakers. Factors that have been found to be important in the gender assignment process are the biological sex of the referent, the TP of the word, and analogical gender.

In the gender assignment process one area that has not been researched in depth is that of words whose TPs are non-typical for Spanish. Spanish words generally end in only a few select phonemes, /l, o, n, e, r, s, a, d/. English words, however, end in many different phonemes that are not acceptable in or typical of

Spanish phonology. This indicates that many words borrowed from English will have these atypical endings. Bull (1965) found examples of words with TPs that would be atypical but there were few of them and in most cases they were established borrowings from other languages. Bull did find that those words with atypical endings did have masculine gender. Clegg et al. (2009) also indicated that nouns with atypical TPs were overwhelmingly assigned masculine gender 91% of the time.

2. Method

The present study is an experiment designed to measure whether English-Spanish bilinguals will consistently assign gender to potential English-origin words with endings, or TPs that are atypical for Spanish.

The participants in this study were asked to provide certain personal information in order to effectively evaluate the data. Of the surveys completed, forty were selected for analysis. The only requirements for this study were that the participants had to be Spanish-English bilinguals who had acquired Spanish as their first language in a predominantly Spanish-speaking environment and then acquired English starting before the age of 12. The purpose of these language requirements was to ensure that the participants could clearly speak Spanish well. The participants ranged in age from sixteen years old to fifty-seven years old. The gender of the participants was controlled for by selecting equal numbers of male and female participants.

Following the suggestion of Natalicio (1983) possible, but not actual, English words were created for the experiment. These potential words were carefully constructed to have terminal graphemes that would be non-typical for Spanish. All of the words are nonsense words that are phonotactically possible in English but did not in fact exist at the time of the experiment. These words can be seen in Table 1 below. The words are paired, that is, there are two words for each non-typical ending. The groups are represented by the TP with which each word ends. The TPs, /t, b, p, k, h, g, and m/, represent endings that are non-typical endings for Spanish and are the group on which the analysis will be performed. There are six words with endings that are phonotactically acceptable for Spanish that serve as distracters. The distracter words are those words in Table 1 ending in /o, a, and l/. Each invented word was provided a definition in order to allow the participants to consider biological gender and analogical gender.

The participants were given a written survey and told that the words were new English words. They were then asked to read the words and choose a gender, *el* or *la*, for each new English word. They were also given the option to write the word the way they would say it, or leave it alone. The surveys were then collected and analyzed to evaluate the results and determine whether or not they show the predicted outcome.

Nonsense words	Definitions
Pligit	A type of glue used to glue metal objects
Nosbert	A new computer game
Brumb	Name of a new animal found in Africa that resembles a horse
Snarb	New type of bicycle tire impervious to flats
Roup	New American sport in which a long stick and two balls are used
Climp	New type of medical instrument used in giving stitches.
Spaltek	A new machine that makes computers function more rapidly
Krunk	Tool used on an automobile
Flam	Part of the engine that purifies combustion
Schlem	Type of food made from chicken, rice, and vegetables.
Kerlog	Machine used for recording and playing sounds
Krang	A new part of a TV that improves colors
Deraph	Visual defect that only allows one to see objects close to the face
Louch	New type of dance
Cavril	New type of nuclear train
Kinistal	Apparatus used to store electricity

Sarlo	A new type of hat used by women
Risho	Machine used in the kitchen to mix dough
Schota	Small firearm
Blicka	Tool used in home construction

Table 1. Nonsense words and their definitions

3. Results

The sample words were placed into groups of two for evaluation. There are a total of twenty words. The distracter groups which include words ending in *-o*, *-a*, and *-l* were not analyzed together with the phonotactically non-typical endings. After removing the distracters seven groups of words with two words in each group remained. These groups can be seen in Table 2. For example, group one is the *-t* group which includes the two words *pligit* and *nosbert*. The remaining six groups followed the same pattern.

The two words in each group were analyzed both independently and together but the differences were minimal. The results indicate that, overall, 90% of the time the respondents chose to assign the new words to the *el*, or masculine category, showing a preference for the masculine gender when dealing with endings that are non-typical for Spanish. This is the expected result based on the tendencies described in Bull (1965), which can be seen the last column of Table 2, and also the findings of Clegg et al. (2009).

Group	Occurrence of <i>el</i>	%	Occurrence of <i>la</i>	%	Bull's Percentages
T Pligit/Nosbert	73	91%	7	9%	94.5%
B Brumb/Snarb	77	96%	3	4%	100%
P Roup/Climp	74	92%	6	8%	100%
K Spaltek/Krunk	66	82%	14	18%	100%

H Louch/Deraph	69	86%	11	14%	100%
G Krang/Kerlog	74	92%	6	8%	100%
M Schlem/Flam	71	89%	9	11%	96.9%
TOTAL	504/560	90%	56/560	10%	98.77%

Table 2. Gender assignment grouping results.

The groups containing words ending in *-t*, *-b*, *-p*, and *-g* had a percentage of masculine occurrences greater than 90%. This number is slightly lower than the findings of Bull. For *-b*, *-p*, and *-g* he predicted 100% masculine occurrences while for the *-t* group he predicted 94.5% masculine occurrences. The slight difference in percentages can be explained by considering that Bull did his study on words that had experienced some level of phonological adaptation into Spanish.

Barkin (1980) found the level of phonological adaptation of a word may influence its gender assignment. Barkin interviewed native Spanish speakers using pictures to elicit loanwords. The loanwords used during the interview were then analyzed. In each case Barkin found that the level of phonological adaptation affected whether the word was assigned masculine or feminine gender. An example cited by Barkin is the English word *swan* which has two forms in different stages of adaptation, *un swan* which is masculine and in an early stage of adaptation, and *la suana* which is feminine and has been more fully adapted. The gender assignment of these two words may be different since their terminal phoneme is different.

Barkin (1980) however, did not consider the method of introduction of the new English-origin word. Borrowed words may be based on a written or oral model (Higa, 1979; Guitart, 1981; and Smead, 1998). An example of this may be found in the word *sweater*. In U.S. Spanish there are two possible forms, *la suera* and *el sueter*. *Sueter* appears to be modeled on the written English form “sweater” and *suera* on the oral form. Speakers will adapt the forms based on first medium of contact. If that contact is oral, the word will likely be adapted differently than if the contact is written. This would indicate that among the words analyzed by Bull, which in order to be in the dictionary are well established and adapted borrowings, not many would retain the original non-

typical endings. As a result, Bull did not have a high number of tokens upon which to base his analysis, which may explain his higher percentages.

The definitions that were assigned to each potential word were also considered in the analysis of the data. None of the definitions arbitrarily assigned a biological sex to the particular item, however, some were associated with use by males or females only. This was done to counter the potential effects of semantic equivalence that Espinosa (1946) found in his study of New Mexican Spanish. Espinosa (1946: 286) claims that words will become masculine if, "...they denote males, masculine articles of clothing, male occupations, or mechanical instruments or new machinery and inventions...". To test the prediction of Espinosa (1946) the *-o* words were given definitions that would normally be associated with use by females and the *-a* words were given definitions that would associate them with men. This was done intentionally to see if this might cause the participants to waver in their assignment of gender to these words which have TPs that are almost categorical in their gender assignment.

The groups *-o*, *-a*, and *-l* were the distracters in the experiment. These endings are common in Spanish and native speakers should be accustomed to dealing with them. Bull found that *-o* was masculine 99.7% of the time, *-a* was feminine 98.9% of the time and *-l* was assigned masculine gender 96.6% of the time. The words in the *-o* group, which is associated with masculine gender, were given definitions associated with females; a feminine article of clothing, and an apparatus used in the kitchen. The *-a* group, which is associated with feminine gender, had words whose definitions are associated with males; a small gun and a tool used in house construction. The *-l* group is associated with masculine gender and the words were given definitions that were associated with males; an electronic device and a type of train, which may be considered as male interests or playthings.

The results for the *-a* group show that the potential words ending in *-a* were assigned a feminine gender 89% of the time in spite of the masculine biased definitions. The *-o* ending words with feminine biased definitions were assigned masculine gender 90% of the time. Interestingly, the *-l* group was only assigned a masculine gender in 91% of the cases even though there were masculine oriented definitions. On the whole the experiment revealed that 90% of the time these distracter words were assigned the correct gender based on the TP regardless of the definition of the word. This indicates that the definition given to the words or

the potentially gender-influencing function of the words does not play a major role in the assignment of gender to words.

In the test nonsense word groups the Groups with words ending in *-k*, *-h*, and *-m* contained masculine occurrences higher than 82% but lower than 90%. Bull's findings indicate that *-k* and *-h* have a masculine assignment 100% of the time while *-m* is masculine 96.9% of the time. The definitions for these words varied, some of the items included computer equipment, tools, and auto parts. There were also definitions that were slightly different such as, a type of food, a new style of dance, and an illness. The uniformity of the results within the category, and across the other categories, shows that the influence of the definitions on gender assignment is not significant.

There were some additional interesting phenomena in the results of the survey. The participants were asked to provide a written version of the new word as they would pronounce it in Spanish. In some cases the participant was not comfortable with the non-typical ending. He or she re-wrote the new word altering it before they were willing to assign it a gender. Some examples include:

Kerlog to El Querlo
 Snarb to El Esnar.
 Flam to El Flamo.
 Roup to El Roupe.

Krunk to El Crun.
 Climp to El Clim.
 Spaltek to El Espaltex.
 Cavrilo to El Cavrilo

In cases where the participant altered the form and then assigned it a gender, these tokens were eliminated in the analysis of the results since the endings were no longer non-typical. Not enough participants chose to alter the words to determine a pattern or a preferred adaptation strategy. It does appear that the participants in the study did not like words that ended in consonant clusters, like the word *krunk*, which ends with two consonants. Spanish has very few words that end in consonant clusters, and words in the survey that ended in consonant clusters were generally altered more frequently than words that did not. It is suggested that a follow-up study be done to evaluate which typical ending would appear most frequently. In most cases the participant would alter the word by either eliminating orthographic letters or consonant clusters until they arrived at one they were comfortable with, or by adding a familiar ending and assigning it the corresponding gender.

4. Conclusions

The results of this study show that noun gender tendencies established by Bull (1965) do concur with native Spanish speaker intuitions. Based on the findings of Bull it was predicted that 97% of words with these non-typical endings would be assigned masculine gender and, as can be seen by the results, 90% of all the non-typical words were assigned masculine gender. This study shows that in the majority of cases loanwords from English with non-typical word endings will be assigned masculine gender.

The participants in this study appear to assign gender based on an internalized system of rules that they hold in common. All participants regardless of place of origin, age, or gender responded in a similar manner. This gender-determining instinct appears to be based on the gender associated with the terminal phoneme, or TP, of the word. Respondents overlooked potentially gender biased definitions and chose the gender that is associated with the TP of the word on a consistent basis. The definitions given each of the potential English words were designed to test if those definitions or their association with one sex or the other would have any affect on the assignment of gender. The results show that regardless of the definition the speakers assigned gender predictably. Words ending in -o were assigned masculine gender and words ending in -a were assigned feminine gender regardless of their definition. This shows that English/Spanish bilinguals have an internalized noun gender system and apply it predictably to nouns based on the TP especially when assigning gender to nouns with endings that are atypical for Spanish.

Bibliography

- BANFIELD, Robert L. (1994): *Gender assignment of English loanwords in the Spanish language*. Masters thesis. Brigham Young University; Provo, Utah.
- BARKIN, Florence. (1980): "The role of loanword assimilation in gender assignment." *Bilingual Review* 7, 02, 105-112.
- BERGEN, John J. (1978): "A simplified approach for teaching the gender of Spanish nouns." *Hispania* 61, 865-876.
- BULL, William. (1965): *Spanish for teachers: Applied Linguistics*. New York: The Ronald Press Co.

- CLEGG, Jens. and WALTERMIRE, Mark. (2009): "Gender assignment to English-origin nouns in the Spanish of the Southwestern United States." *Southwest Journal of Linguistics* 28, 1, 1-17.
- EDDINGTON, David. (2002): "Spanish gender assignment in an analogical framework." *Journal of Quantitative Linguistics* 9, 49-75.
- ESPINOSA, Aurelio M. (1946) [1909]: *Estudios sobre el español de Nuevo México, Parte II: Morfología*. Traducción, reelaboración y notas de Ángel Rosenblat. Biblioteca de Dialectología Hispanoamericana. Buenos Aires: Universidad de Buenos Aires.
- GARCIA, Maryellen. (1998): "Gender marking in a dialect of Southwest Spanish." *Southwest Journal of Linguistics* 17, 1, 49-58.
- GUITART, Jorge. (1981): "On loanword phonology as distinctive feature phonology: An autosegmental approach." In CRESSEY, W. and NAPOLI, D. (eds.), *Linguistic symposium on romancelanguages* 9, 17-23. Georgetown: Georgetown University Press.
- HIGA, Masanori. (1979): "Sociolinguistic aspects of word borrowing." In MACKEY, W. and ORNSTEIN, J. (eds.), *Sociolinguistic studies in language contact: Methods and cases*, 277-292. NY: Mouton.
- JUILLAND, A. and CHANG-RODRÍGUEZ, E. (1964): *Frequency Dictionary of Spanish Words*. The Hague: Mouton.
- NATALICIO, Diana. (1983): "Native speaker intuitions as a basis for determining noun gender rules in Spanish." *Southwest journal of linguistics* 6, 1, 49-55.
- OTHEGUY, Ricardo and LAPIDUS, Naomi. (2001): "An adaptive approach to noun gender in New York contact Spanish." In NUÑEZ-CEDENO, R. and LOPEZ L. and CAMERON, R. (eds.), *A romance Perspective on Language Knowledge and Use*, 209-229. Amsterdam: Benjamins.
- POPLACK, Shana and SANKOFF, David. (1982): "Un estudio comparativo de la asignación de género a préstamos nominales." In ALBA, O. (ed.), *El español del caribe: ponencias del VI simposio de dialectología*, 239-269. Santiago, República Dominicana: Universidad Católica Madre y Maestra.
- SMEAD, Robert N. (1998): English loanwords in Chicano Spanish: Characterization and rationale. *Bilingual Review* 23, 2, 113-123.
- SMEAD, Robert N. (2000): On the assignment of gender to Chicano anglicisms: Processes and results. *The Bilingual Review* 25, 3, 277-297.

18 Jens H. Clegg

TESCHNER, Richard V. and RUSSELL William M. (1984): The gender patterns of Spanish nouns: an inverse dictionary-based analysis. *Hispanic Linguistics* 1, 115-132.