

Harris 1991

1. El Cheapo

- Spanish gender commonly characterized by two rules: [fem] N ends in ^-a, [masc] N ends in 'o'

(1)	<u>muchacho</u>	'boy'	<u>muchacha</u>	'girl'
	<u>abuelo</u>	'grandfather'	<u>abuela</u>	'grandmother'
	<u>nieto</u>	'grandson'	<u>nieta</u>	'granddaughter'
	<u>tío</u>	'uncle'	<u>tía</u>	'aunt'

- But -a and -o belong to a class of seven or more analogous **word markers** (Harris' term for endings of this sort).
- Category error: conflation of three different concepts that are interrelated but ust be kept separate:
 - grammatical gender
 - conceptual gender
 - form class (aka **declension class**) (a class of nouns, adjectives, etc. that take the same inflectional forms)
- Consider that (3a) patterns opposite to (1), while (3b-d) differ altogether.

(3)	<i>Masculine</i>		<i>Feminine</i>	
a.	<u>poeta</u>	'poet'	<u>virago</u>	'virago'
b.	<u>sirviente</u>	'servant'	<u>sirvienta</u>	'servant'
c.	<u>doctor</u>	'doctor'	<u>doctora</u>	'doctor'
d.	<u>cónyuge</u>	'spouse'	<u>cónyuge</u>	'spouse'

- But it is not the case that anything goes. The pairings of [masc] and [fem] forms in (4) is disallowed. The problem is not the well-formedness of these examples but their pairing.

H

(4)	<i>Masculine</i>	<i>Feminine</i>	
	*{ <u>presidento</u>	<u>presidente</u> }	'president'
	*{ <u>lingüista</u>	<u>lingüisto</u> }	'linguist'

- A problem of **negative evidence**: pairs like *{presidento (m)/presidente (f)} never occur in the linguistic experience of a child. Since they don't occur, they can't be coded as unacceptable. How then does the child develop a grammar that excludes these? This must be deducible from a combination of generalizations about the language and universal principles.

- Claims:
 - formal representation of grammatical gender involves one **privative** feature [femine] (f). There is no corresponding [masculine] feature, nor a binary feature [+/- fem].
 - form class is also privative, with -o being an unmarked default class.
 - the fact that Spanish nouns and adjectives canonically end in a vowel is due to a prosodic template that is normally (but not always) satisfied by a word marker.

2. Data

- Word markers:
 - Mark a derivationally and inflectionally complete word
 - Cannot be followed by any other suffix except the plural -s
democrat-a **democrat-a-ica* *democrat-ica*
 - Do not appear in the infinitive denominal verbs
democrat-izar **democrat-a-izar*
 - Occur on Nouns, Adjectives and Adverbs

(5)	Marker	Gender	Example	
a.	-o	m only	muchacho	'boy'
		f only	mano	'hand'
		m or f	testigo	'witness'
		none	dentro	'inside'
b.	-a	f only	muchacha	'girl'
		m only	día	'day'
		m or f	turista	'tourist'
		none	fuera	'outside'
c.	-Vs	m only	Lucas	'Luke'
		f only	síntesis	'synthesis'
		m or f	mochales	'loony, nutty'
		none	lejos	'far'
d.	-u	m only	espíritu	'spirit'
		f only	tribu	'tribe'
e.	-i	m only	bikini	'bikini'
		f only	metrópoli	'metropolis'
		m or f	cursi	'in bad taste'
f.	-s	m only	tóra[k̩s]	'thorax'
		none	quizás	'maybe'
g.	-e	m only	héroe	'hero'
		f only	prole	'progeny'
		m or f	inmune	'immune'
		none	adrede	'intentionally'

- The complete inventory includes all spanish vowels.
- "Nearly all marker shapes are found in words that are only masculine, words that are only feminine, words that can be either gender, and words with no gender at all (adverbs, see below). " (p. 31)
- Some nouns, adjectives, adverbs, have no word marker. These words include all of the gender possibilities seen in words that do have word markers.

(6)	<i>Gender</i>	<i>Example</i>	
a.	m only	padre	'father'
		sol	'sun'
b.	f only	madre	'mother'
		col	'cabbage'
c.	m or f	amante	'lover'
		mártir	'martyr'
d.	none	delante	'ahead'
		atrás	'behind'

- The words ending in -e in these examples do not have a word marker. The -e is a prosodic repair for words that end in unyllabified segments, e.g. *dr*, *nt* are not permissible sequences in Spanish codas.
- Harris says (6), like (1), obfuscates the nature of the system. While not cartoonishly simple like (1), (6) fails to capture systematicity.
- Harris observes two axes of systematicity:
 - generalizations particular to each of the categories of words that bear word markers (nouns, pronouns, determiners, adjectives, and adverb)
 - generalizations particular to subtypes of word markers; " the examples in (5)-(6) are not all on a par; rather, they fall naturally into three hierarchically related classes: an **inner core** of prototypes, illustrated in (1); an **outer core** of slightly deviant cases, namely, the class of markerless words illustrated in (6); and a **motley residue**." (p.32) [emphasis added]

Table 1

REGULAR				IRREGULAR
Inner core		Outer core		Residue
m	f	m	f	
hijo 'son'	hija 'daughter'	padre 'father'	mujer 'woman'	problema (m) 'problem'
cedro 'cedar'	sidra 'cider'	mar 'sea'	liebre 'hare'	lejos 'far'
				tribu (f) 'tribe'
				etc.

- Inner core: -o is invariably attached to masculine stems and -a is invariably attached to feminine ones.
- Outer core: do not have word markers (but may have [e] for syllabicity)
- Residue: all words not in the core ; it is small overall, most subclasses within it are tiny, e.g. *tribu* is the only -u marked word. (But about 600 -a words); it is not productive; there is historic drift from core to residue, but never vice versa.

2.1 Adverbs

- No gender, but they do have word markers: this shows the independence of conceptual/grammatical gender and form class.

(7) a.	dentro	(adverb, (5a))	'inside'
	a + dentr + ar	(verb)	'to penetrate'
	*a + dentro + ar		
b.	fuera	(adverb, (5b))	'outside'
	fuer + eño	(noun)	'outsider'
	*fuera + eño		
c.	lejos	(adverb, (5c))	'far'
	lej + ano	(adjective)	'distant'
	*lejos + ano		

2.2 Adjectives

- No inherent gender, but they do show concord for the gender of the noun they modify.
- Adjectives show independence of conceptual/grammatical gender vs. form class; while they must match gender of the head noun, the form needn't match.

- (8) a. $\text{Mi } \left\{ \begin{array}{l} \text{sobrino (m)} \\ \text{padre (m)} \end{array} \right\} \text{ es } \left\{ \begin{array}{l} \text{inteligente (m)} \\ \text{alto (m)} \end{array} \right\} .$
 ‘My $\left\{ \begin{array}{l} \text{nephew} \\ \text{father} \end{array} \right\}$ is $\left\{ \begin{array}{l} \text{intelligent} \\ \text{tall} \end{array} \right\} .$
- b. $\text{Mi } \left\{ \begin{array}{l} \text{sobrina (f)} \\ \text{madre (f)} \end{array} \right\} \text{ es } \left\{ \begin{array}{l} \text{inteligente (f)} \\ \text{alta (f)} \end{array} \right\} .$
 ‘My $\left\{ \begin{array}{l} \text{niece} \\ \text{mother} \end{array} \right\}$ is $\left\{ \begin{array}{l} \text{intelligent} \\ \text{tall} \end{array} \right\} .$

- Adjectives can be classified into four types:

(9)	<i>Masculine</i>	<i>Feminine</i>	
a.	italiano	italiana	‘Italian’
	bonito	bonita	‘nice’
	crudo	cruda	‘raw’
	zurdo	zurda	‘left-handed’
b.	posible		‘possible’
	verde		‘green’
	azul		‘blue’
c.	anglosajón	anglosajona	‘Anglo-Saxon’
	español	española	‘Spanish’
	holgazán	holgazana	‘lazy’
	grandote	grandota	‘big’ (augmentative)
d.	cosmopolita		‘cosmopolitan’ (*cosmopolito)
	agrícola		‘agricultural’ (*agrícola)
	belga		‘Belgian’ (*belgo)

- inner core words with -o in the masculine and -a in the feminine (9a)
- outer core words lacking word markers (9b)
- "gentilic" adjectives, mostly with national, geographical, or ethnic reference; these have no word marker in the masculine but -a in the feminine (9c)
- residual type where both masculine and feminine have -a marker (9d)

- Missing from (9):
 - the converses of (9c) and (9d) do not occur.
 - converse of (9c) would be pairs where the masculine is inner core, with marker -o, whereas the feminine is outer core, with no word marker (*grandoto (m)|grandote (f)).
 - converse of (9d) would be pairs where marker -o is neutral with respect to concord. In actuality, every adjective with -o is only masculine.
 - there is no adjective that can be only masculine or feminine

2.3 Nouns

- All nouns have gender, either masculine or feminine.
- For human nouns, grammatical gender usually matches conceptual gender, but doesn't have to.
- Gender is arbitrary; it does not correlate with meaning or form.

(11)	<i>Masculine</i>	<i>Feminine</i>		
a.	domicilio_	residencia_	'home'	'residence'
	pájaro_	ave	'bird'	'bird'
	ostión	almeja_	'oyster'	'clam'
	asiento_	silla_	'seat'	'chair'
	ratón_	rata_	'mouse'	'rat'
	sapo_	rana_	'toad'	'frog'
b.	libro_	libra_	'book'	'pound'
	caso_	casa_	'case'	'house'
	paso_	pasa_	'step'	'raisin'
	palo_	pala_	'stick'	'shovel'

- Nonhuman outer core and residue words freely belong to either gender

(12)	<i>Masculine</i>	<i>Feminine</i>		
a.	frente	frente	'front'	'forehead'
	orden	orden	'order'	'order'
			(succession)	(command)
	moral	moral	'blackberry bush'	'morality'
	corte	corte	'cut'	'court'
	pez	pez	'fish'	'pitch'
				(resin)

b.	análisis	síntesis	'analysis'	'synthesis'
	brindis	crisis	'toast'	'crisis'
			(drink)	
	cruce	cruz	'crossing'	'cross'
	pene	ene	'penis'	'(letter) n'

- "There exists, however, a striking asymmetry among residual nouns."
- nearly 600 exemplars: masculine nouns (human and nonhuman) with word marker -a (13a). This is the largest class in the residue.
- only 1 guaranteed: feminine noun with word marker -o in common use.

(13) a. *Masculine*

telegrama	'telegram'
profeta	'prophet'
poeta	'poet'
día	'day'
problema	'problem'
síntoma	'symptom'
drama	'drama'
mapa	'map'
<i>and approximately 600 others</i>	

b. *Feminine*

dinamo/dínamo	'dynamo'
virago	'virago'
nao	'ship'
libido/lívido	'libido'
mano	'hand'
<i>and no others</i> ¹⁴	

"Mating" in nouns

(14) Human nouns are "mated": both a masculine and a feminine form exist for each human noun.

(15)	<i>Masculine</i>	<i>Feminine</i>	
a.	secretario_	secretaria_	'secretary'
	campesino_	campesina_	'peasant'
	cocinero_	cocinera_	'cook'
	criado_	criada_	'servant'
	alumno_	alumna_	'student'
	amigo_	amiga_	'friend'
b.	estudiante		'student'
	intérprete		'interpreter'
	cómplice		'accomplice'
	esquimal		'Eskimo'
	caníbal		'cannibal'
	cónyuge		'spouse'
	mártir		'martyr'
	joven		'young man'/'young woman'
c.	(presid)ente	(presid)enta_	'president'
	(sirvi)ente	(sirvi)enta_	'servant'
	(profes)or	(profes)ora_	'professor'
	colegial	colegiala_	'student in a <i>colegio</i> '
	doncel	doncella_	'lad'/'lass'
	monje	monja_	'monk'/'nun'
	nene	nena_	'child'
	jefe	jefa_	'chief'
d.	duque	duquesa_	'duke'/'duchess'
	poeta_	poetisa_	'poet'/'poetess'
	actor	actriz	'actor'/'actress'
e.	(aristó)crata		'aristocrat'
	(art)ista		'artist'
	(mon)arca		'monarch'
	camarada		'comrade'
	acrobata		'acrobat'
	patriota		'patriot'
	suicida		'(person who commits) suicide'
	policía		'police'
	colega		'colleague'

(b) "No word marker (outer core); each can be either masculine or feminine. This set is large enough to defy exhaustive listing."

(c) "Outer core masculines (plus residual *nene*)' mated to (inner core) feminines with word marker -a. This set is listable"

(d) "Masculine and feminine related by derivational suffixation. This set is listable and relatively small."

(e) "Word marker -a; each can be either masculine or feminine....This set is sizeable and heterogeneous."

	<i>Masculine</i>	<i>Feminine</i>	
	pirata_		'pirate'
	guía_		'guide'
f.	contralto_		'contralto'
	soprano_		'soprano'
	testigo_		'witness'
	modelo_		'model'
g.	macho_	hembra_	'male'/'female'
			(*hembro/*macha)
	hombre	mujer	'man'/'woman'
	padre	madre	'father'/'mother'
	yerno_	nuera_	'son-'/'daughter-in-law'
			(*nuero/*yerna)

(f) "Word marker -o; each can be either masculine or feminine. Contains only the examples shown "

(g) "Suppletive mates. List is exhaustive, barring oversight."

- Exceptions to (14) are rare. Just 4 "solid" cases found:

(16)	<i>Masculine</i>	<i>Feminine</i>		
	marido_	*marida_	'husband'	'wife'
	*criaturo_	criatura_	'baby'	
	*persono_	persona_	'person'	
	*víctimo_	víctima_	'victim'	

- **Sociolinguistic comments:** there is increasingly more dialectal and idiolectal variation with respect to "mating" patterns.
- Nonhuman nouns not marked in same way as human.
 - general case (18a): one gender, arbitrarily masculine or feminine, is used for both male and female of a given species
 - In a few cases (18b) the names of males and females of the same species are related by suppletion or derivational affixation.
 - Only a few familiar animal names (18c), are like human prototypical inner core nouns.

(18)	<i>Masculine</i>	<i>Feminine</i>		
a.	camello_	*camella_	'camel'	
	reno_	*rena_	'reindeer'	
	erizo_	*eriza_	'hedgehog'	
	dinosaurio_	*dinosauria_	'dinosaur'	
	*foco_	foca_	'seal'	
	*cebro_	cebra_	'zebra'	
	*ardillo_	ardilla_	'squirrel'	
	*jirafa_	jirafa_	'giraffe'	
b.	toro_	vaca_	'bull'/'cow'	(*vaco_, *tora_)
	caballo_	yegua_	'horse'/'mare'	(*yeguo_, *caballa_)
	carnero_	oveja_	'ram'/'ewe'	(*ovejo_, *carnera_)
	gallo_	gallina_	'cock'/'hen'	(*gallino_, *galla_)
c.	perro_	perra_	'dog'	
	gato_	gata_	'cat'	

2.4 Summary

Table 3

	Noun	Pronoun	Adjective	Determiner	Adverb
Sex reference	±	±	±		
Lexical gender	+	±			
Gender concord target			+	+	
Form class	(5a–g) (6)	(5a–b) (6)	(5a–d,g) (6)	(5a,b,e) (6)	(5a–c) (6)

3. Analysis

- Goal: Propose solution to the "mating problem": account for the attested associations between conceptual gender (semantics), grammatical gender (syntax) and form class (morphology)

3.1 Basic machinery

- Formal representation of **delension class** (i.e. form class) must be distinct from formal representation of gender. (We saw in (5) an (6) that each form class is mixed with respect to gender).
- Proposed notation for lexical entries:

(19) a. Gender: f = feminine
 Class:]a = words with marker -a
]∅ = words with no marker (phonetic -e ~ ∅ depending on syllable structure)

b.	libro (m)	libra (f)	libre	día (m)
Meaning	'book'	'pound'	'free'	'day'
UR of stem	/libr/	/libr/	/libr/	/di/
Category	N	N	A	N
Gender		f		
Class]∅]a

- **masculine** is entirely unmarked/default, gender is not binary [+/-fem] (see (20))
- form class is non-binary. -o class is unmarked. The other classes are marked by non-binary diacritics

- Form class of [f] words in inner core is entirely predictable and so does not need to be listed in lexical entry. It can be supplied instead by a **redundancy rule** (a late rule (usually morphological or phonological) that is entirely predictable from a feature specification (hence redundnat).. The rule usually enriches the representation.)

(21) *Feminine Marker Rule*

$f \rightarrow]a$

(22) *Marker Realization Rule*

- a. To form the X^0 level of nouns, adjectives, and adverbs, insert suffixal /a/ if the stem is marked]a; otherwise, insert suffixal /o/.

b. $\emptyset \rightarrow \left\{ \begin{array}{l} [+low] \\ [+round] \end{array} \right\} / \left[\left[\left\{ \begin{array}{l}]a \\ (else) \end{array} \right\} \right]_{N/A} \text{ — }]_{X^0} \right.$

[+low] picks out /a/

[+round] picks out /o/

- Being marked with the diacritic] \emptyset exempts a word from the Marker Realization rule; this gives us the set of words that do not take word markers.
- Crucially, the Marker Realization Rule contains no mention of grammatical or conceptual gender. The rule expresses a purely formal generalization.
- Derivation of words in (19b)

(23) *Lexical entries*

libro	libra	libre	día
/libr/	/libr/	/libr/	/di/
N	N	A	N
	f		
] \emptyset]a

- (25) a. No adjective stem is marked lexically for gender (as either exclusively masculine or exclusively feminine).
 b. Gender in adjectives is supplied through concord.
 c. Many adjectives have *-a* for both genders.
 d. No adjective has *-o* for (both masculine and) feminine.

- (25a) captured by (26)

(26) *Constraint on Lexical Entries*

*[Adj, f]

- Generalization (25c) requires many adjectives to be lexically specified for declension class.

(27)	agrícola	belga
Meaning	'agricultural'	'Belgian'
UR of stem	/agricol/	/belg/
Category	A	A
Gender		
Class]a]a

- (25d) is explained by the proposal that *-o* is the unmarked case.

"If there was a lexical diacritic]o parallel to]ø and]a, then adjective stems could bear this diacritic as a lexical peculiarity. In that case there would be no natural way of ruling out the set of forms that does not occur." (p. 48).

- Idiosyncratic class of gentilic adjectives (9c) requires additional machinery. These have no word marker in masculine, but *-a* in feminine. Harris proposes (28), which supplies the]a diacritic to feminine forms and exempts masculine forms from the Marker Realization rule.

(28) *Gentilic Rule*

$$g \rightarrow \left\{ \begin{array}{l}]a / [\text{---}, f] \\]\emptyset \text{ (elsewhere)} \end{array} \right\}$$

(29) Lexical entries	crudo/-a 'raw' /krud/ A	verde 'green' /berd/ A]Ø
Concord	" " "	/krud/ f "]Ø

Gentilic Rule (28)

Feminine Marker (21)

Marker Realization (22)

Syllabification

]a	<i>block</i>
[krud]o	[krud]a	[berd]
kru.do	kru.da	ber.de

Lexical entries

galés/-a
'Welsh'
/gales/
A
g

belga
'Belgian'
/belg/
A
]a

Concord	" " "	/gales/ f g	" " "	/belg/ f]a
---------	-------------	-------------------	-------------	-------------------

Gentilic Rule (28)

Feminine Marker (21)

Marker Realization (22)

Syllabification

]Ø]a		
<i>block</i>			
[gales]	[gales]a	[belg]a	[belg]a
ga.les	ga.le.sa	bel.ga	bel.ga

3.4 Nouns

3.4.1 Core nonhuman nouns

- Machinery developed so far for adverbs and adjectives is sufficient for core non-human nouns.

(30)		<i>UR</i>	<i>Gender</i>	<i>Marker</i>
a.	libro	/libr/		
b.	libra	/libr/	f	
c.	orden	/orden/]Ø
d.	orden	/orden/	f]Ø
e.	drama	/dram/]a

Libro exemplifies the class of masculine inner core nouns. Gender and word marker class are unspecified; default word marker *-o* is supplied by Marker Realization.

Libra exemplifies the class of feminine inner core nouns. Gender is specified; word marker class is not. Rule (21), Feminine Marker, predicts marker class *]a* from marked gender; Marker Realization guarantees the phonological shape */a/* for the word marker.

Orden (30c) exemplifies the class of masculine outer core nouns. Gender is unspecified; word marker class is specified as *]Ø*, which exempts the stem from Marker Realization.

Orden (30d) exemplifies the class of feminine outer core nouns. Gender is marked; word marker class is specified as *]Ø*, which exempts the stem from Marker Realization.

Drama exemplifies the class of residual masculine nouns with word marker *-a*. Gender is unspecified; marker class is idiosyncratically specified as *]a*, which triggers Marker Realization to guarantee word marker *-a*.

p. 50

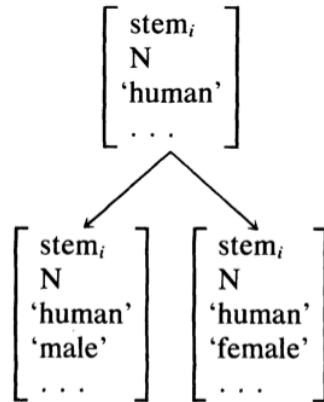
- What about *mano*? It cannot be generated in this system.
 - It must have [f] specified (we know from concord, etc.
 - But [f] cannot be realized as */o/* in this analysis
 - the */o/* must be lexicalized. (Same holds for other residue cases ending with unpredictable V(s)).

3.4.2 Core human nouns

- Two fundamental observations made in 2.3:
 - In human nouns, grammatical gender matches conceptual gender (with rare exceptions)
 - both a masculine and a feminine form exist for each human noun

(32) a. *Human Gender*
 'female' → f / [____, 'human']

b. *Human Cloning*



The effect of Human Cloning is this: for every lexical entry L of a noun stem that contains the semantic specification 'human' without specification of sex, Human Cloning replaces L with a pair of entries L_m and L_f , each identical to L except for the addition of the semantic specifications 'male' and 'female', respectively. Of course, Human Cloning is not deeply explanatory; it simply encodes the observation made in English prose in (14). As a structure-building redundancy rule, however, it has the desirable formal property of blocking when its output would duplicate corresponding elements of an existing lexical entry. Consider, for example, the entries in (33), which exemplify the set of suppletive mated pairs illustrated in (15g):

p. 51

(33)	padre	madre	yerno	nuera
Meaning	'parent'	'parent'	'child's spouse'	'child's spouse'
	'male'	'female'	'male'	'female'
UR	/padr/	/madr/	/yern/	/nuer/
Category	N	N	N	N
Gender				
Class]Ø]Ø		

- Suppletive pairs of mated human nouns are exceptional. The vast majority undergo cloning in Harris' analysis.

(34) Lexical entries	amigo/-a		juez	
	/amig/		/xwes/	
	N		N	
	'friend'		'judge'	
]∅	
Cloning (32b)	/amig/	/amig/	/xwes/	/xwes/
	N	N	N	N
	'friend'	'friend'	'judge'	'judge'
	'male'	'female'	'male'	'female'
]∅]∅
Gender (32a)		f		f
Feminine Marker (21)]a]a
Marker Realization (22)	[amig]o	[amig]a	<i>block</i>	<i>block</i>
Syllabification	a.mi.go	a.mi.ga	xwes	xwes

- The remaining subtypes of nouns can be dealt with the existing machinery

(35) a.	b.	c.	d.	e.
jefe/-a	gato/-a	colega	marido	persona
/xef/	/gat/	/koleg/	/marid/	/person/
'chief'	'cat'	'colleague'	'spouse'	'person'
'human'	'human'	'human'	'human'	'human'
			'male'	
g]a		f

- (35a) uses same]g diacritic as gentilic adjectives
- (35b) accommodates exceptional animal nouns that have mated nouns; they are lexically [human] (exploiting diacritic use of features)
- (35c) is lexically marked for]a but not for gender, so we have -a form for both genders. Derivation is like *belga* (29) except cloning rule, rather than concord, generates gender alternatives.
- (35d) and (e) proceed as per what we've already seen

4. Summary

- Spanish morphology has 7 types of word marker used across multiple grammatical categories, interacting with the expression of grammatical and semantic gender in complex ways.

- Harris models this system by dissociates gender from form class and proposes an abstract feature system consisting of
 - a single grammatical gender marker,
 - three form-class diacritics]a,]ø and]g,
- And a rule system consisting of:
 - the redundancy rules (21) and (28)
 - the Marker Realization rule
 - the Gentilic Rule.