Gender Assignment to New Words in Catalan

Nadia El-Yousseph

Masters research paper
Ciència Cognitiva i Llenguatge

Supervised by Eulàlia Bonet Alsina
Universitat Autònoma de Barcelona
March 26th, 2010

Table of Contents

Nadia El-Yousseph

Abstract

This paper investigates the original debate about gender assignment criteria, including recent explanations within optimality theory and a study of gender assignment for new words in Catalan. Data were collected on speakers' gender assignment preferences to loan words and Catalan-looking made-up words and at least two assignment criteria were found for Catalan: shape-based criteria and meaning-based criteria. Finally, this paper applies Rice's (2005) optimality theoretic approach to gender assignment in Catalan.

Grammatical gender is an aspect of language that has long baffled both linguists and learners alike. Indeed, humorist, David Sedaris, documented his difficulties understanding the purpose of gender while he was in an introductory French class in Paris.

> "The teacher's reaction led me to believe that these mistakes were capital crimes in the country of France.
>
> 'Were you always this *palicmkrexis*?' she asked. 'Even a *fiuscrzsa ticiwelmun* knows that a typewriter is feminine.'
>
> I absorbed as much of her abuse as I could understand, thinking – but not saying – that I find it ridiculous to assign gender to an inanimate object incapable of disrobing and making an occasional fool of itself. Why refer to Lady Crack Pipe or Good Sir Dishrag when these things could never live up to all that their sex implied? (Sedaris 2000)"

Part of Sedaris' problem is that grammatical gender makes no sense if it is thought of as an extension of sex, though historically linguists have also thought so.

Section 1 of this paper investigates the original debate on grammatical gender, exploring the conflict between Grimm (1890) (among others), who thought that gender related to word meaning and Brugmann (1891) and Bloomfield (1933), who thought that gender is arbitrary. Research by Zubin and Köpcke (1984 a, b); Köpcke & Zubin (1986) and Corbett (1991) on the relationship between grammatical gender and word meaning are presented. Section 2 of this paper presents a study conducted by the author on new words in Catalan. It shows that apart from the sex of an animate referent, there are two primary ways that a new word in Catalan can get its gender: hyperonym effect and word shape. Section 3 presents Rice's (2005) *optimal gender assignment theory* and applies it to Catalan.

1. Background

1.1 Introduction

There are two generative views of grammatical gender. Under one approach, the lexical entries must be existing words. Authors such as Aronoff (1976) for English, Scalise (1985) for Italian and Bermúdez-Otero (2007) for Spanish, among others, propose that the word ending forms a part of the lexical entry. Therefore a word like *mesa* 'table' has the lexical entry as in example (1).

(1) /mes-a/ noun 'meaning'

Under this approach, gender need not be stored in the lexicon because the lexical entry has a final vowel, called a 'stem-formative' in Bermúdez-Otero (2007). It is by the stem-formative that gender can be predicted, given that nouns ending in /a/ are feminine in Spanish. Under this approach, gender must be specified in the lexical entry only when it cannot be predicted from its shape or its meaning. This would be the case for a word like *tema* 'topic.$_{MASC}$': even though it ends in *-a*, it is masculine.

In the other approach, lexical entries need not be existing words. Authors like Mascaró (1986), for Catalan, Harris (1991), for Spanish, believe only the root and the gender of the noun to be in the lexical entry, as seen in (2).

(2) /mes/ noun [+F] 'meaning'

According to the second approach, gender is an inherent part of each lexical entry. Redundancy rules assign the unmarked endings such as [+F] > /a/ or [-F] > /o/ in Spanish. Therefore, under the first approach one can say that *mesa* is feminine because it ends in /a/, whereas by the second approach *mesa* ends in /a/ because it is feminine.

If words are lexically marked for gender, then any analysis which claims to find the 'rules' for gender assignment, is actually just discovering patterns that have formed in each gender category. These patterns are the criteria by which most loan words' genders are assigned (Corbett 1991).

The goal of this study is not to evaluate these two approaches, but to examine how gender is assigned by native speakers to loan words and made-up nouns.

1.2 Two schools of thought: The original debate on gender

The discussion on gender and sex dates back to Grimm (1890), who thought that gender had semantic motivation. Essentially, he believed that the grammatical genders that people and (higher) animals are assigned are based on their sex. The neuter gender, he admits on two occasions, is difficult to explain, but appears to have the meaning of an undeveloped sex and is used for children in German e.g. *Kind* 'the child$_{\text{-NEUT}}$'.

By Grimm's analysis, nouns without any 'grammatical' (i.e. shape) reason to be feminine or masculine are assigned their gender born through what he calls 'fantasy'. Kilarski (2000), who presents a thorough summary of early gender debate, interprets this to be personification achieved through "imagination":

> "... das masculinum scheint das frühere, größere, festere, sprödere, raschere, das thätige, bewegliche, zeugende; das femininum das spätere, kleinere, weichere, stillere, das leidende, empfangende; das neutrum das erzeugte, gewirkte, stoffartige, generelle, unentwickelte, collective ... Diese kennzeichen stimmen zu den bei dem natürlichen genus ... aufgestellten." (Grimm 1890: 357).
>
> The masculine seems to be 'earlier, larger, more solid, rougher, quicker, to be active, to conceive; the feminine is later, smaller, weaker, stiller, it suffers, receives; the neuter is that which is concrete, acted upon, made of material, general, underdeveloped, collective. These attributes match those found in natural gender. (translation: mine)

In Grimm (1890), inflection takes a backseat to semantic features in gender assignment and derivation doesn't map the suffix onto a particular gender. Grammatical gender is based on semantic features.

Many linguists and language officials have kept the theory of grammatical gender-sex relation above water against attacks. Recently, the strict sex-grammatical gender correlation has been dropped in favor of other semantic means for determining gender (Zubin & Köpcke 1984 a, b; Köpcke & Zubin 1986, and Corbett 1991). See discussion below for more on this.

In contrast to many semantic-based explanations of gender, formalists Brugmann (1891) and Bloomfield (1933) contend that gender is an arbitrary noun feature. They believe that words don't get assigned their gender based on their meanings as Grimm and other suggest, but that the words first *arbitrarily* arrive at their genders and only when the words already have a certain gender can the personification mentioned above take place. Brugmann, for instance, states that suffixes *-a* and *-i* in Indo-European were not inherently feminine. It was only through the association with **māmā* and **genā* that *-ā* was reinterpreted and subsequently extended to a few

nouns. Bloomfield (1933), like Brugmann, denies that gender has a semantic motivation. He summarizes as such:

> "There seems to be no practical criterion by which the gender of a noun in German, French or Latin could be determined: to define the meaning of the episememe 'masculine' in such a language would be simply to list the markers of masculine nouns and the nouns that belong arbitrarily to the class, and to say whatever is common, in the practical world, to all these objects, is the "meaning" of the masculine gender category." (Bloomfield, 1933, p. 280)

To find out how loan words are assigned their gender, we must look to the gender system of the borrowing language, as "the gender of loan words is determined exactly like that of other nouns" (Corbett 1991).

1.2.1 Semantic-based assignment

One criterion for gender assignment in both native words and loan words that is practically uncontroversial is the relation between gender values and sex. Words referring to adult males are masculine and words relating to adult females are nearly always feminine.[1] The natural gender forms the basis for the semantic core of every language (Corbett 1991).

However irrational Bloomfield may have found it to be, there have been many papers supporting the semantic-grammatical gender relationship; see for example Zubin & Köpcke (1984a), Köpcke & Zubin (1986), DuBord (2004) and Morin (2006).

In a series of studies by Zubin & Köpcke, they purport to have identified several highly structured gender regularities phonologically, morphologically and semantically in German. In their (1984a) study Zubin & Köpcke probed the Last Member Principle (LMP), which states that the last gender-assigning component of compound nouns in German determines the gender of the word for morphologically complex nouns. Therefore, according to the LMP, a word like *Mitleid* 'sympathy$_{\text{-NEUT}}$' is neuter because its final component is *Leid* 'sorrow$_{\text{-NEUT}}$', whereas, *Mitleidigkeit* 'capacity for sympathy$_{\text{-FEM}}$' is feminine because of the feminine suffix *-keit.* The German word *Mut* means 'courage$_{\text{-MASC}}$' but in compounds, *-mut* has a meaning similar to 'mood' in English. According to the LMP, all *-mut* compounds should be masculine, such as in the words *Lebensmut* 'exhilaration$_{\text{-MASC}}$', *Übermut* 'bravado$_{\text{-MASC}}$', *Wagemut* 'daring$_{\text{-MASC}}$', etc.

[1] There are few cases where exceptions to this occur, such as one of German's words for 'woman', *Weib*, which is not feminine, but neuter.

However, they observe some violations to the LMP. There are several complex nouns ending in *-mut* that are not masculine, but feminine: *Armut* 'poverty$_{\text{-FEM}}$' *Anmut* 'gracefulness$_{\text{-FEM}}$', *Demut* 'humility$_{\text{-FEM}}$', etc. The variable gender compounds were listed in several dictionaries as either masculine or feminine (never neuter). In their (1984a) study, they conducted an experiment to check the dictionary assignments. Participants were to fill in the determiner of all the *-mut* compounds that are still in common usage, which revealed that there is much higher variance in gender assignment than dictionaries reflect.

In the first part of their study, Zubin & Köpcke (1984a) observe a pattern between the gender of the *-mut* compounds and the word's relative *Extroversion-Introversion*, whereby masculine words have a high *Extroversion* Affect[2] and feminine words have a high *Introversion* Affect[3].

In the second part of their study, participants were presented with 6 adjectives and had to mark where on an Introversion-Extroversion scale each of the *mut*-compounds lie. The scale included five boxes between two polar introversion-extroversion adjectives (e.g. passive-active, sad-happy, beautiful-ugly). They expected to find the following relationships: 1) feminine-Introversion, 2) masculine-Extroversion, 3) a middle rating for words of varying gender.

They found that the compounds previously rated to be feminine were often associated with Introversion adjectives and compounds previously rated to be masculine were found associated with Extroversion adjective. Words that lie in the middle were the words of variable gender assignment (see (3)).

[2] By their definition, 'Extroversion' suggests "types of conduct or attitude which are directed toward controlling the outside world or view it as controllable, or which protect the self from the outside control."
[3] Introversion is a label for words that "suggest conduct or attitude which place the self under outside control or view it as controllable, or which open the self to outside influence."

(3)

		Gender listed in dictionary	Proportion of feminine-gender	Semantic Differential Rating
Lebensmut	'exhilaration'	M	0	
Übermut	'bravado'	M (V)[4]	0	3.4
Wagemut	'daring'	M	.05	
Hochmut	'arrogance'	M	.10	3.6
Unmut	'bad temper'	M	.15	
Wankelmut	'vacillation'	M (V)	.15	
Kleinmut	'small-mindedness'	V	.20	
Freimut	'frankness'	M	.25	
Gleichmut	'equanimity' ,	V	.45	2.7
Grossmut	'generosity' ,	V	.60	3.2
Schwermut	'melancholy'	F	.80	
Sanftmut	'tenderness'.	F	.85	
Demut	'humility'	F	.95	
Wehmut	'sadness'	F (V)	.95	2.4
Anmut	'gracefulness'	F	1	2.5
Armut	'poverty'	F	1	

Reproduction of Zubin & Köpcke's experimental results. The semantic differential rating range is 1-5, where 1 is a polar *Introvertedness* adjective and 5 is a polar *Extrovertedness* adjective.

In many ways, this is reminiscent of Grimm's early report that the masculine gender contains words that are more agile, rough, active, etc. whereas the feminine gender has words that are weaker, sadder and smaller Zubin & Köpcke stress, however, that their findings do not endorse sex-specific associations. "…the Introversion-Extroversion affective polarity goes much deeper, both historically and psychologically, than present-day attitudes about sex-specific behavior." Parallel findings have been found with the endings *-nis* and *-sal*[5], whereby feminine variants have a higher degree of Introversion than their neuter counterparts.

In a later study Köpcke & Zubin (1986) find a taxonomic system of dividing objects into superordinate, base level and subordinate level (e.g. *furniture - chair - love seat*). They observe that many superordinate terms are neuter. They find the correspondence between superordinate terms and neuter for many semantic groups: instruments, tools, vegetables, etc. Consider the table in (4).

[4] The (V) means that the dictionaries unanimously gave one gender and that only one gave a secondary gender. All the other V's mean that primary gender varied across dictionaries.

[5] These endings originated in Old High German as neuter *(-nassi, -nessi, -nissi,* and *–nussi*) and feminine *(-nissa* and *-nissi*) but collapsed into the same suffix when Middle High German lost its final vowels. Similarly, *-sal* was originally neuter and *-sala* originally feminine. Therefore neither of these suffixes is attributable to a historically constant determined gender.

(4).

Superordinate	**Base level**
Instrument 'instrument$_{\text{-NEUT}}$'	*Guitarre*, guitar$_{\text{-FEM}}$' *Trompete*, trompete$_{\text{-FEM}}$'
Gemüse 'vegetable$_{\text{-NEUT}}$'	*Spinat*, spinach$_{\text{-MASC}}$' *Kohl*, cabbage$_{\text{-MASC}}$'
Werkzeug 'tool$_{\text{-NEUT}}$'	*Hammer*, hammer$_{\text{-MASC}}$' *Säge*, saw$_{\text{-FEM}}$' *Schraubenzieher*, screwdriver$_{\text{-MASC}}$'

Examples of superordinate neuter correlation.

In the animal and plant kingdoms they find the same results, but with an interesting new variance. As expected, several superordinate animal terms are also neuter, such as *Tier* 'animal$_{\text{-NEUT}}$', *Säugertier* 'mammal$_{\text{-NEUT}}$', *Reptil* 'reptile$_{\text{-NEUT}}$', *Insekt* 'insect$_{\text{-NEUT}}$ and *Weichtier* 'mollusk$_{\text{-NEUT}}$'. As superordinate terms, they are generic and do not specify a particular base level member of the semantic category that they head. That is, one cannot imagine a generic 'mammal'; they must rather imagine a base level term such as 'elephant' or 'bear' which fits under the 'mammal' superordinate. Köpcke & Zubin tie the idea of 'imaginability' with gender: superordinate terms which are generic enough to not be able to be imagined are neuter.

There are some superordinates, however, that can be generically imagined. Both birds and fish can be generically imagined. Birds all have tails, feathers and beaks, whereas fish all have gills and scales and fins. See (5) below:

(5)

			Imaginable generic member?
Neuter gender superodinates	Säugertier	'mammal$_{\text{-NEUT}}$'	no
	Reptil	'reptile$_{\text{-NEUT}}$'	no
	Insekt	'insect$_{\text{-NEUT}}$'	no
	Weichtier	'mollusk$_{\text{-NEUT}}$'	no
Non-neuter gender superordinates	Vogel	'bird$_{\text{-MASC}}$'	yes
	Fisch	'fish$_{\text{-MASC}}$'	yes
	Krebs	'crustation$_{\text{-MASC}}$'	yes
	Wurm	'worm$_{\text{-MASC}}$'	yes

Superordinate-imaginability correlation.

In their (1986) study, Zubin & Köpcke conclude that the neuter gender is often associated with categorical *neutrality*. Therefore, superordinate terms that cannot be a specific member of

their category, such as 'mammal' and 'reptile' are neuter. However, some superordinate level terms, such as 'bird', 'fish', 'crustation' and 'worm' *are* imaginable and are correspondingly not neuter, as can be seen in the last four rows in (5) above. The non-neuter superordinates are not categorically neutral, but differentiable. The sex-associated genders, masculine and feminine, refer to fully differentiated taxa, whose shape and specifiable parts are fully imaginable and distinguishable from other taxas. Moreover, they observe that new words, such as *Flugzeug* 'airplane$_{\text{-NEUT}}$', are falling into their predicted gender and build their structure around neuter-specifying suffixes when necessary. The connection between superordinate level words and gender will be examined further for Italian (Thorton 2007) in section 3.1.

1.2.2 Shape-based assignment

There has been research that investigates to what extent phonological regularities play a part in the gender outcome. There are several studies that conclude that gender rules are not just influenced by meaning, but also are a consequence of word shape.

In their (1977) study on French, Tucker, Lambert & Rigault identify strong correspondences between shape and gender, shaking the assumption that the French gender system is one of the most opaque. By analyzing only the final phones, they are able to correctly assign gender to 84.5% of the 31, 619 nouns in the *Petit Larousse.* These findings are quite impressive, given that, as Corbett (1991) points out, they view any semantic factors, such as sex, to be an exception to their shape based system.

They find, for example, that 94.2% of words with final /ʒ/ are masculine, such as *ménage* 'housekeeping'. There is a 90 % correspondence between words with final /z/ and the feminine, such as *église* 'church'. There is a 100% masculine correspondence with words ending in /œ̃/. Such phonological rules in conjunction with semantic rules based on sex are able to account for the majority of French nouns.

Zubin & Köpcke (1983) and Köpcke (1982) also present evidence for a range of gender assignment rules for German based on phonologic shape. While some of these rules are absolute, such as diminutive endings *-lein* and *-chen* are nearly always neuter[6], others are more probabilistic. For example, words that begin *#/ʃ/+C,* tend to be masculine (e.g. *Schuh*

[6] Robinson (2007) correlates variance in gender in Grimm & Grimms' fairy tales rooted in the behavior of girls in their stories. It is expected, and substantiated in later versions of their stories that *Mädchen 'girl$_{\text{-NEUT}}$'* be referred to

'shoe$_{\text{-MASC}}$') and words that end *(C) {f, x, ç}+t#* tend to be feminine (e.g. *Nacht* 'night $_{\text{-FEM}}$', *Frucht* 'fruit$_{\text{-FEM}}$)

Zubin & Köpcke and Tucker, Lambert & Rigault's findings show a strong correlation between phonological shape and gender. They contend that it is the shape of the word that determines the gender.

1.3 Recent analyses of gender

Roca observes gender from the other way around. In his two part essay, Roca (2005) offers a full analysis of the Spanish gender system. Like Mascaró (1986) for Catalan and Harris (1991) for Spanish, he assumes the approach wherein nouns have idiosyncratic gender. Shape is therefore a consequence of the inherent gender of each noun. Shape becomes relevant for gender only to new words because they do not have inherent gender.

Roca puts forth evidence against Grimm's (among others) idea that gender is a reflection of sex. If this were so, then *mesas* 'tables', *flores* 'flower' and *plumas* 'feathers' would all be "females" while *papeles* 'papers' and *árbol*es 'trees' would be "males", which, in his words "is absurd." The naturalness of gender must be sought outside of sex. A word like 'mesa' is the feminine gender not because it is a woman, but because, having inherent gender, it shows feminine concord. In Roca's account, gender is a binary feature and intrinsic to each word as (+/- F).

Furthermore, word endings are not responsible for the gender of words. Therefore, although *-o* is a typical Spanish masculine ending, it can also appear in feminine words, such as in *mano* 'the hand$_{\text{-FEM}}$'. Similarly, words with the typical feminine word ending, *-a*, can be masculine, such as *tema* 'topic$_{\text{-MASC}}$'. Moreover, some words can refer to both males and females with the same, feminine shaped ending; e.g. *la Belga* 'the Belgian woman' and *el Belga* the Belgian man'.

Mascaró (1986), Harris (1991)and Roca (2005), among others believe that the direction is not word ending > gender, as is postulated by Köpcke & Zubin (1983) but precisely the opposite; the direction is gender > word ending. Gender is an idiosyncratic property of every nominal morpheme; it forms a part of each lexical entry.

---

by the neuter pronoun *es*. In earlier versions of their stories are sometimes referred to with the feminine pronoun *sie*. He observes an association between badness and sexual availability and the use of *sie*.

Roca, among others, says that gender is marked as a lexical feature. Therefore, both feminine (*casa* 'house$_{-FEM}$') and unfeminine shaped words (e.g. *pared* 'wall$_{-FEM}$') alike must be marked as feminine. In parallel, masculine shaped (e.g. *sol* 'sun$_{-MASC}$') as well as feminine shaped words (e.g. *poeta* 'poet$_{-MASC}$') must be marked as [-F]. The default agreement in Spanish is masculine [-F] and the word ending, unless otherwise specified, is /o/. Similarly, if a noun is lexically specified to be feminine, its word ending, unless lexically specified, is /a/. Under this view, the question that arises is how gender is assigned to nouns without a specified gender because they are new words. This is the issue that is addressed in section 2 for Catalan.

In his (2005) paper, Rice sidesteps the issues Roca brings up by suggesting that the problem of gender can be resolved resorting to the constraint system of Optimality Theory. He demonstrates that his *Optimal Gender Assignment Theory (OGAT)* can resolve gender conflicts in Russian and German. Following Rice, Thorton (2007) reworks Rice's constraints for Italian and shows that there is a hyperonym effect for all members of the category 'car'. That is, all words that belong to the 'car' category (e.g. Toyota, Mercedes, etc.) inherit their gender from their hyperonym, 'car'.

1.4 Assigning Gender: Loan Words

Corbett (1991) says that as loan words enter the language, each one is assigned a gender based on the semantic and phonological regularities of the borrowing language, but there are several earlier analyses that suggest that loan words are sometimes treated differently.

In an early analysis by Ibrahim (1973), differing criteria for gender assignment are presented depending on the type of loan word. If the loan word refers to an *animate* being, semantic gender assignment takes precedence over all other possible factors. Males are assigned the masculine gender; females are assigned the feminine gender. If the loan word is *inanimate* there are several principles that have been observed which determine gender assignment. Ibrahim's gender assignment criteria are summarized in (6).

(6)
a. **semantic gender**: biological sex
b. **analogical gender**: loan word takes the gender of a near-synonym in borrowing language)
c. **gender of the suffix**: if it resembles any gender assigning suffixes in the borrowing language

d. **homophony and rhyme**: if a borrowed word sounds like a word in the borrowing language, it assumes that word's gender
e. **default gender assignment**
f. **gender of the word in the source language**
g. **differentiation of meaning**: such as *das Band* 'the rubber band', *die Band* 'the military band'

In a subsequent summary of gender assignment criteria to loan words, Hock (1986) trims it down to only four assignment parameters for the gender of loan words. Hock's gender assignment criteria are summarized in (7)

(7)
a. **semantic-based** assignment (e.g. sex)
b. **shape-based** assignment
c. **analogical gender**
d. **default gender**, resorted to, only if the other criteria cannot be employed. He ignores Ibrahim's (6g).

Hock's data came primarily from borrowings into German, which has a three gender system, masculine, feminine and neuter, while French has a two gender system, feminine and masculine. English, on the other hand has a 'natural' or 'sex' gender system, but has no 'grammatical' gender division; therefore the alignment of gender across languages, (6f) is not always possible. Gender assignment to loan words by influence of the original gender, thus, is not the norm, but rather, as Ibrahim (1973) also reports, a very rare exception. Nevertheless, it has been observed on some occasions.

German has many borrowings from Greek and Latin ending in *-a*, such as *die Prima* 'top or senior class at school' and *die Aula* 'the great hall', which were feminine in their native languages. When they were adopted into German, the feminine gender remained, setting a precedence for the correspondence between words ending in *-a* and the feminine gender. These were conveniently able to retain their feminine gender. Likewise, the masculine Sanskrit word *vēda* 'traditional/sacred knowledge', tends to be feminine in German because it ends in *-a*. However, Sanskrit-savvy speakers assign *vēda* the masculine gender, reflecting their knowledge of the original gender of the word.

Shape based assignment (7b) occurs when a part of the morphology or phonology influences the gender assignment. Such is the case for the word *computer*. It was easily placed

into the agent or instrument noun category for words ending in *-er*, as in *Bäcker* 'baker' and *Schaber* 'scraper'. These types of nouns are masculine in German, so *computer* is assigned the masculine gender.

Parameter (7d) can be observed in the borrowing from English, *Beatnik.* The term *Beatnik* is usually used to refer to males, allowing influence from parameter (7a), but also, the masculine gender assignment is the usual gender for words referring to people whose sex is not specified.

The term *Trend* is assigned its gender by parameter (7d). Hock says that because of the shape of *Trend,* feminine native near-synonyms like *Anlage* or *Tendenz* can automatically be discounted. Words ending in consonants in German are usually masculine and neuter, but there are no formal criteria to guide us to which one. This is when, according to Hock, parameter (7d), goes into action and a semantically related native word is sought. There are two masculine native near-synonyms, congruent with the shape of *Trend* that can be used: *Zug* and *Hang.* Because the masculine gender is the unmarked in German, *Trend* is assigned the masculine gender.

As noted above, sex has been found to be the overriding factor in the prediction of grammatical gender in loan words; see for example, Ibrahim (1973) Poplack, et al. (1982), Corbett (1991), DuBord (2004), Roca (2005), and Morin, (2006). What is less clear is how meaning and word shape interact in other cases.

Poplack, et al. (1982) carried out an early study on English loan words in Puerto Rican Spanish in New York and Montreal French in Canada. They found that word shape overrides meaning in gender assignment to loan words in both languages. They also conclude that masculine cannot be named the unmarked form in loan words in Spanish. This is because very few borrowings from English end in *-a*, the feminine marker in Spanish, so there isn't sufficient data to support an unmarked masculine gender.

In German, shape has also been found to be a significant factor. With a set of rules: 24 phonological, 5 morphological and 15 semantic, Köpcke (1983) was able to predict the gender of 90% of the 1466 monosyllabic words in a volume of *Duden,* a German dictionary. Köpcke & Zubin (1983) substantiate these findings with an experimental study. In a forced-choice task, made-up words were presented with two determiners. For example, die Pucht 'Pucht$_{\text{-FEM}}$' and

das Pucht 'Pucht-$_{NEUT}$'. Words of this type were assigned the feminine gender in 73% of the cases in accordance with the previously mentioned, feminine assigning rule, *(C) {f, x, ç}+t#.* This research supports the view based on phonological and morphological cues.

Morin (2006) agrees that sex is the overriding factor in gender assignment to animate referents in her study of computer and internet-related loan words into Spanish. This assertion follows Roca's (2005), among others, observation that the masculine is the unmarked gender of animates. In contrast, Poplack et al. (1982), finds that word shape is not the primary consideration in the assignment of gender to loan words. By itself, shape can only account for 39% of the masculine words. Word shape can work in conjunction with semantic considerations to bolster a masculine assignment.

The 'semantic' assignments in her study are based on the analogy to Spanish language synonyms (native or loan translations). Therefore, words like *Playstation* ('consola$_{-FEM}$' in Spanish) and *toolbar (*'barra$_{-FEM}$ de herramientas' in Spanish*)* acquire their gender from their Spanish synonyms. The analogical gender was attributable to 87.5% of the feminine assignments and 61% of the masculine assignments.

In words of variable gender in her study, the gender comes from the underlying Spanish noun: el (archivo) mp3, but *una (página) web* vs. *un (sitio) web*. When *web* means *página,* it is feminine and when it means *sitio* it is masculine. This is called metonymic gender and accounts for 29% of the gender assignments found in part 2 of this study.

If the semantic criteria are feminine and the shape is masculine, the two assignment criteria conflict. Morin does not offer an explanation of how the conflict is resolved in Spanish, but instead points out that when these conflicts arise, the data reflect these ambiguities. Take, for example, *las/los cookies* (*galletas*), both the masculine and feminine variants can be found in different dialects of Spanish. The un-Spanish spelling gives rise to the masculine variants while the analogy with *galletas* gives rise to the feminine variants.

1.5 Early attempts to order constraints

In early attempts to offer gender assigning constraints, there was not much attempt to rank them. Both Ibrahim (1973) and Hock (1986) agree that sex is the overriding factor for animate beings. Hock says that the default is turned to, only when the other parameters give no definitive assignment. Hock's seminary analysis leaves a very important question unanswered:

in which circumstances are each of these 4 paradigms activated. The only answer he offers is the default gender is used when words "are not classifiable by other criteria." But there is no analysis of the semantic vs. shape conflict.

An early formal attempt to sort out the discrepancies between semantic- versus shape-based gender assignment appeared in Steinmetz (1985) and (1986) work on German. He suggests a hierarchy for all the rules put forth by Köpcke & Zubin outlined above. According to his theory there are two processes: 1) *Gender Tally*, which counts the number of times each gender is assigned and selects the gender of the highest value and 2) *Gender Eclipsis,* which orders the genders in a hierarchy, so that if two genders have the same *Gender Tally*, the higher gender wins. He suggests German's hierarchy is ***masculine > feminine > neuter,*** but this varies by language. This means that if there is no marking for feminine or neuter gender, the gender of the word will be the unmarked masculine.

According to *Gender Tally***,** there are several semantic rules, morphological and subcategorization rules. His subcategorizaton rules apply within the domain of semantic rules. For example, although it is a hyperonym, the semantic group *fruit* does not correspond to the neuter gender. Therefore if there are no criteria vying for another gender, the gender is masculine, as is seen in (8e) below. However, a subcategorization of *fruit, tropical fruit,* corresponds to feminine gender; that is all words for *tropical fruits* are feminine. Section 3.1 expands on this with Thorton's (2007) introduction of a hyperonym effect in optimality theory. In the (8) below, Ø is a zero-category rule, which specifies that *fruit* does not entail a specific gender. Zero-category rules are the most common in Steinmetz's theory. The criteria are all rank evenly and the gender with the highest tally wins. For an example of *Gender Eclipsis,* see (9b).

(8)

Type of rule	**Type of rule**	**Gender correspondence**
semantic	superordinates	neuter
	male	masculine
	female	feminine
	fruit	Ø
Morphological	*-e*	feminine
	ge-	neuter
	-nis	feminine or neuter
	-ucht	feminine
Subcategorization	tropical fruit	feminine

Exemplary gender assigning rules.

(9) (a) *Gefängnis* 'jail'
-nis =f, n
ge- = n
0m 1f 2n = n

(b) *Obst* 'fruit'
superordinate = n
0m 0f 1n = n

(c) *Frucht* 'fruit'
superordinate = n
-ucht = f
0m 1f 1 n = f

(d) *Dattel* 'date'
fruit = Ø
SC: tropical = f
0m 1 f 0n = f

(e) *Apfel* 'apple'
fruit = Ø
0m 0f 0n = m (f)

(f) *Pflaume* 'plumb'
-e = f
fruit = Ø
0m 1f 0n = f

In example (9a), Gender Tally can be observed. There are more criteria vying for the neuter than for the feminine, so the neuter is the winner. Example (9c) shows the dominance hierarchy; both feminine and neuter have 1 criteria vying for the gender, but because the feminine is higher, it wins.

1.6 Recent attempts at constraints

DuBord (2004) offers a simple answer to the question of assignment criteria dominance, which was left open by Morin (2006): what happen when gender assigning criteria clash? DuBord investigated the gender of English loan words in spontaneous Southern Arizonan Spanish, agreeing with that biological sex, when available, is the best predictor of grammatical

gender. But in cases where phonological and semantic (analogical) criteria for gender assignment clash, the phonological form prevails.

Section 3.1 of this paper examines Rice's 2005 *Optimal Gender Assignment Theory* (OGAT), which draws together Steinmetz's (1985) *Gender Tally* and *Gender Eclipsis* with Prince, Alan, Smolensky and Paul's (1993) Optimality Theory. He proposes gender feature constraints be crucially un-ranked.

Thorton (2007) analyzes Italian's gender system, looking for a dominance relation between shape- and semantic-based gender assignment criteria. She rejects Rice's claim that gender feature constraints should be unmarked, citing cases in Italian where it produces the wrong optimal form. She finds that semantic rules override all others in Italian. She also observes many cases of the hyperonym effect, such as in car names, all of which are feminine in Italian. For more on Thorton's analysis, see section 3.1.

2. Gender Assignment in Catalan

2.1 Introduction

There is a strong relationship between morphology and gender in Catalan. The morphs that correspond to gender in Catalan are similar to those of Spanish. The masculine has allomorphs Ø (e.g. *dit* 'finger$_{\text{-MASC}}$'), [u] (e.g. *suro* 'cork$_{\text{-MASC}}$'), and for Eastern dialects, [ə] (e.g. *pare,* 'father$_{\text{-MASC}}$' *poeta* 'poet$_{\text{-MASC}}$'). There are also a few very marked ones (Mascaró 1986).

The feminine ending has the allomorphs [ə], *(e.g. roba* 'clothes$_{\text{-FEM}}$'*),* Ø (e.g *sal* 'salt$_{\text{-FEM}}$'). There are many derivational suffixes that are assigned the feminine gender {*–tat/-dad}, -ió, -(it)ud}.* Central Catalan, orthographic *-e* does not always correspond to the feminine gender (Mascaró 1986), as is discussed below.

Most pertinent to this study is the pronunciation of word final vowels. Word final [o] is normally reduced to [u] as in *trasto* 'junk$_{\text{-MASC}}$' and *musclo* 'mussle$_{\text{-MASC}}$.' In most cases, word final [a] and [e] are both reduced to [ə], as in *cadira* 'chair$_{\text{-FEM}}$' and *llibre* 'book$_{\text{-MASC}}$'. However, as has been pointed out (see, for instance; Bonet & Lloret 1998 Cabré Monné 2002) in some cases in Central Catalan these reductions don't take place, as is discussed later in this section.

Cabré Monné (2002) discusses various gender assignment criteria. She cites analogy with a native Catalan word (e.g. *dream team* is masculine because of the association with *equip*

'team$_{\text{-MASC}}$', whereas *trattoria* 'restaurant$_{\text{-FEM}}$' is feminine because it looks like other feminine Catalan words that carry the gender marking *-a*, such as *cerveseria* 'pub$_{\text{-FEM}}$' and *carnisseria* 'butcher´s shop$_{\text{-FEM}}$'. Moreover, foreign word endings are often re-interpreted as Catalan word endings. Therefore words such as *cassette* 'cassette$_{\text{-MASC}}$', and *suèter* 'sweater$_{\text{-MASC}}$' are interpreted to have the masculine Ø morph.

The incorporation of loan words into the Catalan phonology is another question entirely. The pronunciation of loan words in Catalan depends on several factors. One is how the word enters the language: whether it be through written or spoken language. If it enters through written language, speakers often use spelling pronunciation (Cabré Monné 2002). If it enters through spoken language (e.g. the media), it mimics the original phonologic form of the word as closely as possible. This has orthographic repercussions when eventually codified (Cabré Monné 2002).

In Central Catalan, generally only [i], [u] and [ə] appear in unstressed position. Cabré (2006) states that many loan words ending in *-o* that have entered Catalan from Spanish have been fully phonetically adapted. Therefore, words like *pago* 'pay$_{\text{-MASC'}}$', *regalo* 'present$_{\text{-MASC'}}$', *barco* 'ship$_{\text{-MASC}}$' end in [u] in Catalan.

However, it has been observed that, in Central Catalan, orthographic *-e* is often not reduced to [ə] and *-o* is not reduced to [u] in words borrowed from a recent ancestor, such as *classe* 'class', *base* 'base', *frase* 'sentence' *cràter* 'crater', *túnel* 'tunnel', *examen* 'exam' and *soprano* 'soprano', *ego* 'ego' , *zoo* 'zoo', *esperanto* 'Esperanto', *libido* 'libido', and *colon* 'colon'(see Mascaró 2002 and references therein). The same phenomenon is present in many proper names such as *Balmes, Pedralbes, Oslo, Boston,* loan words such as *vàter* 'toilet' and *bàsquet* 'basket ball', and *clàxon* 'horn' and in acronyms like *RENFE, BOE, UNESCO, CIO* and truncated forms such as *cine* 'cinema' , *video* 'film', *moto* 'motorcycle'.

It has been noted that languages often have default genders (see for example, Steinmetz (1985) for German and Icelandic, Wechsler (2009) for Serbo-Croatian, French and Icelandic, Roca (2005) for Spanish). As Bonet and Cabré (2002) note, Catalan's default gender is masculine. This is observable when gender conflict in coordination gives rise to masculine concord as in (10a-b)

10 (a) La paret i el sostre són blancs.

The wall$_{\text{-FEM}}$ and the ceiling$_{\text{-MASC}}$ are white$_{\text{-MASC}}$

(b) *La paret i el sostre són blanques.
*The wall$_{\text{-FEM}}$ and the ceiling$_{\text{-MASC}}$ are white$_{\text{-FEM}}$

2.2 The study

The process of gender assignment can only be observed when speakers come across new words. These words can be borrowings from other languages or coined for specific purposes.

To investigate the interaction of the 3 gender assignment criteria: default gender, assignment based on shape, assignment based on semantics, a questionnaire was used to collect data from Catalan speakers.

To gather data about gender assignment for this study, a group of participants were presented with words and asked to pick between feminine and masculine adjectives, articles and pronouns. Because gender is observable through concord, we are able to discover what gender participants thought each word to be by their selection. Concord is a consequence of the noun's lexical gender.

2.2.1 Methodology

Data were collected from thirty-four native Central Catalan speakers who completed the questionnaire. Participants included 22 females and 12 males between the ages of 14 and 80. Participation was voluntary.

The questionnaire was composed of 103 questions divided into three sections: made-up words with Catalan word form (e.g. *pras, suna, truse)*, product names of cars (e.g. *Mazda*), toothpaste (e.g. *Colgate*) and soft drinks (e.g. *Sprite*), and loan words from Turkish, Chinese, Arabic, English, German, Japanese, Italian, Hindu, French, Spanish. The made-up Catalan words referred to "new inventions", whose patents, participants were told, were still pending. As crazy as some of the inventions may sound, many of the new inventions actually do exist and were taken from the website www.WackyInventions.com.

The words were embedded in sentences, which were given to provide a context for the words. They appeared as questions, statements and exclamations. In each sentence, the participants had to select the gender of the word through concord with possessive pronouns, articles or adjectives. In every question, both gender options were given. In half of the questions

the feminine option came first and in the other half, the masculine option came first. The order of the questions in each section was randomized so that the same word shape was not consistently followed by a similar one. A representative example question is given below:

(7) Han portat **el / la** Panda al desguàs.

2.2.2 Product names in Catalan

This part of the questionnaire included three distinct semantic groups: cars, soft drinks and toothpaste. It has been observed by Steinmetz (1985) Köpcke & Zubin (1986) and Thorton (2007), among others, that some semantic groups have a hyperonym effect. As we will see, car names and most likely motorcycle names have the hyperonym effect. Toothpaste and soft drink brand names appear to have their gender assigned primarily by their shape.

If it is the shape of the word that is the main influencer of gender, then we should expect that participants prefer to pair feminine shaped words (e.g. *Honda* and *Fanta*) with feminine adjectives and articles.

Many Catalan commercial product names have come from foreign languages but do not adopt the original pronunciation of the word. This is due to phonological constraints of the native sound system, but also the environment in which the word is learned. As mentioned in 2.1 above, proper nouns ending in *-e* are often pronounced as [e] and not [ə] as Central Catalan dictates. This ending is investigated further in 2.2.4 on made-up words.

Another possibility that may affect the gender of these groups of words is the gender of their hyperonym (*cotxe* 'car$_{\text{-MASC}}$', *pasta de dents* 'toothpaste$_{\text{-FEM}}$', *beguda* 'beverage$_{\text{-FEM}}$'). That is, the hyponyms such as *Honda* would inherit the gender from their hyperonym *cotxe* 'car$_{\text{-MASC.}}$ If this is the case, we should see car names consistently paired with a masculine adjective or article, regardless of their seemingly feminine endings.

2.2.2.1 Car names in Catalan

The *c*ar names used in this study come from other languages. A total of 10 car words were used: *Opel, Fiat, Saab, Mercedes, Audi, Ford, Honda, Mazda, Nova, Panda.* With the

exception of *Honda*, all of these brand names are used exclusively for cars. *Honda* is often used to refer to motorcycles as well. See appendix 1 for the full questionnaire.

Results

The cars with masculine shaped names were found to be consistently masculine by nearly every participant. Participants paired 93% of the masculine shaped car makes with masculine concord. Likewise, cars with feminine shaped names were also very often paired with masculine options, averaging 89% masculine. *Mazda* and *Panda* both had a 94% masculine concord rate. Both *Honda* and *Nova* have slightly lower masculine ratings, 86% and 80% respectively.

Discussion

It appears that the hyperonym 'car' affects the grammatical gender in cars. The masculine shaped words were overwhelmingly paired with masculine concord. Overall, the feminine shaped words had a slightly lower masculine rating, but still only 4% lower than their masculine shaped counterparts. This may suggest that word shape also exerted an influence. Car names with percentages of gender assignment are listed in (11).

(11)

Masculine-shaped car names	**Percent masculine**	**Feminine-shaped car names**	**Percent masculine**
Opel	94%	Honda	86%
Fiat	91%	Mazda	94%
Saab	91%	Nova	80%
Mercedes	94%	Panda	94%
Audi	91%		
Ford	94%		
average	93%	average	89%

Car words and their masculine percentages divided by shape.

Both *Nova* and *Panda* exist in the Catalan lexicon outside of car names. *Nova* 'new' is a feminine adjective. *(Os) Panda* 'panda bear' is a masculine noun. *Honda* and *Mazda* are not lexical entries outside the realm of cars.

Honda averaged 86% masculine pairings and it is possible that the slightly lower masculine rating might be due to the high frequency of *Honda* being used to refer to motorcycles, which in Catalan is *motocicleta* 'motorcycle-$_{\text{FEM}}$'. This suggests that *motorcycle*

also induces a hyperonym effect. The case for hyperonym effect in motorcycles is explored below.

Nova had the lowest masculine rating of all, at 80%, which is not surprising because, as a feminine adjective, it only ever appears with feminine concord. Residual effects of agreement in other contexts are likely playing a part in *Nova*'s low masculine rating.

That leaves two feminine shaped car makes: *Mazda* and *Panda,* both of which were among the highest of car words to be paired with masculine concord at 94%. There is no pull from the *-a* towards the feminine gender and thus no shape effect.

Although it was not tested in this study, there seems to be a hyperonym effect for *motorcycles* as well. The word for motorcycle in Catalan, *motocicleta,* is feminine and in common usage, all motorcycles names are also feminine, such as in (12).

(12)

BMW	Derbi
Suzuki	Vespa
Yamaha	Harley Davidson
Kawasaki	Ducati
Honda	

This is true regardless of any masculine shape. Moreover, even when there is a possible sex referent available compounded with masculine shape, such as with Harley's *Fat Boy,* the hyperonym effect still overrides, giving the feminine gender.

Summarizing, word shape (i.e. the final *-a*) does not seem to have a much of an effect on concord in car or motorcycle names. The combination of the Catalan's default masculine gender, possibly in combination with the hyperonym effect of *'car'* resulted in a very high masculine rating. The gender of motorcycles also seems to be governed by the hyperonym effect. The hyperonym effect for cars has also been attested for in Italian. See section 3.1 for Thorton's (2007) analysis.

2.2.2.2 Soft drink names

This study also tested 12 soft drinks for the hyperonym effect: *Aquarius, Cacaolat, Colacao, Sprite, Schweppes, Nestea, Seven Up, Red Bull, Bitter Kas, Coca-Cola, Fanta, Trina.* In Catalan, the word for 'beverage' is feminine, *beguda.* Therefore, if there is a hyperonym effect, we should expect beverages be feminine across the board, despite any masculine form, just as is observed feminine shaped cars be assigned the masculine gender in the section above. See appendix 2 for the full questionnaire.

Results

Words for these beverages generally followed shape criteria**.** Of the twelve product names, seven ended in consonants (including *Sprite* [əsprájt], whose original word final [-t] has been preserved despite the orthographic *-e*). Data with percentages of gender assignment are listed in (13).

(13)

Masculine-shaped soft drink names	**Percent masculine**	**Feminine-shaped soft drink names**	**Percent masculine**
Sprite	89%	Coca-Cola	0%
Aquarius	77%	Fanta	3%
Cacaolat	94%	Trina	46%
Colacao	86%		
Nestea	94%		
Schweppes	17%		
Seven Up	91%		
Red Bull	89%		
Bitter Kas	77%		
average	79%	average	16%

Soft drink names and their masculine percentages divided by shape.

Schweppes is a seeming outlier and with a masculine shape but very low masculine rating of just 17%, it doesn't fit into the masculine shaped group.

Coca-Cola and *Fanta* were all treated as feminine (except for one person for *Fanta),* suggesting that *-a* was used as a criterion for feminine concord. *Trina*, on the other hand, was evaluated by 46% to be masculine despite its feminine shape.

Discussion

The ambiguous gender of *Trina* is likely due to a historic change in its name. Originally, it was called *Trinaranjus* and was later truncated to *Trina.* This change is reflected in the data. The majority of the 22 participants in the youngest group, age 14-24, assigned *Trina* the feminine gender. In fact, only 33% selected the masculine. In the next age group, the four participants aged 25-39 assigned mostly the masculine gender. Only one person in that group selected feminine concord. The masculine rating in the group aged 40-54 went up slightly to 80%. In the oldest group, which contained only 3 people, the masculine rating went down again to 33%. These results appear in (14).

(14)

Age group	**% masculine**
14-24	33%
25-39	75%
40-54	80%
55-80	33%

Masculine ratings of Trina divided by age.

The results for *Trina* suggest that the youngest group accepts the truncated form of *Trinaranjus* to be the actual word, whereas the older group does not. It is likely that the youngest group doesn't even know the original word.

People often say '*Tònica Schweppes*', and 'tònica' is a feminine word in Catalan. This can shed light on why *Schweppes* has a low masculine rating, despite the trend for soft drinks' gender to be assigned based on shape. Speakers are actually referring to a feminine word.

To sum up, beverages in Catalan are not subject to the hyperonym effect, but instead are influenced by their phonetic shape.

2.2.2.3 Toothpaste brands

There were 9 toothpaste brands tested for gender assignment criteria: *Colgate, Signal, Close Up, Gingilacer, Kemphor, Profidén, Sensodyne, G.U.M.* and *Binaca.* The goal of this part was the same as the others: to determine which gender assignment rules take precedence and whether the hyperonym effect is present. See appendix 3 for the full questionnaire.

Results

Toothpaste brands were more often than not assigned masculine gender, on average 73%. The only feminine shaped toothpaste brand, *Binaca,* had a masculine rating a bit lower than average at 46%. The word final *-e* in *Colgate* and *Sensodyne* did not seem to have a clear effect. Colgate had the highest masculine rating of the toothpaste group at 86% while *Sensodyne* had a much lower 68%. These results appear in (15).

(15)

Masculine shaped toothpaste names	**Percent masculine**	**Feminine shaped toothpaste names**	**Percent masculine**
Colgate	86%	Binaca	46%
Signal	77%		
Close up	68%		
Gingilacer	77%		
Kemphor	66%		
Profidén	71%		
Sensodyne	69%		
G.U.M	68%		
average	73%	average	46%

Toothpaste names and their masculine percentages divided by shape.

Discussion

Most of the soft drink names end in segments other than *-a*, which is consistent with the Ø masculine gender marker in Catalan (Mascaró 1986), and might explain the overall tendency towards the masculine gender in this semantic group. However, unlike with cars, the hyperonym 'toothpaste', which in Catalan is *pasta de dents*, seemed not to dictate that all its hyponyms be feminine. Interestingly though, we do see that as a group, the toothpaste brands have a lower masculine rating than cars, despite both having 6 Ø morph items. It is possible that 'toothpaste' is indeed exerting its hyperonym effect, but it is competing against two factors that are vying for the masculine gender: default gender and word shape.

Indeed, we do see that the sole feminine shaped item, *Binaca*, has a lower masculine rating than the rest. However, if it has two factors for the feminine gender, hyperonym effect as well as the feminine shape, why do we not see a lower masculine rating? That is not immediately clear. It could be due to the effect of the one remaining masculine factor: masculine as default.

2.2.3 Loan words

The existing loan words used in the questionnaire were very low frequency for the general public and have come from Turkish, Chinese, Arabic, English, German, Japanese, Italian, Hindu, French, and Spanish[7]. The majority of the words have prescribed genders cited in Neoloteca, a neologism library, which includes all the terms normalized by the supervising board of Termcat. Some words used were not found in Neoloteca, but through Google searches. No word used in the study was listed in Neoloteca as having variable gender. In total 63 existing loan words are analyzed.

This part of the study investigated whether there is a semantic effect for foods, beverages and musical instruments. The words were presented first in sentences and their definitions were given underneath in smaller print. All instruments' definitions started with "instrument…". Drinks' definitions all started with "una beguda…". Those semantic groups would have a fixed hyperonym. However, food's definitions started in various ways "plat…" "menjar…" "formatge…" "sopa…" and so on. When available, definitions were taken from Neoloteca. Occasionally the longer definitions were condensed. A representative example question is given below in (16). See appendix 4 for the full questionnaire of foods and drinks and appendix 5 for full questionnaire on instruments.

[7] For example, *bongos* from Afro-Cuban Spanish.

(16) El falafel l'hi va donar amb pita **dur / dura.**
☐ *conec la paraula* ☐ *no la conec*

Pita - Pa de forma rodona o ovalada, molt prim i buit, que s'obre com una butxaca i pot omplir-se amb ingredients diversos, típic de la Mediterrània oriental i del Pròxim Orient.

2.2.3.1 The data

Semantic gender assignment did not seem to play a role in the food and drink group. There was not any hyperonym effect for *beverage* on any of the drinks. In fact, none of the beverages fell into the feminine category (i.e. under 50% on the masculine scale). The closest was *boza* at 58% masculine. Full results can be found in Appendix 6.

When the food and drink and the instrument groups are combined, there are 12 words with final orthographic *-a***.** Ten of the words were paired with feminine concords more than 50% of the time. There were four interesting exceptions.

Neoloteca prescribes that *tahini* be feminine, but was rated to be masculine by 86% of participants. On the other hand, *tambura* and *biwa* are listed in Neoloteca as masculine, but were given a very low masculine percentage: 17% and 28%, respectively. Sitting right at 50%, *taiko* was also proscribed to be masculine, but did not have a convincingly masculine showing.

Discussion

Participants were not consistent in their answering criteria. For the five loan words ending in *-o*, (*asiago, miso, koto, taiko* and *bongos)* there was intra-speaker variation. That is, most, but not all answers, were consistent with masculine concord. For every person, there was an average of 15% answers that were feminine.

Of the twelve feminine shaped words, there was also variance in answers. About 30% of the answers per person varied from their normal answering pattern, which was feminine concord.

Words ending in *-e* also had irregular answering criteria. That is, participants averaged about 37% intra-speaker variation from the feminine answers out of the two words ending in an orthographic *-e*.

To conclude, we have seen that loan words that refer to foods, drinks and instruments seem to be influenced more by word shape than by semantic factors. With substantial intra-speaker variation however, we cannot assume that the shape of the word is not the definitive criterion.

2.2.4 Made-up words in Catalan

The made-up words used were created specifically for this study. A total of 28 words were included in this section, all of which have possible native Catalan shape, but are not identical to any existing members of Catalan vocabulary. The presentation of each question is as follows: first, a definition was given of a "new invention", which was said to have just made its debut in Catalonia. Then, like in the rest of the questionnaire, the word was embedded in an affirmative sentence, question or expression (see appendix 7 for the full questionnaire).

(17) Aparell que neteja la pols de la prestatgeria dels llibres: suna
Quina pena! **Aquella / Aquell** suna es va espatllar.

This section of the study investigated two gender assignment criteria: semantic gender assignment through hyperonym and shape-based gender assignment by the form of the word. To test for the hyperonym effect, the inventions were presented as either a *màquina* 'machine $_{\text{-FEM}}$' or *aparell* 'apparatus$_{\text{-MASC}}$'. If the hyperonym makes an impact on gender assignment in these made- up words, we should expect to see a relationship between the type of invention (either a feminine *màquina* or a masculine *aparell*) and the gender participants assigned the word.

If shape has an effect on the gender assignment of the word, we should expect words ending in a consonant or a stressed vowels (e.g. *sablè*, *rou*) to be assigned the masculine gender and words ending in *-a* to be assigned the feminine gender. Other words of 'ambiguous' gender are those ending in *-e*. As mentioned above, *-e* is often pronounced [e] in new words and has no gender connection. It is frequently observed in both feminine words (e.g. *classe* 'class $_{\text{-FEM}}$') as well as masculine words (e.g. *mate* 'type of drink-$_{\text{MASC}}$'). The role that the ending *-oma* will play is unclear from the onset. It is always masculine in Catalan (e.g. *fibroma* 'fibroid$_{\text{-MASC}}$' and *glaucoma* 'glaucoma$_{\text{-MASC}}$'). However, it is thought that participants might not look further than the word final *-a*, a definitive feminine marker. These words were therefore analyzed separately. Participants had access to only the written form of the made-up words.

The data

There were 7 types of word combinations in the questionnaire: (20a) feminine-shaped words paired with the feminine word *màquina* 'machine', (20b) feminine-shaped words paired with the masculine word *aparell* 'appliance,' (20c) masculine-shaped words paired with *màquina,* (20d) masculine-shaped words paired with *aparell,* (20e) questionable-shape words paired with *màquina,* (20f) words ending in *-oma* paired with *màquina* (20g) words ending in *-oma* paired with *aparell.* See (18) for results of made-up words section of the questionnaire.

(18)

	Feminine-shaped words	**% masculine**	**Masculine-shaped words**	**% masculine**	**Ambiguous shaped**	**% masculine**
Paired with *màquina*	*selupa*	4%	*nuntan*	70%	*truse*	65%
	bitnela	29%	*cro*	83%	*prame*	26%
	rona	14%	*lorum*	78%	*pule*	30%
	riatula	0%	*pras*	70%	*truse*	52%
			estalo	74%		
			romo	75%		
			traix	100%		
average		**12%**		**79%**		**43%**
Paired with *aparell*	*carsa*	13%	*rou*	91%		
	suna	17%	*sablè*	87%		
	pima	13%	*crió*	96%		
	rolces	13%	*brit*	65%		
	tatupa	15%	*serat*	100%		
			pamill	96%		
			casso	88%		
average		**15%**		**90%**		

Breakdown of made-up words used in questionnaire and their hyperonym.

The hyperonyms 'machine' and 'apparatus' seemed to have little effect on words ending with the suffix *-oma,* as shown in (19). In fact, we observe that the words that were embedded in sentences with the feminine hyperonym, *màquina* were assigned the masculine gender more often than the word that was embedded in a sentence with *aparell.*

(19)

	-oma word	% masculine
Paired with *màquina*	xegoma	26%
	cassiroma	30%
average		**28%**
Paired with *aparell*	tuloma	17%
average		**17%**

Made up words with *–oma* suffix.

The types of hyperonym-word combination did seem to exert an effect on the masculine assignment rate, (20). Words with two reasons to be masculine like in (20a) (i.e. masculine hyperonym and masculine word shape had the highest masculine rating. Words with two reasons to be feminine, like in (20e) (i.e. feminine hyperonym and feminine-shaped word) had the lowest masculine rating.

(20)

Type of combination	**% masculine**
a. *Aparell* with masculine-shaped word	90%
b. *Màquina* with masculine-shaped word	79%
c. *Màquina* with word of ambiguous shape	43%
d. *Aparell* with feminine-shaped word	15%
e. *Màquina* with feminine-shaped word	12%
f. *Aparell* with *-oma* word	17%
g. *Màquina* with *-oma* word.	28%

Word groups listed by combinations of semantic groups and word shape.

When all the words of like shapes in this section are combined, the hyperonym effect can be observed. The masculine shaped words, (21a) have by far the highest masculine rating at 85%, the words of ambiguous shape, (21b) lie in the middle at 48% and the feminine words, (21c) are again the lowest at 13%.

(21)

Shape	**% masculine**
a. Masculine	85%
b. Ambiguous	48%
c. Feminine	13%

Shape-based division of words.

A similar, but very weak effect is observed with data organized only by its hyperonym, as in (22). That is, words whose definitions begin with *aparell*, (22a) have a higher masculine rating, at 52%, while words whose definitions begin with *màquina* (22b) are rated lower, at 46%.

(22)

Semantic group	**% masculine**
(a) Aparell	52%
(b) Màquina	46%

Semantic-based division of words.

There was one word, *truse* that was repeated in two different sentences. In one sentence, it was assigned the masculine gender at a rate of 65% and in the other, it was assigned the masculine gender at a substantially lower 52%. In both sentences, *truse* appeared with *màquina*.

Discussion

The results of this section of the study suggest that both the semantic effect of the hyperonym as well as the shape of the word affect the gender assignment. It appears that they work in conjunction to enhance the gender effect when they agree as in cases (20a) and (20e). When they don't agree, as in cases (20b), (20c), and (20d) the effect is weaker, tending towards the shape of the word. Words ending in *-oma* had a low masculine rating, but not as low as combinations with *màquina*. It would seem that while some participants looked only at the final letter to make their gender choice, others looked to the pen-ultimate letter, thereby identifying the masculine-assigning *-oma* suffix.

When observed individually, both the semantic and shape effects have an impact on gender assignment. Shape has the strongest impact, but doesn´t appear to be the sole influencer of gender.

Masculine-shaped words, as in (21a) have a high masculine rating, but when combined with a masculine semantic group, as in (20a), the masculine gender assignment increases. A parallel effect was found for the feminine shaped words, but to a very small degree. Feminine shaped words, as in (21b) have a low masculine rating, but when combine with a feminine semantic group, as in (20e), they drop lower still.

2.3 *Conclusions*

It has been observed by Zubin & Köpcke (1984a), Corbett (1991) that semantic factors play a significant role in the assignment of gender. The semantic factor has been found in the assignment of gender to new words such as in DuBord (2004) and Morin (2006) for Spanish; Köpcke & Zubin (1986) for German; Thorton (2007) for Italian. This study substantiates those findings, particularly the hyperonym effect reported in Thorton (2007) and Köpcke & Zubin (1986). While the hyperonym effect was not always available, when it was available, it was used such as with car and motorcycle names.

It has also been observed by Tucker, Lambert & Rigault (1977) for French, Köpcke & Zubin (1983) for German, that shapes play an important role in the assignment of gender. The findings in this study support those claims: soft drink names, toothpaste brands, loan words and made-up words all demonstrate this instance. However, like many previous studies, no formalized constraint system has been identified. For this, we turn to Rice's (2006), who uses an optimality theoretic approach to develop answers to these questions.

3. Optimizing Gender

3.1 *Optimizing Gender*

Both meaning and shape have been identified as criteria for gender assignment. Until recently, there have been two camps: one that maintains that semantic criteria universally override shape-based criteria (see Corbett 1991 and Corbett and Fraser 2000). The other advocates that shape is the most influential criteria, such as in Tucker, Lambert & Rigault (1977).

The interaction of the two criteria has been investigated by authors such as Steinmetz (1985, 1986). The present section investigates the applicability of Rice's (2006) *optimal gender assignment theory* (OGAT). There must be a reason why, in the face of conflict, gender resolution is sometimes won by semantic factors, other times won by shape factors and still others due to a default gender. Rice fits together Steinmetz's *Gender Tally* and *Gender Eclipsis* with Prince and Smolensky's (1993) Optimality Theory.

Under Rice's approach, there are two types of constraints: gender markedness constraints and gender feature constraints. Rice posits that while gender markedness constraints are ranked, gender features constraints are crucially unranked. As we will see, when two constraints are in a

balanced conflict, the decision is left to the markedness constraints. For brevity, this section only presents his work on German.

Because German has a 3-gender system, there need to be three gender markedness constraints, one to constrain each gender, represented in (23).

(23) * MASCULINE: a noun is not assigned the masculine gender.
* FEMININE: a noun is not assigned the feminine gender.
* NEUTER: a noun is not assigned the neuter gender.

The markedness hierarchy gives a ranking of these constraints. In German, masculine is the least marked, followed by feminine and then neuter, as is seen in (24). As is often the case in OT, the constraints occur as prohibitions. The prohibition against neuter is the highest ranked, so if neuter receives a violation, it is out of the competition first.

(24) MARKEDNESS HIERARCHY
*neuter >> *feminine >> *masculine

When there are no factors to motivate for a neuter or feminine gender assignment, German nouns are masculine. Take, for instance, the German word *Herbst* 'autumn'; it has no gender principles vying for any of the three genders. Therefore, the three potential surface forms, *Herbst*$_{\text{-MASC}}$, *Herbst*$_{\text{-FEM}}$ *and Herbst*$_{\text{-NEUT}}$ all violate one constraint. As in (25), all candidates have one violation, but because the neuter candidate (c) and feminine candidate (b) each violate one constraint that is more highly ranked constraints than the masculine, they are out of the competition. Their dismissal from the competition is a fatal violation marked with an exclamation mark.

(25)

Herbst 'autumn'	*NEUT	*FEM	*MASC
☞ a. *Herbst*$_{\text{-MASC}}$			*
b. *Herbst*$_{\text{-FEM}}$		*!	
c. *Herbst*$_{\text{-NEUT}}$	*!		

Illustration of German's markedness constraints.

Rice points out that all the rules that Köpcke (1982) offers for masculine gender assignment could be simplified to just one: masculine is the default, all others are marked. If the

constraints in German were limited to those in (24) only, then all nouns would be masculine, which is not true. There have to be constraints that can override the markedness hierarchy. These are the gender feature constraints. Here we will consider three of them.

Firstly, nouns ending with the suffix *-e* [ə] in German tend not to be masculine or neuter, but feminine, such as in (27) below. This motivates the constraint in (26).

(26) Words ending in *-e* cannot be masculine or neuter.
*-E# → M, N.

(27) Blume 'flower$_{\text{-FEM}}$'
Schule 'school$_{\text{-FEM}}$'
Bühne 'theater stage$_{\text{-FEM}}$'
Pfanne 'pan$_{\text{-FEM}}$'
Puppe 'doll$_{\text{-FEM}}$'

Secondly, nouns beginning with the prefix *ge-* tend not to be feminine or masculine, but neuter, such as in (28)

(28) Gelüst 'longing$_{\text{-NEUT}}$'
Geheul 'howling$_{\text{-NEUT}}$'
Geschwätz 'chatter'
Gespräch 'conversation$_{\text{-NEUT}}$'

This motivates the constraint (29).

(29) Words beginning with *ge-* cannot be masculine or feminine.
#ge- → M, F

There are some words that take their gender from their meaning. As seen in section 1.5 superordinates in German tend not to be feminine or masculine, such as in (30).

(30) Obst 'fruit$_{\text{-NEUT}}$'
Glied 'limb$_{\text{-NEUT}}$'
Vieh 'livestock$_{\text{-NEUT}}$'
Metall 'metal$_{\text{-NEUT}}$'
Object 'object$_{\text{-NEUT}}$'
Instrument 'musical instrument$_{\text{-NEUT}}$'

This motivates the constraint (31).

(31) Superordinate terms cannot be masculine or feminine.
*SUP→ M, F

However, not all superordinates are neuter. Some can be feminine with typical feminine shapes such as *-e* (e.g. Pflanze 'plant$_{-FEM}$',Waffe 'weapon$_{-FEM}$' and Wette 'bet$_{-FEM}$') and *-ucht*, as seen in (32). There are other words ending in *-ucht* as well that are feminine, such as *Nacht* 'night$_{-FEM}$'.

(32) Frucht '(citrus) fruit$_{-FEM}$'

This motivates constraint (33).

(33) Words ending in *-ucht* cannot be masculine or neuter.
* +*ucht#* → M, N

The relevant gender features for German are summarized below in (34)

(34) **German**
a. 3 genders: neuter, masculine, feminine
b. Markedness hierarchy: * N >> *F >> * M
c. Semantic constraint:
- Nouns denoting superordinates are neuter
= *SUP→ M, F

d. Formal constraint:
- Nouns ending in +e are feminine
=*-E# → M, N
- Nouns ending in *-ucht* are feminine
= * +*ucht#* → M, N
- Nouns beginning with *ge-* are neuter
= *#ge- → M, F

e. Constraint hierarchy: *SUP→ M, F, *-E# → M, N, * +*ucht#* → M, N, *#ge- → M, F >> * N >> *F >> * M

A tableau of nouns with the final suffix *-e* can be observed in (35) below. The masculine candidate (a) and neuter candidate (c) are eliminated by the gender feature *-E →M, N.

(35)

Straße 'Street'	GENDER FEATURES *-E →M, N *GE-→M, F *SUP→ M, F	*NEUT	*FEM	*MASC
a. Straße$_{-MASC}$	*!			*
☞ b. Straße$_{-FEM}$			*	
c. Straße$_{-NEUT}$	*!	*		

When a noun doesn't have any relevant gender features, no violation of gender features can be incurred. In words like 'road', there is no ge- prefix, there is no -e suffix and the word does not correspond to a superordinate. In these cases, the markedness constraints decide the gender.

(36)

Weg 'road'	GENDER FEATURES *-E →M, N *GE-→M, F *SUP→ M, F	*NEUT	*FEM	*MASC
☞ a. Weg$_{-MASC}$				*
b. Weg $_{-FEM}$			*	
c. Weg$_{-NEUT}$		*		

There are some words with two gender features in conflict, such as in (37) below. In such instances, a balanced gender conflict can be observed. Nouns which have both initial *ge-* and final *–e* experience a balanced conflict because words beginning with *ge-* are neuter (therefore they violate *#ge- → M, F), and words ending in *-e* are feminine (therefore they violate *-E# → M, N). It is in balanced conflicts that gender markedness becomes decisive. Rice offers examples of words that have both the neuter assigning *ge-* and the feminine assigning *-e*, such as in (38). Each of these words has one criterion militating for a neuter assignment and one militating for feminine assignment.

(37) *Gebärde* 'gesture $_{-FEM}$',
Gemeinde 'congregation $_{-FEM}$'
Geschichte 'history $_{-FEM}$'.

These conflicts can be resolved with Rice's crucially unranked gender feature constraints, such as in tableau (38) below.

(38)

Geschichte 'story'	GENDER FEATURES *-E →M, N *GE-→M, F	*NEUT	*FEM	*MASC
a. Geschichte $_{-MASC}$	* *!			*
☞ b. Geschichte $_{-FEM}$	*		*	
c. Geschichte $_{-NEUT}$	*	*!		

Balanced gender conflict.

The masculine candidate (a) incurs two violations, while candidates (b) and (c) incur one each. Because candidates (b) and (c) incur the same number of violations, the gender is decided

by the markedness constraints, which are marked. Therefore, candidate (b) wins because *NEUT is more highly ranked than *FEM.

There are also instances where 1 constraint militates for a gender and 2 constraints militate for another gender, such as in (39). This is what Rice calls an 'imbalanced conflict'.

(39) Gemüse 'vegetable$_{\text{-NEUT}}$'
Gewerbe 'trade, occupation$_{\text{-NEUT}}$'
Gewebe 'fabric$_{\text{-NEUT}}$'
Gebaüde 'building$_{\text{-NEUT}}$'

All of words in (39) are neuter, despite the feminine being the least marked of the two conflicting categories. This is because there is the superordinate, which is also taken into consideration. That is, there are two constraints militating for neuter, the ge- beginning and the superordinate status, and one militating for feminine, the *-e* ending. In Rice's words, "when the conflict is imbalanced, the noun is assigned to the category that is most vigorously advocated—in other words, majority rules." Observe tableau (40).

(40)

Gemüse 'vegetable'	GENDER FEATURES *-E →M, N	*GE-→M,F	*SUP→ MF	*NEUT	*FEM	*MASC
a. Gemüse$_{\text{-MASC}}$	*	*!	*			*
b. Gemüse$_{\text{-FEM}}$		*	*!		*	
☞ c. Gemüse$_{\text{-NEUT}}$	*			*		

Imbalanced gender conflict.

The masculine candidate (a) violates all three gender features, while the feminine candidate (b) violates two. This leaves the neuter candidate (c) as the winner. While it is not perfect, as it violates one constraint, *-E →M, N, it's the best. Because the neuter violates one constraint, any candidate which violates two is discarded, as is the case for candidates (b) and (c), indicated by the exclamation marks after their second violations.

In response to the overwhelming assertions that the sex is the key factor in the assignment of gender to words, Rice (2005) examines the Russian word for 'uncle', *djadja.* He states that sex must not be the overriding factor when it is available, but can also be put into his crucially un-ranked constraint system. A summary of the relevant constraints for Russian are found in (41) below.

(41) **Russian**

a) 3 genders: neuter, masculine, feminine
b) Markedness hierarchy: *n>> *f>> *m
c) Semantic constraint:
   - Nouns denoting biological males are masculine
     =* [+ MALE] → F, N
d) Formal constraint:
   - Nouns ending in +a are feminine
     =*+A# → M, N

By formal criteria, *djadja* should be feminine because of its final *-a*, but by semantic criteria it should be masculine, because it refers to a man.

(42)

Djadja 'uncle'	GENDER FEATURES * [+ MALE] → F, N *+A# → M, N	*NEUT	*FEM	*MASC
☞ a. djadja$_{-MASC}$	*			*
b. djadja$_{-FEM}$	*		*!	
c. djadja$_{-NEUT}$	* *!	*		

Rice's unranked constraints show that semantic features do not override shape-based ones.

Candidates (b) and (c) both violate the first constraint because *djadja* denotes a male. The second constraint is violated by (a) and (c) because words ending in *-a* should be feminine, which is not the case. Whereas candidates (a) and (b) violate only one constraint, candidate (c) violates both of the equally ranked constraints and is therefore eliminated. Candidate (a) outranks candidate (b) in the markedness hierarchy, which leaves (a) as optimal. Rice (2005) concludes that semantic features such as sex do not outweigh other gender features.

Thorton's (2007) paper confirms that Rice's *optimal gender assignment theory* also works for Italian nouns referring to men in Italian. In words such as, *sherpa* 'Tibetan baggage carrier', which end in *-a* but refer to a man, OGAT correctly assigns the masculine gender.

In Italian, like in Catalan, there is a hyperonym effect for the word for 'car' *macchina.* In Italian, 'car' is feminine and all of its subordinate terms are also feminine, regardless of their shape. That is, car names ending in a consonant, /i/ or /o/ are feminine, despite having masculine-shaped endings, as is the case for *Fiat, Ford, Mercedes, Torpedo, Uno, Tipo, Ritmo, Panda, Bravo.* Even though car names like *Uno, Tipo, Ritmo Panda* and *Bravo,* which are

homophonous with masculine nouns (*uno*, article/adjective 'one$_{\text{-MASC}}$'; *tipo*, 'type$_{\text{-MASC}}$'; *ritmo*, 'rhythm$_{\text{-MASC}}$'; *panda*, 'panda$_{\text{-MASC}}$'; *bravo*, adjective 'good$_{\text{-MASC}}$') are feminine.

Thorton has therefore included a semantic hyperonym constraint:

(43) Car words in Italian cannot be masculine.
* CAR →M.

In her (2007) paper on Italian, she identifies several gender constraints, the relevant ones are listed in (44) below:

(44) **Italian**

a) 2 genders feminine and masculine
b) Markedness hierarchy: *f>> *m
   - (masculine is less marked than feminine)
c) Semantic constraints:
   - *CAR →M
d) Phonological constraints:
   - *-a# →M
   - *-o# → F

By Rice's (2005) analysis that gender features are crucially unranked, the car name *Torpedo* is assigned the wrong gender. Observe (45).

(45)

Torpedo	GENDER FEATURES *CAR → M	*/O/# → F	*FEM	*MASC
☜ a. torpedo$_{\text{-MASC}}$	*			*
b. torpedo$_{\text{-FEM}}$		*	*!	

Rice's crucially unranked constraints.

Both candidate (a) and (b) incur one gender feature violation and one markedness violation. That means that the optimal candidate for *Torpedo* is (a), which is wrong. This finding works against Rice's claim that gender feature constraints are crucially unranked and in favor of theories such as Corbett's that semantic rules dominate over shape-based ones. The correct tableau for *Torpedo* is (46).

(46)

Torpedo	*CAR → M	*/O/# → F	*FEM	*MASC
a. torpedo-$_{MASC}$	*!			*
☞ b. torpedo$_{-FEM}$			*!	

Thorton's revisions to Rice's crucially unranked constraint system.

Summarizing, we have seen that crucial non-ranking can work for words like Russian's 'uncle', which Thorton confirms with words like 'Tibetan baggage carrier'. However, there crucial non-ranking cannot account for feminine car names with masculine shape.

3.2 A look at Catalan

This paper has identified several phonological and semantic factors that correspond to gender in Catalan. This section has observed how these factors work in both Rice's crucially unranked constraint system and in Thorton's ranked system.

As seen in section 2.1, the masculine is the unmarked gender in Catalan. Therefore the markedness hierarchy is as is seen in (47).

(47) Markedness hierarchy: masculine is the marked, feminine is the unmarked.
*FEM >> *MASC

Section 2.1 also states that Ø morph corresponds to the masculine gender. Therefore, constraint (48) and (49) is posited.

(48) Words ending in Ø morph cannot be feminine
* Ø# → F

There is a correspondence between [ə] and the feminine gender. Therefore constraint (49) can be posited.

(49) Words ending in [ə] cannot be masculine
*[ə]# →M

As has been discussed in section 2.2.1 of this paper, Catalan car names inherit their genders from their hyperonym. The word for 'car', *cotxe,* is masculine in Catalan, so car names such as the following are masculine: *Honda, Mazda, Nova, Panda.*

(50) Car names cannot be feminine.
*CAR → F

Although not tested in this study, it has also been observed that there is a hyperonym effect for motorcycle names in Catalan: they are all feminine. This motivates (51).

(51) Motorcycle names cannot be masculine.
*MOTORCYCLES →M

The following table, (52), summarizes the relevant gender constraints in Catalan:

(52) **Catalan**

a) 2 genders: feminine and masculine
b) Markedness hierarchy: *FEM >> *MASC
(masculine is less marked than feminine)
c) Semantic constraints:
- *CAR → F
- *MOTORCYCLES → M

d) Phonological constraints:
- *Ø# → F
- *[ə]# →M

Sometimes both Rice's and Thorton's constraint system works for Catalan. Tableau (53) explores the Catalan's gender constraints. By Rice's crucially unranked constraints candidates (a) and (b) incur one gender feature violation and one markedness violation. Because the masculine is the less marked, *Panda* is correctly assigned the masculine gender.

(53)

Panda	GENDER FEATURES *CAR → F	*/ə/# → M	*FEM	*MASC
☞ a. Panda$_{\text{-MASC}}$		*		*
b. Panda$_{\text{-FEM}}$	*		*!	

Rice's crucially unranked constraints

As show in (54), Thorton's ranked constraints also correctly assign the masculine gender to *Panda.* The outcome for *Panda* is identical according to Thorton's revised ranked constraints. Candidate (b) is assigned two violations as is candidate (a). However, because candidate (a) is the least marked, it wins.

(54)

Panda	*CAR → F	*/ə/# → M	*FEM	*MASC
☞ a. Panda$_{\text{-MASC}}$		*		*
b. Panda$_{\text{-FEM}}$	*!		*	

Thorton's revisions to Rice's crucially non-ranked constraint system.

For car words, there is no compelling reason to adopt one approach over the other. For more evidence, we look to motorcycle names.

Catalan motorcycle names are consistently feminine. They take their names from their feminine hyperonym *motocicleta*. Therefore, words such as the following are feminine: *Yamaha, Honda, Vespa,* as well as names with masculine shape: *BMW, Suzuki, Kawasaki, Derbi, Harley, Davidson, Ducati.* In (55) Rice's non-ranked constraints are tested; in (56) Thorton's ranked constraints are tested.

(55)

Harley Davidson	GENDER FEATURES * MOTORCYCLES → M	* /Ø/# →F	*FEM	*MASC
☜ a. Harley Davidson$_{\text{-MASC}}$	*			*
b. Harley Davidson$_{\text{-FEM}}$		*	*!	

Rice's crucially non-ranked constraints

According to Rice's crucially non-ranked gender features, both candidate (a) and (b) incur one violation. This results in the winning of the least marked gender, which in Catalan is masculine. In contrast to tableau (53) for car names, tableau (55) is untenable for Catalan motorcycle names. This parallels the problems Thorton pointed out for *Torpedo* in (45). Unranked gender feature constraints produce the wrong candidate.

(56)

Harley Davidson	*MOTORCYCLES → M	* /Ø/# →F	*FEM	*MASC
a. Harley Davidson$_{-MASC}$	*!			*
☞ b. Harley Davidson$_{-FEM}$		*	*	

Thorton's revisions to Rice's crucially non-ranked constraint system.

According to Thorton's revised, ranked constraint system, *Harley Davidson* is assigned the correct gender. Other semantic groups for Catalan discussed in this paper, such as toothpaste and soft drink names, do not have a hyperonym effect; therefore the constraints on shape are decisive on their own, such as in (57) and (58).

(57)

Fanta	GENDER FEATURES *A →MASC	*FEM	*MASC
a. Fanta$_{-MASC}$	* !		*
☞ b. Fanta$_{-FEM}$		*	

Ranked constraints with no semantic criteria for Catalan.

(58)

Kemphor	GENDER FEATURES *Ø →FEM	*FEM	*MASC
☞ a. Kemphor $_{-MASC}$			*
b. Kemphor $_{-FEM}$	*!	*	

Ranked constraints with no semantic criteria for Catalan.

This has reviewed recent attempts to constrain gender assignment Criteria. Rice's (2005) OGAT ranks only markedness, leaving gender features, such as word shape and semantic effects crucially unranked. Thorton's revisions to Rice's OGAT finds that gender features must indeed be constrained. It has been observed that in some cases (e.g. cars), both Rice and Thorton's approaches work. For other cases (i.e. motorcycles) the ranking of semantic criteria over shape-based criteria is necessary.

4. Conclusions

This paper has explored the correspondence between the gender of new words and their meanings and shapes. While it is practically undisputed that sex is an overriding factor in gender outcome, the significance of remaining factors have been more controversial.

The study presented in this paper asked native Catalans to decide the gender of a variety of words: car names, toothpaste brands, soft drink names, low-frequency loan words and made-up words. It was found that the gender of some semantic categories, i.e. soft drinks and tooth paste brands have no correspondence to their hyperonym. Therefore the words' genders of two groups are decided not based on the relationship to 'toothpaste' or 'soft drink', but to the shape of the word. Therefore, feminine shaped words, such as *Coca-Cola* and *Binaca* were found to be feminine, whereas masculine shaped words, such as *Red Bull* and *Kemphor* were found to be masculine. It was found that both word meaning and shape based factors have an effect on the gender of the words.

For made-up words, the hyperonym (either *màquina* 'machine$_{\text{-FEM}}$ or *aparell* 'apparatus$_{\text{-MASC}}$') seemed to have a slight effect. There was a small tendency for words presented as 'machines' to be feminine (46% on a masculine scale). Words presented as 'apparatuses' had a slightly higher masculine rating (52% on a masculine scale). It was shape that had a clear effect on the assignment of gender to these made-up words.

A strong hyperonym effect was found for car names. The masculine assignment rate was very high. Words ending in the typical masculine shape, stressed vowels and consonants, such as *Opel* and *Audi*, were predictably assigned the masculine gender at a high rate. The feminine-shaped car names, such as *Honda* and *Nova* were also assigned the masculine gender despite their shape.

Although it was not tested in this study, it was observed that motorcycle names also exhibit a hyperonym effect. There are motorcycle names with masculine shape, such as *BMW* and *Harley Davidson* and motorcycle names with feminine shape, such as *Yamaha*. Regardless of their shape, all are assigned the feminine gender, even when there is a possible male referent, i.e. Harley Davidson's *Fat Boy.*

Another interesting topic in Catalan gender that could be explored in the future is the relationship between the sex of the referent and the gender of the word. It is usually taken for granted that the sex correlates to the gender. It would be interesting to explore the extent of this

in new words, such as *guru*, which originally had referred to men, but can now refer to women as well. Do these originally-masculine words become feminine when they refer to a woman, or do they retain their original gender?

In the final section, this paper investigated gender assignment in an optimality theoretic framework. It was found that Rice's (2005) claim that gender features should be crucially unranked is untenable for Catalan. Instead, as Thorton (2007) work on Italian claims, gender feature constraints in Catalan must be ranked, with semantic features, such as hyperonyms overriding shape-based features.

References

Bermúdez-Otero, Ricardo. 2007. Morphological Structure and Phonological Domains in Spanish Denominal Derivation. In *Optimality-Theoretic Studies in Spanish Phonology*, by Sonia Colina and Fernando Martínez-Gil (eds). Amsterdam: John Benjamins. 278-311.

Bloomfield, Leonard. 1933. *Language.* New York: Holt, Reinhard and Winston, Inc.

Bonet, Eulàlia, & Maria-Rosa Lloret. 1998. *Fonologia Catalana.* Barcelona: Ariel

Bonet, Eulàlia, & Teresa Cabré. 2002. *El Gènere Gramatical*. Bellaterra: UAB.

Brugmann, Karl. 1891. Zur Frage der Entstehung des grammatischeen Geschlechts. *Beträge zur Geschichte der Deutschen Sprache und Literatur 15*: 523-531.

Cabré Monné, Teresa. 2002. Altres Sistemes de Formació de Mots. In *Gramàtica del Català Contemporani*, by Joan Solà, Maria-Rosa Lloret, Joan Mascaró and Manuel Pérez Saldanya (eds), 889-932. Barcelona: Empúries.

Cabré, Teresa. 2006. Els Sistema Vocàlic del Català i l'Adaptació dels Manlleus. *Actes del XIV Col.loqui Internacional de Llengua i Literatura Catalalnes.* Publicacions de l'abadia de Montserrat.

Cabré, Teresa & Eulàlia Bonet. 2002. *El Gènere Gramatical*. Bellaterra: Universitat Autònoma de Barcelona.

Corbett, Greville. 1991. *Gender*. Cambridge: Cambridge University Press.

DuBord, Elise. 2004. Gender Assignment to English Words in the Spanish of Southern Arizona. *Divergencias: Revista de estudios lingüísticos y literarios 2 (2)*: 27-39.

Grimm, Jakob & Wilhelm Grimm. 1984-1952. *Deutsches Wörterbuch.* Leipzig: Bibliographisches Institut.

Grimm, Jacob. 1878. *Deutsche Grammatik.* Berlin: Ferd. Dümmler.

Hock, Hans Henrich. 1986. *Principles of Historical Linguistics.* Berlin: Mouton de Gruyter.

Ibrahim, Muhammad Hasan. 1973. *Grammatical Gender: Its origin and development.* The Hague: Mouton.

Kilarski, Marcin. 2000. Grimm vs. Brugmann on Gender: Analogies in Ancient, Medieval and Modern Linguistics." *Filozofska Fakulteta Univerze v Ljubljani, Volume Linguistics and Language Studies: Exploring Language from Different Perspectives. Proceedings of the 32nd Annual Meeting of Societas Linguistica Europaea.* Ljubljana, 87-96.

Köpcke, Klaus-Micheal & David Zubin. 1984 Sechs Prinzipien für die Genuszuweisung im Deutschen: Ein Beitrag zur natürlichen Klasifikation. *Linguistische Berichte, 93*: 25-50.

Köpcke, Klaus-Micheal & David Zubin. 1983. Die Kognitive Organisation der Genuszuweisung zu den Einsilbigen Nomen der Deutschen Gegenwardsprache *Zeitschrift für germanistische Linguistik*, 11: 166-182.

Köpcke, Klaus-Micheal, & David Zubin. 1986. Gender and Folk Taxonomy: the Indexical Relation between Grammatical and Lexical Categorization. In *Noun Classes and Categorization. Proceedings of a symposium on categorization and noun Classification, Eugene, Oregon*, by in Craig G. Colette (ed.), 139. Amsterdam: John Benjamins.

Mascaró, Joan. 1986. *Morfologia*. Barcelona: Impalsa.

Mascaró, Joan. 20002. Es Sistema Vocàlic. Reducció vocálica. In *Gramàtica del Català Contemporani*, by Joan Solà, Maria-Rosa Lloret, Joan Mascaró and Manuel Pérez Saldanya (eds) 89-123.

Morin, Regina. 2006. Evidence in the Spanish Language Press of Linguistic Borrowings of Computer and Internet-Related Terms. *Spanish in Context 3:2*: 161-179.

Poplack, S., A. Pousada, & D. Sankoff. 1982. Competing Influences on Gender Assignment. *Lingua, 57*: 1-28.

Rice, Curt. 2005. Optimizing Gender. *Lingua 116*: 1394-1417.

Robinson, Orrin W. 2007. Does Sex Breed Gender?: Pronominal Reference in the Grimms' Fairy Tales. *Marvels & Tales, 21:1*: 107-121.

Roca, Ignacio M. 2005 La Gramática y la Biología en el Género del Español. *Revista Española de Lingüística*: 17-44.

Sedaris, David. 2000. Me Talk Pretty One Day. In *Me Talk Pretty One Day*, by David Sedaris, 166-173. Boston: Little, Brown and Company.

Steinmetz, Donald. 1986. Gender in German and Icelandic: Inanimate nouns. *German linguistics. Papers from a Symposium at the University of Chicago*. Bloomington, IN: IULC, 1985. 10-28.

Steinmetz, Donald. Two Principles and Some Rules for Gender in German: Inanimate Nouns. *Word, 37:*189-217.

Thorton, Anna M. 2007. Constraining Gender Assignment Rules. *Language Sciences* 31: 14-32.

Tucker, G. R., W. E. Lambert, & A. A. Rigault. 1977 *The French Speaker's Skill with Grammatical Gender: an Example of Rule-Governed Behavior*. The Hague: Mouton.

Viaplana, J. & Lloret M.R. 1992. Les Vocals Finals dels Elements Nominals en Català. In *Miscel·lània: Estudis de llengua i literatura a cura d'Antoni Ferrando i d'Albert G. Hauf*, by Joan Fuster, 419- 445. Barcelona: Publicacions de l'Abadia de Montserrat.

Wechsler, Stephen. 2009. 'Elsewhere' in Gender Resolution. In *The Nature of the Word - Essays in Honor of Paul Kiparsky*, by Kristin Hanson & Sharon Inkelas (eds). Cambridge, MA: MIT Press.

Zubin, David, & Klaus Micheal-Köpcke. 1984a. Affect Classification in the German Gender System. *Lingua 63*: 41-96.

Zubin, David, & Klaus-Micheal Köpcke. Sechs Prinzipien für die Genuszuweisung im Deutschen: ein Beitrag zur natürlichen Klassifikation. *Linguistische Berichte* 93, 1984b: 26-50.

5. APPENDICES

1. Product names: Reproduction of questions in questionnaire for car words with instructions.

Estem investigant la regularització dels noms dels cotxes en català. Encercla l'opció de gènere que et sembli millor.

1. Compraries **un / una** Opel de segona mà?
2. Pots apagar l'aire condicionat **de la teva / del teu** Honda?
3. El Marc va aparcar **el / la** Mazda al costat de Correus.
4. Vaig comprar-me **la / el** Fiat a Badalona.
5. Has vist **una / un** Nova que passava per aquí?
6. Han portat **el / la** Panda al desguàs.
7. Em deixes **el teu / la teva** Saab per tot el dia?
8. Sempre menja al cotxe, i això fa que **el seu / la seva** Mercedes sigui un desastre.
9. Saps algun lloc on pugui rentar el cotxe? L'Audi és **brut / bruta**.
10. M'han regalat **un / una** Ford pel meu aniversari!

2. Product names: Reproduction of questions in questionnaire for soft drinks with instructions.

Estem investigant la regularització dels noms de begudes en català. Encercla l'opció de gènere que et sembli millor.

1. Beus Sprite amb gel? És més **fresc / fresca.**
2. Al final de la marató van donar Aquarius **caducats / caducades** als corredors.
3. Cada dia prenc **un / una** Coca-Cola al migdia.
4. Molts joves prefereixen posar unes gotetes de vodka **al / a la** Fanta
5. El que més m'agrada **del / de la** Trina és que no té gas.
6. Encara que sigui hivern, als matins prenc Cacaolat **freda / fred**.
7. Ara ja han acabat de vendre Cola Cao a les escoles perquè no és **nutritiva / nutritiu**.
8. Un entrepà i **un / una** Nestea, si us plau.
9. La venda **de la / del** Schweppes ha pujat un 5%.
10. La cosa que valoro de Seven Up és que és **descafeïnat / descafeïnada.**
11. Passa'm **la / el** Red Bull.
12. No m'agrada **el / la** Bitter Kas perquè quan en prenc tinc molt mal de panxa.

3. Product names: Reproduction of questions in questionnaire for toothpaste brands with instructions.

També estem investigant la regularització dels noms de pastes de dents en català. Encercla l'opció de gènere que et sembli millor.

1. **Aquesta / aquest** Colgate ha caducat.
2. Ahir hi havia una oferta de Binaca . Així que en vaig comprar **molts / moltes**.
3. Signal és **bonissim / bonissima** i em deixa les dents blanquíssimes.
4. No et recomanaria Close Up. És **cara / car** i té un gust horrible.
5. Per genives que sagnen, Gingilacer és **fabulós / fabulosa.**
6. Em solc comprar Kemphor perquè és **baratíssima / baratíssim.**
7. No li agrada gens **el / la** Profidén, però la seva dona en compra sovint.
8. **Aquella / aquell** Sensodyne l'hi va comprar la seva mare.
9. La dentista dóna a tots els clients **un / una** G.U.M. de regal.

4. Loan words: Reproduction of questions in questionnaire with instructions for foods and drinks.

Els últims anys s'han introduït al català moltes paraules d'altres llengües. Aquí n'apareixen unes quantes. Encercla, a les frases, el determinant que et sembli millor. A continuació indica amb una creu si coneixes la paraula. Si no la coneixes, pots mirar la definició que apareix en lletra més petita.

1. Quan vam anar al restaurant japonès, ens van donar **un / una** sake que no es podia beure.
   ☐ *conec la paraula* ☐ *no la conec*

Sake- Beguda alcohòlica pròpia del Japó, de 12 a 17 graus, obtinguda per mitjà de la fermentació de grans d'arròs cuits.

2. L'eggnog s'ha de mantenir **refrigerat / refrigerada**. Si no, es fa malbé al cap de poques hores.
   ☐ *conec la paraula* ☐ *no la conec*

Eggnog- Beguda nord-americana d'ous amb nata, llet, sucre, nous i canyella. Servit calent al hivern.

3. **La / El** Glühwein s'ha de bullir abans que es pugui beure.
   ☐ *conec la paraula* ☐ *no la conec*

Glühwein- Beguda alcohólica que es compon principalment de vi calent amb espècies i és molt popular a Àustria i Alemanya.

4. Has provat proja amb mantega? És **deliciós / deliciosa**, no?
   ☐ *conec la paraula* ☐ *no la conec*

Proja –Menjar serbi de blat de moro amb ous que es fa al forn.

5. T'agrada **la / el** sahleb?
M'encanta!
   ☐ *conec la paraula* ☐ *no la conec*

Sahleb - Beguda turca de llet, pols d'arrel d'orquídia i canyella.

6. M'ha caigut boza **calent / calenta** a la cama i m'he cremat.
   ☐ *conec la paraula* ☐ *no la conec*

Boza- Beguda turca de cigrons, canyella i sucre. És de color groc i es beu a l'hivern.

7. Com més gelat hi posis, més **bona / bo** és l'apfelstrudel.
   ☐ *conec la paraula* ☐ *no la conec*

Apfelstrudel- Menjar alemany de pomes, mantega i espècies obert amb una. Se serveix calent amb gelat.

8. A la muntanya es menja **molt / molta** asiago.
   ☐ *conec la paraula* ☐ *no la conec*

Asiago- Formatge fet amb llet de vaca, de pasta ferma semicuita i maduració mitjana, que se sol presentar en peces grosses, originari de la ciutat italiana d'Asiago.

9. **La / El** biscota és millor si es pren amb llet.
   ☐ *conec la paraula* ☐ *no la conec*

Biscota- Pa elaborat principalment amb farina, malt, sucre, greixos animals i vegetals i un substitutiu de la llet en pols, cuit en motlles amb tapa, tallat a llesques, torrat i envasat.

10. He intentat fer balotina moltes vegades, però sempre em queda **fastigós / fastigosa.**
    ☐ *conec la paraula* ☐ *no la conec*

Balotina- Plat originari de la cuina francesa consistent en una peça de carn o de peix desossada, enrotllada i generalment farcida, que, un cop cuita, se sol servir tallada a rodanxes, freda o calenta.

11. Cada dia pren **una / un** bíter abans de dinar.
    ☐ *conec la paraula* ☐ *no la conec*

Bíter- Beguda gasosa de color vermellós, obtinguda per maceració, extracció o percolació en una solució hidroalcohòlica.

12. **Els / Les** brètzels es poden menjar amb formatge.
    ☐ *conec la paraula* ☐ *no la conec*

Brètzel- Pa empolsat amb sal i comí, que té forma de llaç.

13. Quan va anar al Marroc el va posar malalt **aquella / aquell** harira.
    ☐ *conec la paraula* ☐ *no la conec*

Harira- Sopa espessa i consistent típica del Magrib que se sol elaborar amb cigrons, llenties, carn i verdures, i que es menja especialment durant el ramadà.

14. L'harissa que fa la meva mare no és tan **aigualit / aigualida** com al restaurant.
    ☐ *conec la paraula* ☐ *no la conec*

Harissa- Salsa espessa i picant que es prepara amb bitxo, oli, all i herbes aromàtiques, i que s'utilitza especialment com a condiment per al cuscús, típica del nord d'Àfrica i del Pròxim Orient.

15. He comprat kaixkavl **búlgara / búlgar**.
    ☐ *conec la paraula* ☐ *no la conec*

Kaixkaval- Formatge fet amb llet de vaca o d'ovella, de pasta dura, textura granulosa i gust suau quan és jove, i fort i picant quan és vell, originari de Bulgària.

16. **El / La** kebe sense coure no està tan malament com et pensaves.
    ☐ *conec la paraula* ☐ *no la conec*

Kebe - Mandonguilla, generalment de forma ovalada, elaborada amb una pasta de búrgul i carn capolada de xai condimentada amb espècies, ceba i pinyons, que es menja fregida, crua o bé cuita al forn, típica del Pròxim Orient.

17. Ha menjat kefta **fornejada / fornejat** al tandoori.
☐ *conec la paraula* ☐*no la conec*

Kefta- Pilota feta de carn picada, formatge o verdures, condimentada amb espècies, típica del Pròxim Orient i de l'Índia.

18. **El / La** làban amb molt all combina estupendament amb l'arròs.
☐ *conec la paraula* ☐*no la conec*

Làban- Iogurt obtingut a partir de llet acidificada, típic del Pròxim Orient i del nord d'Àfrica

19. **La / El** miso es menja sovint als millors restaurant de Tòquio.
☐ *conec la paraula* ☐*no la conec*

Miso -Pasta fermentada elaborada amb una barreja d'aigua, soia i ordi o arròs, de gust salat i amb un contingut proteic elevat, molt utilitzada en la cuina japonesa.

20. **Aquell / Aquella** mussaca va caure a terra.
☐ *conec la paraula* ☐*no la conec*

Mussaca -Menja cuita al forn, elaborada amb albergínia tallada a rodanxes, fregida i disposada en forma de capes, les quals alternen amb capes de carn picada i assaonada de vedella.

21. El menú del dia té **una / un** mousse d'ous i taronja.
☐ *conec la paraula* ☐*no la conec*

Mousse- Plat cuinat d'origen francès fet amb diferents ingredients que es quallen després de barrejar-los amb clara d'ou batuda, nata batuda o gelatina.

22. El falafel l'hi va donar amb pita **dur / dura.**
☐ *conec la paraula* ☐*no la conec*

Pita - Pa de forma rodona o ovalada, molt prim i buit, que s'obre com una butxaca i pot omplir-se amb ingredients diversos, típic de la Mediterrània oriental i del Pròxim Orient.

23. Es pot mantenir **la / el** tahini de sèsam a la nevera durant 10 mesos.
☐ *conec la paraula* ☐*no la conec*

Tahina - Mantega de sèsam, típica del Pròxim Orient.

24. Et va posar malalt **el / la** xauarma d'aquest restaurant?
☐ *conec la paraula* ☐*no la conec*

Xauarma - Carn marinada, generalment de be, que es cou en un ast vertical que es fa girar, i que se serveix tallada a tires, sovint com a farciment d'entrepà, típica del Pròxim Orient.

25. Quan van anar al Japó, van provar wakame. M'han dit que és **deliciósa / deliciós**.
☐ *conec la paraula* ☐*no la conec*

Wakame -Alga bruna de grans dimensions, rica en calci, molt utilitzada en la cuina japonesa.

5. Loan words: Results of gender assignment task for instruments.

Els últims anys s'han introduït al català moltes paraules d'altres llengües. Aquí n'apareixen unes quantes. Encercla, a les frases, el determinant que et sembli millor. A continuació indica amb una creu si coneixes la paraula. Si no la coneixes, pots mirar la definició que apareix en lletra més petita.

1. Ja has après a tocar **el / la** guquin?
   ☐ *conec la paraula* ☐ *no la conec*

   Guquin- Instrument cordòfon pinçat de la família de les cítares constituït per una caixa de ressonància de fusta de forma rectangular, d'uns 125 cm de llargada i uns 20 cm d'amplada, amb un extrem més estret que l'altre i amb la taula harmònica lleugerament corbada.

2. **La / El** koto és un instrument molt exòtic.
   ☐ *conec la paraula* ☐ *no la conec*

   Koto- Instrument cordòfon pinçat de la família de les cítares constituït per una caixa de ressonància de fusta de forma allargada.

3. Tothom gaudeix del so **del / de la** taiko.
   ☐ *conec la paraula* ☐ *no la conec*

   Taiko- Instrument membranòfon de membrana doble i de dimensions variables constituït per una estructura de fusta en forma de bóta i amb lligadures de corda, que es toca amb les mans o bé amb unes baquetes de fusta.

4. On t'has comprat **aquesta / aquest** biwa?

   ☐ *conec la paraula* ☐ *no la conec*

   Biwi- Instrument cordòfon pinçat de la família dels llaüts, que pot fer fins a 100 cm de llargada, constituït per una caixa de ressonància de fusta en forma de pera i de fons bombat, amb la taula harmònica plana i amb dues obertures acústiques en forma de mitja lluna .

5. Li va trencar **els / les** bongos d'un cop de martell.
   ☐ *conec la paraula* ☐ *no la conec*

   Bongos - Instrument membranòfon de percussió directa, format per dos tambors curts i lleugerament cònics fixats l'un amb l'altre, que es col·loca entre les cames i es toca amb els dits o amb el palmell de les mans.

6. Vaig trobar uns cucs **a la / al** didjeridú del meu pare.
   ☐ *conec la paraula* ☐ *no la conec*

   Didjeridú - Instrument aeròfon de columna de la família dels corns, que fa entre 120 i 180 cm de llargada, amb el tub recte i sense broquet i amb una embocadura de resina.

7. El Chang toca cada matí **el seu / la seva** dizi al jardí**.**
   ☐ *conec la paraula* ☐ *no la conec*

   Dizi - Instrument aeròfon de columna de la família de les flautes travesseres, tradicionalment de bambú, amb sis forats d'entonació que es tapen amb els dits, un mínim de dos forats acústics i un forat que es cobreix amb

una membrana, generalment de paper d'arròs, la qual vibra per simpatia i confereix un so característic a l'instrument.

8. Has de buscar sempre una bona afinació **per a la teva / per al teu** erhu.
   ☐ *conec la paraula* ☐ *no la conec*

   Erhu- Instrument cordòfon d'arquet de la família de les cítares, que fa uns 70 cm de llargada i uns 15 cm d'amplada, constituït per una caixa de ressonància de forma hexagonal sense obertures acústiques, amb la taula harmònica de pell de serp i amb un mànec llarg sense claviller ni trasts que travessa la caixa.

9. Va trobar a Internet un curs molt interessant sobre **el / la** gamelan.
   ☐ *conec la paraula* ☐ *no la conec*

   Gamelan- Conjunt instrumental propi d'Indonèsia, especialment de Java i Bali, constituït essencialment per instruments idiòfons i membranòfons, que sol acompanyar les danses tradicionals, la música ritual i el teatre.

10. Per a tocar **aquella / aquell** qin cal un escalfament previ.

    ☐ *conec la paraula* ☐ *no la conec*

    Qin- Instrument cordòfon pinçat de la família de les cítares constituït per una caixa de ressonància de fusta de forma rectangular, d'uns 125 cm de llargada i uns 20 cm d'amplada, amb un extrem més estret que l'altre i amb la taula harmònica lleugerament corbada, generalment amb set cordes de diferent gruix separades de la taula amb dos ponts fixos, que es toca en posició horitzontal.

11. Al Joan se li ha trencat **el / la** zheng.

    ☐ *conec la paraula* ☐ *no la conec*

    Zheng-Instrument cordòfon pinçat de la família de les cítares constituït per una caixa de ressonància de fusta de forma rectangular i corbada, d'uns 120 cm de llargada i uns 25 cm d'amplada, amb obertures acústiques laterals i amb una vintena de cordes habitualment afinades seguint l'escala pentatònica, que es toca en posició horitzontal, amb l'ajuda de diversos ponts mòbils i amb plectres en forma de didal.

12. Si vols aprendre a tocar **la / el** sanxian, hauries d'anar a l'Índia.

    ☐ *conec la paraula* ☐ *no la conec*

    Sanxian- Instrument cordòfon pinçat de la família dels llaüts, constituït per una caixa de ressonància de fusta de forma aproximadament ovalada i de fons pla, amb la taula harmònica feta de pell de serp, i un mànec llarg, claviller, i tres cordes, que es toca amb un plectre o amb les ungles de la mà dreta.

13. Ja no té prou capacitat pulmonar i ha hagut de deixar **el / la** shakuhachi .

    ☐ *conec la paraula* ☐ *no la conec*

    Shakuhachi-Instrument aeròfon de columna amb el bisell tallat a l'extrem superior del tub, de buf directe i tradicionalment de bambú, d'uns 55 cm de llargada i uns 4 cm de diàmetre, amb cinc forats que es tapen amb els dits de les dues mans.

14. El meu avi em va deixar **aquesta / aquest** shamisen.

    ☐ *conec la paraula* ☐ *no la conec*

Shamisen- Instrument cordòfon pinçat de la família dels llaüts, que fa uns 100 cm de llargada i uns 20 cm d'amplada, constituït per una caixa de ressonància de fusta de forma quadrangular i de fons pla, amb la taula harmònica plana i generalment de pell de gos o de gat.

15. Has de canviar les cordes **del teu / de la teva** tambura.

☐ *conec la paraula* ☐*no la conec*

Tambura- Instrument cordòfon pinçat de la família dels llaüts, que pot fer fins a 130 cm de llargada, constituït per una caixa de ressonància semiesfèrica feta amb una carabassa o, modernament, amb fusta, amb la taula harmònica corbada, amb el mànec llarg i sense trasts, amb pont, celleta i amb quatre cordes metàl·liques.

16. S'ha investigat una tècnica nova per tocar **la / el** tombak .

☐ *conec la paraula* ☐*no la conec*

Tombak -Instrument membranòfon tubular en forma de copa constituït per un ressonador de fusta que fa entre 20 i 30 cm de diàmetre al qual s'enganxa una pell de cabra gruixuda, que es toca percudint-lo amb les mans.

17. **Aquest / Aquesta** xenai és de roure.

☐ *conec la paraula* ☐*no la conec*

Xenai - Instrument aeròfon de columna constituït per un tub cònic de fusta, amb set forats equidistants, una llengüeta doble fixada sobre un guardallavis en un tudell independent.

6. Loan words: Results of gender assignment task foods, drinks and musical instruments. Results given for percentage assigned the masculine concord

Foods and beverages	
sake	94%
eggnog	83%
glühwein	86%
proja	33%
sahleb	89%
boza	58%
apfelstrudel	78%
asiago	83%
biscota	31%
balotina	25%
bíter	86%
brètzel	78%
harira	11%
harissa	17%
kaixkavl	86%
kebe	47%
kefta	31%
làban	89%
miso	86%
mussaca	6%
mousse	28%
pita	42%
tahini	86%
xauarma	69%

Musical Instruments	
wakame	72%
guquin	92%
koto	94%
taiko	50%
biwa	28%
bongos	89%
didjeridú	83%
dizi	69%
erhu	83%
gamelan	75%
qin	86%
zheng	86%
sanxian	92%
shakuhachi	77%
shamisen	89%
tambura	17%
tombak	91%
xenai	83%

7. Made-up words: Reproduction of questions in questionnaire with instructions.

A continuació trobaràs la descripció d'una colla de nous invents que encara no s'han patentat, juntament amb el nom que s'ha pensat per a designar-los. Encercla, per a cada cas, el determinant que et sembli més adient en les frases en què apareix utilitzat aquest nou nom.

1. Màquina que anul·la el soroll del tren o del carrer perquè puguis parlar tranquil·lament pel mòbil. Es col·loca a sobre del cap: traix

No pot ser d'un altre color, **la / el** traix?

2. Aparell que fa forats en un paper automàticament. Només cal que hi col·loquis una pila de papers: rou

Aquest / Aquesta rou no funciona gaire bé.

3. Aparell que et posa el tap al tub de pasta de dents**:** carsa

Sort que m'han regalat **una / un** carsa. Sempre m'oblido de posar el tap al tub.

4. Aparell que t'asseca després de dutxar-te sense que hagis de fer res: sablè

La meva / El meu sablè no asseca bé.

5. Màquina que es col·loca a la finestra del cotxe per evitar que el gos tregui el cap, però no impedeix que el vent li arribi a la cara: estalo

El Nevat es va fer mal ahir amb **el / la** estalo, i vam haver d'anar al veterinari.

6. Aparell que neteja la pols de la prestatgeria dels llibres: suna

Quina pena! **Aquella / Aquell** suna es va espatllar.

7. Aparell que obre cartes automàticament: crió

M'he fet un tall al dit amb **la** / **el** crió.

8. Màquina que et fa treure el barret cada vegada que et trobes amb algú i l'has de saludar. Es col·loca a sota de la jaqueta: nutan

Amb **el / la** nutan pots saludar tantes vegades com faci falta sense cansar-te..

9. Aparell que planxa la roba pitjant un botó: brit.

Estalvio molt de temps amb **la / el** brit.

10. Màquina que et protegeix el bigoti perquè no es mulli quan beus. Així prevens les infeccions: xegoma

Aquest / Aquesta xegoma pot durar fins a 5 anys.

11. Màquina que, col·locada a la forquilla, compta quantes vegades mastegues i t'avisa quan ho has fet 15 vegades. Així controles que mastegues prou: selupa

A **aquesta / aquest** selupa se li han acabat les piles.

12. Aparell que esborra els traços del bolígraf sense deixar cap marca al paper**:** pamill

Aquest / aquesta pamill funciona fantàsticament.

13. Màquina que fa pujar la tapa del vàter amb un motor perquè no l'hagis de tocar: truse

Podries haver utilitzat **la / el** truse; la tapa és bruta!

14. Aparell que et barreja la sal i el pebre amb una corda que s'estira: casso

El Pol s'ha equivocat. Ha ficat el sucre dins **el** / **la** casso i ara les patates braves són plenes de sucre.

15. Màquina que deixa anar una descàrrega elèctrica al gos cada vegada que s'apropa al final del jardí. Així no cal cap corretja ni cap reixa: cro

No m'agrada **aquesta / aquest** cro per al meu gos perquè és un invent cruel.

16. Aparell que manté la mantega tova per poder untar bé el pa: tuloma

Has de comprar **un / una** tuloma! Estan d'oferta al supermercat.

17. Màquina que treu la bosseta del te de la tassa just quan toca: prame

El meu / La meva prame sap quants minuts necessita el te negre, el te de fruites i el te verd.

18. Màquina que t'eixuga l'orella després de nedar. Així no t'hi queda aigua: lorum

He trobat **una / un** lorum assequible i també té garantia.

19. Màquina que fa girar el gelat perquè el puguis llepar sense bellugar-lo manualment: truse

Em vaig comprar **un / una** truse perquè hi havia una promoció.

20. Aparell que passa els fulls del llibre que llegeixes: pima

Quan tens les mans ocupades, fer servir **un / una** pima és molt pràctic.

21. Màquina que conté un tub amb forats on s'hi col·loquen diverses cigarretes per poder-les fumar alhora: pras

Els nois han ficat 7 cigarretes a la vegada **al / a la** pras .

22. Màquina que et desperta amb cops gràcies a unes peses. Estan col·locades en un bastiment sobre el cap de la persona que dorm i baixen quan s'ha de despertar: romo

Sembla un sistema perillós, **el / la** romo, però no ho és. A més, em va molt bé per llevar-me.

23. Aparell que et lliga els cordons de les sabates: turoc

El meu / La meva turoc m'estalvia molt de temps.

24. Màquina que fa que el llibre no es tanqui. És molt ampla i es col·loca al dit gros d'una mà. Així et queda una mà per fer una altra cosa: cassiroma

M'encanta **aquesta / aquest** cassiroma.

25. Aparell que es col·loca a les sabates i serveix per netejar el terra al mateix temps que camines per casa: rolça

La Mònica ha sortit de la dutxa i s'ha posat **els / les** rolces.

26. Aparell que talla les ungles mitjançant un gallet. Aquest aparell porta un mirall que evita que hagis d'inclinar-te quan et talles les dels peus: serat

Li he comprat **una / un** serat a la meva tia, que és paralítica.

27. Aparell que es posa a un plàtan pelat perquè no es torni marró: tatupa.

El plàtan s'ha podrit en quatre hores. Això vol dir que **aquesta / aquest** tatupa no ha anat bé.

28. Màquina que silencia els crits. Perfecta per a una pel·lícula de por: pule

Si anem al cine porta **aquesta / aquest** pule, que no vull haver-te de sentir tota l'estona.

8. Made-up words: Results of gender assignment task.
*Results given for percentage assigned the masculine concord

traix	100%	pule	30%
rou	91%	truse	52%
carsa	13%	pima	13%
sablè	87%	pras	70%
estalo	81%	romo	75%
suna	17%	turoc	96%
crió	96%	cassiroma	30%
nuntan	70%	rolces	13%
brit	65%	serat	100%
xegoma	26%	tatupa	17%
selupa	4%	bitnela	29%
pamill	96%	riatula	0%
truse	65%	rona	14%
casso	88%	bitnela	29%
cro	83%	rona	14%
tuloma	17%	riatula	0%
prame	26%		
lorum	78%		