The Third Phi Feature: Gender
Day 4, A .Nevins, `egg.auf.net`

Research Programmatic Ideas Explored Today:

- The autosegmental nature of operations on the geometry is displayed by re-association
- The markedness relationships that obtain between gender and number can be determined by the geometry

1 Delinking and Relinking

(1) El libro, a ellos, quien se los prestó?
The book, to them, who ??? ??? loaned?
"who lent the book to them"

(2) Si ella me quiere comprar el caballo, yo se la venderé
if she dat-1sg wants buy the horse, I ??? ??? will-sell
"If she wants to buy my horse, I will sell it to her"

Suppose that the delinking rule removes *everything*[1] under the node representing the clitic (except the argument terminal). The complex that included [Dat][Masc][Grp] will be realized by the default item *se*. However the surprise is that the neighboring accusative clitic, otherwise doubling a singular argument (*el libro*) shows up as plural. This happens only and exactly when the plural feature is delinked from the Dative clitic (which would be otherwise realized as *les*).

The same thing appears to happen with the feminine feature in the second example. The dative clitic is doubling a feminine argument; when it is delinked, the feminine feature reassociates with the accusative clitic, which should otherwise be masculine (as it is doubling *el caballo*). The even more interesting fact here is that gender not even overtly realized in the Dative when it stands alone, as the case feature conditions insertion of the *-e* theme usually. The facts here indicate not only that the feminine feature is present in the dative despite its nonovert appearance, but that the delinking operation is more than merely the failure to realize a feature, as that feature shows up on a clitic that syntactically it should not belong with.

Combinations of both (for example *si las muchachas me quieren...* show transfer of both feminine and plural to the (otherwise masc.sg.) accusative clitic. The autosegmental relinking is thus not limited to displacement of one feature.

[1] We will discuss how Case might be integrated into the feature geometry tomorrow.

1.1 Delinking vs Licensing

2 Sex vs Gender vs Class

(3) "Genders are classes of nouns reflected in the behavior of *associated words*" (Hockett 1958): in other words the determining criteria for gender is agreement.

The El Cheapo Problem: Is there Markedness among gender[2] features? How do we capture the fact that *-o* is overwhelmingly the predominant ending for Spanish *adverbs, adjectives, nouns, pronouns, and determiners.* A pervasive idea in linguistics textbooks is that endings such as *-o* and *-a* are the exponents (phonological realization of; i.e. they *expose*) gender. One needs to keep apart: 1) biological *sex*, 2) grammatical *gender* (controlling agreement, for example) and 3) theme *class.* Each of these deserve their own formalizations: 1) in biology/semantics, 2) in the syntax of agreement, and 3) in the morphophonology of individual lexical items. Of course, there are preferences for sex and grammatical gender to correlate, and for gram.gender and class. The first of these is implemented by the Harris's "Cloning Procedure" and the second by the nature of the Theme redundancy rule for feminine gender.

There are many attested pairings of opposed gender:

(4)

muchacho	muchacha	'boy'/'girl'
sirviente	sirvienta	'servant'
doctor	doctora	'doctor'
cónyuge	cónyuge	'spouse'

We'll call *-o* class 1 and *-a* class 2. There's a negative evidence problem: how does a learner know that not all pairings are possible? How does a Spanish speaker know that a pair like *tindra/tindre* could never exist? Must be something in UG!

How do we know masculine is default? Take the use of a metalinguistically mentioned preposition *para* ('for'):

(5) Tienes demasiados "paras" en ese párrafo.
You have too-many(masc) "fors" in that paragraph.

-e shows up after phonologically unsyllabifiable sequences such as *dr* (*padre*) and *nt* (*amante*). It is the case of having no theme class.

(6) Cloning procedure: for a pair of two biological-sex opposed human nouns, the male-sex will have male gender and the female sex will have female gender

Note that (6) does not apply in the case of nonhuman nouns, which have inherent gender.

[2]NB: Classifiers never trigger agreement, so we ignore them in a discussion of phi-features.

(7) Class 2 Redundancy: If unspecified, feminine gender → class 2.

(8) o-a (given by cloning procedure and (8))
e-a (given by cloning procedure and (8))
0-a(given by cloning procedure and (8))
e-e (inherently specified noun)
a-a (inherently specified noun)
o-o (inherently specified noun)
o-e
a-o
a-e

How do we know masculine is default? Take the use of a metalinguistically mentioned preposition *para* ('for'):

(9) Tienes demasiados "paras" en ese párrafo.
You have too-many(masc) "fors" in that paragraph.

The Mating Problem Ruling out the impossible noun pairs: *a-o and *a-e: the noun would have to be inherently specified for class 2, which would then be on both nouns. *e-o and *o-e: the noun would have to be inherently specified for class 3 (no theme/*-e* when syllabification needs it), which would be on both nouns.

(10) The solution to both "El Cheapo" and "The Mating Problem" lay in 1) understanding *-o* as the default and 2) a consistent set of spell-out principles involving either uniform inherent specification or a lack thereof which then depended on principles of redundancy rules and default.

These class marking suffixes appear on every Xo category; adverbs have no grammatical gender or intrinsic sex but still always have theme vowels: *dentro, fuera*.

(11) "The class-marking suffixes have no meaning or function; they obey no higher semantic or syntactic authority". (Harris 1991:59)

3 Verbal Pieces: Variations on a Theme

Oltra-Massuet extended Harris's work to an analysis of the Catalan verbal system, building on the notion of vowels in Romance as formatives that are contextually sensitive and bear their own principles of markedness.

Catalan Theme Vowel Allomorphy: There are Four Classes of Verbs:

(12)

Tense	Agr	I	II	IIIb	IIIa
PresInd	sg,3pl	0	0	0	ES
PresInd	1pl,2pl	E	E	i	i
ImpSubj	all	e	e	i	i
Fut/Cond	all	a	0	i	i
ImpInd	all	a	i	i	i
SimpPast	1sg	i	i	i	i
SimpPast	non1sg	a	e	i	i
Infin	n/a	a	0	i	i
Gerund	n/a	a	e	i	i
Participle	n/a	a	u	i	i

(13) Oltra-Massuet's Theme Vowel Markedness Hierarchy:
(where alpha, beta and gamma represent "pure degrees of markedness")
Class I = -alpha (*cantar*)
Class II = +alpha, +beta (*témer*)
Class IIIa = +alpha, -beta, -gamma (*unir*)
Class IIIb = +alpha, -beta, +gamma (*sortir*)

(14) Theme spell-outs:
/u/ ↔+beta in env [participle] [past]
/0/ ↔+beta in env [future]
/e/ ↔+beta elsewhere
/ES/ ↔-gamma in env [past]
/i/ ↔+alpha
/E/ ↔-alpha in env [participle][plural]
/a/ ↔-alpha elsewhere
Th → +alpha in env [future] [past] (conditional)
Th → +beta in env [subjunctive][past]
Th → +alpha in env [subjunctive]

OM provides a theme list such that themes become more marked in the environment of more marked tense/aspect/mood features.

(15) All functional heads require a theme projection

(16) *aguditzaries* 'you would sharpen' (2sg conditional):

(17) $\sqrt{}$agud + [v -alpha] + [mood(fut) +alpha] + [Tense(past) -alpha] + [Agr(Addr,Min)]
agud + [idz a] + [r i] + [0 a] + [z]

How would a purely syntactic story account for all of these pieces?

The "Alpha-alpha" change in the Spanish subjunctive:

(18) /a/ ↔-alpha
/e/ ↔+alpha
/i/ ↔+alpha + beta

In the subjunctive, alpha → -alpha:

(19) habla → hable
come → coma
escribimos → escribamos

This can't be a morphological apophony as the target of the vowel change is always another theme vowel and never /u/ or /o/.

3.1 Semitic Vowels as themes?

d**i**b**e**r → ye-d**a**b**e**r
d**i**b**e**r → dib**a**r-ti
Regularity throughout binyanim: **i,e** → **a** in env of neighboring [Pers] terminal.

Phonologists: "Imperfect derived from Perfect"; but if we view phonological pieces as constructed by the *syntax* rather than through opportunism, an alternative possibility is that there is a root form (+binyan), with zero derivation to perfect and apophonic derivation to imperfect as distinct steps.

Incidentally OT proponents have been denying everything we have learned about Semitic morphophonology including even the existence of the root, saying it is an epiphenomenal analogy over semantically related forms and that in some cases mere consonantal extraction suffices to provide the input for derivational morphology. A knockdown 30 second argument that this can't be right comes from hypocoristics in Jordanian Arabic, which are formed based on the root of the full name. Thus *Mohamed* (from $\sqrt{}$HMD) is *Haamuud*. Guess what the hypocoristic is for *?Ibtisam* (from $\sqrt{}$BSM): *Baasuum*.

4 The Relation of Gender and Number

One of Greenberg's Universals was that any language that agrees in gender agrees in number as well. The organization of Class under the Individuation node makes this inevitable, as even a bare Individuation node (e.g. with no [Min] or [Grp] dependents) will have an interpretation for number.

Moreover the cyclicity hypothesis bears consequences, even if we understand Pers, Num, and Gender as distinct syntactic terminals. Paradigm advocates claim that a few universals can be captured by constraints on the geometric organization of cells and how they are filled. One such claim regards fusion; thus Wunderlich & Fabri 1994 note that person and gender cannot be fused to the exclusion of number. But given a feature hierarchy, this should be a natural result. Suppose insertion of a VI will knock out everything below it on the hierarchy. Then if gender is inserted before number before person, there will be three affixes. If gender is inserted, then nothing matches at number, then a fused person-number affix is inserted, that pattern will obtain. If nothing matches at gender, then a fused number-gender affix is inserted, then person, then that pattern will obtain. On a cyclic insertion procedure, then, these consequences

are natural; no person-gender fusion to the exclusion of number should be able to obtain.

4.1 Case Study: Swahili

Swahili noun class prefixes:

(20)

1	m-tu	person
2	wa-tu	people
3	m-ti	tree
4	mi-ti	trees
5	gari	car
6	ma-gari	cars
7	ki-atu	shoe
8	vi-atu	shoes
9	n-yuma	house
10	n-yumba	houses
11	u-bao	board
14	u-kweli	truth
15	ku-soma	to read

(12 and 13, from Meinhoff's proto-Bantu numbering system, are absent from Swahili). The noun class triggers agreement on adjectives and verbs:

(21) mtoto huyu wangu mzuri a-me-anguka
1child 1this 1my 1good 1agr-perf-fall
'this my good child has fallen down'

(22) watoto hawa wangu wazuri wa-me-anguka
2child 2this 2my 2good 2agr-perf-fall
'these my good children have fallen down'

(23) mti huu wangu mzuri u-me-anguka
3tree 3this 3my 3good 3agr-perf-fall
'this my good tree has fallen down'

(24) miti hii yangu mizuri i-me-anguka
'these my good trees have fallen down'

..And so forth. But there is nothing particularly unusual about these prefixes; the noun prefixes can be analyzed as fused gender and number spellout (or gender sensitive number spellout), while the agreement on adjectives and verbs is number/gender agreement. Carstens (1991) argues against the preceding view in the literature that the prefixes *supply* the class information, proposing instead that the class information in an inherent property of the noun stem, and is indirectly reflected by the context-sensitivity of the number prefixes. The systematic generalizations that class 1 singulars will have class 2 plurals, class 3 singulars will have class 4 plurals, etc, follows straightforwardly. Carstens suggests that there are 5 noun classes in Swahili: A, B, C, D, E. The VI entries

would then be as follows:

(25) /m-/ ↔ singular in env. class A
/wa-/ ↔ plural in env. class A
and so forth..

As was evident above, most class 5 nouns take no overt prefix, e.g. *gari*. However, there is morphophonolgical ("inward") contextual allomorphy: when the stem is monosyllabic, singular is spelled out as *ji-*:

(26) /ji-/ ↔ singular in env. [class C, monosyllabic]
/0/ ↔ singular in env. [class C] /ma/ ↔ plural in env. [class C]:

(27) gari/ma-gari (car/cars)
ji-cho/ma-cho(eye/eyes)

For classes 9 and 10, there is no overt difference in prefixes. Rather than analyzing this as the result of homophony, we can posit an impoverishment rule that deletes plural[3] on the noun stem in of class E. This kind of impoverishment must be postsyntactic, as there is distinct agreement on verbs:

(28) ndizi hii i-me-iva
banana 9this 9agr-perf-ripen

(29) ndizi hizi zi-me-iva
banana 10these 10agr-perf-ripen

There are interesting phenomena related to class changes under diminutive and augmentative formation ,and locatives. But the point should be clear: Swahili is just like Spanish is just like Hebrew: noun classes/"genders" are an inherent property of roots, and trigger allomorphy on number affixes, as well as agreement on verbs, that match their class.

4.2 Polarity in Somali

We can think more about negative values by considering the definite article in Somali:

(30)

	sg	pl
masc	kii	tii
fem	tii	kii

Surely /kii/ and /tii/ both being accidentally homophonous is implausible. Alpha rules creeping up again?

(31) kii ↔alpha(Fem), alpha(Plural)
tii ↔elsewhere

[3]The implementation could involve /n/ as the spellout of a bare individuation node, among other alternatives.

We return to the interpretation of bare nodes. Suppose the lg-(not-so-)particular interpretation of the Individuation node is [Masc] [Min].

(32) kii ↔[Ind] (strictly defined with no dependents)
tii ↔elsewhere

The operative impoverishment rule simplifies the most marked structure, namely that with both [Fem] and [Grp] instantiated.

4.3 Frequent Neutralizations in Marked Number

The most marked number often shows impoverishment of gender: the neuter and feminine are syncretic in the dual in Slovenian.

5 Variation in the Marked Gender

HR's geometry contains no inherent preference for markedness of gender, since it is often subject to variation. The best way to tell may be by looking at what is default agreement in cases of, say, agreeing with an infinitive – it is neuter in Russian, inanimate (vs. animate) in Menomini.

5.1 Conjunct Resolution

Another instance that can inform markedness is what happens in conjunct resolution. For example, a conjoined feminine and neuter in Serbian result in a masculine agreement:

(33) znanje i intuicija su kod njega sarajdival-i i dopunjaval-i se...
knowledge-neut and intuition-fem are in him worked-together-masc-pl and supplemented-masc-pl. each other

5.2 A Functionalist Explantion for Gender in the 3rd?

It has been noticed that many languages show gender distinctions only in the third person. There is nothing inherent in the geometry (besides perhaps the fact that [Addr] and [Auth] involve more nodes than third person) that would predict this persay. A functionalist explanation can be called in, perhaps, for the rampant impoverishment of gender in env of [Auth] and [Addr], namely there is much less ambiguity about the identity of the discourse participants, who are present. However, this only goes so far (and not far, really) since there is commonly adjectival concord with [Auth] and [Addr]. Since neither I nor many others have a theory of concord (short of some speculations about autosegmental spreading) we may have to leave this aside for the moment.

6 Exercise: Hindi Agreement

Describe what's going on in each of the following sentences. What do the agreement features mark?

(34) Meyn roTii khaa-un-g-ii
I-fem bread will-eat(with agreement)

(35) Ham roTii khaa-en-g-ee
We-masc bread will-eat(with agreement)

(36) Sumita roTii khaa-e-gii
Sumita-fem-sg bread will-eat(with agreement)

(37) Ramesh roTii khaa-e-gaa
Ramesh-masc-sg bread-fem-sg will-eat(with agreement)

(38) Ve laRke roTii khaa-en-g-ee
Those boys-masc-pl bread-fem-sg will-eat(with agreement)

(39) Ramesh ne kelaa khay-aa
Ramesh-masc-sg banana-masc-sg ate(with agreement)

(40) Sumita ne kelaa khay-aa
Sumita-fem-sg banana-masc-sg ate(with agreement)

(41) Ve laRke kelaa khay-aa
Those boys-masc-pl banana-masc-sg ate(with agreement)

(42) Ramesh ne roTii khay-ii
Ramesh-masc-sg bread-fem-sg ate(with agreement)

(43) Ramesh ne kelee khay-ee
Ramesh-masc-sg bananas-masc-pl ate(with agreement)

(44) Ramesh ne kelaa-ko khay-aa
Ramesh-masc-sg the-banana-masc-sg-specific ate(with agreement)

(45) Ramesh ne roTii-ko khay-ii
Ramesh-masc-sg the-bread-fem-sg-sepcific ate(with agreement)

(46) Ramesh ne kelee-ko khay-ee
Ramesh-masc-sg the-bananas-masc-pl-specific ate(with agreement)