From Latin to Romance: Computational Modeling of Syncretism

Tyler Lau*, Maria Polinsky†, and Jake Seaton†

*Department of Linguistics, University of California at Berkeley, USA
†Deparment of Linguistics, Harvard University, USA

Overview

Questions:

- What factors in Late Latin led to the heavy reshaping of the nominal system?
- What minimal information does a connectionist model need to predict syncretism in the correct direction?

Background:

- Analogy driven by factors such as *frequency*, *markedness*, *morpheme length*, etc. (Kurylowicz 1947, Manczak 1958, Bybee 1985, Hock 1991, Albright 2008)
- From Latin to Romance
  - Declension: 5 > 3∼2 (I, II, (III)): frequency, sound change
  - Gender: 3 > 2 (M, F): sound change, contact
  - Case: 6 > 2∼1 (ACC, (NOM/GEN)): sound change, periphrastic constructions (preposition+ACC)
- Fate of the Neuter
  - N.SG ended in same endings as M.SG so many became M
  - N.PL ended in *-a* and as plural inanimates were seen as collectives, reinterpreted as F.SG: ex. Lat. *folia* 'leaves N.PL' > Sp. *hoja* 'leaf F.SG' (Herman 1967)
- Romanian has an ambigeneric system
  - "Neuter" class takes M morphology in singular and F in plural
  - Falls out from same principles as other Romance languages (with N.PL being reinterpreted as F.PL)
  - Many M's migrate to N class via analogy (Lat. *campus* ∼ *campi* 'field M' > Rom. *cîmp* (M.SG) ∼ *cîmpuri* (F.PL) , likely via *tempus* ∼ *tempora* 'time N')

Objective: To use a connectionist simulation of generational learning providing minimal phonological and semantic information and see whether the changes that are actually attested in Romance can be reproduced

Latin Declension System

		I	II	IIIa	IIIb	IV	V
	Root	silv**a-**	ann**o-**	color**-**	ign**i-**	lac**u-**	fid**e-**
	Gloss	'forest'	'year'	'color'	'fire'	'lake'	'faith'
Sg.	Nom.	silv**a**	ann**us**	color	ign**is**	lac**us**	fid**ēs**
	Gen.	silv**ae**	ann**ī**	colōr**is**	ign**is**	lac**ūs**	fid**eī**
	Acc.	silv**am**	ann**um**	colōr**em**	ign**em**	lac**um**	fid**em**
Pl.	Nom.	silv**ae**	ann**ī**	colōr**ēs**	ign**ēs**	lac**ūs**	fid**ēs**
	Gen.	silv**ārum**	ann**ōrum**	colōr**um**	ign**ium**	lac**ūm**	fid**ērum**
	Acc.	silv**ās**	ann**ōs**	colōr**ēs**	ign**īs**/ ign**ēs**	lac**ūs**	fid**ēs**

Figure 1: The Latin Declension Classes

Structure of the Connectionist Model

Results

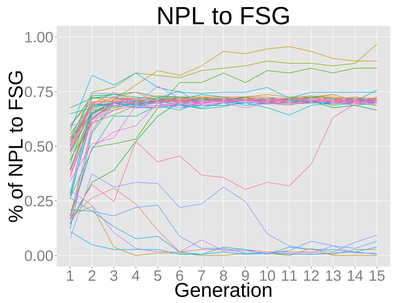

Figure 2: *With the genitive dropped, neuter plurals almost consistently migrate to the feminine singular class due to phonological similarity alone–while collective semantics can be invoked, they are not absolutely necessary to account for the facts.*

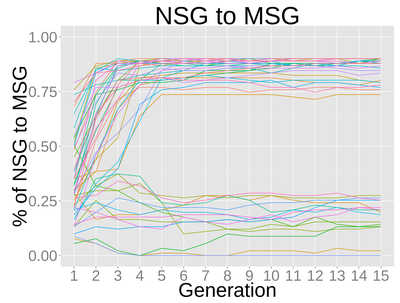

Figure 3: *With the genitive dropped, neuter singulars undergo a bifurcation–they either merge almost completely with masculines or stay very much neuter while drawing masculines to their class (see Figure 6).*

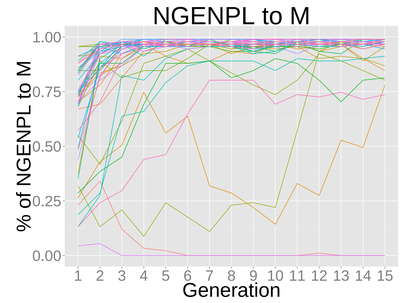

Figure 4: *Without the genitive dropped, the genitive plurals are largely drawn to the masculine, as they do* **not** *end in -a. This prevents the neuter plurals from migrating almost categorically to F.SG.*

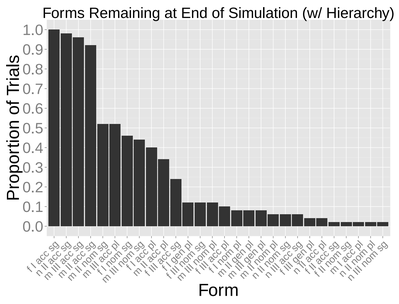

Figure 5: *With case hierarchy in play, the accusative is very robust, the genitive singular falls out completely, and the genitive plural survives in few cases. The most robust forms accord with history.*

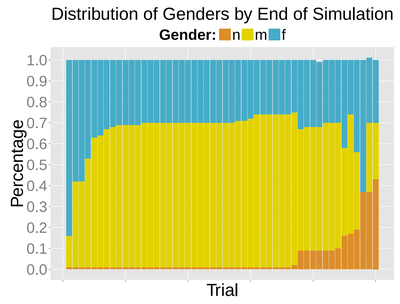

Figure 6: *In most of the trials, the neuter survives in at most a few words (this occurs in Italian, for example (dito M.SG ∼ dita F.PL 'finger'). In the cases where it is more robust, M nouns migrate to the N class, as happened in Romanian.*

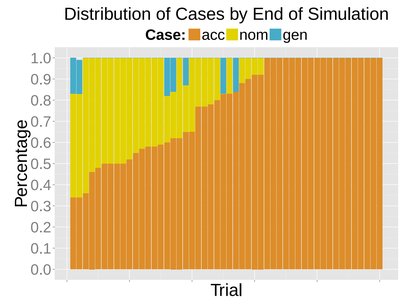

Figure 7: *With case hierarchy in play, the accusative becomes the dominant case in almost every simulation and the only case in almost half. The genitive survives in hardly more than 10% of trials*

Discussion

- With *phonology*, *frequency*, & *animacy semantics*
  - Declensions IV & V fall out in *every* simulation
- With *case hierarchy* added, final forms converge more
  - Genitive singular drops out *completely*
  - Genitive plural hardly survives (only example is oblique 3PL pronoun–Fr. *leur*, It. *loro*, Rom. *lor*)
  - Forms remaining in ≥90% of simulations
    - *-am* > *-a* F.SG ending in all Romance (> *-e* in Fr.)
    - *-um* > *-u* M.SG ending in all of Romance (> *-o* in Sp., It. etc.)
    - *-em* > *-e* SG ending for M/F nouns in all of Romance
  - Forms remaining in 25-90% of simulations
    - *-∅* SG ending for M/F nouns in all of Romance
    - *-ēs* PL ending in western Romance, maybe > *-i* in eastern
    - *-ōs* M.PL ending in western Romance, maybe > *-i* in eastern
    - *-ās* F.PL ending in western Romance, maybe > *-e* in eastern
    - M/F.NOM.SG *-us* & *-as*: in E-Romance., final *-s* falls out; in W-Romance, NOM persists in older Sp. & Fr.
  - Case system converges to accusative in almost half (as in western Romance), eastern Romance shows alternate history where nominative plural may have survived (see D'hulst (2005) on Romance plurals)
  - Neuter rarely survives–when it stays, a sizeable chunk of the masculine class migrates over (as in Romanian)
- With *genitive* dropped, N.SG > M.SG, N.PL > F.SG
  - Otherwise, N.PL GEN > M.PL because of phonology
  - Supports popularity of periphrastic construction view

Acknowledgements

Many thanks to Kevin Ryan, James Kirby, Andrew Garrett, Terry Regier, Mairi McLaughlin, and Yang Xu for comments and guidance throughout the project, to Ezra Van Everbroeck for providing the code for the simulation in Polinsky and Van Everbroeck (2003), and to Edwin Ko for consultation on data visualization.

Contact Information

- Tyler Lau: tylerlau@berkeley.edu
- Maria Polinsky: polinsky@fas.harvard.edu
- Jake Seaton: jseaton@college.harvard.edu