Αναζητώντας την ιστορία των γλωσσών.

Μία διεπιστημονική πρόκληση

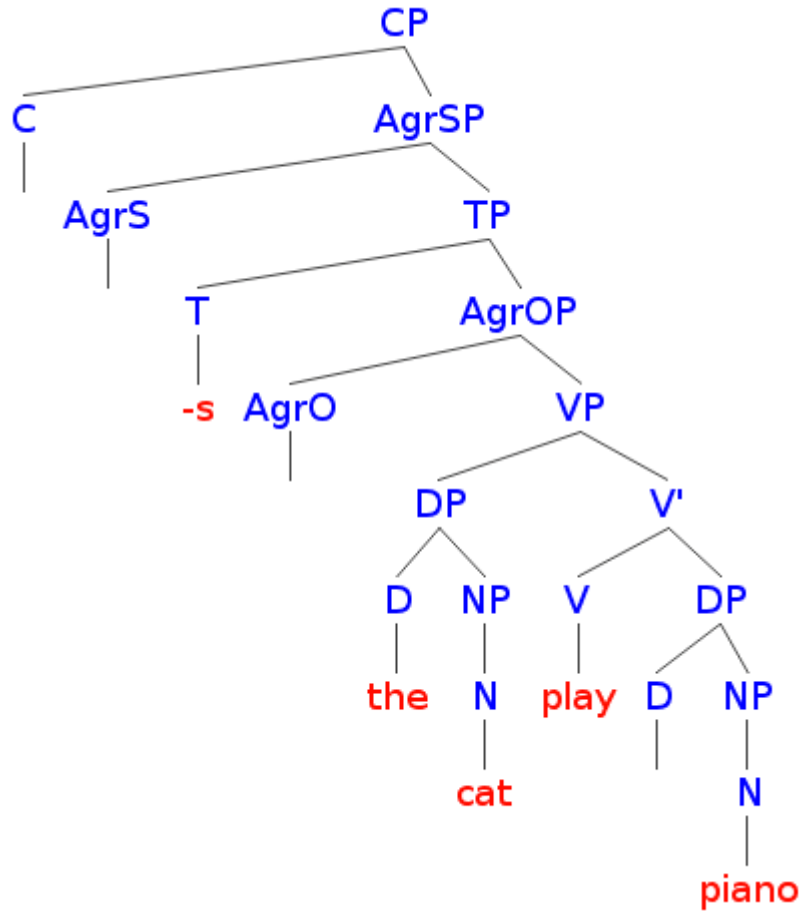
Έλενα Αναγνωστοπούλου

Πανεπιστήμιο Κρήτης, Ηράκλειο, 3 Ιουνίου 2022

Η γλώσσα ως ενδιάθετη γνώση του ανθρώπινου νου – γλωσσική ικανότητα

•Η σύγχρονη γλωσσολογία οφείλει το προσανατολισμό της στον Noam Chomsky, που εστιάζει σε 3 πλευρές της γλώσσας: Α)**Νοητική (**μη συνειδητή γνώση**)**, Β) **Τυπική (**πεπερασμένος αριθμός αφηρημένων μονάδων και κανόνων συνδυασμού τους παράγει απεριόριστο αριθμό προτάσεων**) ,** Γ) **Βιολογική (**η γλωσσική ικανότητα είναι βιολογικά προκαθορισμένη**)**.

Αφηρημένες γλωσσικές παραστάσεις στη μορφοσύνταξη: δέντρα ιεραρχικών σχέσεων



Η μεταβολή της γλώσσας

- Η γλώσσα **μεταβάλλεται διαρκώς**. Το βλέπουμε σε γλώσσες με μακρά γραπτή ιστορία, όπως τα Ελληνικά.

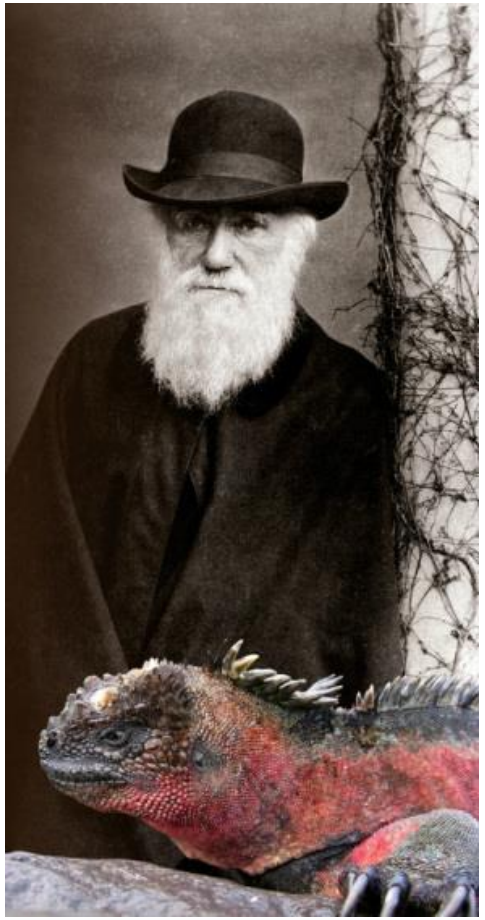
- Είμαστε μάρτυρες στη διαδικασία των αλλαγών. Τα **«γλωσσικά λάθη»** του σήμερα είναι δυνάμει αλλαγές που θα συμβούν αύριο. Χαρακτηριστικά της γραμματικής που παρουσιάζουν **ποικιλομορφία** λόγω έλλειψης σταθερότητας, η οποία θα οδηγήσει σε μεταβολή του συστήματος.

- Από την άλλη πλευρά, **πολλά χαρακτηριστικά της γλώσσας** μεταβάλλονται αργά μέσα στο χρόνο και μένουν **σταθερά για χιλιετίες** (π.χ. 6000 χρόνια, 8000 χρόνια).

Γλωσσολογία και Βιολογία

- Τα τελευταία χρόνια, η γλωσσολογία έχει αρχίσει να χρησιμοποιεί για την **επανασύνθεση της ιστορικής εξέλιξης των γλωσσών** τα εργαλεία που χρησιμοποιεί η **εξελικτική βιολογία** για την **ανακατασκευή της βιολογικής εξέλιξης των ειδών (φυλογένεση)**.

Δύο ευκαιρίες για επαφή και διεπιστημονική συνεργασία στο Πανεπιστήμιο Κρήτης



- Μια συνεργασία που ξεκίνησε από τις **Δαρβινικές Δευτέρες**:

  https://darwinianmondaysgr.weebly.com/

- Και από την πρωτοβουλία του Παν/μίου Κρήτης να ενισχύσει **συνέργειες ανάμεσα στο Ρέθυμνο και το Ηράκλειο** μέσω της υποστήριξης διεπιστημονικών ερευνητικών προγραμμάτων.

Δύο διεπιστημονικά ερευνητικά προγράμματα με τη συνεργασία Γλωσσολογίας- Βιολογίας

- **“Η χρήση των εργαλείων της εξελικτικής βιολογίας στη μελέτη της γλωσσικής ιστορίας και ποικιλίας”.**

  [Έλενα Αναγνωστοπούλου & Μανόλης Λαδουκάκης, Χρηματοδοτούμενο από τον ΕΛΚΕ του Πανεπιστημίου Κρήτης, Research proposals Type C, Interdisciplinary Projects, ΚΑ 10457, 2019-2023]

- **“Μοντελοποιώντας τη γλωσσογένεση. Βιολογικά εργαλεία στη μελέτη της γενετικής και γεωγραφικής σταθερότητας των μορφοσυντακτικών παραμέτρων”.**

  [Χρηματοδοτούμενο από το Ελληνικό Ίδρυμα Έρευνας & Καινοτομίας, ΕΛ.ΙΔ.Ε.Κ., ΚΑ 10506, 2020-2023, http://modelgloss.philology.uoc.gr/]

ModelGloss

MODELLING GLOSSOGENY

ΕΛ.ΙΔ.Ε.Κ.

Ελληνικό Ίδρυμα Έρευνας & Καινοτομίας

Το πρόγραμμα υποστηρίζεται από το Ελληνικό Ίδρυμα Έρευνας και Καινοτομίας (ΕΛ.ΙΔ.Ε.Κ.) στο πλαίσιο της Δράσης «1η Προκήρυξη ερευνητικών έργων ΕΛ.ΙΔ.Ε.Κ. για την ενίσχυση των μελών ΔΕΠ και Ερευνητών/τριών και την προμήθεια ερευνητικού εξοπλισμού μεγάλης αξίας» (Αριθμός Έργου: HFRI-FM17-44)

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
UNIVERSITY OF CRETE

ΑΡΙΣΤΟΤΕΛΕΙΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ

Η Ερευνητική Ομάδα

Μανόλης Λαδουκάκης

Έλενα Αναγνωστοπούλου (ΕΥ)

Δημήτρης Μιχελιουδάκης

Χρήστος Ζιούτης

Παύλος Παυλίδης

Μαργαρίτα Μακρή

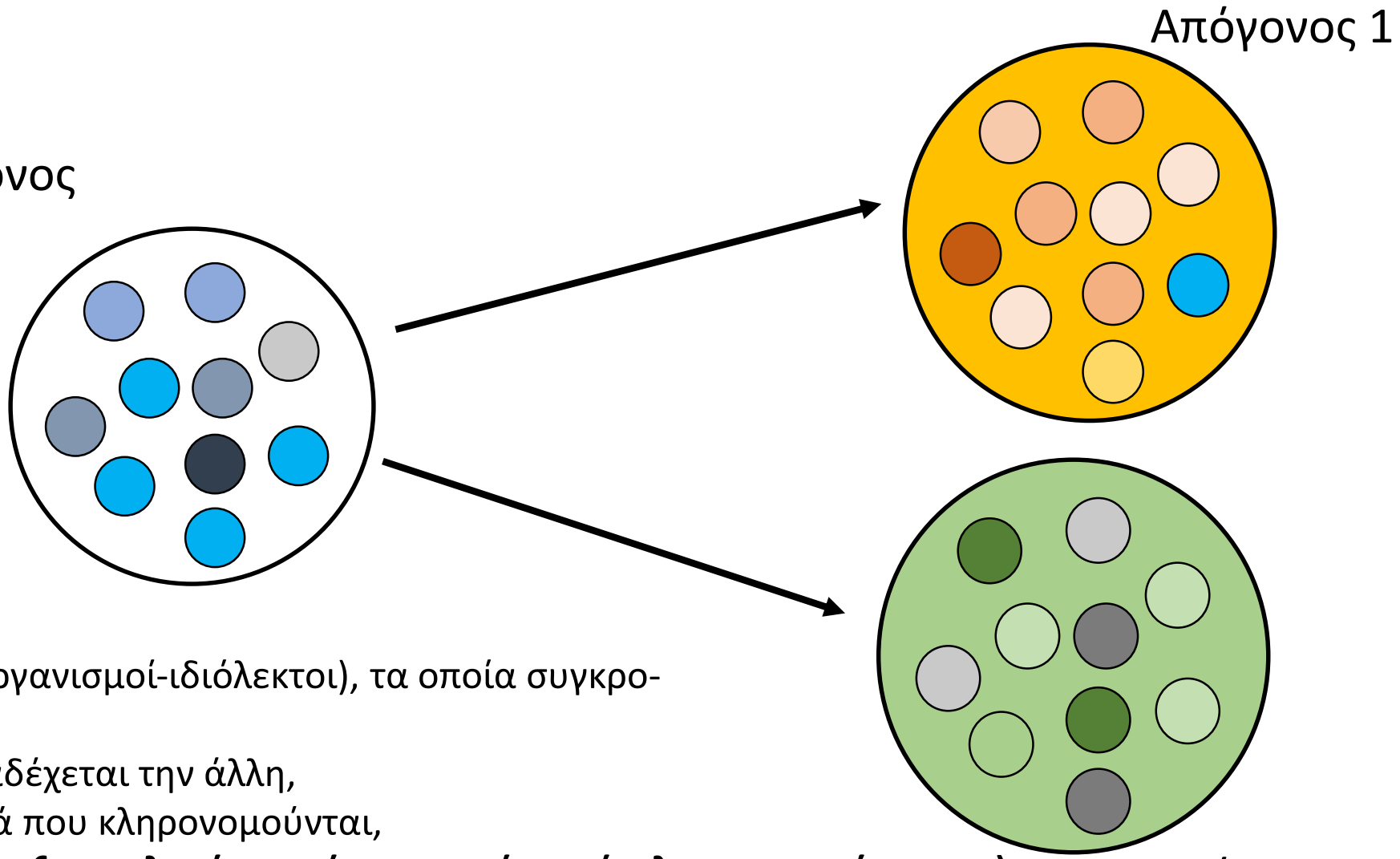
Ruby Sleeman

Ιστορική καταγωγή των γλωσσών

- **Οι γλώσσες**, όπως και τα **βιολογικά είδη**, έχουν κοινούς προγόνους και εξελίχθηκαν στο χρόνο και στο χώρο καθώς οι ανθρώπινοι πληθυσμοί εξαπλώθηκαν στον πλανήτη.

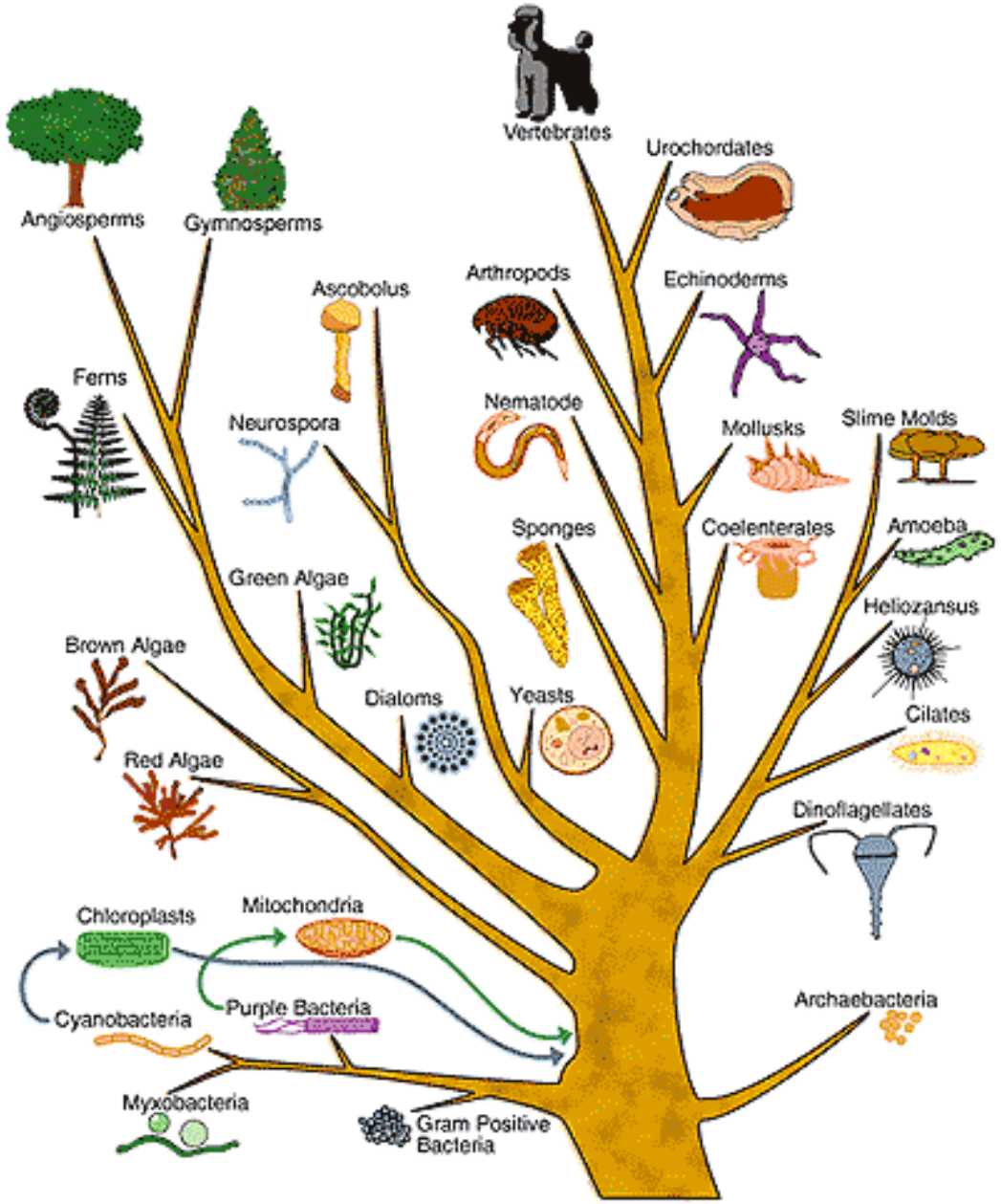
- Όπως και με τα βιολογικά είδη, οι **συστηματικές ομοιότητες** μεταξύ γλωσσών συχνά αντανακλούν την **κοινή καταγωγή τους** και οι **διαφορές** μεταξύ των γλωσσών μεγαλώνουν όσο μεγαλώνει η απόσταση από τους κοινούς προγόνους τους.

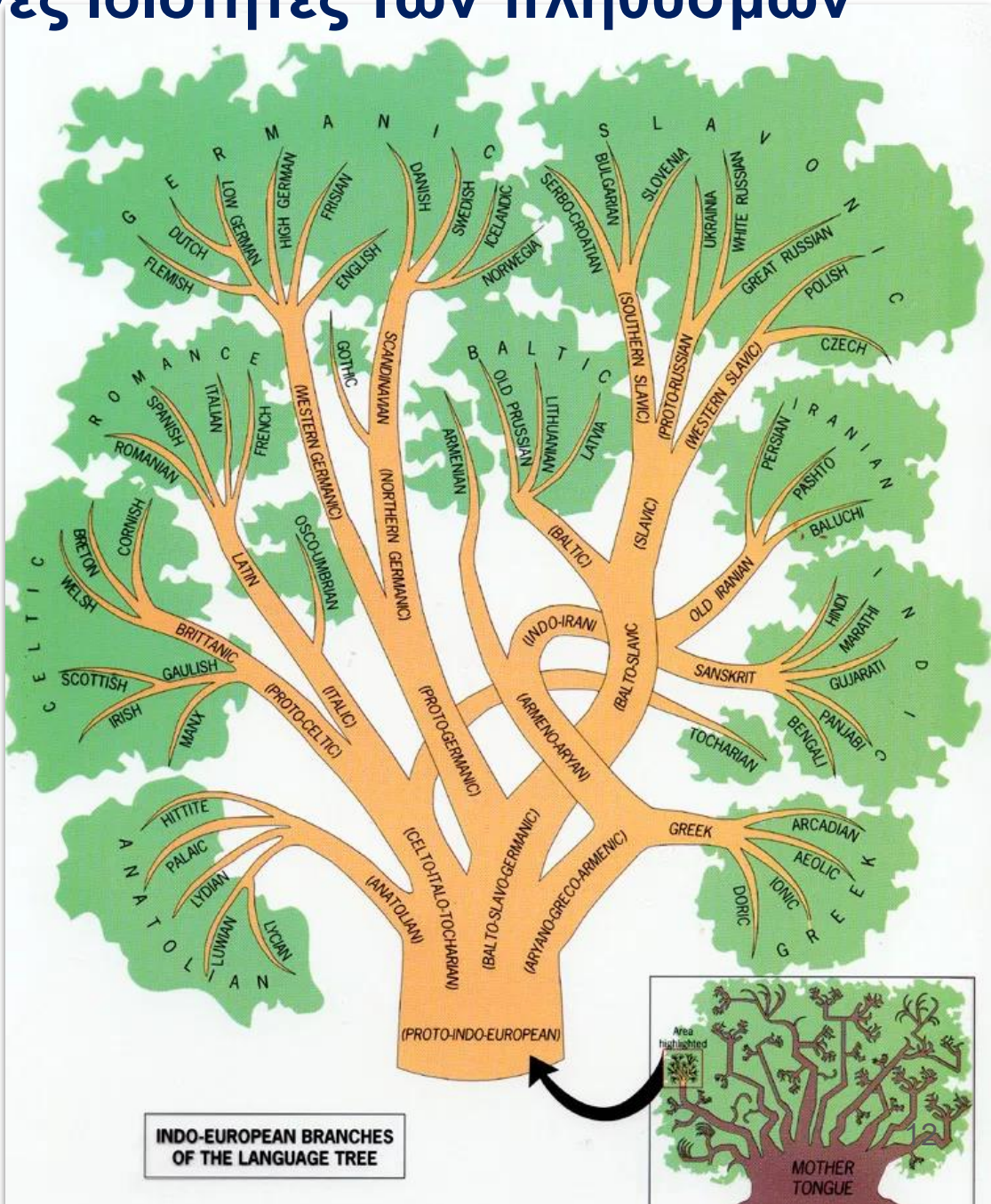
Δύο εξελίξιμα συστήματα: οργανισμικά είδη και γλώσσες



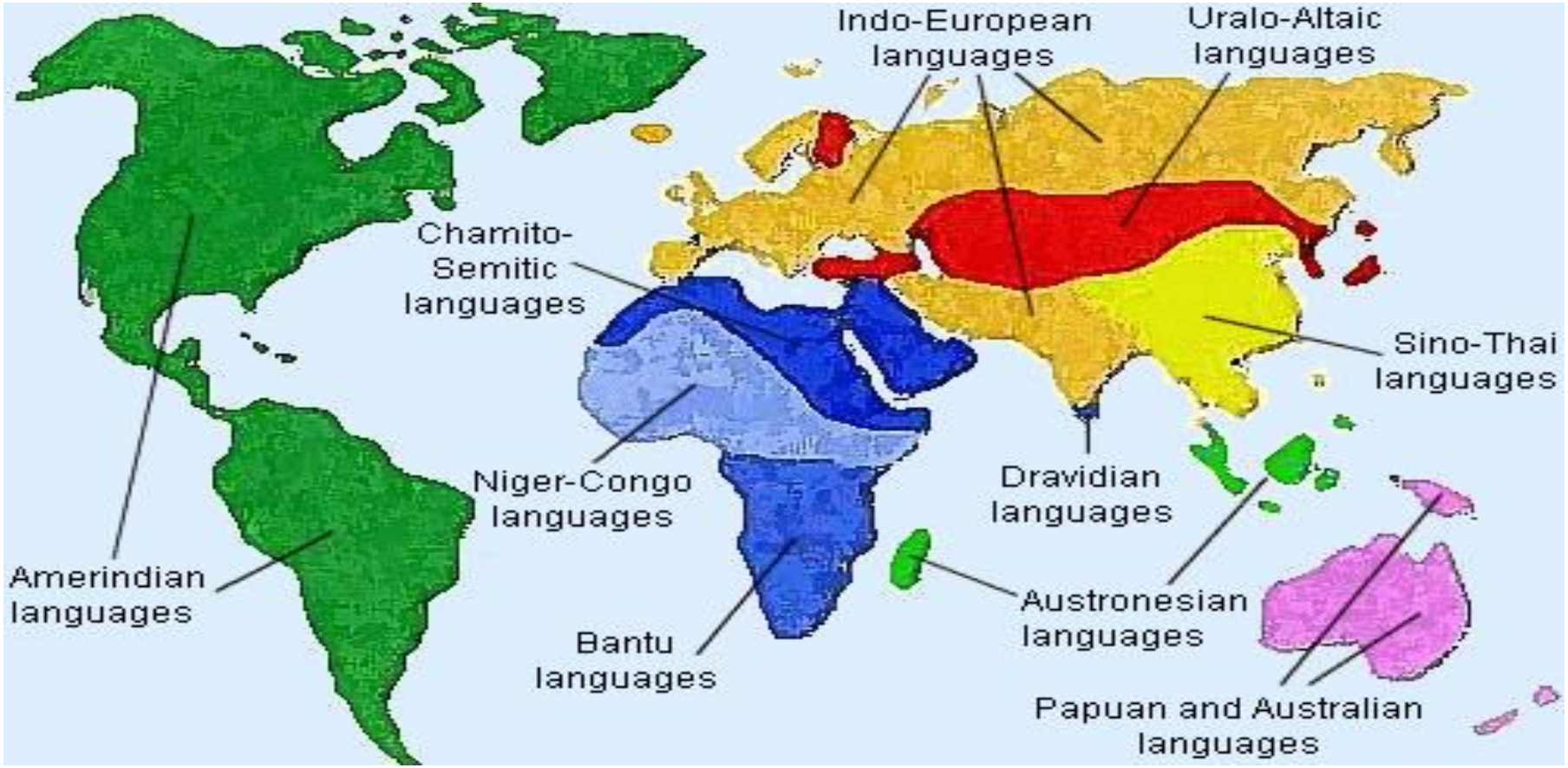
- Αποτελούνται από «άτομα» (οργανισμοί-ιδιόλεκτοι), τα οποία συγκροτούν πληθυσμούς,
- μέσα στο χρόνο η μια γενιά διαδέχεται την άλλη,
- τα άτομα έχουν χαρακτηριστικά που κληρονομούνται,
- παρόμοιες διαδικασίες **(μετάλλαξη, επιλογή, τυχαία γενετική παρέκκλιση, μετανάστευση)** διέπουν την εξέλιξη και των δύο συστημάτων. [Ladoukakis, Michelioudakis & Anagnostopoulou 2022)

Εξέλιξη είναι η αλλαγή στις κληρονομούμενες ιδιότητες των πληθυσμών





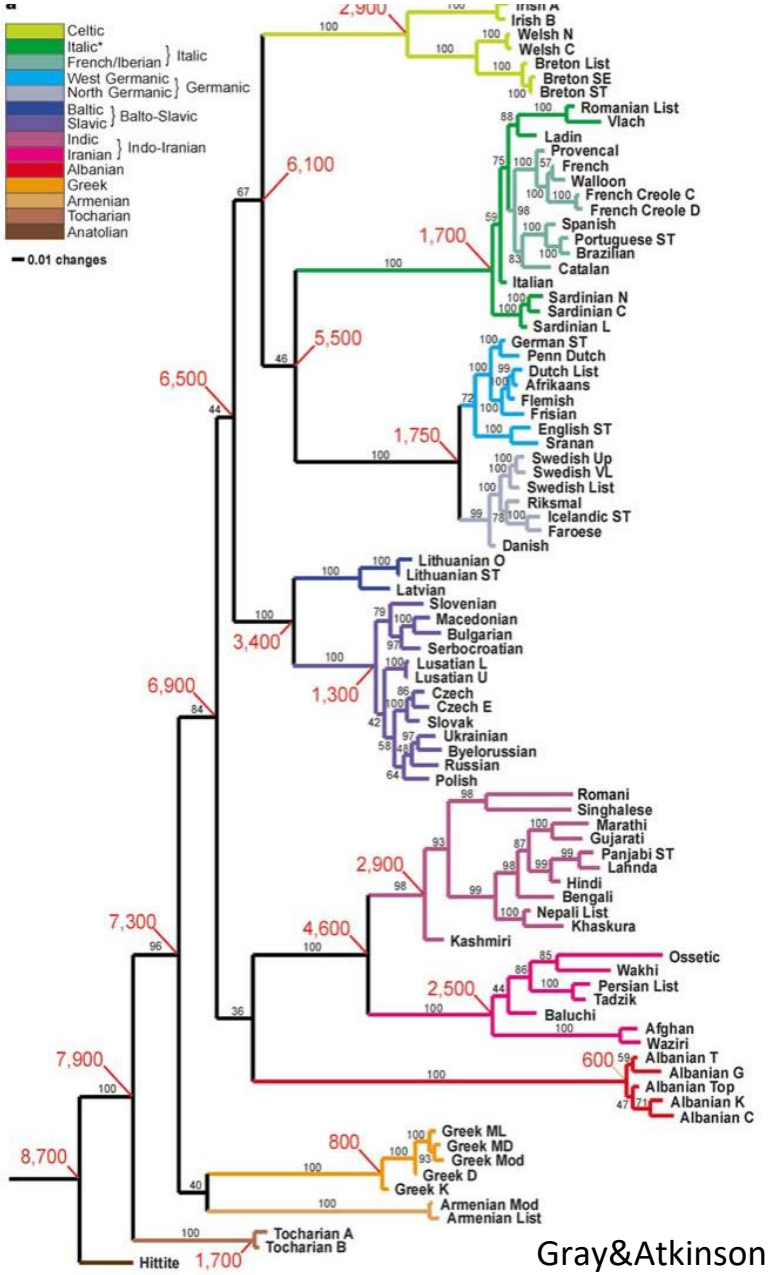
7000 γλώσσες περίπου και οι γλωσσικές τους οικογένειες



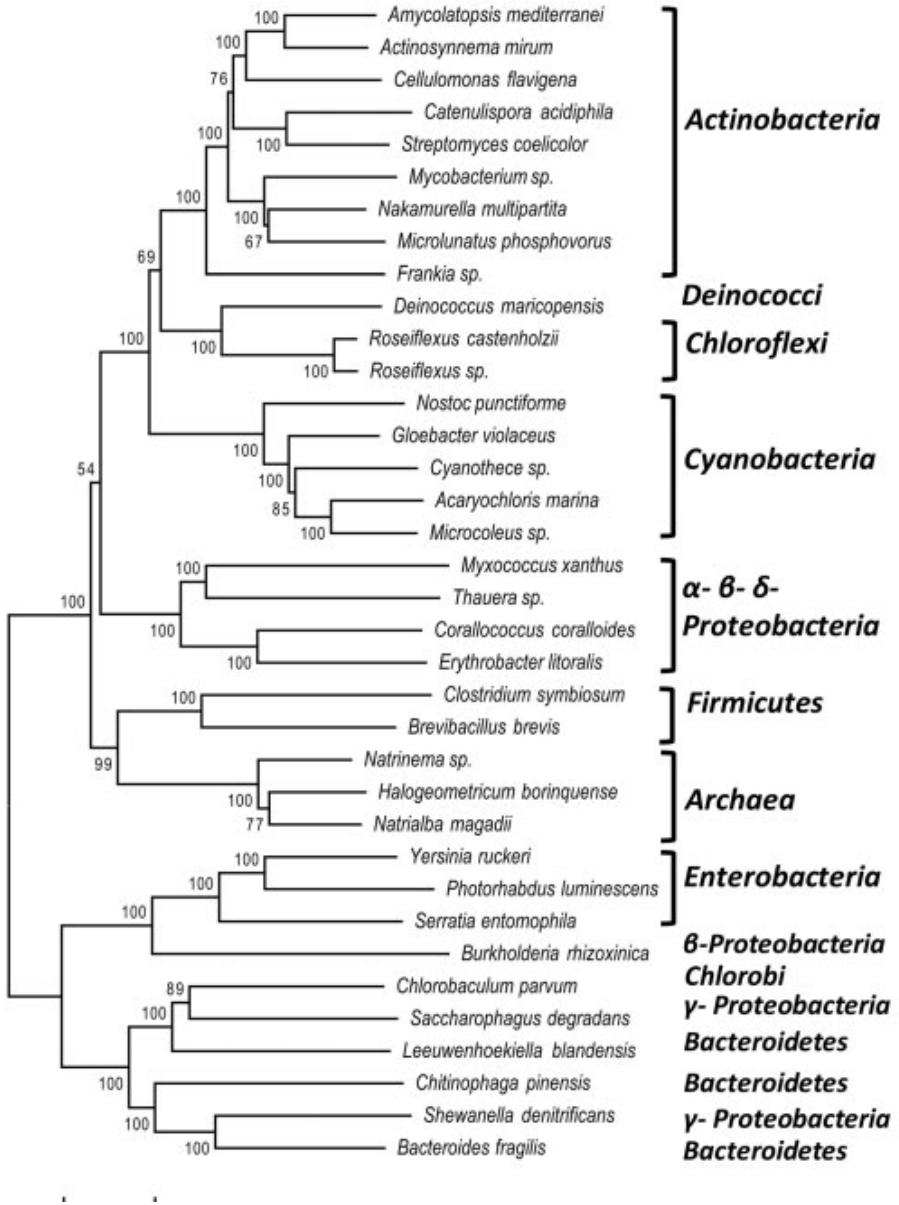
Η γλώσσα ως εξελίξιμο σύστημα: ερωτήματα

- Πώς συγκρίνουμε τις γλώσσες;
- Ποια δεδομένα και ποιες μεθόδους χρησιμοποιούμε;
- Πολύ συνηθισμένη μέθοδος, δανεισμένη από τη Βιολογία είναι τα **δέντρα**.

Εκτίμηση γενετικών σχέσεων με φυλογενετικές μεθόδους



Gray&Atkinson 2003, Nature



Saris, Ladoukakis et al. 2014 GBE

Πρόκληση: γλωσσική επαφή → οριζόντια μεταφορά

(Ladoukakis, Michelioudakis & Anagnostopoulou 2022)

Έννοια	Βιολογία	Γλώσσα
Κληρονόμηση/ μεταφορά χαρακτηριστικών	κάθετη	Κάθετη και οριζόντια
Γονείς	1-2	1-πολλοί
Κληρονόμηση επίκτητων χαρακτηριστικών	όχι	ναι
Ανταλλαγή πληροφορίας	Μεταξύ ειδών όχι	Μεταξύ γλωσσών ναι

Για να είναι αξιόπιστες οι φυλογενέσεις στη γλώσσα θα πρέπει να στηριχθούμε σε χαρακτηριστικά που μεταφέρονται κάθετα.

Ομόρριζα δεδομένα: μόνο μέχρι την Πρωτογλώσσα μιας οικογένειας

- Στη σύγχρονη ιστορική έρευνα που στηρίζεται σε φυλογενετικές μεθόδους, τα δεδομένα που χρησιμοποιούνται προέρχονται από το **συντηρητικό λεξιλόγιο των γλωσσών**, το οποίο έχει αποδειχθεί ότι αντιστέκεται στο δανεισμό (Swadesh 1955).

- Σύμφωνα με την κρατούσα άποψη στην ιστορική γλωσσολογία, αυτή η μέθοδος μπορεί να φτάσει **σε ένα ιστορικό βάθος 6000-10000 χρόνων, στην Πρωτογλώσσα της οικογένειας.**

Αγγλικά	Ιταλικά	Ρωσικά	Ουγγρικά
one	uno	odin	egy
two	due	dva	két
three	tre	tri	három
mother	madre	mat’	anya
brother	fratello	brat	testvér
sister	sorella	sestrá	növér

Μορφοσυντακτικά δεδομένα: πιο πίσω στο χρόνο

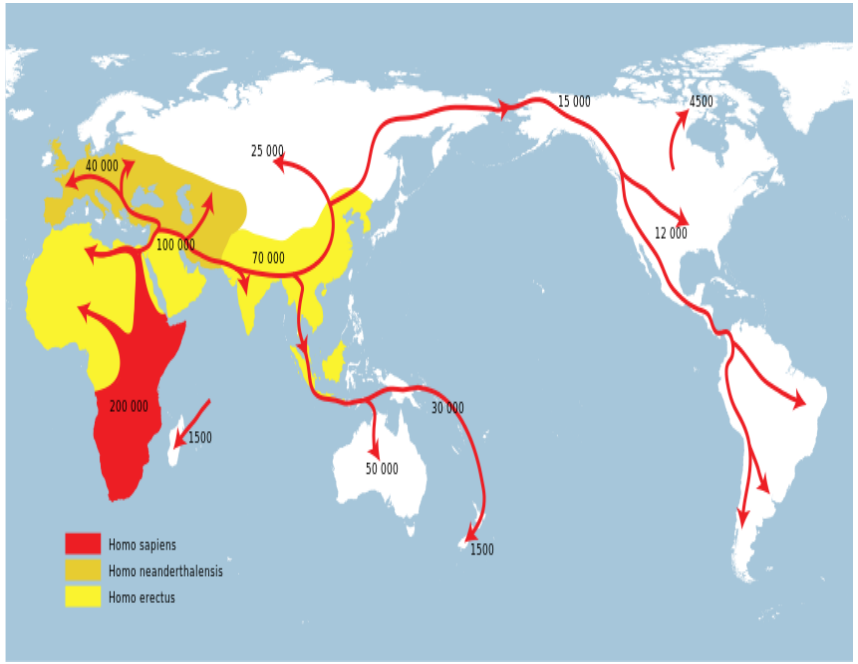
- Η Nichols (1992) πρότεινε ότι **τα γραμματικά/ μορφοσυντακτικά χαρακτηριστικά** μπορούν να μας βοηθήσουν να ανιχνεύσουμε **βαθύτερες ιστορικές σχέσεις**, τουλάχιστον 15.000 ετών ενδεχομένως έως και 50.000 ετών.

Παραδείγματα πολύ σταθερών και σταθερών μορφοσυντακτικών χαρακτηριστικών στον πίνακα.

Γλώσσα	Άλλη ρίζα στον ενικό των αντωνυμιών α' προσώπου (ONOM/AIT) *εγώ-εμένα/ I-me*	Γένη στα ουσιαστικά	Κλιτικές τάξεις στα ουσιαστικά
Αγγλικά	Ναι	Όχι	Όχι
Γερμανικά	Ναι	Ναι	Ίχνη
Λιθουανικά	Ναι	Ναι	Ναι
Ρώσικα	Ναι	Ναι	Ναι
Γαλλικά	Ναι	Ναι	Όχι
Αρμένικα	Ναι	Όχι	Ίχνη

Η περιπλάνηση του ανθρώπου

- Η κατανομή χαρακτηριστικών **από τα δυτικά προς τα ανατολικά** συμβαδίζει με ανεξάρτητες ενδείξεις για τις μεταναστεύσεις του προϊστορικού ανθρώπου (**Παλιός – Κόσμος, Ωκεανία/ Νέα Γουινέα, Νέος-Κόσμος**)
- Αυτό συνάδει με την επικρατέστερη σήμερα θεωρία για το πώς εξαπλώθηκε ο σημερινός άνθρωπος. Εξελίχθηκε βιολογικά στην Αφρική πριν από 50.000-100.000 (μπορεί και 200.000-80.000) χρόνια και από εκεί εξαπλώθηκε.
- Η θεωρία στηρίζεται σε γενετικό και αρχαιολογικό υλικό (Liu et al. 2006, Stringer 2003).



Γενετικές και μορφοσυντακτικές αποστάσεις

- **Πρόσφατα ευρήματα** των Matsumae *et al*. 2021 επιβεβαιώνουν ότι μόνο με τα μορφοσυντακτικά χαρακτηριστικά μπορούμε να πάμε σε **βάθος χρόνου** και να ανιχνεύσουμε **σχέσεις μεταξύ οικογενειών, επειδή μόνο οι μορφοσυντακτικές αποστάσεις συσχετίζονται συστηματικά με τις γενετικές αποστάσεις.**

- Οι ερευνητές μελετούν και συγκρίνουν γενετικό υλικό από **14 πληθυσμούς** της Άπω Ανατολής με γλώσσες **από διαφορετικές γλωσσικές οικογένειες** και δείχνουν ότι:

  - Οι **γενετικές αποστάσεις** μεταξύ των πληθυσμών **συσχετίζονται συστηματικά** με τις αποστάσεις των γλωσσών τους ως προς τα **μορφοσυντακτικά/ γραμματικά χαρακτηριστικά** τους.

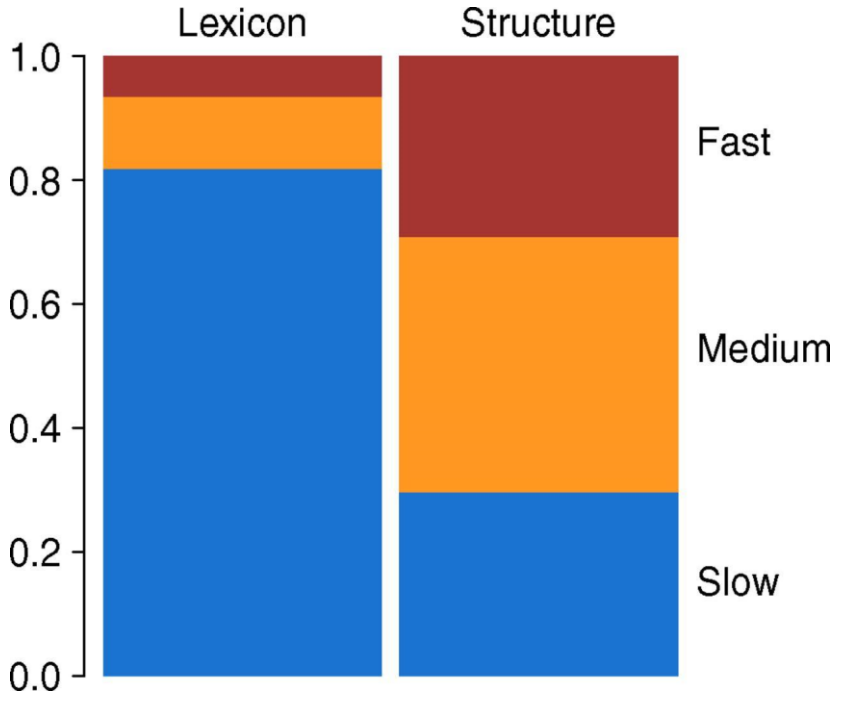
  - Οι **γενετικές αποστάσεις** μεταξύ των πληθυσμών **δεν συσχετίζονται συστηματικά** με τις αποστάσεις των γλωσσών τους ως προς τα **λεξικά χαρακτηριστικά τους** (ομόρριζα), τα **φωνολογικά χαρακτηριστικά τους** (τους φθόγγους και τον τονισμό) και τη **μουσική τους.**

Ρυθμοί μεταβολής χαρακτηριστικών

- Από την άλλη πλευρά, οι Greenhill et al. (2017) βρίσκουν ότι τα λεξικά χαρακτηριστικά είναι περισσότερο αξιόπιστα από τα γραμματικά χαρακτηριστικά για δύο πιθανούς λόγους:

  i. Πολλά γραμματικά χαρακτηριστικά **αλλάζουν γρηγορότερα** από το συντηρητικό τμήμα του λεξικού

  ii. τα δομικά δεδομένα εμφανίζουν υψηλότερα επίπεδα **αντιφατικής πληροφορίας** λόγω
  - παράλληλης εξέλιξης και
  - γεωγραφικής επαφής/ διάχυσης.



Greenhill et al 2017, PNAS

Τι επιδιώκει το ModelGloss

- Στόχος μας είναι να διερευνήσουμε και να αποκωδικοποιήσουμε την **κάθετη** και **οριζόντια** μεταφορά χαρακτηριστικών στη μορφοσύνταξη, διακρίνοντας:

1. Χαρακτηριστικά με **γενεαολογική σταθερότητα** που χαρακτηρίζουν οικογένειες και γένη.
2. Χαρακτηριστικά που είναι ευαίσθητα στην **οριζόντια μεταφορά λόγω γλωσσικής επαφής.**
3. Χαρακτηριστικά που είναι κοινά λόγω **καθολικών χαρακτηριστικών** της ανθρώπινης γλώσσας.

Βάση δεδομένων 60 IE γλωσσών βασισμένη στα χαρακτηριστικά του World Atlas of Language Structures (WALS)

Σύγχρονες Γλώσσες (53)

- Ευρωπαϊκές Γλώσσες (44)
  - Αλβανικά
  - Αρμένικα
  - ΚΝΕ, Κυπριακά, 3 Ιταλιώτικες & 5 Μικρασιατικές ποικιλίες
  - Ρομανικές γλώσσες (7)
  - Κέλτικές Γλώσσες (4)
  - Βαλτο-σλαβικές Γλώσσες (12)
  - Γερμανικές Γλώσσες (9)
- Ινδο-Ιρανικές Γλώσσες(9)
  - Ινδικές (7)
  - Ιρανικές (2)

Παλαιότερα συστήματα (7)

- Κλασική Ελληνική
- Ελληνικά Καινής Διαθήκης
- Λατινική
- Παλαιά Εκκλησιαστική Σλαβονική
- Γοτθικά
- Παλαιά Αγγλικά
- Σανσκριτικά

Δεδομένα του ModelGloss

Τα δεδομένα του ModelGloss περιλαμβάνουν:

- Πίνακα χαρακτηριστικών του WALS με τιμές που συλλέχθηκαν για το ModelGloss για 60 Γλώσσες της Ινδο-Ευρωπαϊκής.

		Characters					
		37A	38A	39A	40A	41A	42A
Languages	Italian	1	2	3	3	2	1
	Spanish	1	2	3	3	3	1
	Catalan	1	2	3	3	3	1
	French	1	2	3	3	1	2
	EU Portug.	1	2	3	3	3	1
	Br. Portug.	1	2	3	3	3	1
	Romanian	3	2	3	3	2	1
	Latin	5	5	3	3	4	1
	Cl. Greek	1	4	3	3	3	1
	New Test.	1	4	3	3		1

Δυαδικοί χαρακτήρες

- Από τα 156 χαρακτηριστικά του WALS, επετεύχθη πληρότητα ≥75% για 92.
- Μετατρέψαμε τις τιμές αυτών των 92 χαρακτηριστικών σε ερωτήσεις ολικής αγνοίας που ρωτούσαν εάν η κατάσταση που περιγράφεται από την τιμή του WALS είναι αληθής ή όχι για την εκάστοτε υπό εξέταση γλώσσα.
- Αυτή η μετατροπή απέδωσε 425 δυαδικούς χαρακτήρες.

		Characters				
		37A-1	37A-2	37A-3	37A-4	37A-5
Languages	Italian	1	0	0	0	0
	Spanish	1	0	0	0	0
	Catalan	1	0	0	0	0
	French	1	0	0	0	0
	EU Portug.	1	0	0	0	0
	Br. Portug.	1	0	0	0	0
	Romanian	0	0	1	0	0
	Latin	0	0	0	0	1
	Cl. Greek	1	0	0	0	0
	New Test.	1	0	0	0	0

Μέθοδοι αναπαράστασης της εξελικτικής ιστορίας με δέντρα

- Μέθοδοι αποστάσεων (Distance-based).
- Μέθοδοι μέγιστης φειδωλότητας (Maximum Parsimony).
- Μπεϋζιανές φυλογενετικές μέθοδοι.

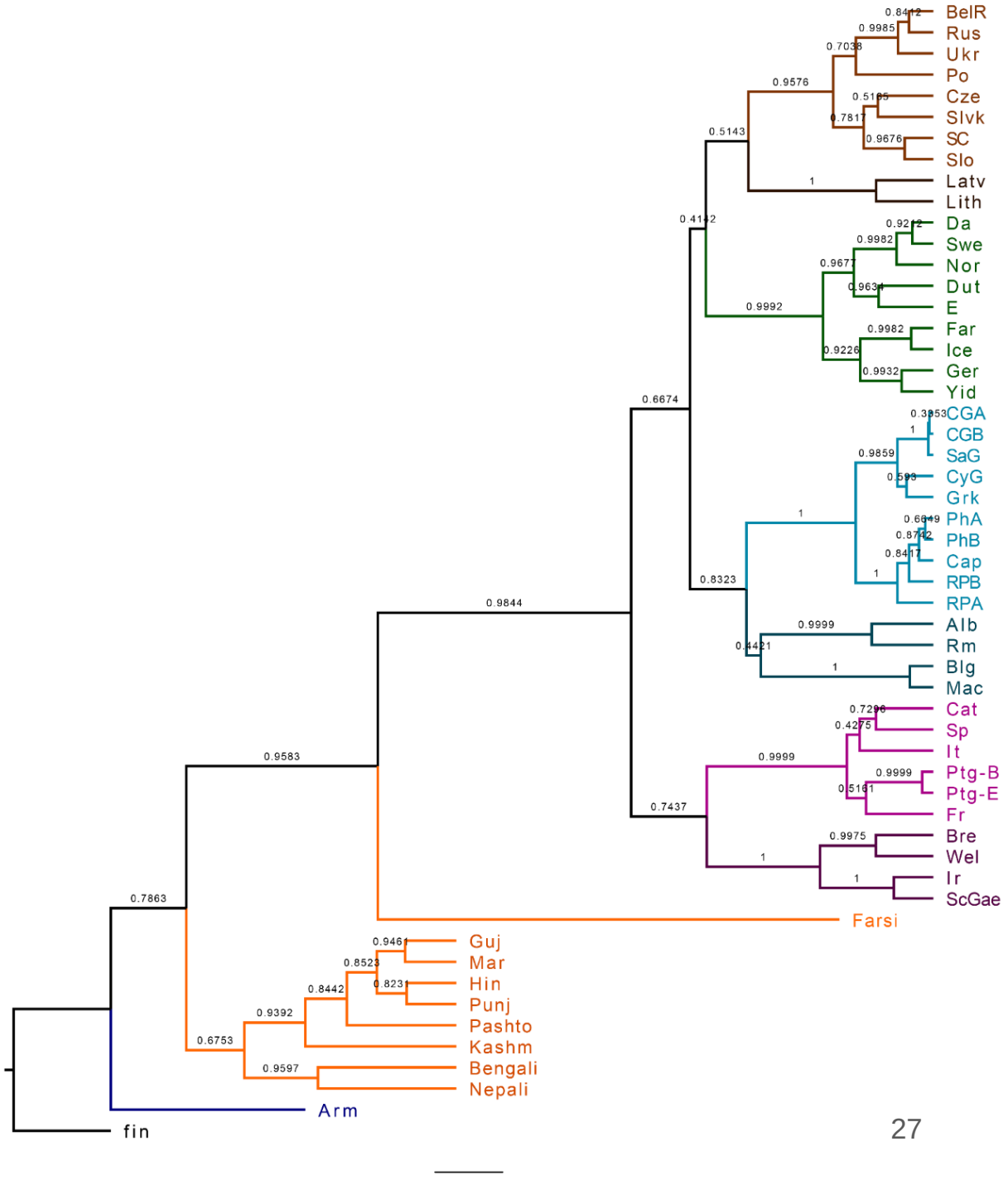
Τα καλύτερα δέντρα μας

Το δέντρο που είναι πιο κοντά στην παραδεδεγμένη γνώση για την IE, προέκυψε από μπεϋζιανή ανάλυση:

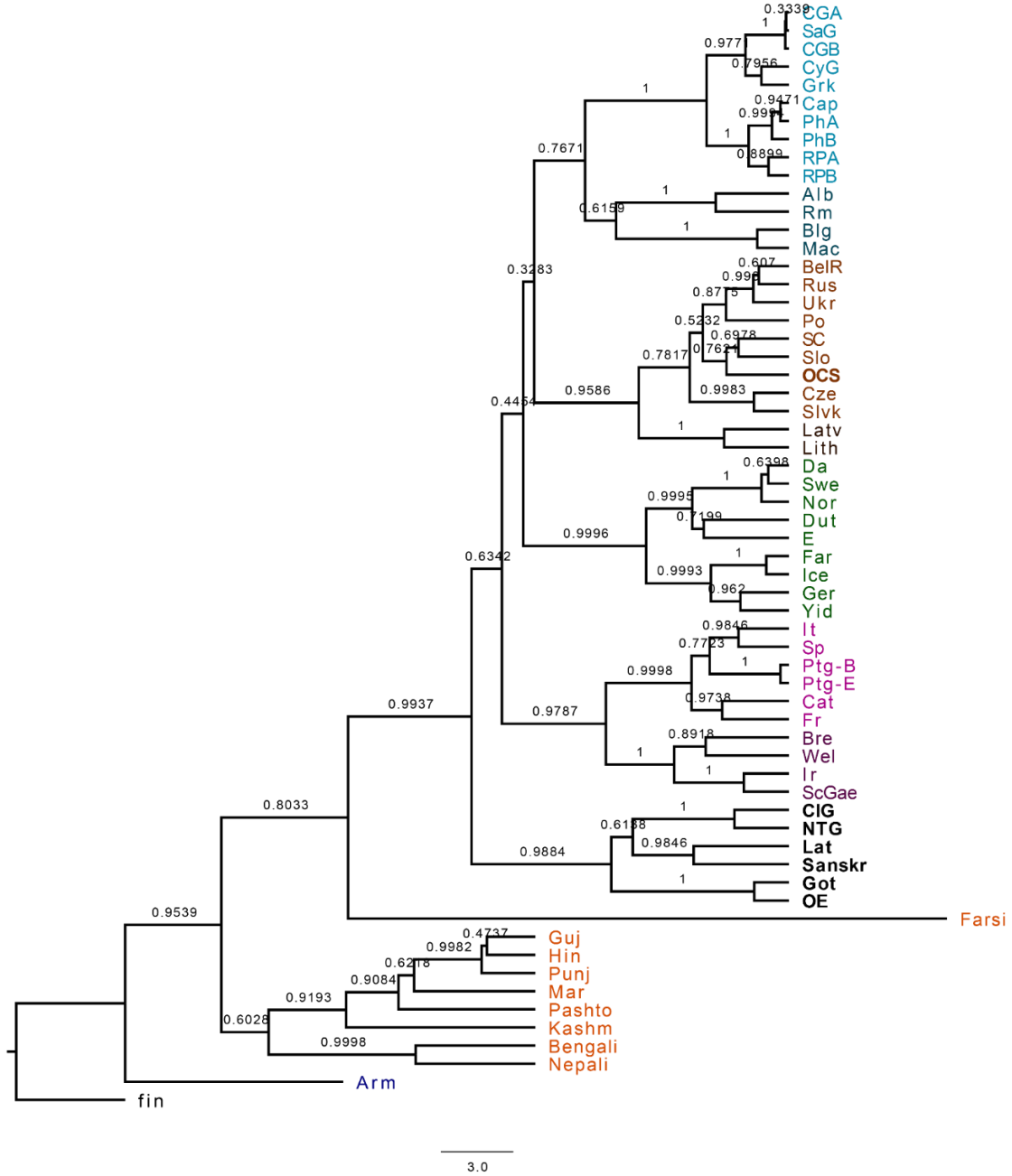
(i) ενός δείγματος που περιλάμβανε μόνο τις σύγχρονες γλώσσες

(ii) μοντέλο αυστηρού ρολογιού;

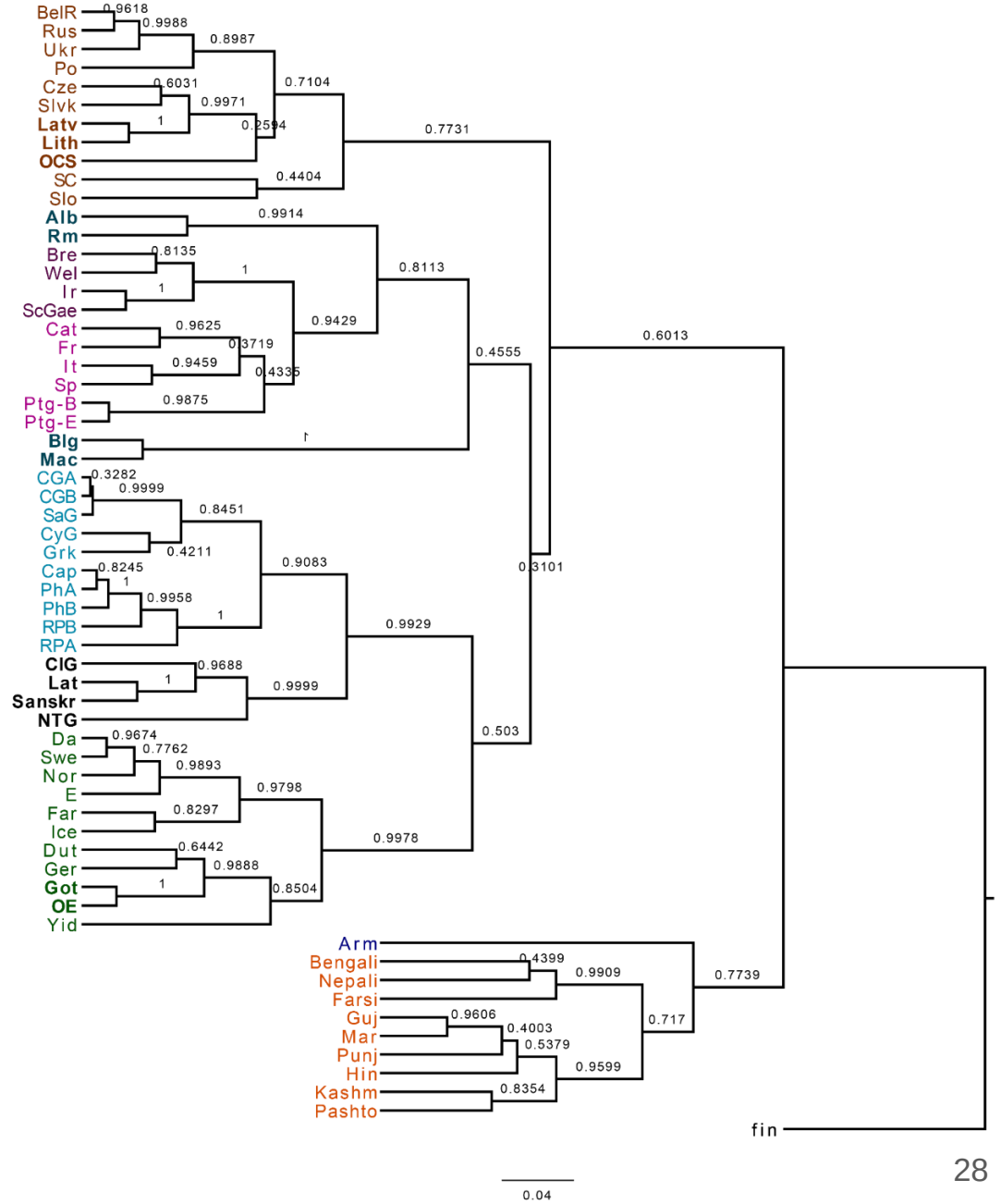
(iii) πολύ σταθερούς, σταθερούς και ασταθείς χαρακτήρες (δηλαδή χωρίς τους πολύ ασταθείς χαρακτήρες).



Strict Clock



Relaxed Clock



Ευρήματα: ιστορικές και γεωγραφικές πληροφορίες λόγω κάθετης και οριζόντιας μεταφοράς χαρακτηριστικών

1. Τα μορφοσυντακτικά χαρακτηριστικά παρέχουν **χρήσιμες ιστορικές πληροφορίες. Επανασυνθέτουν τα γένη**, π.χ. Ρομανικές, Κελτικές, Ελληνικά (με τις διαλέκτους τους), Σλαβικές, Γερμανικές, Ινδο-ιρανικές.

2. Σημαντική εξαίρεση**:** ομαδοποίηση γλωσσών με **μακροχρόνια γλωσσική επαφή,** γλώσσες του **Βαλκανικού Sprachbund** (π.χ. Αλβανικά, Ρουμάνικα, Βουλγάρικα, Ελληνικές ποικιλίες).

3. Τα δέντρα που βασίζονται σε **ομόρριζα δεδομένα από το συντηρητικό λεξιλόγιο** και παράγονται ακριβώς με τις ίδιες μεθόδους δεν εμφανίζουν αντίστοιχα λάθη στις τοπολογίες τους.

Νέες μέθοδοι για την ανάλυση των χαρακτηριστικών

- Αναπτύσσουμε **νέες ποσοτικές φυλογενετικές μεθόδους** που θα μας επιτρέψουν, σε συνδυασμό με **ποιοτικές γλωσσολογικές μεθόδους,** να μελετήσουμε σε βάθος τον ρόλο του **κάθε χαρακτηριστικού** και **των συνδυασμών τους** στην παραγωγή δέντρων και να κατανοήσουμε πώς παράγονται οι **σωστές και λανθασμένες** ομαδοποιήσεις γλωσσών.

Τι θα πετύχουμε

1. Με **τη βοήθεια των φυλογενετικών δέντρων και των μεθόδων μας** θα εντοπίσουμε και θα ερμηνεύσουμε τα χαρακτηριστικά με **γενεαλογική σταθερότητα** διακρίνοντάς τα από χαρακτηριστικά που **μεταφέρονται οριζόντια** ή **εξελίσσονται παράλληλα** στις γλώσσες.
2. Θα χαρτογραφήσουμε τις **δυνάμεις γλωσσικής αλλαγής** στο επίπεδο των σχέσεων μεταξύ **ποικιλιών της ίδιας γλώσσας** (μικρο-ποικιλότητα), σχέσεων μεταξύ **γλωσσών και κλάδων της ίδιας οικογένειας** (μεσο-ποικιλότητα), και **σχέσεων μεταξύ οικογενειών** (μακρο-ποικιλότητα).
3. Θα δώσουμε **λίστα συντηρητικών μορφοσυντακτικών χαρακτηριστικών** αντίστοιχη με την **λίστα συντηρητικών ομόρριζων** του Swadesh (1955).
4. Αυτό θα επιτρέψει να **ξεπεράσουμε το φράγμα της Πρωτογλώσσας** και να μπορέσουμε να διερευνήσουμε με φυλογενετικές μεθόδους **σχέσεις μεταξύ οικογενειών**.

Σας ευχαριστώ πολύ!