Proportional analogies in typed systems (HPSG?)

Matías Guzmán Naranjo - Universität Düsseldorf

EW-HPSG
02.06.2018

Proportional analogies

What are PA models

In PA (proportional analogy) models of morphology, the different inflected forms of a word are not produced by inflectional rules, but simply associated by analogical relations.

A proportional analogy is expressed by the formula:

$$A : B :: C : X \tag{1}$$

meaning A is to B like C is to X. Where if we know A, B and C, we can deduce X .

What are PA models

As an example:

(1) a. compr**o**
buy.1sg
b. compr**as**
buy.2sg

(2) a. par**o**
stop.1sg
b. X
stop.2sg

compro:compras::paro:X

⇒ X=*paras*

What are PA models

But also:

(3)

amár	:	amará	:	amára
parár	:	parará	:	parára
LOVE.INF		LOVE.3SG.FUT.IND		LOVE.1/3SG.PRES.SUBJ

For which there is no obvious morphemic segmentation, or function/semantic relation.

What are PA models

We can express analogies between paradigm cells by abstracting away the phonological material common to those cells:

(4)

Ý	:	Xá	:	Ýa
.INF		.3SG.FUT.IND		.1/3SG.PRES.SUBJ

These are not ‘rules’ because the *X* and the segment *a* does not correspond to a morpheme and there is no feature composition between these cells.

I will call this the ‘X-notation’

Advantages of PA models

There are several advantages of PA models:

- no need for morphemes
- (thus) no segmentation problem
- no transformations,
- no special rules,
- no rule orderings,
- no feature decompositions.

In short: they are very **simple** and intuitive.

Problems with PA models

And then there are problems (at least with the X-notation):

- there is no clear interface between the morphology and the rest of the grammar,
- only one level of abstraction (e.g. no abstraction across inflection classes),
- their status is unclear (what exactly are proportional analogies and how do they work!?),
- not properly formalized,
- hard(er) to implement as a complete system,
- incompatible with theories that require lexemes (or not? hard to tell).

Problems with PA models

A concrete example with some Spanish verb cells:

	-ar		-er		-ir	
	regular	o-ue	regular	o-ue	regular	o-ue
infinitive	cant-ár	prob-ár	com-ér	dol-ér	sub-ír	mor-ír
1sg present	cánt-o	pruéb-o	cóm-o	duél-o	súb-o	muér-o
2sg (tú) present	cánt-as	pruéb-as	cóm-es	duél-es	súb-es	muér-es
2sg (vos) present	cant-ás	prob-ás	com-és	dol-és	sub-ís	mor-ís
3sg future	cant-ará	prob-ará	com-erá	dol-erá	sub-irá	mor-irá
1sg imp.subj	cant-ára	prob-ára	com-iéra	dol-iéra	sub-iéra	mur-iéra
	‘sing’	‘taste’	‘eat’	‘hurt’	‘go up’	‘die’

We can analyze the Spanish verbal inflection system as being composed of a stem (process), a stress pattern and a suffix or set of suffixes.

Problems with PA models

We can define concrete analogical relations between any two cells in the previous examples.

(5) a. amár:ámas::cantár:X
b. Yr:Ys
c. X = cántas

However, we cannot properly specify analogies covering more general patterns. For example, the analogy in (5) only covers *ar* and *er* verbs, but not *ir* verbs: *subir-subes*
Howerver, we miss the fact that for all three classes the second person singular present indicative has an *-s* marker.

Problems with PA models

The difficulty with analogies is that it is not clear how to express independent partial patterns (stress, stem alternations, suffix markers), and then combine these individual patterns together.
A naive approach would not work. The stem alternations alone represent a problem:

(6) a. Xar:Xo
b. XoYar:XueYo

(6)a cannot capture stem alternating verbs, while (6)b cannot capture non-alternating verbs. Similarly, a partial analogy only for stem alternations as in (7) could not be unified with a partial analogy like in (6)a.

(7) a. XoY:XueY

More over, (7) is exclusively making reference to stems (it does not mention suffix makers), but PA approaches do not assume stems.

Problems with PA models

Some generalizations we would like to express, but cannot:

- *-o* marks first person singular present indicative,
- stress patterns are identical across inflection classes,
- stem-*Vs* marks second person singular,
- some verbs share exactly the same stem alternation process even though they belong to different thematic vowels.

Problems with PA models

Without proper formalization it is not clear how we can solve these issues, and whether these are limitations of the notation or of PAs themselves.

Objectives

Objectives

This talk has aims to present a path to the formalization of a PA model which:

- retains he simplicity characteristic of analogical models (no rules, orderings, stems, morphemes, etc.),
- is completely surface oriented,
- allows for partial analogies,
- allows for lexemes,
- could potentially interact with syntactic theories,
- and lets us test the limits of analogy.

Basic assumptions

Basic assumptions

The basic assumptions are:

- Lexemes bundle together all their possible realizations.
- Analogical relations hold across phonological strings, not words.
- Analogical relations are constrains on phonological strings.
- Analogical relations are organized in an inheritance hierarchy according to degree of abstraction.

The basics

$$\begin{bmatrix} \textit{lexeme} & \\ \textsc{features} & \textit{fts-obj} \\ \textsc{paradigm} & \left\{ \begin{bmatrix} \textit{cell} & \\ \textsc{features} & \textit{fts-obj} \\ \textsc{phon} & \textit{phon-obj} \end{bmatrix}, \begin{bmatrix} \textit{cell} & \\ \textsc{features} & \textit{fts-obj} \\ \textsc{phon} & \textit{phon-obj} \end{bmatrix} \right\} \\ & \ldots \end{bmatrix}$$

Where FTS contains whatever features are associated with any given cell in the paradigm (syntax, agreement, semantics, pragmatics, etc.), and PHON a phonological string (whatever its representation may be).

Phonological representation?

$$\begin{bmatrix} \textit{phon-obj} \\ \textsc{word}\ \textit{list}(\textit{syllable-obj}) \end{bmatrix} \begin{bmatrix} \textit{syllable} \\ \textsc{onset}\ \ldots \\ \textsc{rime} \begin{bmatrix} \textsc{nucleus}\ \ldots \\ \textsc{coda}\ \ldots \end{bmatrix} \end{bmatrix} \begin{bmatrix} \textit{vowel} \\ \textsc{high}\ +/- \\ \textsc{round}\ +/- \\ \textsc{back}\ +/- \end{bmatrix}$$

cantar ‘to sing’

$$\left[\begin{array}{l} \textit{lexeme-cantar} \\ \text{PARADIGM} \left\{\begin{array}{ll} \left[\text{FTS } \textit{1sg.pres.ind} & \text{PHON } \textit{/\#canto\#/}\right], \\ \left[\text{FTS } \textit{2sg.pres.ind} & \text{PHON } \textit{/\#cantas\#/}\right], \\ \left[\text{FTS } \textit{3sg.pres.ind} & \text{PHON } \textit{/\#canta\#/}\right], \\ \left[\text{FTS } \textit{1pl.pres.ind} & \text{PHON } \textit{/\#cantamos\#/}\right], \\ \left[\text{FTS } \textit{2pl.pres.ind} & \text{PHON } \textit{/\#cantáis\#/}\right], \\ \left[\text{FTS } \textit{3pl.pres.ind} & \text{PHON } \textit{/\#cantan\#/}\right], \\ \ldots & \end{array}\right\} \end{array}\right]$$

And similarly for *comer* and *subir*.

Ideally, however, we would like to define objects between the fully abstract *lexeme* type, and the fully specified *lexeme-cantar* type.

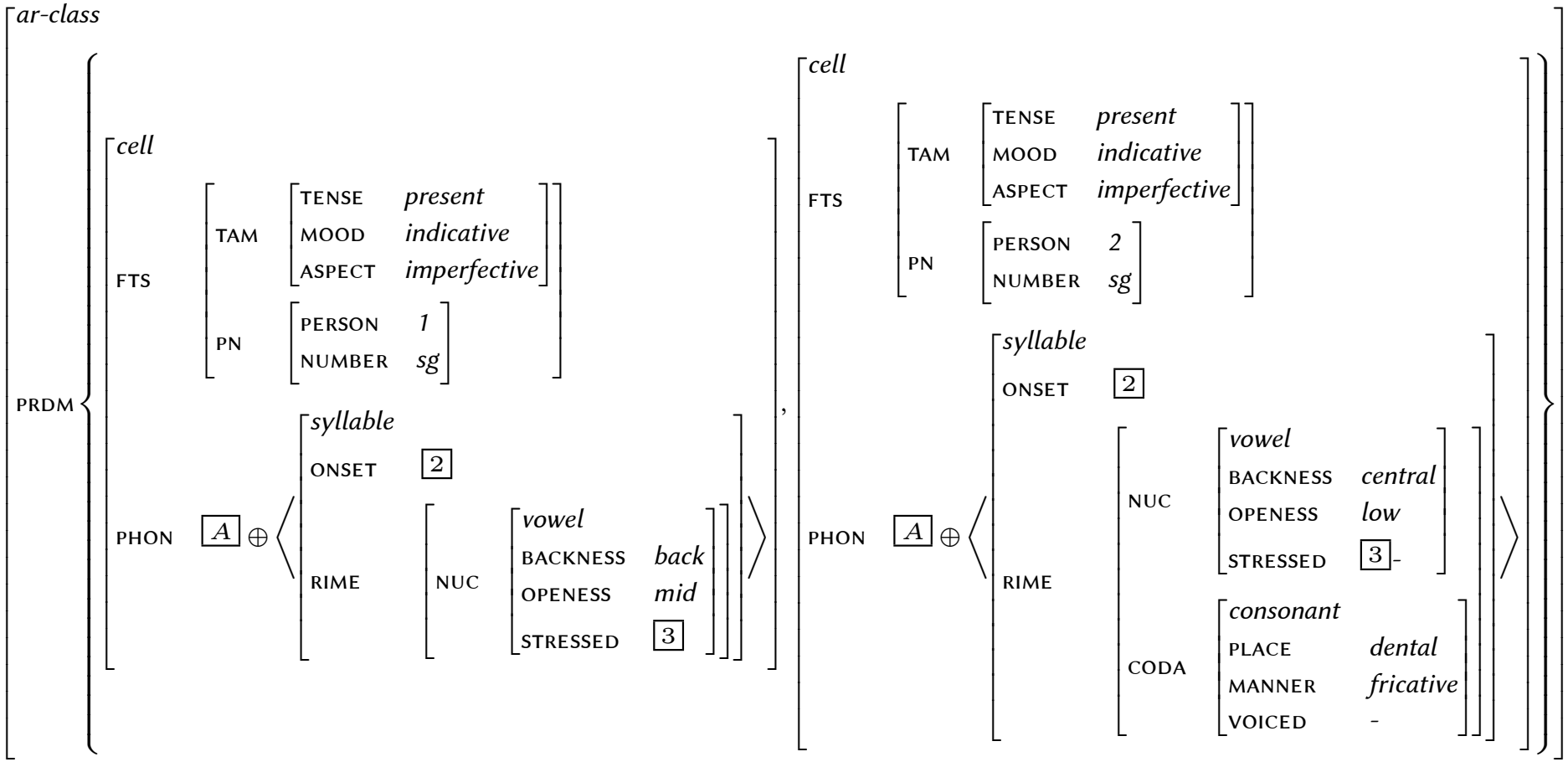
The idea of PA models is that we can reconstruct missing forms in a paradigm.

We do this by abstracting away the proportional analogies across all phonological strings in the paradigm of *cantar*, *comer*, *subir*, etc.

Analogies

$$\left[\begin{array}{l} \textit{ar-class} \\ \textsc{p}\left\{\begin{array}{ll} \left[\textsc{fts}\ \textit{1sg.prs} & \textsc{ph}\ /X^1\,o\#/\right] \\ \left[\textsc{fts}\ \textit{2sg.prs} & \textsc{ph}\ /X^1\,as\#/\right] \\ \left[\textsc{fts}\ \textit{3sg.prs} & \textsc{ph}\ /X^1\,a\#/\right] \\ \left[\textsc{fts}\ \textit{1pl.prs} & \textsc{ph}\ /X^2\,amos\#/\right] \\ \left[\textsc{fts}\ \textit{2pl.prs} & \textsc{ph}\ /X^2\,ais\#/\right] \\ \left[\textsc{fts}\ \textit{3pl.prs} & \textsc{ph}\ /X^1\,an\#/\right] \\ \ldots & \end{array}\right\} \end{array}\right] \Vert \left[\begin{array}{l} \textit{er-class} \\ \textsc{p}\left\{\begin{array}{ll} \left[\textsc{fts}\ \textit{1sg.prs} & \textsc{ph}\ /X^1\,o\#/\right] \\ \left[\textsc{fts}\ \textit{2sg.prs} & \textsc{ph}\ /X^1\,es\#/\right] \\ \left[\textsc{fts}\ \textit{3sg.prs} & \textsc{ph}\ /X^1\,e\#/\right] \\ \left[\textsc{fts}\ \textit{1pl.prs} & \textsc{ph}\ /X^2\,emos\#/\right] \\ \left[\textsc{fts}\ \textit{2pl.prs} & \textsc{ph}\ /X^2\,eis\#/\right] \\ \left[\textsc{fts}\ \textit{3pl.prs} & \textsc{ph}\ /X^1\,en\#/\right] \\ \ldots & \end{array}\right\} \end{array}\right] \Vert \left[\begin{array}{l} \textit{ir-class} \\ \textsc{p}\left\{\begin{array}{ll} \left[\textsc{fts}\ \textit{1sg.prs} & \textsc{ph}\ /X^1\,o\#/\right] \\ \left[\textsc{fts}\ \textit{2sg.prs} & \textsc{ph}\ /X^1\,es\#/\right] \\ \left[\textsc{fts}\ \textit{3sg.prs} & \textsc{ph}\ /X^1\,e\#/\right] \\ \left[\textsc{fts}\ \textit{1pl.prs} & \textsc{ph}\ /X^2\,imos\#/\right] \\ \left[\textsc{fts}\ \textit{2pl.prs} & \textsc{ph}\ /X^2\,is\#/\right] \\ \left[\textsc{fts}\ \textit{3pl.prs} & \textsc{ph}\ /X^1\,en\#/\right] \\ \ldots & \end{array}\right\} \end{array}\right]$$

Analogies


ar-class
PRDM
cell
FTS
TAM
TENSE present
MOOD indicative
ASPECT imperfective
PN
PERSON 1
NUMBER sg
PHON A ⊕ ⟨
syllable
ONSET 2
RIME
NUC
vowel
BACKNESS back
OPENESS mid
STRESSED 3
⟩
,
cell
FTS
TAM
TENSE present
MOOD indicative
ASPECT imperfective
PN
PERSON 2
NUMBER sg
PHON A ⊕ ⟨
syllable
ONSET 2
RIME
NUC
vowel
BACKNESS central
OPENESS low
STRESSED 3-
CODA
consonant
PLACE dental
MANNER fricative
VOICED -
⟩

Analogies

$$\left[\begin{array}{l} \textit{stress} \\ \textsc{p}\left\{\begin{array}{l} \left[\textsc{fts}\ \textit{1sg-pres}\ \textsc{phon}\ /\acute{\sigma}\sigma\#/\right], \\ \left[\textsc{fts}\ \textit{2sg-pres}\ \textsc{phon}\ /\acute{\sigma}\sigma\#/\right], \\ \left[\textsc{fts}\ \textit{3sg-pres}\ \textsc{phon}\ /\acute{\sigma}\sigma\#/\right], \\ \left[\textsc{fts}\ \textit{1pl-pres}\ \textsc{phon}\ /\acute{\sigma}\sigma\#/\right], \\ \left[\textsc{fts}\ \textit{2pl-pres}\ \textsc{phon}\ /\acute{\sigma}\#/\right], \\ \left[\textsc{fts}\ \textit{3pl-pres}\ \textsc{phon}\ /\acute{\sigma}\sigma\#/\right], \\ \ldots \end{array}\right\} \end{array}\right]$$

But since this pattern applies to all (*) verbs, we can think of it as a supper type of the individual patters for markers.

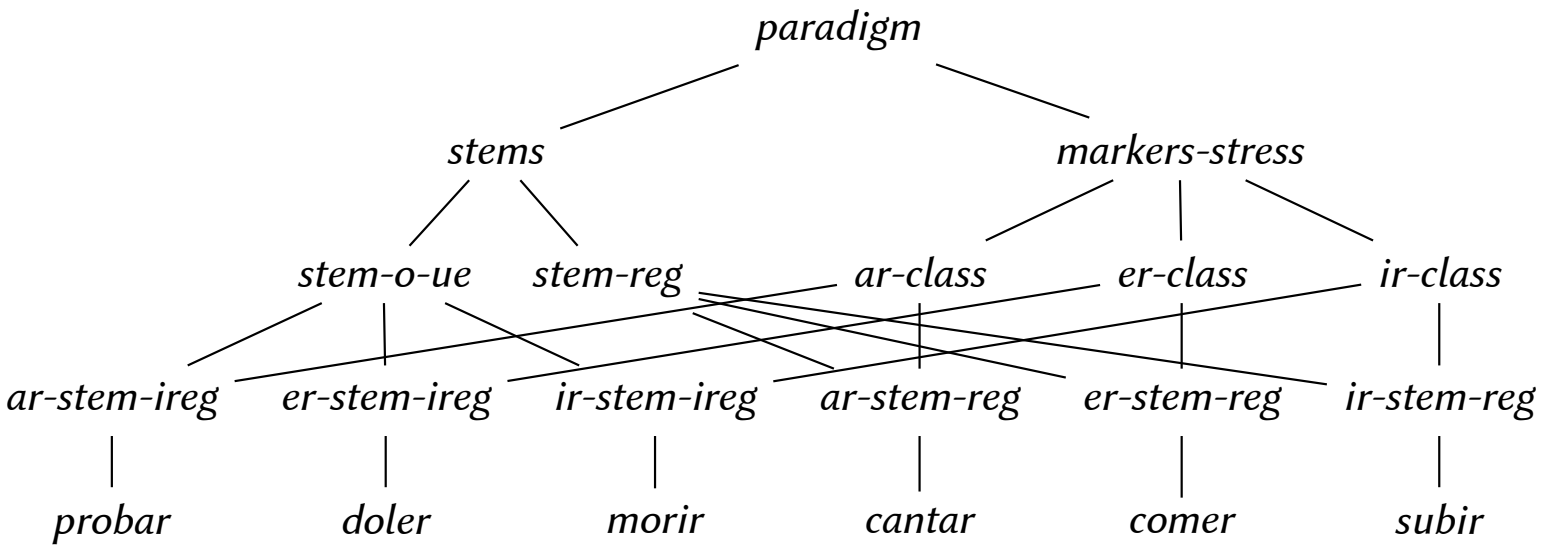
Analogies

$$\left[\begin{array}{l}\textit{stem-o-ue} \\ \text{P}\left\{\begin{array}{ll}\left[\text{FTS } \textit{1sg-p} \text{ PHON} & /\#X^1\sigma^2_{[ue]}\sigma_{[O^1]}\#/\right] \\ \left[\text{FTS } \textit{2sg-p} \text{ PHON} & /\#X^1\sigma^2_{[ue]}\sigma_{[O^1]}\#/\right] \\ \left[\text{FTS } \textit{3sg-p} \text{ PHON} & /\#X^1\sigma^2_{[ue]}\sigma_{[O^1]}\#/\right] \\ \left[\text{FTS } \textit{1pl-p} \text{ PHON} & /\#X^1\sigma^2_{[o]}\sigma_{[O^1]}\sigma\#/\right] \\ \left[\text{FTS } \textit{2pl-p} \text{ PHON} & /\#X^1\sigma^2_{[o]}\sigma_{[O^1]}\#/\right] \\ \left[\text{FTS } \textit{3pl-p} \text{ PHON} & /\#X^1\sigma^2_{[ue]}\sigma_{[O^1]}\#/\right] \\ \ldots & \end{array}\right\}\end{array}\right]$$

$$\left[\begin{array}{l}\textit{stem-regular} \\ \text{P}\left\{\begin{array}{ll}\left[\text{FTS } \textit{1sg-p} \text{ PHON} & /\#X^1\sigma_{[O^1]}\#/\right] \\ \left[\text{FTS } \textit{2sg-p} \text{ PHON} & /\#X^1\sigma_{[O^1]}\#/\right] \\ \left[\text{FTS } \textit{3sg-p} \text{ PHON} & /\#X^1\sigma_{[O^1]}\#/\right] \\ \left[\text{FTS } \textit{1pl-p} \text{ PHON} & /\#X^1\sigma_{[O^1]}\sigma\#/\right] \\ \left[\text{FTS } \textit{2pl-p} \text{ PHON} & /\#X^1\sigma_{[O^1]}\#/\right] \\ \left[\text{FTS } \textit{3pl-p} \text{ PHON} & /\#X^1\sigma_{[O^1]}\#/\right] \\ \ldots & \end{array}\right\}\end{array}\right]$$

Analogies

Finally, we need a hierarchy to put everything together.



But we can do more...

The issue is that these examples are easily captured with any theory. There are two aspects of PA systems, however, which are rarely if ever captures by constructive approaches. These are predictability/information and relations between cells.

The first has to do with the fact that knowing one cell of a paradigm might give us partial or complete information about the other cells in the paradigm. Similarly, knowing part of a cell in a paradigm might give us information about which cell it is.

The second issue has to do with the fact that some generalizations are better expressed as relations between two fully inflected cells, and not as a derivation starting from a stem.

Partial predicatability

$$\left[\begin{array}{l} \textit{o-marker} \\ \textsc{paradigm} \left\{ \left[\begin{array}{ll} \textsc{fts} & \textit{1.present} \\ \textsc{phon} & /\sigma_{[\textsc{nuc}\ /o/]}\#/ \end{array}\right] \right\} \end{array}\right]$$

$$\left[\begin{array}{l} \textit{1-marker} \\ \textsc{paradigm} \left\{ \left[\begin{array}{ll} \textsc{fts} & \textit{1sg.present} \\ \textsc{phon} & /o\#/ \end{array}\right], \left[\begin{array}{ll} \textsc{fts} & \textit{1pl.present} \\ \textsc{phon} & /mos\#/ \end{array}\right] \right\} \end{array}\right]$$

	love	eat	feel
1SG.	am**o**	com**o**	sient**o**
1PL.	amam**os**	comem**os**	sentim**os**

	love	eat	feel
1PL.	am**o**	com**o**	sient**o**
1PL.	ama**mos**	come**mos**	senti**mos**

Cell relations

	cantar	eat	go up
INF	cantár	comér	subír
1SG.COND	cantaría	comería	subiría
1SG.FUT	cantaré	comeré	subiré
1SG.IMP.SUBJ	cantára	comiéra	subiéra
3SG.FUT	cantarás	comerás	subirás
3SG.IMP.SUBJ	cantáras	comiéras	subiéras
3SG.PRES	cánta	cóme	súbe
2SG.IMP/*tú*	cánta	cóme	súbe
2SG.IMP/*vos*	cantá	comé	subí
2SG.IMP/*usted*	cánte	cóma	súba
1SG.PRET	canté	comí	subí
1SG.PRES	cánto	cómo	súbo
3SG.PRET	cantó	comió	subió

Cell relations

	cantar	eat	go up
1SG.IMP.SUBJ	**cantára**	**comiéra**	**subiéra**
2SG.IMP.SUBJ	**cantára**s	**comiéra**s	**subiéra**s
3SG.IMP.SUBJ	**cantára**	**comiéra**	**subiéra**
1PL.IMP.SUBJ	**cantára**mos	**comiéra**mos	**subiéra**mos
2PL.IMP.SUBJ	**cantára**is	**comiéra**is	**subiéra**is
3PL.IMP.SUBJ	**cantára**n	**comiéra**n	**subiéra**n
1SG.COND	**cantaría**	**comería**	**subiría**
2SG.COND	**cantaría**s	**comería**s	**subiría**s
3SG.COND	**cantaría**	**comería**	**subiría**
1PL.COND	**cantaría**mos	**comería**mos	**subiría**mos
2PL.COND	**cantaría**is	**comería**is	**subiría**is
3PL.COND	**cantaría**n	**comería**n	**subiría**n

Morphemes

In the imperfect:

	cantar	comer	subir
1.SG.IMP	cant-ába	com-ía	sub-ía
2.SG.IMP	cant-ába-s	com-ía-s	sub-ía-s
3.SG.IMP	cant-ába	com-ía	sub-ía
1.PL.IMP	cant-ába-mos	com-ía-mos	sub-ía-mos
2.PL.IMP	cant-ába-is	com-ía-is	sub-ía-is
3.PL.IMP	cant-ába-n	com-ía-n	sub-ía-n

Morphemes

The same patter repeats for all but the indefinido:

	cantar	comer	subir
1SG.IND	cant-é	com-í	sub-ío
2SG.IND	cant-asté	com-isté	sub-isté
3SG.IND	cant-ó	com-ió	sub-ió
1PL.IND	cant-a-**mos**	com-i-**mos**	sub-i-**mos**
2PL.IND	cant-asté-**is**	com-isté-**is**	sub-isté-**is**
3PL.IND	cant-aró-**n**	com-ier**ó-n**	sub-ieró-**n**

Which nonetheless retains the same markers for the plural (*mos, is, n*).

Morphemes

$$\left[\textit{tam-markers}\ \left\{\left[\textsc{fts}\ \textit{1pl};\ \textsc{phon}\ /\#X^1\,\textit{mos}\#/\right],\left[\textsc{fts}\ \textit{2pl};\ \textsc{phon}\ /\#X^1\,\textit{is}\#/\right],\left[\textsc{fts}\ \textit{3pl};\ \textsc{phon}\ /\#X^1\,\textit{n}\#/\right]\right\}\right]$$

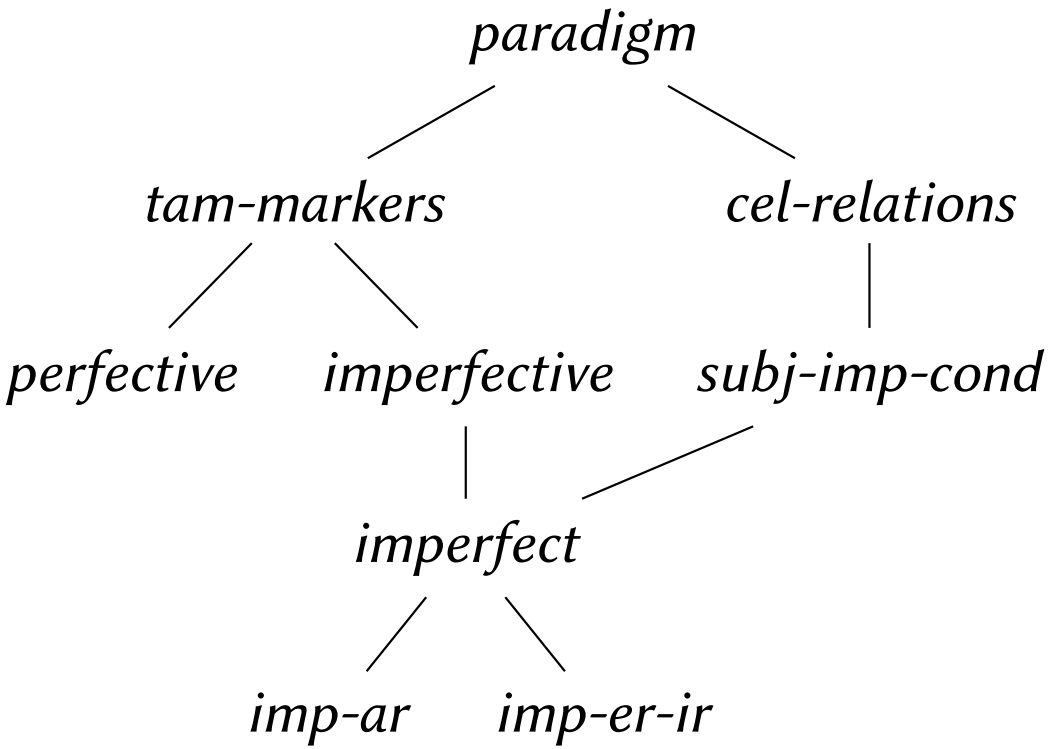
$$\left[\textit{subj-imp-cond}\ \left[\textsc{fts}\ \textit{1sg.(subj}\vee\textit{ind.imp}\vee\textit{cond)};\ \textsc{phon}\ /\#X^1\,\#/\right],\left[\textsc{fts}\ \textit{2sg.(subj}\vee\textit{ind.imp}\vee\textit{cond)};\ \textsc{phon}\ /\#X^1\,\textit{s}\#/\right],\left[\textsc{fts}\ \textit{3sg.(subj}\vee\textit{ind.imp}\vee\textit{cond)};\ \textsc{phon}\ /\#X^1\,\#/\right],\left[\textsc{fts}\ \textit{pl.(subj}\vee\textit{ind.imp}\vee\textit{cond)};\ \textsc{phon}\ /\#X^1/\right]\right]$$

$$\left[\textit{imperfective}\ \left\{\left[\textsc{fts}\ \textit{2sg.imperf};\ \textsc{phon}\ /\textit{s}\#/\right]\right\}\right]$$

$$\left[\textit{imp-ar}\ \left\{\left[\textsc{fts}\ \textit{3/1sg.imp};\ \textsc{phon}\ /\textit{ba}\#/\right]\right\}\right]$$

$$\left[\textit{imp-er-ir}\ \left\{\left[\textsc{fts}\ \textit{3/1sg.imp};\ \textsc{phon}\ /\textit{ía}\#/\right]\right\}\right]$$

Morphemes


paradigm
tam-markers
cel-relations
perfective
imperfective
subj-imp-cond
imperfect
imp-ar
imp-er-ir

Morphemes and cell relations

We can express both ‘morphemes’ and cell relations with the same technique, at the same time, and at the same level of abstraction.

A few issues

Some difficult issues I can see so far:

- It is hard to model partial overabundance
- Morphemes with variable positions are tricky

Complete overabundance is easy

The imperfect subjunctive in Spanish has two possible realizations: *-se* and *-ra*. This applies to **all** verbs:

(8)
a. comiera - comiese ‘eat’
b. supiera - supiese ‘know’
c. amara - amase ‘love’
d. ...

We can model this by simply having several cells with identical morphosyntactic features but different phonological constraints:
{[FTS 1sg.imp.subj PHON /#Xra#/],
[FTS 1sg.imp.subj PHON /#Xse#/] }

Partial overabundance is harder

The issue arises with systems in which some items, but not all, are overabundant:

The Spanish plural system is relatively trivial:

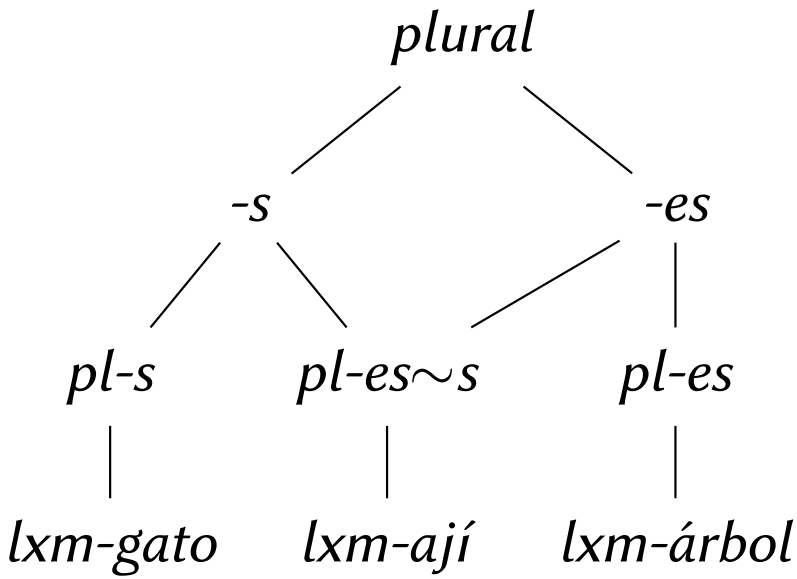
(9)
a. gat**o** - gato**s** (‘cat’)
b. perr**o** - perro**s** (‘dog’)
c. árbo**l** - árbol**es** (‘tree’)
d. baú**l** - baul**es** (‘chest’)

With some exceptions:

(10)
a. aj**í** - ají**s** ∼ ají**es** (‘chili pepper’)
b. colibr**í** - colibrí**s** ∼ colibrí**es** (‘hummingbird’)

Overabundance as multiple inheritance

An intuitive way to think about overabundance is as multiple inheritance:



But with normal unification we cannot grow the paradigm.

What to do?

- A way around this is unification as in CxG (with set union), but...?
- maybe an exception just for paradigms...?
- having three independent plural classes seems **wrong** because it misses the fact one class is the combination of the other two.

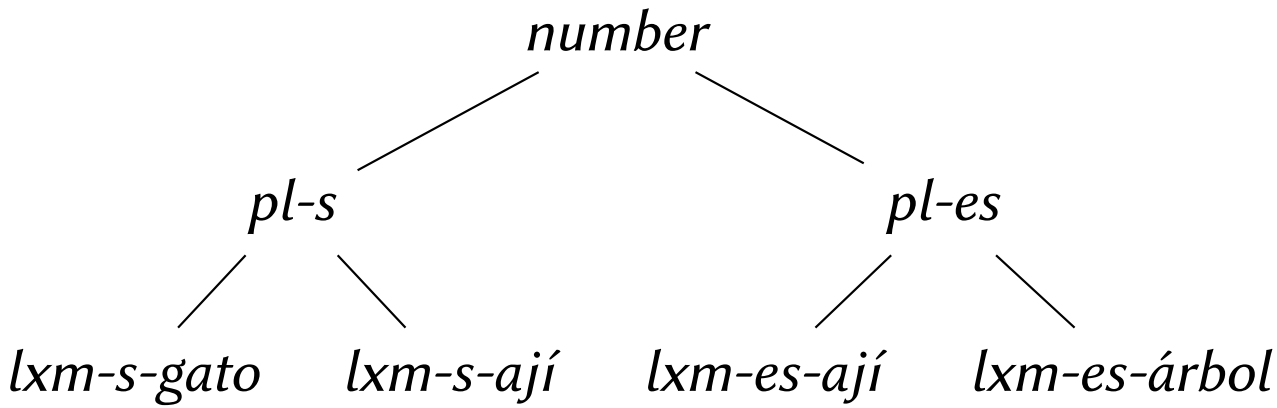
Naming cells?

Something like:

$$\begin{bmatrix} \textsc{paradimg} & \begin{bmatrix} \textsc{singular} & \begin{bmatrix} \textsc{fts} & sg & \textsc{phon} & \ldots \end{bmatrix} \\ \textsc{plural-1} & \begin{bmatrix} \textsc{fts} & pl & \textsc{phon} & \ldots \end{bmatrix} \\ \textsc{plural-2} & \begin{bmatrix} \textsc{fts} & pl & \textsc{phon} & \ldots \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Also seems quite wrong.

But maybe:



Variable morphotactics?

To stay on language, a lame example of variable morphotactics:

(11)

a.	tóma drink.IMP.2SG 'drink (it)!'	a. tómas drink.IMP.2SG '(you) drink'
b.	tóma-**te** (un jugo) drink.IMP.2SG-1.DAT (a juice) drink a juice (for your self)!	b. **te**-tómas (un jugo) 1.DAT-drink.IND.2SG (a juice) '(you) drink '
c.	tóma-**lo** (*un jugo) drink.IMP.2SG-2.ACC 'drink-it!'	c. **lo**-tómas (*un jugo) 2.ACC-drink.IND.2SG '(you) drink it'
d.	tóma-**te**-**lo** drink.IMP.2SG-1.DAT-2.ACC 'drink it up!'	d. **te**-**lo**-tómas 1.DAT-2.ACC-drink.IND.2SG '(you) drink it up'

(Baby steps towards Neo Aramaic or Swahili)

Floting morphemes will not work

What could not work, is to try to define “floating” morphemes, which then attach at the right place:

$$\begin{bmatrix} \textit{lo-clitic} \\ \text{PRDM} \left\{ \begin{bmatrix} \textit{cell} \\ \text{FTS} \begin{bmatrix} \text{OBJ} \begin{bmatrix} \text{PERSON} & 2 \\ \text{NUMBER} & sg \end{bmatrix} \\ \text{ARG-ST} \left\langle \text{NP} \right\rangle \end{bmatrix} \\ \text{PHON} \quad /\#(\sigma)?lo|lo\#/ \end{bmatrix} \right\} \end{bmatrix}$$

Because it stops being a relational system (and it also probably does not technically work)

What could work

Build complex cells incrementally:

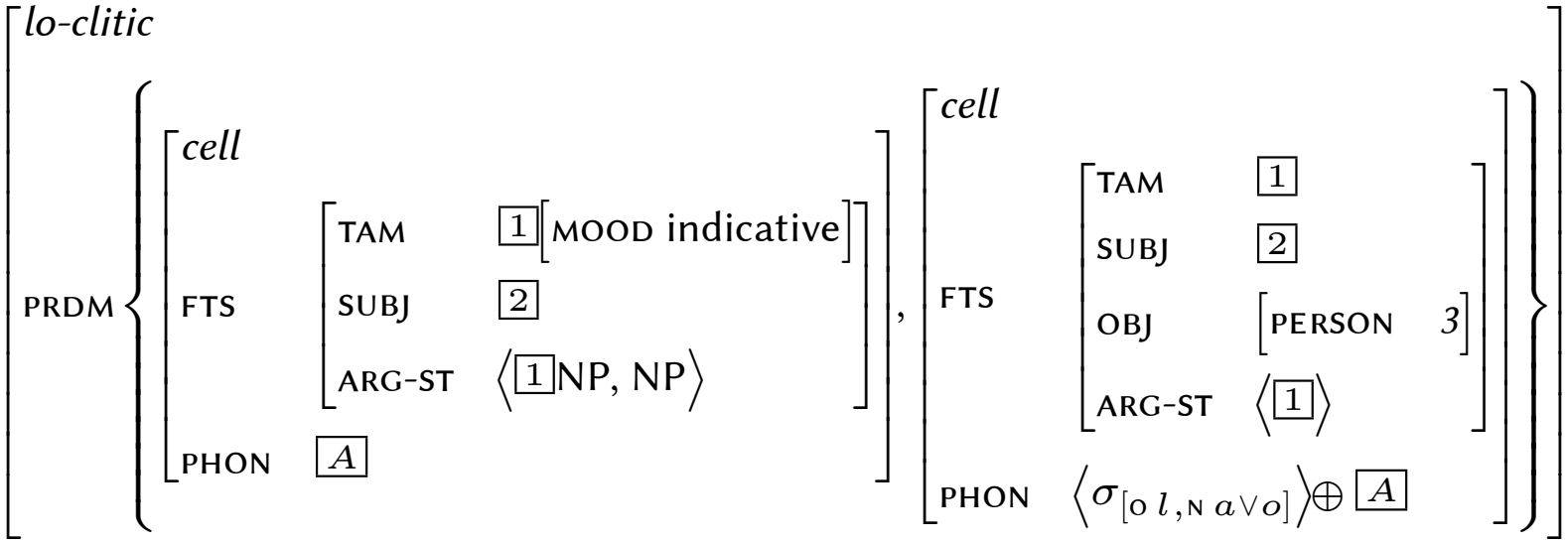
$$\left[\textit{lo-clitic}\ \text{PRDM}\left\{\left[\textit{cell}\ \text{FTS}\left[\text{TAM}\ \boxed{2}\left[\text{MOOD}\ \textit{imperative}\right]\right]\ \text{SUBJ}\ \boxed{3}\ \text{ARG-ST}\ \boxed{C}\left\langle\boxed{1}\text{NP, NP}\right\rangle\right]\ \text{PHON}\ \boxed{A}\right],\left[\textit{cell}\ \text{FTS}\left[\text{TAM}\ \boxed{4}\left[\text{MOOD}\ \textit{indicative}\right]\right]\ \text{SUBJ}\ \boxed{3}\ \text{ARG-ST}\ \boxed{C}\right]\ \text{PHON}\ \boxed{B}\right],\right.\right.$$
$$\left.\left.\left[\textit{cell}\ \text{FTS}\left[\text{TAM}\ \boxed{2}\ \text{SUBJ}\ \boxed{3}\ \text{OBJ}\ \boxed{5}\left[\text{PERSON}\ 3\ \text{NUMBER}\ sg\ \text{GENDER}\ masc\right]\ \text{ARG-ST}\ \left\langle\boxed{1}\right\rangle\right]\ \text{PHON}\ \boxed{A}\oplus\left\langle\sigma_{[lo]}\right\rangle\right],\left[\textit{cell}\ \text{FTS}\left[\text{TAM}\ \boxed{4}\ \text{SUBJ}\ \boxed{3}\ \text{OBJ}\ \boxed{5}\ \text{ARG-ST}\ \left\langle\boxed{1}\right\rangle\right]\ \text{PHON}\ \left\langle\sigma_{[lo]}\right\rangle\oplus\boxed{B}\right]\right\}\right]$$

$$\left[\textit{la-lo-clitics}\ \text{PRDM}\left\{\left[\textit{cell}\ \text{FTS}\left[\begin{array}{ll}\text{TAM} & \boxed{1}\\ \text{SUBJ} & \boxed{2}\\ \text{OBJ} & \left[\begin{array}{ll}\text{PERSON} & \boxed{3}\,3\\ \text{NUMBER} & \boxed{4}\\ \text{GENDER} & \textit{masc}\end{array}\right]\\ \text{ARG-ST} & \langle\boxed{1}\rangle\end{array}\right]\ \text{PHON}\ \ldots\oplus\langle\sigma_{[\text{O }l,\text{N }o]}\rangle\oplus\ldots\right],\left[\textit{cell}\ \text{FTS}\left[\begin{array}{ll}\text{TAM} & \boxed{1}\\ \text{SUBJ} & \boxed{2}\\ \text{OBJ} & \left[\begin{array}{ll}\text{PERSON} & \boxed{3}\\ \text{NUMBER} & \boxed{4}\\ \text{GENDER} & \textit{fem}\end{array}\right]\\ \text{ARG-ST} & \langle\boxed{1}\rangle\end{array}\right]\ \text{PHON}\ \ldots\oplus\langle\sigma_{[\text{O }l,\text{N }a]}\rangle\oplus\ldots\right]\right\}\right]$$

$$\left[\textit{lV-lVs-clitics}\right.$$
$$\left.\textsc{prdm}\left\{\left[\textit{cell}\ \left[\begin{array}{ll}\textsc{tam} & \boxed{1}\\ \textsc{subj} & \boxed{2}\\ \textsc{obj} & \left[\begin{array}{ll}\textsc{person} & \boxed{3}\,3\\ \textsc{number} & sg\\ \textsc{gender} & \boxed{4}\end{array}\right]\\ \textsc{arg-st} & \langle\boxed{1}\rangle\end{array}\right]\ \textsc{phon}\ \ldots\oplus\langle\sigma_{[\textsc{o}\ l,\textsc{c}\ nil]}\rangle\oplus\ldots\right],\left[\textit{cell}\ \left[\begin{array}{ll}\textsc{tam} & \boxed{1}\\ \textsc{subj} & \boxed{2}\\ \textsc{obj} & \left[\begin{array}{ll}\textsc{person} & \boxed{3}\\ \textsc{number} & \textsc{pl}\\ \textsc{gender} & \boxed{4}\end{array}\right]\\ \textsc{arg-st} & \langle\boxed{1}\rangle\end{array}\right]\ \textsc{phon}\ \ldots\oplus\langle\sigma_{[\textsc{o}\ l,\textsc{c}\ s]}\rangle\oplus\ldots\right],\right\}\right]$$

What could work:

Underspecified relations between cells:



Concluding remarks

The take home message:

- In order to talk about analogy, we need a proper formalization of analogy.
- I have presented a possible path, others are possible.
- The X-notation is **not** good enough.

With the model I propose, we can:

- express partial analogies,
- express abstract analogies,
- use underspecification to express general patterns (e.g. ‘morphemes’)

Questions

One tricky issue I am aware of:

- Possition classes (e.g. Swahili)

Some important open questions:

- **Can all and any generalization be expressed with this system?** (i.e. what can analogy **not** do?)
- How should we represent phonological strings? what is the right level of abstraction? how much phonetics do we want to include?
- How much phonology can we get rid of? (e.g. we can do harmony without phonological process, using underspecification)
- How should we think about redundant constraints?
- What about derivation? (can all derivational cells be listed for a lexeme? would this be required for the model to work?)

That's it...

Derivation

	INF	1SG	SG		SG
‘to populate’	poblar	pueblo	pueblo	‘town’	población
‘to tell’	contar	cuento	cuento	‘tale’	
‘to dream’	soñar	sueño	sueño	‘dream’	
‘to sing’	cantar	canto	canto	‘song’	canción
‘to distinguish’	distinguir	distingo		‘distinction’	distinción

ATR harmony

Kasem number system:

singular	plural	gloss
yʊkwala	yʊkwalɪ	headscarf
yukwələ	yukwəli	small skull
valʊ	vala	farmer
vəlu	vələ	traveller
sugu	sum ∼ suni	guinea-fowl
sʊgʊ	sʊm ∼ sʊnɪ	knife
peeli	peelə	shovel
pɛɛlɪ	pɛɛla	sardine

$$\left[\begin{array}{l} \textit{lexeme–harmony-atr} \\ \textsc{par}\left\{\left[\textsc{phon}\ /\left[\begin{array}{l}\textit{vowel}\\ \textsc{atr}\ \boxed{1}\end{array}\right]\ldots\left[\begin{array}{l}\textit{vowel}\\ \textsc{atr}\ \boxed{1}\end{array}\right]\#/\right]\right\} \end{array}\right]$$

$$\left[\begin{array}{l} \textit{lexeme–class-ıa} \\ \textsc{par}\left\{\begin{array}{l}\left[\textsc{fts}\ \textit{sg}\quad \textsc{phon}\ /X\left[\begin{array}{l}\textit{vowel}\\ \textsc{height}\ \textit{high}\\ \textsc{backness}\ \textit{front}\end{array}\right]\#/\right],\\ \left[\textsc{fts}\ \textit{pl}\quad \textsc{phon}\ /X\left[\begin{array}{l}\textit{vowel}\\ \textsc{height}\ \textit{low}\\ \textsc{backness}\ \textit{center}\end{array}\right]\#/\right]\end{array}\right\} \end{array}\right]$$