



# UNIVERSITAT DE BARCELONA

## Població i Ordre de Constituents

Mikel Fernández Georges

**ADVERTIMENT.** La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX ([www.tdx.cat](http://www.tdx.cat)) i a través del Dipòsit Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

**ADVERTENCIA.** La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR ([www.tdx.cat](http://www.tdx.cat)) y a través del Repositorio Digital de la UB ([diposit.ub.edu](http://diposit.ub.edu)) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

**WARNING.** On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX ([www.tdx.cat](http://www.tdx.cat)) service and by the UB Digital Repository ([diposit.ub.edu](http://diposit.ub.edu)) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

# Població i Ordre de Constituents

**Mikel Fernandez Georges**

Tesi presentada per optar  
al grau de **Doctor en Lingüística**  
al programa de doctorat de  
*Lingüística i Comunicació bienni, 2006-2008*  
Departament de Lingüística,  
Universitat de Barcelona,

sota la direcció de

**Dra. Carme Junyent Figueras**  
Universitat de Barcelona



Desembre de 2015

El meu agraïment més sincer a tothom qui m'ha ajudat i ha estat amb mi durant la realització d'aquesta tesi.

A la doctora Carme Figueras, directora de la tesi, pel suport que de del principi i en tot moment m'has donat per realitzar la tesi, a partir d'una idea inicial que ha acabat en aquest volum. Moltes gràcies Carme.

Als tots els professors dels cursos de doctorat de la Universitat Autònoma de Barcelona,

Als companys i professors del meu primer intent de fer un doctorat, la Rosa, El David, la Meritxell, la Cristina, la Rosa, la Sofia, la Montse, la Beatriz, l'Alberto, la Cesca i el Josep.

A la Roser, per escoltar i ajudar-me sempre que li he demanat.

A l'Àlex, per haver revisat algunes parts i haver-me resolt dubtes amb el seus coneixements enciclopèdics.

Al Jordi i l'Anna, la Maria José, en Carlos, l'Eduardo, la Montse, per escoltar durant anys com remugava sobre la tesi.

A l'Eduard, pels seus consells i ajuda en molts dubtes i converses.

A l'Iain per ajudar-me amb es dubtes d'anglès i tants altres.



A la meva mare i al meu germà



Aquesta tesi és una recerca sobre els factors, sobre tot els de tipus social com la població, que poden ajudar explicar la distribució geogràfica i filogenètica de l'ordre d'alguns dels constituents més àmpliament usats a la bibliografia sobre tipologia lingüística.

Donat que aquest ordre no té cap motiu per estar sotmès a una pressió selectiva, ja que cap autor ha trobat que un canvi d'ordre sintàctic pugui afavorir l'expansió d'una llengua, es fa necessari trobar una explicació pel fet que la distribució sembla força capriciosa en molts casos.

L'estudi comenta els principals treballs que han suposat un avenç en aquest camp o que han fet aportacions que han ajudat a la realització de la tesi.

El treball ha requerit la realització d'una ampla base de dades en la que s'han recollit el valor de nou dels constituents que més s'han fet servir a la bibliografia. A partir d'aquesta base de dades s'han realitzat proves estadístiques, s'han confeccionat mapes de distribució i creat arbres filogenètics per inferir la història dels diferents valors a nivell diacrònic.

El principal resultat és la detecció de set àrees del món on aparentment s'ha produït un canvi des d'un ordre de nucli final al seu contrari en quatre de les variables estudiades. Aquestes àrees serien les responsables de l'existència de la pràctica totalitat de llengües de nucli inicial. Per la resta de variables no s'ha proposat cap patró concret. La unitat d'estudi en una majoria de casos no ha estat la llengua sinó el *genus* per raons d'eficàcia.

Un segon resultat rellevant és el fet que les set àrees precedents es troben totes en zones on coincideix la presència d'una gran riquesa lingüística a nivell de *genus* al costat d'una gran densitat de població, lligada a l'ús de l'agricultura o piscicultura.

Un tercer resultat és el fet que la variància de de les diferents variables depèn principalment d'un únic factor que sembla arrossegar als altres. També s'han trobat algunes correlacions interessants entre diversos factors.





**Title:** Population and Constituent word orders

**Author:** Mikel Fernandez Georges

**Advisors:** Carme Junyent Figueras

**Date:** December 12, 2015

**Abstract:**

This dissertation is work of research into the factors, especially those of a social nature such as population, which may help explain the geographic distribution and phylogeny of some of the syntactic constituents used widely in the literature on linguistic typology.

Since syntactic order has no reason to be subjected to selective pressure, and no author no author has found that a change of syntactic order may favor the expansion of a language, it is necessary to find an explanation for what seems to be quite a capricious distribution in many cases.

The study comments on previous works on word order and the spread of language families which have furthered knowledge in this field or have reached conclusions that have aided the writing i of the thesis.

The work has required the creation of a comprehensive database that has brought together information on nine of the constituents that have been used widely in the literature. Using this database statistical tests were have carried out distribution maps created and phylogenetic trees were constructed to indicated the different values diachronically to infer the history of different variables.

The main result is the detection of seven areas of the world where apparently there has been a shift from left-branching to a right-branching pattern in the case of four of the variables studied. These areas are responsible for the existence of almost all the right-branching languages. For the remaining variables no specific pattern has emerged. In the majority of cases the framework of the study has not been individual languages but, for reasons of efficacy, the genus.

A second important finding is that the preceding seven areas under study are ones characterised by dense linguistic variety at the level of the genus coinciding with a high population density and a society based on farming.

A third result is that the variance of the different variables depends primarily on a single factor that seems to drag the others along. We also found some interesting correlations between different factors.



## ÍNDIX

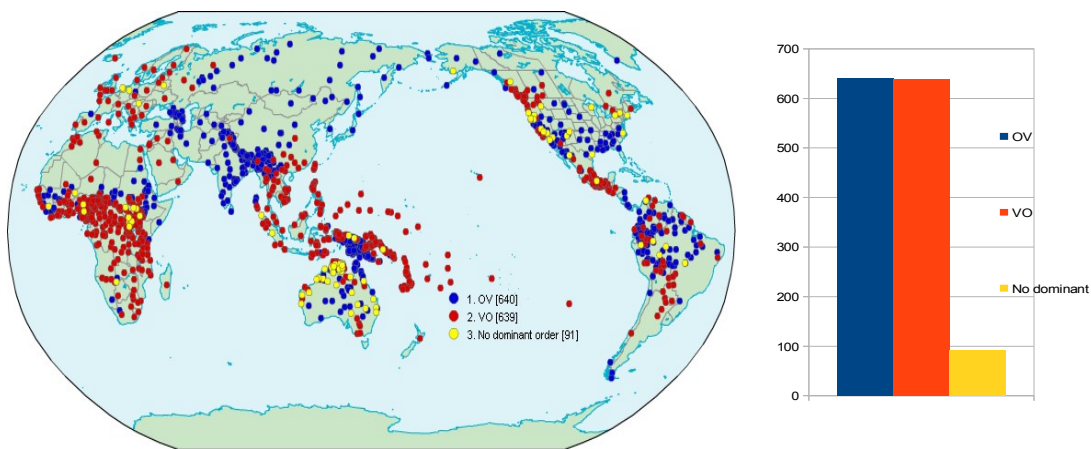
<b>1 INTERÈS I OBJECTIUS</b> .....	1
<b>2 INTRODUCCIÓ</b> .....	3
2.1 Antecedents bibliogràfics .....	4
2.1.1 Ordre de constituents .....	4
2.1.2 Dispersió i propagació de les famílies lingüístiques .....	11
2.1.3 Correlacions entre factors lingüístics i socials .....	15
2.2 Les societats humanes .....	24
2.2.1 Dispersió de l' <i>Homo sapiens</i> .....	24
2.2.2 Demografia històrica .....	25
2.2.3 Gens i llengües .....	26
2.2.4 Neolític i domesticació de plantes .....	28
2.2.5 Neolític i famílies lingüístiques .....	29
2.2.6 Llengües de llarga tradició literària .....	32
2.3 Canvi lingüístic i factors socials .....	34
2.3.1 Pídgins .....	34
2.3.2 Criolls .....	35
2.3.3 Llengües en vies d'obsolescència .....	35
2.3.4 Llengües franques .....	36
2.3.5 Complexitat .....	36
<b>3 MATERIAL I MÈTODES</b> .....	39
3.1 Mostres .....	39
3.1.1 Problemes de mostratge .....	39
3.1.2 <i>Genus</i> i llengües .....	40
3.1.3 Recollida de dades .....	41
3.1.4 Tractament de dade .....	42
3.1.5 Ordre de constituents .....	43
3.1.6 Població dels <i>genus</i> .....	44
3.1.7 Programaris específics .....	44
3.2 Correlacions i diagrames de barres .....	45
3.3 Anàlisi de Components Principals .....	45
3.3.1 Presentació del mètode .....	45
3.3.2 Aplicació .....	46
3.4 Mapes .....	46
3.5 Arbres.....	47
<b>4 RESULTAT ACP</b> .....	49
4.1 Resultats ACP WAL5 i interpretació .....	49
4.2 Comparació amb WAL8 .....	51
4.3 Comparació amb la resta de bases de dades .....	52
4.4 Conclusions de la prova ACP .....	52

<b>5 RESULTATS: CORRELACIONS</b>	53
5.1 Adposició i variables socials	53
5.2 Adposició i població	52
5.3 Ordre de constituents i població	54
5.4 Ordre de constituents i geografia	56
5.5 Adposició per continents	58
5.6 Adposició i ramificació	60
5.6.1 Ramificació de <i>genus</i>	60
5.6.2 Ramificació de famílies	61
5.7 Altres factors	62
5.7.1 Adposició i dependència de l'agricultura	62
5.7.2 Població, fonemes i ordre de constituents	63
5.7.3 Geografia del nombre de fonemes, WLI, i distància a l'Àfrica	64
<b>6 DISTRIBUCIÓ DE LES VARIABLES</b>	67
6.1 Subjecte	68
6.2 Objecte	70
6.3 Adposició	72
6.4 Genitiu	74
6.5 Adjectiu	76
6.6 Demonstratiu	78
6.7 Numeral	80
6.8 Possessiu	82
6.9 Relatiu	84
<b>7 ÀREES ESPECIALS</b>	87
7.1 Àfrica	88
7.2 SE Àsia	90
7.3 Colúmbia Britànica	92
7.4 Mesoamèrica	94
7.5 Regió Euromediterrània	96
7.6 Amazònia boliviana	98
7.7 Nova Guinea	100
7.8 Les set àrees: dades comparatives	102
7.9 Visió diacrònica	104
<b>8 DISCUSSIÓ I MODEL EXPLICATIU</b>	109
<b>9 CONCLUSIONS</b>	113
BIBLIOGRAFIA	115
ANNEXOS	123

# 1 INTERÈS I OBJECTIUS

Ja a la primera meitat del segle XVIII, abans de l'aparició de la lingüística com a ciència i també de la comparatística, hi havia autors que, com Gabriel Girard [1], constataren que podia distingir-se entre les llengües en que el verb va davant dels seus complements, i aquelles amb l'ordenació contrària. Aquesta primera intuïció s'ha vist confirmada i s'ha ampliat posteriorment a d'altres constituents sintàctics donant lloc a la moderna distinció entre llengües de nucli inicial i llengües de nucli final.

Al veure la distribució geogràfica d'aquests dos tipus lingüístics i quantificar-los, com es pot veure al mapa 1 i al seu corresponent gràfic que mostren la distribució d'ordenacions del Verb i l'Objecte, agafat com exemple, queda clar que cap ordre predomina sobre l'altre ni a nivell geogràfic ni numèric. També sembla evident que es reparteixen les àrees geogràfiques i que les llengües sense un ordre dominant són molt minoritàries.



**Mapa 1.** Mostra la distribució de llengües segons l'ordre del Verb i l'Objecte (Font: WALS [2])

**Gràfic 1.** Quantificació de les llengües del mapa 1.

El fet que el verb vagi davant o darrera de l'objecte no té perquè tenir cap efecte positiu ni negatiu en la comprensió d'una oració ni molt menys pot afavorir que una llengua es propagui per sobre d'altres.

Al mateix temps, la lingüística històrica deixa clar que moltes llengües han canviat la seva gramàtica al llarg dels segles de forma important, com és el cas del català mateix respecte del llatí. També és un fet que aquests canvis s'han observat amb més o menys intensitat a tot arreu on l'existència de textos antics ha fet possible la comparació.

Dels dos fets anteriors cabria esperar que les llengües vagin canviant lliurement sense cap sentit especial ni sense que els canvis siguin particularment més nombrosos en un lloc o època que en qualsevol altre. Però justament la tipologia ha demostrat que la situació real és completament diferent. Els trets lingüístics, i especialment els sintàctics que s'examinen en aquesta tesi, es reparteixen geogràficament d'una manera molt ordenada i en àrees compactes i cohesionades.

Justament aquest és l'interès que ha motivat aquesta tesi. Com explicar que llengües sense parentiu que es parlen a través de milers de kilòmetres i que han anat canviant històricament, han acabat totes coincidint en una mateixa alternativa en una mateixa àrea? I perquè en altres continents la coincidència es dona en el sentit contrari?

La situació és encara més intrigant si tenim en compte que no hi ha dubte que totes les

llengües humanes procedeixen d'una única llengua primigènia, amb la seva gramàtica i el seu ordre sintàctic. Sigui quin sigui l'ordre que tenia aquesta llengua *Proto-World*, resulta evident que el nombre de llengües que li han restat fidels és similar al d'aquelles que l'han canviat pel seu contrari.

Per explicar-ho, cal descartar la comparació amb l'evolució dels éssers vius, malgrat que és un fet recurrent en lingüística, ja que ningú ha trobat cap gramàtica més *ben adaptada* que d'altres. Sempre podria quedar el dubte raonable sobre si existeixen tipologies lingüístiques realment problemàtiques i que dificultin la comunicació, però si hagués existit realment un cas així, seria impensable que el canvi lingüístic o el manlleu no hagués solucionat el problema en un parell de generacions.

Aquest treball es fonamenta en el fet empíric de que totes les llengües són igualment vàlides a nivell comunicatiu i en que no existeixen llengües millors ni més ben adaptades que d'altres. La substitució d'unes llengües per altres ens podria donar, en tot cas, idea del potencial militar o demogràfic dels pobles que les parlen, mai de les seves virtuts comunicatives.

Així, la pregunta bàsica a la qual vol respondre aquesta tesi és:

- Per què aproximadament un 50% de les llengües del món han canviat el seu ordre sintàctic original en determinades zones del planeta, si no hi ha cap avantatge evolutiu conegut per fer-ho?

Altres objectius d'aquesta tesi també són:

- Investigar si els ordres dels diferents constituents oracionals responen a múltiples factors o a només uns pocs que arrosseguen la resta.

Per dur-ho a terme s'aplicaran diferents metodologies a la informació continguda en una col·lecció de bases de dades tan completa com ha sigut possible.

- En cas de trobar algun factor determinant que expliqui la distribució, estudiar-lo per veure de quina manera varia a nivell geogràfic i filogenètic, i si està lligat a altres variables lingüístiques o no. És cercaran patrons de distribució i propagació.

Per portar-ho a cap, es realitzaran mapes i arbres filogenètics on els colors distingiran les diferents opcions de cada variable i es consultarà la bibliografia per comprovar les diferents hipòtesis relatives al tema proposades per autors anteriors.

- S'intentarà realitzar una proposta de quin ha sigut l'històric de canvis sintàctics dels diferents grups lingüístics des de la seva aparició fins avui.

- Es proposarà una hipòtesi explicativa de les tipologies i patrons dominants al món.

## 2 INTRODUCCIÓ

El llenguatge és un sistema natural d'organització, empaquetament, intercanvi i emmagatzematge d'informació. El seu alt grau d'eficiència permet als humans aprofitar una quantitat d'informació fora de l'abast de qualsevol altre espècie animal que sens dubte ha resultat fonamental en la seva supervivència i expansió per tot el planeta.

La immensa majoria de llengües del món són sensibles a l'ordre de les paraules, sent justament aquest ordre el que vehicula una gran part de la informació transmesa en un discurs. Però no existeix cap llengua en que les diferents paraules individualment puguin ocupar lliurement qualsevol lloc d'una oració. Les paraules s'organitzen sempre de manera jeràrquica en unitats que són alhora funcionals, categorials i semàntiques anomenades components (o sintagmes segons algunes propostes gramaticals). Quan dos constituents s'uneixen, formen un nou constituent de nivell superior que hereta la funció i la categoria d'un d'ells (l'anomenat *nucli*), mentre que l'altre actua com a complement. Serà, doncs, l'ordre dels constituents que fan de nucli i complement el que determinarà el significat final de l'oració.

Al 1963, Greenberg va deixar clarament establert [3] que certs parells de constituents gramaticals es correlacionen entre ells de manera que el fet que en una llengua el nucli segueixi al complement en un determinant tipus de constituent gramatical, permet intuir-ne l'ordre d'altres, encara que s'ha posat en qüestió per Dunn [4]. L'exemple més evident és l'acreditada afirmació que a les llengües on el verb precedeix l'objecte també és freqüent trobar -hi preposicions, mentre que en cas contrari, el que hi trobem són postposicions.

Posteriors investigacions han deixat clar que els diferents ordres observats no són igualment freqüents ni es distribueixen a l'atzar. N'hi ha que són molt més corrents que d'altres i la distribució a les diferents àrees geogràfiques i grups filogenètics mostra patrons força clars. Els lingüistes han esmerçat molts esforços en identificar quins són els diferents constituents que correlacionen entre si i les possibles causes, així com el perquè de la seva peculiar distribució.

La majoria de propostes i línies de recerca per explicar les diferents freqüències que trobem d'ordre de constituents es poden enquadrar en dues vies principals d'aproximació. Una és suposar que els patrons estan codificats genèticament com a part d'una Gramàtica Universal. L'altra és proposar principis lingüístics generals o tendències tipològiques degudes a factors funcionals i d'eficiència. Encara s'hi podria afegir com tercera alternativa l'existència d'uns hipotètics principis de canvi lingüístic que donarien com a resultat final, després d'una llarga història de canvi diacrònic, la situació present. A les següents pàgines es comentaran els principals autors que han intervingut en el debat per cadascuna d'aquestes línies.

Aquesta tesi intenta fer una aproximació multidisciplinària al tema. Es vol fer servir les dades disponibles a nivell tipològic, històric, geogràfic, filogenètic i estadístic per tal de donar dimensió temporal a la foto fixa que veiem avui dia per aconseguir una pel·lícula que permeti veure quin ha estat l'història del canvis succeïts i acostar-se a una explicació del fenomen.

## 2.1 ANTECEDENTS BIBLIOGRÀFICS

En bona part, aquesta tesi s'ha fet a partir de la informació lingüística de desenes de milers de referències bibliogràfiques, però filtrada i tabulada en forma de bases de dades *online* i de lliure accés. La tesi s'ha vist molt condicionada i limitada per factors metodològics i de qualitat d'aquestes dades, de manera que la tesi ha resultat un compromís entre allò que es volia fer i el era factible.

La producció bibliogràfica sobre tipologia és enorme i la història d'aquesta ciència és llarga. A l'hora de redactar una secció sobre l'estat de la qüestió calia fer una selecció de les aportacions més rellevant i restringir-la als temes tractats a la tesi. Això vol dir que s'han obviat les prolífiques discussions sobre les categories gramaticals, la classificació de les llengües, el limit entre llengua i dialecte, els universals lingüístics, les polèmiques entre generativistes i funcionalistes o l'evolució del llenguatge. Totes elles, i moltes altres, tenen a veure amb la tesi, però només de forma tangencial.

Així que només es farà una revisió dels principals autors que s'han centrat en el temes més directament implicats com el de l'ordre de constituents, el de la propagació de les famílies lingüístiques i la correlació hipotètica entre factors socials i lingüístics. Aquests temes són els que han incidit directament en el desenvolupament de la tesi.

I de cada autor, només es comentaran les principals aportacions, ja que alguns d'ells tenen una bibliografia de dotzenes d'articles amb diferent grau interès sobre el tema, fruit en molts casos d'una llarga carrera i a cops contradictoris els uns amb els altres.

### 2.1.1 ORDRE DE CONSTITUENTS

Lehman [5] va fer una contraposició entre la tipologia del segle XIX, interessada principalment en la morfologia, i la del segle XX més interessada en la sintaxi, que és la més interessant per aquesta tesi. Malgrat aquesta generalització, ja al segle XVIII hi havia autors com Gabriel Girard [1], que l'any 1747 va distingir entre idiomes *anàlegs* i *transpositius*, basant-se en la sintaxi per definir les dues categories: sent els primers els d'ordre Subjecte-Verb-Objecte i els segons els d'ordre Subjecte-Objecte-Verb. Els noms que va escollir feien referència a si és seguia o contradeia el que ell pensava que era l'ordre natural i lògic del pensament (evidentment el de la seva llengua materna).

També un lingüista tan complet com Humbolt va avançar al 1836 [6] la idea que si tots els humans tenien la clau per entendre qualsevol llenguatge humà es podia deduir que el que ell anomenava la *forma* era la mateixa, un clar precedent del concepte de Gramàtica Universal. Al 1891 va ser Gabelentz [7] el que va definir la paraula *Tipologia* per referir-se a la disciplina i va ser el primer en deixar de considerar el canvi lingüístic com una degeneració i en proposar l'ús de l'estadística per fer de la tipologia una ciència predictiva. També va ser molt moderna la seva idea que el canvi lingüístic anava de l'estadi A al B, per tornar a altre cop a un estadi molt pròxim al l'A, però lleugerament diferent del primer, de manera que el resultat final seria una espiral.

Ja entrant en la qüestió pròpiament dita de l'ordre de constituents, al 1844 Weil [8] va distingir les construccions *ascendents* on les paraules qualificadores precedeixen les qualificades, de les *descendents* amb l'ordre invers. Tesnière [9] fa servir la terminologia de Weil, però hi afegeix com a sinònim, els de *centrifugues* per les construccions descendents i *centrípetes* per les ascendents.

Hawkins [10] defineix l'article de 1963 de Greenberg anomenat "*Some Universals of*



*Grammar with Particular Reference to the Order of Meaningful Elements*" com el treball seminal sobre el que s'han basat els posteriors treballs sobre els universals de l'ordre de constituents. Greenberg només va fer servir una mostra de 30 llengües representatives de diferents continents i grups filogenètics i unes 15 variables morfològiques o sintàctiques. Les que més joc van donar van ser l'ordre del subjecte, verb i objecte, el de les adposicions, el genitiu amb el nom i l'adjectiu amb el nom. A partir d'aquí va establir 45 Universals morfològics i sintàctics. Va representar els diferents ordres amb una codificació basada en representar els components segons el seu ordre i les seves inicials. D'aquesta manera GN volia dir que el genitiu va davant del nom i SVO que el verb va entre el subjecte i l'objecte.

Greenberg va donar la màxima importància a la variable de l'ordre del verb amb el subjecte i l'objecte, de la que va trobar tres alternatives possibles: SVO, VSO i SOV. Després va suposar que aquests ordres es correlacionaven amb la resta de variables. Els seus Universals eren majoritàriament implicatius i molts absoluts, de l'estil:

*Universal 3: Les llengües amb ordre dominant VSO són sempre preposicionals.*

Molts, però presenten contraexemples i els va haver d'enunciar com estadístics:

*Universal 17: Amb una freqüència significativa les llengües VSO són AN*

Greenberg també va formular dos principis, el de Dominància i el d'Harmonia. Un ordre dominant, en contraposició a un de recessiu, era el que no estava sotmès a condicions concretes, com per exemple, el fet que les preposicions són dominants perquè es poden trobar a llengües amb qualsevol ordre, mentre que les postposicions serien recessives perquè només les va trobar en llengües SOV.

El concepte d'*harmònic* o *disharmònic* fa referència al fet que hi ha grups de variables en que un parell d'ordenacions particulars de dues variables van invariablement junts o no. Per exemple les postposicions són *harmòniques* amb els ordres SV i OV, però *disharmòniques* amb VS i VO.

El treball de Greenberg 1963 [3] ha resultat transcendent, i els estudis posteriors amb mostres de milers de llengües no han suposat un canvi dels patrons principals. El que sí que han fet es trobar moltes excepcions als suposats universals absoluts i s'han trobat també ordres molt minoritaris que han complicat el panorama.

També ha quedat sense canvi, des del primer treball amb la mostra de 30 llengües fins els posteriors, el fet que els diferents ordres no es distribueixen a l'atzar sinó que dels 24 ordres possibles que surten de permutar les 4 variables bàsiques (Subjecte-Verb-Objecte, Adposició, Adjectiu i Genitiu) n'hi ha vuit que acumulen la major part de les llengües de la mostra, mentre que els altres setze estan buits o només contenen unes poques llengües. També s'ha mantingut la nomenclatura de Greenberg. La taula 1 de la pàgina 8 llista aquests 24 tipus, recollits per Hawkins per fer una mostra més ample.

Lehman (1973) [11] va proposar reduir la variable principal Subjecte-Verb-Objecte de tripartita a una de binària: Objecte-Verb (OV) i Verb-Objecte (VO), i en tot cas deixar l'ordre entre el Subjecte i el Verb per una altra variable. Això li permetia equiparar-la a la resta i proposar una explicació global a la que va anomenar *Fundamental Principle of Placement* (FPP). Segons aquest principi, els modificadors es col·loquen sempre al costat oposat d'un constituent bàsic respecte el seu concomitant. Així en un llenguatge OV els modificadors verbals aniran a la dreta del verb i els modificadors nominals a l'esquerra del nom. A les llengües VO passa just el contrari. Els modificadors en cap cas aniran entre el verb i l'objecte. Així hi hauria dos ordres bàsics en una oració:

Llengües VO amb ordre (adjectius, genitius i relatius)- Objecte- Verb- (negació, causatiu i reflexius)

Llengües OV amb ordre (negació, causatiu i reflexius)- Verb- Objecte- (adjectius, genitius i relatius)

Una proposta simultània és la Vennemann (1974) [12], en molts aspectes similar a la de Lehman i que com aquesta posa el focus en intentar trobar els patrons subjacents que explicarien les dades de Greenberg. En aquest cas la proposta també passa per ignorar el subjecte i centrar-se a la dicotomia VO/OV. Aquest autor elabora una llista de parelles de categories, al primer element de la qual anomena *operador* i al segon *operant*. I que serien:

-

<u>OPERADORS (Complements)</u>	<u>OPERANTS (Nuclis)</u>
Objecte	Verb
Adverbi	Verb
Verb principal	Verb auxiliar
Adjectiu	Nom
Genitiu	Nom
Oració de relatiu	Nom
Numeral	Nom-Adjectiu
Determinant	Nom
Sintagma nominal	Adposició

A partir d'aquí enuncia el *Principle of Natural Serialization* (PNS): L'ordre dels operadors i dels operants tendeix a ser serialitzat en una direcció única, o els operadors davant dels operants o els operadors darrera dels operants.

Vennemann també troba que només 3 de les 24 combinacions definides per Greenberg complirien el PNS:

1. VSO/PrN/NG/NA
2. SVO/PrN/NG/NA
3. SOV/NPo/GN/AN

Tant el FPP com el PNS són models molt reduccionistes que evidentment no donen compte de bona part de la tipologia que es troba a les mostres. Segons Hawkins [10] només ho fan de la meitat aproximadament de la mostra, i es va fer necessari formular models més complexos.

Hawkins [10] representa una nova visió del tema, després d'ampliar la base de dades de les llengües estudiades a més de 350. Retorna a la variable tripartida Subjecte-Verb-Objecte, però donant prioritat com a millor indicador a les adposicions.

Type Nr	Values of the 4 parameters	Number of languages
1	V-1/Pr/NG/NA	38
2	V-1/Pr/NG/AN	13
3	V-1/Pr/GN/AN	1
4	V-1/Pr/GN/NA	0
5	V-1/Po/NG/NA	0
6	V-1/Po/NG/AN	0
7	V-1/Po/GN/AN	1
8	V-1/Po/GN/NA	0
9	SVO/Pr/NG/NA	56
10	SVO/Pr/NG/AN	17
11	SVO/Pr/GN/AN	7
12	SVO/Pr/GN/NA	4
13	SVO/Po/NG/NA	0
14	SVO/Po/NG/AN	0
15	SVO/Po/GN/AN	12
16	SVO/Po/GN/NA	13
17	SOV/Pr/NG/NA	10
18	SOV/Pr/NG/AN	0
19	SOV/Pr/GN/AN	2
20	SOV/Pr/GN/NA	0
21	SOV/Po/NG/NA	11
22	SOV/Po/NG/AN	0
23	SOV/Po/GN/AN	96
24	SOV/Po/GN/NA	55

**Taula 1.** És la mostra de Hawkins de 350 llengües distribuïdes segons els 24 tipus de Greenberg. Font: [10]

Hawkins va proposar implicacions absolutes (no estadístiques) i unilaterals del tipus:

a)  $Po \rightarrow (AN \rightarrow GN)$     b)  $Pr \rightarrow (NA \rightarrow NG)$

Després ho modula introduint el *Heaviness Serialization Principle* (HSP) segons el qual els modificadors més pesats acostumen a col·locar-se a la dreta dels més lleugers per raó d'un millor processament de la informació ja que un ràpid reconeixement del nucli és necessari per entendre el sentit de l'oració. Així els relatius anirien a la dreta del genitiu i aquest a la dreta de l'adjectiu i aquest de demostratius i numerals.

$Rel < Gen < Adj < \{ Dem, Num \}$

Per últim invoca el *Mobility Principle* (MP) segons el qual els més lleugers (Demonstratius, Adjectius i Numerals) són més mòbils que els més pesats (Relatius i Genitiu) de manera que tenen facilitat per adoptar posicions diverses respecte el substantiu al que determinen.

Tomlin [13] es va proposar explicar la distribució peculiar dels 6 ordres bàsics de la variable Subjecte-Verb-Objecte. Calia una explicació pel fet que els ordres SOV i SVO superen el 40% cadascú, VSO ronda el 10%, mentre que els altres 3 pràcticament no estan representats (taula 2) a les dades de 4 autors diferents.

TABLE 2

Word Order	Ruhlen		Mallinson and Blakc		Hawkins		Tomlin	
	No.	%	No.	%	No.	%	No.	%
SOV	222	51	41	41	174	52	180	44.78
SVO	155	35	35	35	109	32	168	41.79
VSO	47	11	9	9	45	13	37	9.20
VOS	8	2	2	2	8	2	12	2.99
OVS	2	0.5	1	1	—	—	5	1.24
OSV	1	0.25	1	1	—	—	—	—

Taula 2. És mostren les freqüències atribuïdes a cada ordre per diferents autors Font: [13]

Tomlin intenta explicar els percentatges amb 3 principis:

1) *The Theme First Principle (TFP)*: La informació més nova segueix a la ja coneguda. Això es tradueix a la pràctica en què els noms definits precedeixen als indefinits i els pronoms als sintagmes nominals sencers.

2) *The Animated First Principle (AFP)*: A les oracions transitives simples, els sintagmes nominals que fan referència a un ésser animat precedeixen als referits a un d'inanimat segon una doble jerarquia d'animacitat i rols semàntics.

a) Jerarquia d'animacitat: Primer els éssers humans, després la resta d'animals i finalment els inanimats.

b) Jerarquia de rols semàntics: Primers els agents, després els instruments, després els beneficiaris i per últim els pacients.

En cas de contradicció, prima el rol semàntic sobre l'animacitat.

3) *The Verb-Object Bonding Principle (VOB)*: L'objecte està més unit al verb que el subjecte.

WO	TFP	AFP	VOB	Score
SOV	+	+	+	3
SVO	+	+	+	3
VSO	+	+	-	2
VOS	-	-	+	1
OVS	-	-	+	1
OSV	-	-	-	0

Taula 3. Es mostra el grau de compliment dels principis de Tomlin dels ordres de constituents, tret de [13]

A la taula superior es mostra com al provar els 3 principis en els 6 ordres possibles que adopten el Subjecte, l'Objecte i el Verb es pot veure com els ordres més majoritaris són justament els que compleixen un major nombre de principis. De manera més senzilla es pot explicar el mateix si es considera que la posició de l'objecte és indiferent, però que el subjecte tendeix a precedir clarament al verb. D'aquí, fent la taula de probabilitat, si  $Pr(VO)=Pr(OV)$  i  $Pr(SV) \gg \gg Pr(VS)$  s'obtidrien uns resultats similars als de la mostra.

Amb Dryer 1992 [14] s'intenta corregir els problemes de la majoria de propostes efectuades fins el moment. D'una part el fet de treballar amb mostres de llengües massa petites i per l'altre no donar compte de totes les excepcions que en nombre considerable no segueixen els principis enunciats.

La mostra de Dryer és de 625 llengües i 24 variables de les que en selecciona 9 per examinar-ne la distribució dels ordres a 6 àrees geogràfiques (Àfrica, Euràsia, SE Àsia & Oceania, Austràlia-Nova Guinea, Amèrica del nord i de sud). L'autor torna a usar com a variable principal l'ordre del verb i l'objecte i en lloc de principis absoluts es conforma en trobar significació estadística a causa de la gran quantitat d'excepcions.

Cada variable és una parella d'elements gramaticals dels quals es mira si es correlacionen amb l'ordre del Verb i l'Objecte. Als que covarien amb el verb els anomena *verb patterner (o nonphrasals)* i als que ho fan amb l'objecte *object patterne (o phrasals)*.

El que troba és que només 17 de les 24 variables escollides es correlacionen, però n'hi ha cinc que no i dues són dubtoses. Això el porta a rebutjar la hipòtesi que l'ordre del Nucli i el Complement sempre ha de coincidir amb el del Verb i l'Objecte, que ell anomena *Head-Dependent Theory* (HDT) i formular-ne una de nova que anomena *Branching Direction Theory* (BDT). Segons aquesta, en una llengua OV un parell d'elements només seguirà l'ordre XY significativament si X és *nonphrasal* i Y és *phrasal*. En canvi, li permet preveure una dependència entre la posició de l'objecte, l'adposició, el genitiu i el relatiu, respecte el verb.

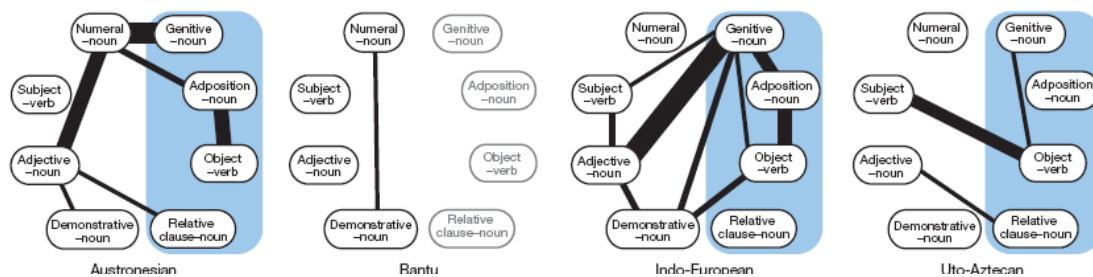
Aquesta formulació permet excloure i justificar les variables que refuten l'HDT. Aquests són l'Adjectiu, el Demonstratiu, l'Intensificador i l'Adjectiu, la Partícula negativa amb el Verb, i per acabar, les Partícules aspectuals amb el verb. En tots cinc casos ambdós element de cada parella serien *nonphrasals* i això justificaria la falta de correlació.

Ehara [15], [16] va provar d'aplicar l'Anàlisi de Components Principals (ACP), una tècnica estadística força habitual en altres disciplines, a les variables d'ordre de constituents. El primer treball del 1995 es limita a una base de dades de 126 llengües i 20 variables sintàctiques, el de 2005 amplia la mostra amb les 625 llengües de Dryer [14] però redueix les variables a set.

Les conclusions del treball és que hi ha dos factors que expliquen prop del 75% de la variació total de les dades. El primer factor acumularia prop d'un 50% i coincidiria amb l'ordre entre Verb i Objecte, mentre que el segon, que representaria un 25%, encara que de forma dubtosa, podria estar vinculat a l'adjectiu.

Aquesta metodologia serà reproduïda a l'apartat 3.3 de la tesi, però es treballarà amb cinc bases de dades diferents per poder fer-ne un estudi segons el nombre de variables.

Dunn et al. [17], van esvalotar el panorama de la tipologia al concloure en un estudi que les correlacions entre diferents ordres de constituents no eren universals sinó que eren particulars de cada llinatge. El seu treball es va centrar en quatre grans grups lingüístics (indoeuropeu, bantu, austronèsic i utoasteca) amb una nova metodologia estadística anomenada Mètode de Montecarlo per cadenes de Markov (MCMC) que resulta absolutament inusual entre els lingüistes i que ha fet que no hagi rebut pràcticament cap crítica seriosa tret de Croft et al [18], que qüestiona l'adequació del mètode a la mostra tractada.



**Fig. 1 Dependències en l'ordre de constituents de quatre famílies lingüístiques.** La dependència entre variables es representa pel gruix de la línia. El color blau distingeix les 4 variables que segons Dryer [14] haurien de mostrar dependència. El color blau no apareix a les bantus perquè s'han exclòs de les . Font: [17]

Com es veu a la figura 1, les correlacions entre un grup lingüístic i els altres són completament diferents, la qual cosa refutaria tots els treballs anteriors. Així que no existiria ni una base genètica ni funcional comuna que pugui explicar els patrons trobats a les diferents àrees.

Els autors esmentats prèviament han fet servir metodologies diferents, i alguns les han pogut anar canviant amb el temps, i han arribat a conclusions contradictòries. Malgrat això, tots ells, tret d'Ehara i Dunn, poden ser vistos de forma conjunta com graons d'una mateixa escala. Primer amb Greenberg es va fer un tast amb una mostra petita de llengües i variables lingüístiques i ho va intentar explicar amb una llista molt llarga d'Universals. Posteriorment, la resta d'autors va progressivament millorant la mostra i va reduint el nombre de principis necessaris i el d'excepcions. Així amb Dryer s'arriba a un únic principi que intenta explicar tota la casuística, excepcions incloses.

Però aquest camí arrossega dos problemes des d'un principi. Un és la simplificació que suposa parlar de models rígidament binaris o tripartits quan moltes llengües tenen ordres molt més flexibles i el segon és el tema de les mostres. Efectivament, el fet que les llengües bantus i oceàniques representin un 20% del total de llengües i siguin a la vegada preposicionals i VO, per exemple, ens porta a pensar que si els pobles que les parlaven fa 2000 anys no haguessin protagonitzat fortes expansions territorials en temps històrics, les mostres amb que van treballar els autors esmentats en aquest capítol i les seves conclusions potser no serien les mateixes.

Però segons Evans 2009 [19], davant d'aquesta visió compartida per tants autors dels universals (i/o els principis tipològics) com fenòmens sincrònics que constreixen la manera com les llengües canvien amb el temps, hi ha la postura contrària que els veu com fenòmens diacrònics que a la llarga han donat com a resultat la situació actual.

Un dels precursors d'aquesta via és Givón 71 [20] que enuncia la llavors provocativa frase «La morfologia d'avui és la sintaxi de demà» que a Givón 84 [21] es converteix en un força més aclaridor «Les correlacions entre l'ordre VO amb les preposicions i entre l'ordre OV i les postposicions reflecteix el freqüent origen de les adposicions a partir de verbs amb el seu respectiu objecte. Una conclusió més de Givón és que l'ordre ancestral de les llengües humanes va ser SOV, però això ja és matèria de l'apartat següent.

## 2.1.2 DISPERSIÓ I PROPAGACIÓ DE LES FAMÍLIES LINGÜÍSTIQUES

Tan aviat els colonitzadors europeus van entrar en contacte amb les poblacions indígenes de la resta del món, a partir del segle XV, va resultar evident que hi havia territoris en que llengües força similars s'estenien al llarg de centenars o milers de kilòmetres, com és el cas de les llengües cree, polinèsies o bantus, mentre que en altres zones era possible trobar llengües completament diferents convivint en àrees molt petites, variant pràcticament de vall en vall, com és el cas de Nova Guinea.

Només recentment, sobre tot els darrers 25 anys, diversos autors han proposat models universals per explicar la distribució global, més enllà de reconèixer que hi ha famílies que s'han estès victoriosament sobre d'altres que han quedat arraconades.

Johanna Nichols és una autora que en més d'una dotzena de publicacions, especialment a partir de l'any 1992 [88], ha anat desenvolupant un model per explicar l'expansió i distribució de les famílies lingüístiques més enllà del mètode comparatiu. La seva metodologia es basa en una mostra de 200 *llinatges* lingüístics entre famílies i llengües aïllades, de les que estudia la seva distribució areal i les afinitats estructurals dels trets estudiats. Suposa que l'existència de trets amb una clara estabilitat gramatical, areal o filogenètica pot ajudar a detectar afinitats on el mètode comparatiu falla.

Conclou que es poden distingir dos tipus d'àrees lingüístiques a les que nomena *zones de propagació* i *zones d'acreció*. Aquestes últimes, també dites residuals són zones amb una gran diversitat filogenètica i estructural com el Caucas, l'Himalaia, Etiòpia, Califòrnia, l'Amazònia, Nova Guinea o el Nord d'Austràlia. En canvi les zones de propagació són àrees amb baixa diversitat i on una única llengua o família ocupa una gran extensió de territori fruit d'una recent expansió, com per exemple la major part d'Austràlia, el Nord d'Àfrica, o les praderies dels Estats Units.

Igual que Nichols, R.M.W Dixon [22] en diferents articles i llibres intenta trobar la manera d'esbrinar les relacions entre les famílies lingüístiques i deduir-ne la història de les expansions i diversificacions, més enllà dels 7.000 anys o 8.000 anys [23] que permet el mètode comparatiu. El treball de Dixon, que ha inspirat a altres autors com Belwood, Renfrew o Nettle, consisteix en una adaptació personal a les llengües de la noció d'equilibri puntuat definit per Eldredge & Gould [24] per a la biologia.

Segons l'autor, les llengües passen per llargs períodes d'equilibri en què les llengües de les diferents zones tendeixen a convergir cap a *prototips comuns*. Aquests períodes es veuen interromputs per successos catastròfics que anomena *puntuacions*, durant les quals es produeixen les expansions i ramificacions de les famílies lingüístiques. La validesa dels arbres filogenètics com models per explicar la història de les llengües es desdibuixa durant els equilibris degut als manlleus i la influència mútua.

Un tercer autor és David Nettle. La seva aportació gira sobre el que anomena *risc ecològic* que desenvolupa en diversos treballs a partir de 1996 [25] i que defineix com el conjunt de factors ambientals que actuen sobre obtenció d'aliment. Donada la dificultat per mesurar aquest risc, el que fa és treballar amb el nombre de mesos a l'any en que hi ha prou humitat per plantar i cultivar (a més mesos amb pluja, menys risc ecològic). Després compara aquesta informació amb el nombre de llengües que es parlen a cada estat dels que te dades. A partir d'aquí obté una serie de generalitzacions:

- La distribució de llengües és similar a la d'espècies biològiques: la diversitat lingüística disminueix cap els pols i és major en selves tropicals que en deserts.
- Quan més baix és el risc ecològic (Fig. 2), a cada estat hi ha més llengües i l'àrea de cada llengua és més petita.
- Quan més difícil és un terreny per circular-hi, menors són els grups lingüístics.

L'explicació darrera totes aquestes conclusions és que quan més gran és el risc ecològic, més gran ha de ser la xarxa social necessària per garantir-ne la supervivència. Així, un territori molt fèrtil permet convida-hi molta gent de varies comunitats lingüístiques, mentre que en un desert, difícilment hi haurà aliments per assegurar la supervivència de poblacions de més d'una única xarxa. També postula que les llengües molt grans frenen el canvi lingüístic i canvien més lentament que les llengües petites.

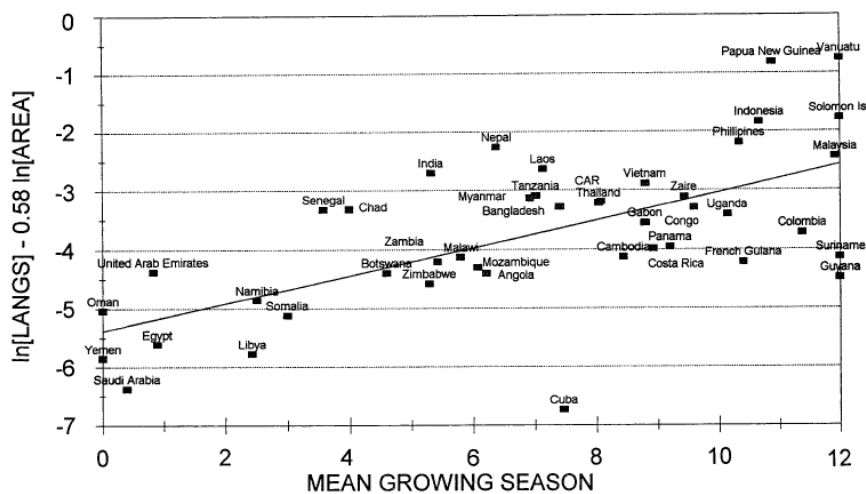


Fig. 2 Correlació entre la densitat de llengües en un àrea i els mesos amb pluja. Font: [2]

Un article de Holman del 2004 [26] intenta explicar perquè les famílies són més grans en unes regions que en altres. Els extrems més evidents serien Sibèria on la família altaica domina amb poques excepcions i Nova Guinea on hi ha centenars de llengües impossibles de reduir a unes poques famílies amb el mètode comparatiu clàssic.

L'autor comença fent un recorregut per les explicacions que troba a la bibliografia per tal de comprovar-les i troba les següents: (i) Polèmica entre *lumpers* i *splitters*. Hi ha dos tendències entre els lingüistes de manera que allà on un veu una sola llengua, altres en poden veure moltes perquè valoren de manera distinta les variacions que veu. (ii) A les famílies europees les cognades s'han estudiat d'acord a lleis de canvi regular i seguint rigorosament el mètode comparatiu mentre que a les d'altres àrees sovint n'hi ha prou una certa semblança fonètica i semàntica per acceptar un origen comú. (iii) Potser les llengües han canviat a ritmes diferents arreu del planeta. (iv). La situació d'Amèrica es podria explicar perquè va ser colonitzada molt recentment.

Després descartar les tres primeres al comprovar que la barrera llengua-dialecte està universalment situada al 70% de cognades malgrat diferències d'autors, afirma haver trobat tres patrons geogràfics segons el continent:

- Amèrica: famílies petites i poc diverses
- Àfrica i Euràsia: famílies grans i molt diverses
- Austràlia i Nova Guinea: famílies petites i molt diverses



The *farming/language dispersal* model és el que ha anat desenvolupant Renfrew al llarg de més de 30 anys, primer en solitari i a partir de principis dels anys noranta amb el suport de Bellwood. [27]. Al 2002 [28] l'explicaven afirmant que «La dispersió de l'agricultura, generalment a través de l'expansió de les poblacions dels agricultors per un procés de colonització o *difusió dèmica* és responsable de la distribució i extensió superficial de moltes de les famílies de llengües del món.

El model va néixer a Renfrew 1973[29], al 1987 [30] s'afegeix la noció d'*ones d'avanç* segons la qual l'agricultura es va propagar en forma d'*ones dèmiques* o sigui causades pel gradual creixement poblacional que porta a la població excedent que no disposa de terres a establir-se una mica més enllà del territori de la generació precedent per cultivar. Al 2000 accepta el concepte de zones de propagació de Nichols i incorpora aspectes de l'equilibri puntuat de Dixon [31] tot fent encaixar l'aparició de l'agricultura com una puntuació especialment important.

Renfrew a més del sistema ja comentat, també contempla altres tipus de fenòmens que explicarien la propagació de les llengües i de les famílies lingüístiques. Per exemple, la migració a territoris encara no colonitzats per l'home, o de zones on un canvi favorable del clima permet introduir-hi l'agricultura o el canvi lingüístic simultani a l'adopció de la ramaderia per part de caçadors-recolectors [32].

	Plus agriculture	Minus agriculture (mostly)
<i>Significantly spread families</i>	Austronesian Bantu (Niger-Congo) Indo-European Semitic Dravidian Sino-Tibetan (Chinese) Tai Chibchan Cariban Tupian Otomanguean Arawakan Cushitic(?) (pastoralists)	Tungusic Uralic (Samoyed) Pama-Nyungan Salishan Uto-Aztecan Eskimo-Aleut Athabaskan Algonquian Siouan Yuman Chon (Tehuelche, Ona)
<i>Relatively non-spread families</i>	Some 25+ Papuan families Nakh-Dagestanian Kartvelian Munda Mixe-Zoquean Mayan Totonacan Xinkan Keresan Tanoan (Kiowa-Tanoan) Panoan Isolates: Zuni, Basque, Huave, Cuitlatec, Tarascan, Chitimacha, Tunica, Natchez, Burushaski Japanese, Korean, Sumerian, Etruscan	Some 25 N. Australian families Wakashan Tsimshian Chumashan Maiduan Pomoan Yukian Wintuan Khoi, San Chinookan Takelman Isolates: Kutenai, Haida, Alsea, Siuslaw, Washo, Yana, Esselen, Beothuk, etc.

**Taula 4.** Principals famílies lingüístiques agrupades segons l'extensió i dependència de l'agricultura. Font:[35]

La taula 4 mostra les principals famílies lingüístiques ordenades d'acord a la dependència de l'agricultura dels seus parlants i nivell de propagació geogràfica.

Per tal de justificar l'existència de llocs amb famílies agrícoles petites al 2000 [31] introdueix el concepte de *zones mosaic* on segons l'autor hi hauria hagut una colonització durant el plistocè tardà i s'han mantingut l'equilibre fins ara, Bellwood [33] afina el model proposant les anomenades *zones de fricció* equivalents a les *zones residuals* que serien zones poc productives com semideserts, boscos de coníferes, estepes o zones costaneres amb alta densitat de caçadors-recolectors.

Una darrera aportació rellevant [34] pot ser la proposta de classificació de les famílies lingüístiques en dos grups: Una Classe A que comprendria les famílies originades en la primera colonització durant el Plistocè, sempre abans del 10.000 AC i una Classe B amb les propagades segons el model agrícola de la *difusió dèmica*.

Campbell [35] realitza una severa crítica de la majoria d'aquests models després d'una exhaustiva exposició dels mateixos. Del model de Nichols, es critica l'ús d'algunes famílies polèmiques, encara no ben establertes, i també la mostra de llengües escollides. Tampoc està d'acord amb la manca de criteris clars per definir els dos tipus de *zones*, ja que molts dels suposats trets estables no ho són en realitat, ni en el fet que el model de Nichols només és descriptiu, però no arriba a mencionar les raons que explicarien la distribució present.

Cambell tampoc està d'acord amb el model de l'equilibri puntuat de Dixon. La principal crítica és l'afirmació que durant els períodes d'equilibri coincideixen amb fenòmens de difusió i convergència, mentre que les puntuacions anirien associades a la divergència i ramificació de les famílies, ja que es poden trobar contraexemples de cada cas. La convergència i la divergència entre famílies es produeixen arreu amb independència que hi hagi o no equilibri.

Les idees de Nettle tampoc convencen a Campbell, entre altres raons perquè troba massa simple la idea que una comunitat econòmica ha de correspondre a una llengua, ja que hi ha altres possibilitats com les llengües franques o el multilingüisme d'una part de la població. També és un problema el fet que en molts casos hi ha convivència entre famílies molt grans amb altres de petites, malgrat viure en zones amb un mateix *risc ecològic*. Per últim, moltes llengües s'han estès per causes diferents a les del model.

Per acabar, sobre el model de Renfrew fa una crítica força coincident amb la que ha fet de la de Dixon. Per una banda hi ha força famílies grans que no són agrícoles i també petites que ho són, tal com es veu a la taula 4 de la pàgina anterior. Però també perquè agricultura i llengua no tenen perquè difondre's de forma conjunta en tots els casos.

Les crítiques de Campbell als models globals estan molt ben documentades i raonades, però no deixen de caure sempre en un mateix error. Que un model tingui excepcions i contraexemples, no vol dir que no pugui ser vàlid allà on funciona. Que un petit poble caçador-recolector adopti l'agricultura o convisqui amb un de molt gran no refuta l'expansió bantu ni fa que el Kalahari sigui una zona de refugi per les llengües khoisanides. La tipologia no és com la física i les excepcions es poden explicar sense necessitat de posar en dubte el model sencer.

### 2.1.3 CORRELACIONS ENTRE FACTORS LINGÜÍSTICS I SOCIALS

El matrimoni Fenk autor de diversos treballs [36], [37], [38] on han trobat correlacions entre variables lingüístiques com el nombre de síl·labes per paraula, paraules per frase, síl·labes per frase, fonemes per síl·laba, l'ús de preposicions i nombre de casos. Encara que el grau de correlació i d'ajustament pot variar força tal com es veu a la figura 3.

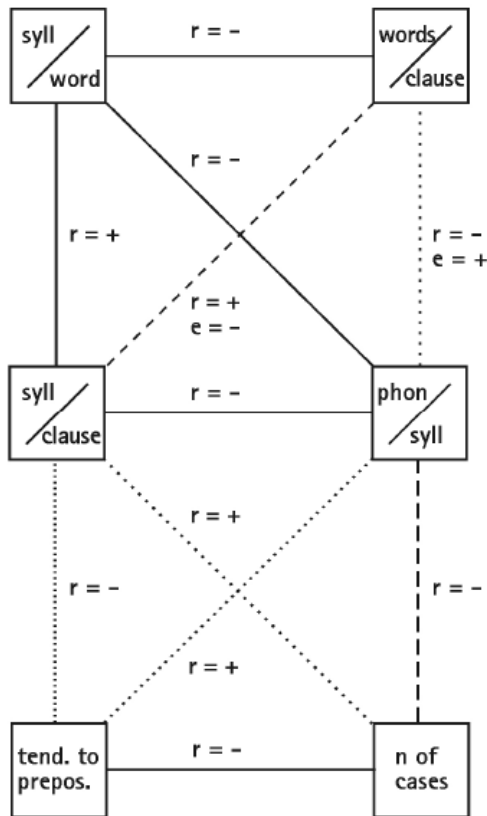


Figure 1: A correlational model connecting metric properties (in the upper part of the figure) with the two non-metric properties “tendency to prepositions” and “number of cases”. Significant correlations: solid lines  
Non-significant coefficients > 0.32: broken lines  
Non-significant coefficients < 0.32: dotted lines  
e = expected sign differing from the sign obtained

**Fig. 3 Model correlacional entre variables lingüístiques.** Les línies contínues són correlacions significatives. Extret de [36]

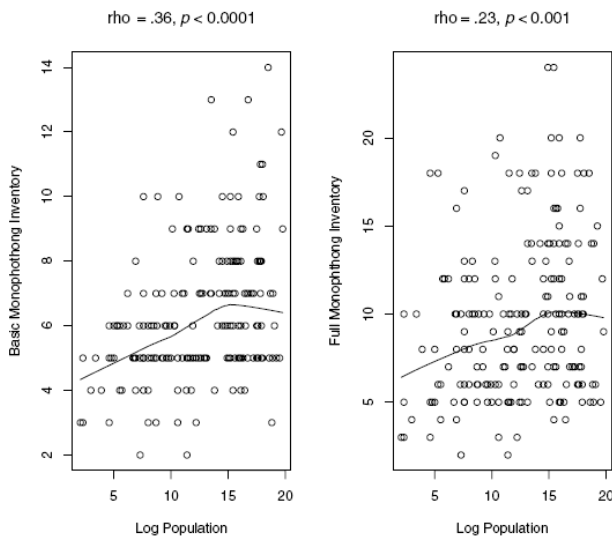
La conclusió a la que arriba és que hi ha una tendència a l'equilibri entre fonologia, morfologia i sintaxi en cada llengua. Les dades les obté dels texts analitzats d'una quarantena de llengües que fa servir per fer els recomptes.

Tal com es veu al gràfic 3, troba una correlació significativa (línies sòlides) negativa entre síl·labes/paraula i paraules/frase, d'una banda i amb fonemes/síl·laba de l'altre. La troba, però, positiva amb el nombre de síl·labes/frase. També la troba significativa i negativa, respectivament, entre síl·labes/frase i fonemes/síl·laba. Per últim entre la tendència a l'ús de preposicions i el nombre de casos la correlació és significativa i negativa. En altres casos troba que la correlació no és significativa, però per poc (línies discontinues) entre altres variables, però amb resultats més confusos.

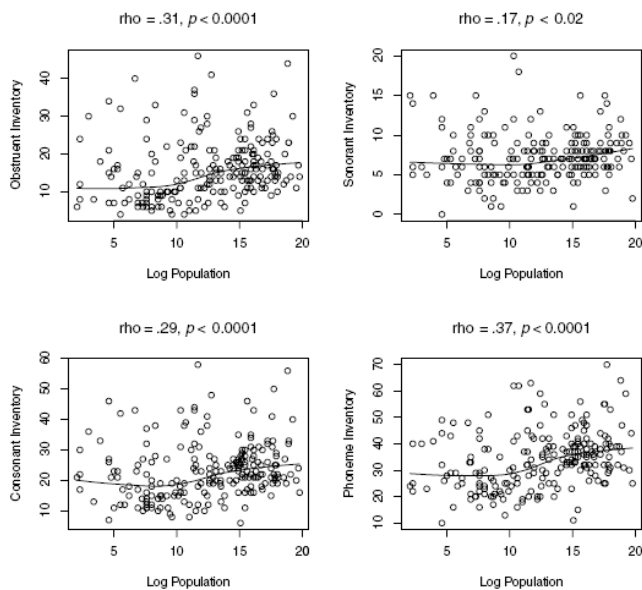
En resum, les llengües amb potsposicions tendeixen a tenir molts casos i fer servir paraules llargues amb moltes síl·labes senzilles. En canvi, les llengües preposicionals tendrien a no tenir casos i fer servir paraules curtes amb unes poques síl·labes, però més complexes. Als primers treballs també hi barreja la tradicional divisió en llengües flexives, aglutinants i aïllades, però als treballs posteriors abandona aquesta opció.

Al treball *Phoneme inventory size and population size* del 2007, J.Hay i L.Bauer [39] van trobar correlacions entre el nombre de fonemes i el nombre de parlants de les llengües, tant vocàlics com consonàntics. A l'article s'aprofundeix fins el punt de distingir entre timbres de monoftongs i del nombre de monoftongs totals per les vocals i entre sonorants i obstruents per les consonants. També s'analitza per llengües i per famílies.

Excepte per les sonorants, on els resultats són poc clars, a la resta de casos troba correlacions significatives en diferent grau. La millor correlació la troba amb el nombre de timbres i consonants a nivell de família. Curiosament no troba correlació entre vocals i consonants, malgrat que ambdues variables correlacionen amb el nombre de parlants i també ho fa així el nombre de fonemes totals. A sota es reproduïxen els gràfics que correlacionen població amb el factors esmentats.



**Fig. 4** Correlació entre nombre de monoftongs i població, per llengua, a la dreta el nombre total i a l'esquerra el nombre de timbres, extret de [39].



**Fig. 5** Correlació entre fonemes i població, per llengua. De dalt a baix, i de dreta a esquerra, obstruents, sonorants, consonant i total. Font: [39]

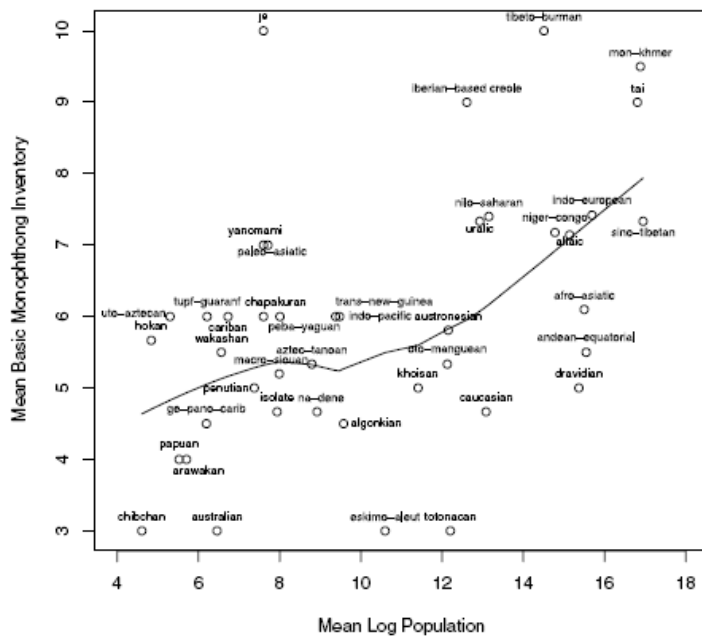


Fig. 6 Correlació entre nombre de monoftongs i població, per famílies. Font [39]

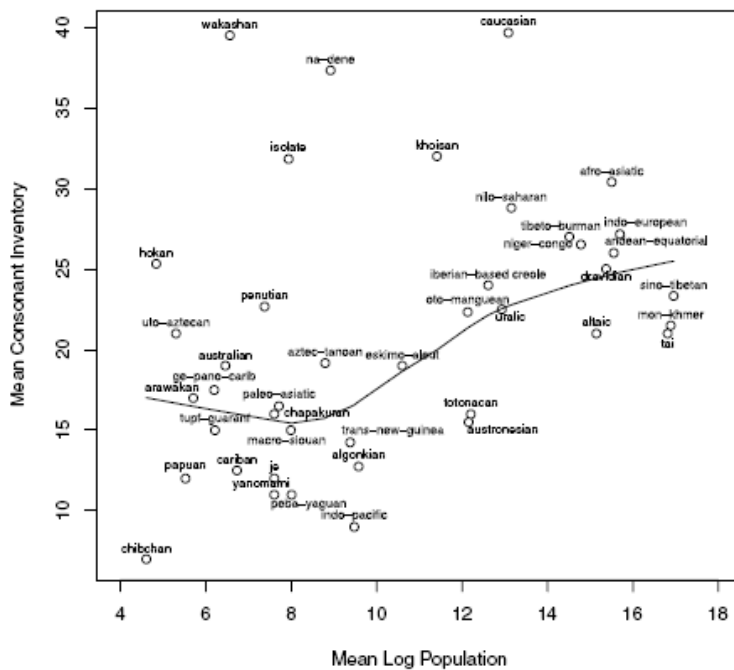


Fig. 7 Correlació entre nombre consonants i població, per famílies..Font [39]

Tots els gràfics d'aquesta pàgina i de l'anterior mostren com a mesura que augmenta la població, s'incrementa la complexitat de l'inventari fonològic. El resultat és consistent i es repeteix tant si es treballa amb llengües (figures 4 i 5) com amb famílies lingüístiques (figures 6 i 7) en que s'ha treballat amb les mitjanes de les llengües que en formen part.

Al 2010, Gary Lupyan i Rick Dale [40] van publicar un article on afirmen haver trobat correlacions fortament significatives entre factors relacionats amb la complexitat morfològica i altres demogràfics i sociohistòrics com el nombre de parlants, l'àrea on es parla i el nombre de llengües veïnes. Aquests tres factors (parlants, àrea i veïns), que potser no són més que reflexos d'un únic factor lligat al *cosmopolitisme* de la llengua, el que es podria definir com *promiscuïtat lingüística*, comporten una morfologia més simple que es compensa amb solucions lèxiques analògiques, com es mostra al gràfic següent. A més nombre de parlants, menys complexitat morfològica.

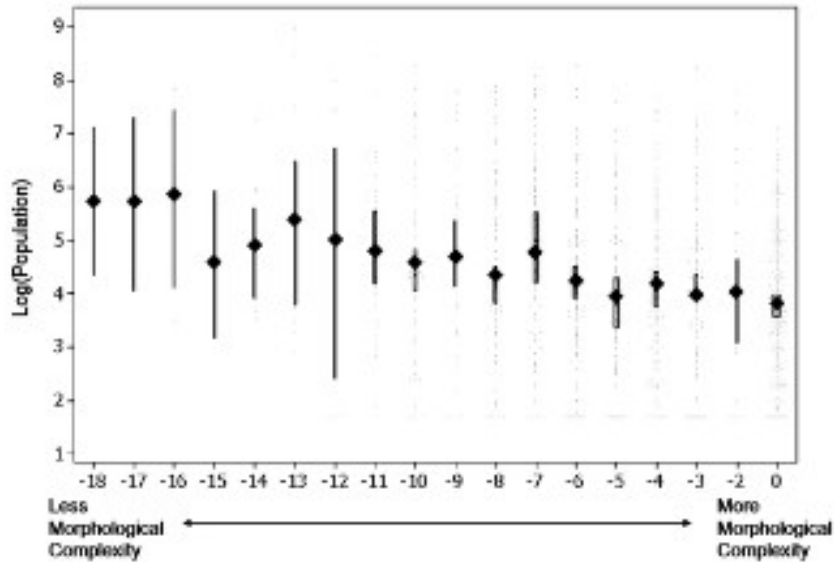
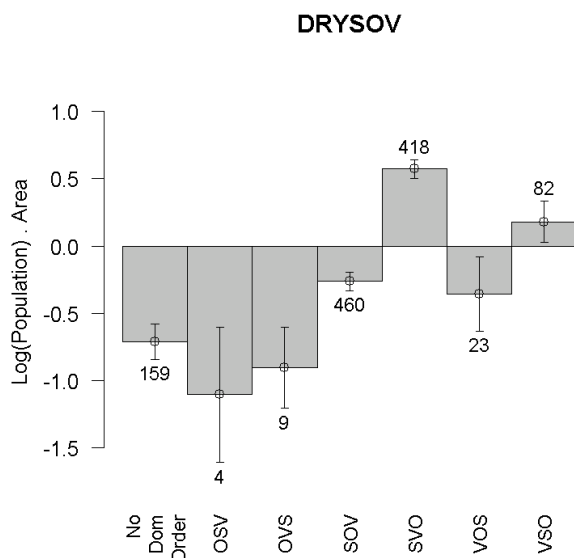


Fig. 8 Gràfic que mostra la correlació negativa entre població i complexitat morfològica. Font:[40]

També proposen la *Linguistic Niche Hypothesis* (LNH) segons la qual les estructures lingüístiques estan sotmeses a pressions diferents segons els ambients socials, igual que passa amb els organismes vius als nínxols ecològics, i d'aquesta manera la llengua s'adapta modificant la morfologia. El treball és basa en una mostra de 3000 llengües i 28

variables. No totes es veuen afectades d'igual manera per la suposada simplificació.



Defensa que quan els adults aprenen una llengua, els trets que els resulten menys senzills d'adquirir passaran més difícilment a la descendència. En canvi, a les llengües usades en petites comunitats s'incrementa la redundància que pot facilitar l'aprenentatge pels part del nens.

També es va mirar l'ordre de Subjecte-Verb-Objecte i troba que VSO es correlacionen positivament amb el nombre de parlants i la resta d'ordres negativament tal com mostra la Fig 9.

Fig. 9 Població i ordre de constituents de Subjecte, Objecte i Verb. Font [40]

Al 2011, Atkinsón [41] va proposar, fent un paral·lel amb la diversitat genètica, que la diversitat de fonemes no només correlaciona amb el nombre de parlants (gràfics 10 i 11 per llengües i famílies) sinó amb la proximitat a l'Àfrica.

L'article comença defensant una correlació entre el nombre de parlants d'una llengua i el que anomena diversitat fonològica total (figura 10). Aquesta mesura no coincideix amb el total de fonemes ja que es dona igual pes al nombre de consonants, de vocals i de tons.

Els resultats es repeteixen si es fa servir la família com unitat, figura 11.

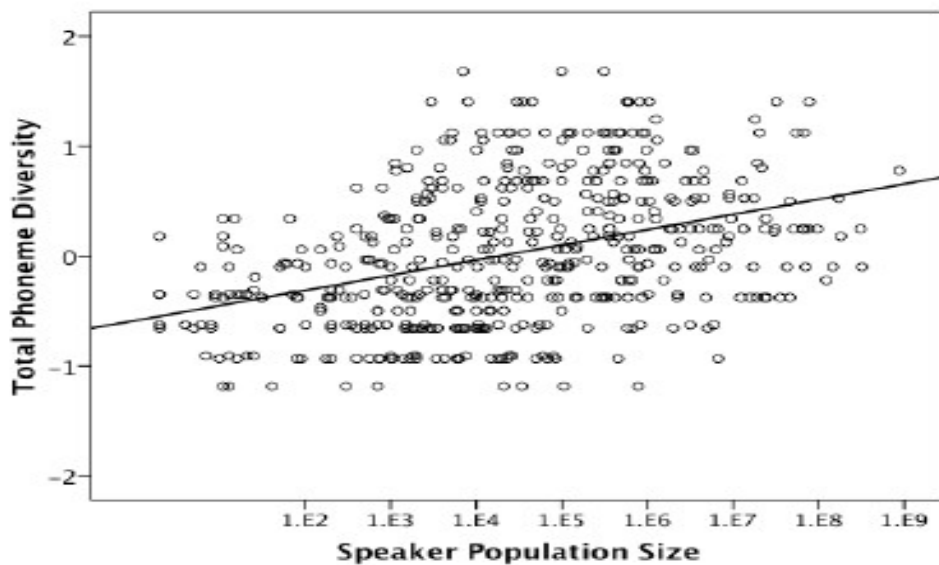


Fig. 10 Correlació entre diversitat fonèmica i població, per llengua. Font [41]

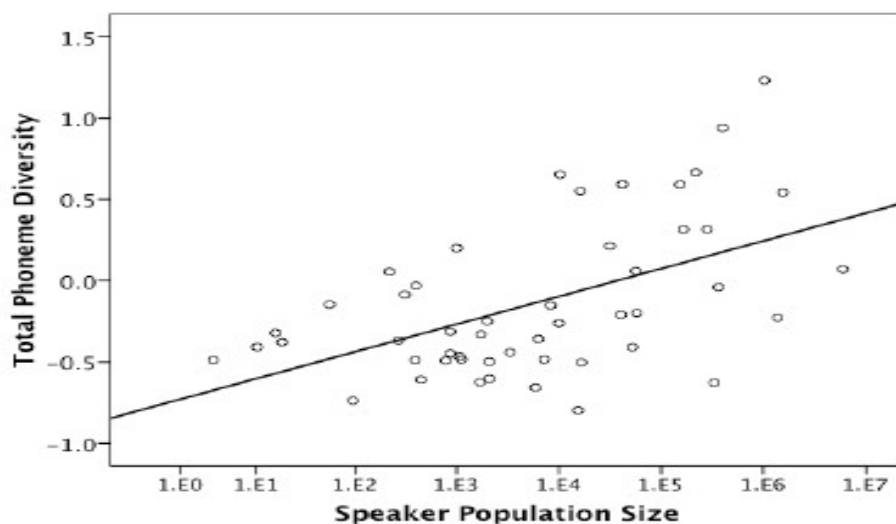
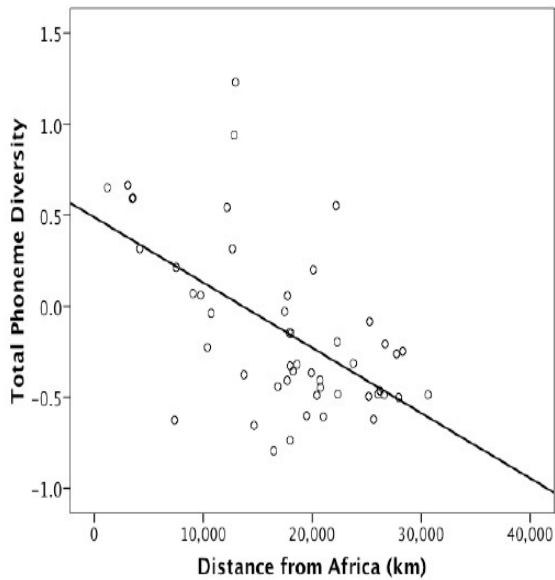
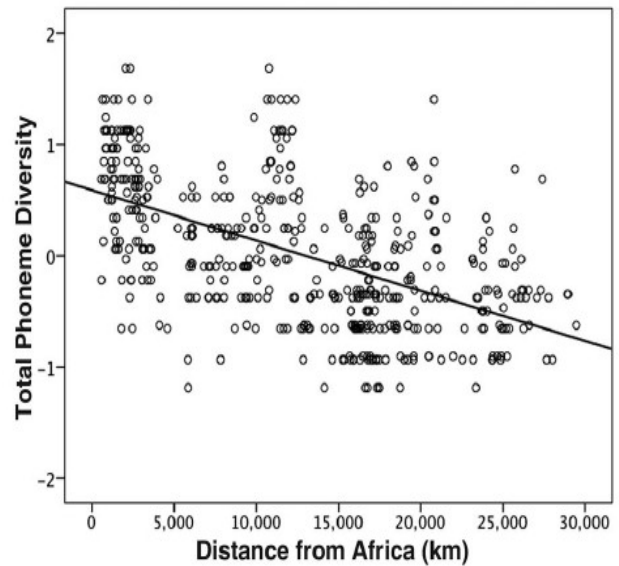


Fig. 11 Correlació entre diversitat fonèmica i població, per família lingüística. Font [41]

Segons l'autor, també a mesura que les poblacions humanes es van anar expandint més i més lluny, la diversitat de fonemes de les llengües i famílies (gràfics 13 i 12) també s'anava reduint més i més. La revelació era força sorprenent i s'ha intentat comprovar en aquesta tesi amb els resultats que es mostren a l'apartat 5.7.3 i que semblen corroborar el fet. La causa estaria en un suposat efecte fundador. Quan un grup marxa a colonitzar noves terres només s'enduria una part de la diversitat fonètica de l'àrea de partida. El mapa 2 en dona una idea del fenomen.



**Fig. 12** Relació entre diversitat fonèmica i distància Àfrica, per família lingüística, Font: [41]



**Fig. 13** Relació entre diversitat fonèmica i distància a Àfrica, per llengua, Font: [41]



**Mapa 2** Diversitat fonèmica segons la distància a Àfrica. Font: [41]

El mapa 2 confirma un màxim de diversitat fonològica total entre les llengües de l'Àfrica més sud-occidental, coincidint aproximadament amb les llengües khoisanides i un mínim entre llengües americanes o les polinèsies.

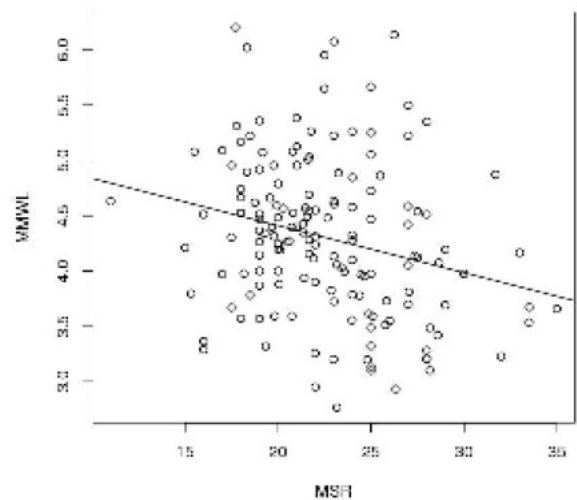
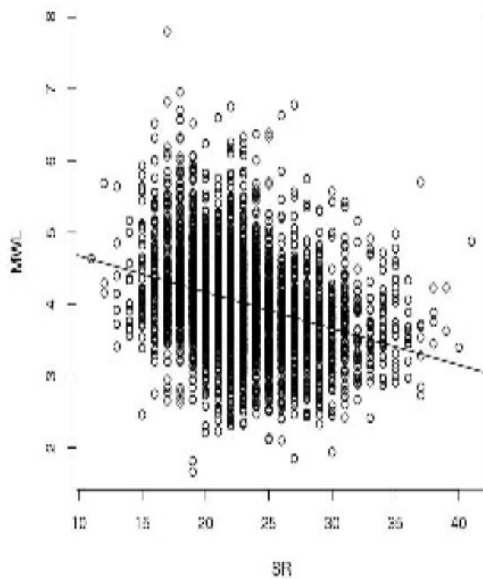


Søren Wichmann, al capdavant d'un grup d'investigadors vinculats al projecte AJSP (Automated Similarity Judgment Program ) ha tirat endavant el projecte de fer una macrobase de dades de vocabularis de llengües, descrita a l'apartat 3.1.2, que li ha permès obtenir resultats que han permès elaborar una classificació del conjunt de llengües del món i diversos articles sobre els resultats obtinguts a nivell tipològic, cronològic i filogenètic. En alguns d'ells es proposa data [42] i lloc d'aparició més probable de cada família [43] i un munt de correlacions interessants.

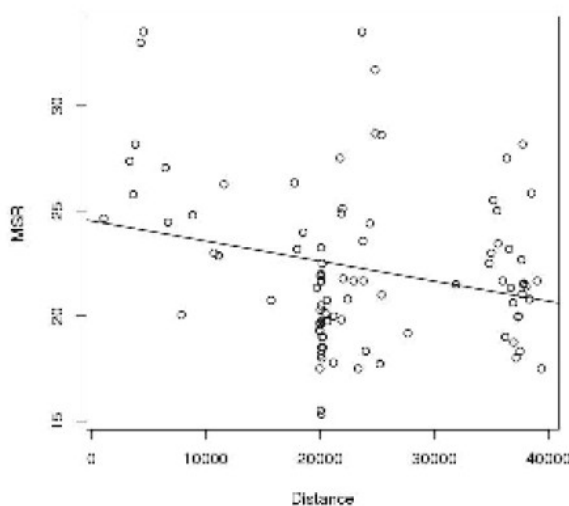
A Wichman 2011[44] es correlacionen la longitud mitjana de les paraules (MWL), el nombre de fonemes (SR) i el nombre de parlants de més de 3000 llengües.

Els resultats (fig 14) mostren una correlació lineal de la longitud mitjana de paraules) amb els segments representats (SR) a l'AJSP. Amb  $r = 0.31$  . Entre famílies (fig 15), la mitjana de cada família (MMWL) correlaciona linealment amb la mitjana dels segments representats (MSR) a l'AJSP. Amb  $r = 0.23$ . Els coeficients de correlació són baixos i l'autor ho justifica per l'ús d'una transcripció fonètica , perquè les famílies no són independents i per la influència entre llengües veïnes.

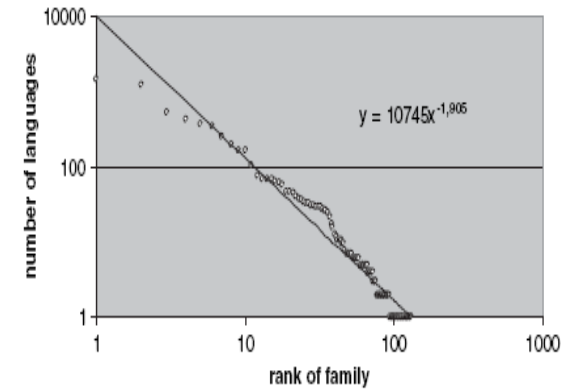
El fet de treballar amb segments representats i no fonemes és perquè la base de dades es fa en una transcripció fonètica simplificada en que un signe pot tenir diversos fonemes.



**Fig. 14** Relació entre MWL i SR, per llengua, Font: [44] **Fig. 15** Relació entre MWL i SR, per família, Font: [44]



**Fig. 16** Mitjana de fonemes/família i distància a Àfrica  
Font [44]



**Figure 2**  
Language family sizes in *Ethnologue* (Grimes 2000) plotted on a log-log scale  
( $R^2=0.957$ )

**Fig 17.** Nombre de llengües per rang de famílies. Font [45]

També troba una correlació igualment débil, figura 16, però que qualifica com a significativa entre fonemes i distància a l'Àfrica, corroborant les dades d'Atkinson amb una metodologia completament diferent.

També [45] troba que el nombre de llengües per família a la classificació de l'*Ethnologue* segueix una llei potencial amb un coeficient d'ajustament del 95%, figura 17. Això implicaria que totes les famílies s'han ramificat sense diferències qualitatives. Al repetir l'experiment pel nombre de parlants, proposa que hi ha diferències qualitatives que podrien explicar-se per la relació amb l'agricultura.

Amb Holman [46] mira les diferents famílies lingüístiques i es troba amb un ventall que va des de les llengües aïllades fins a les grans famílies amb dotzenes o centenars de llengües. Aquest segon tipus invariablement presenta una filogènia complexa amb arbres de diversos nivells. Al mateix temps són també les famílies amb més parlants i la seva expansió coincideix amb sistemes de producció agrícoles. Una altra diferència entre famílies grans i petites és la correlació entre herència genètica i lingüística. En un extrem tindriem les famílies petites (sobretot d'Amèrica, Nova Guinea, Austràlia o l'oest asiàtic) de caçadors recol·lectors on la correlació és pràcticament total i de l'altre les grans famílies d'Europa, pròxim orient, l'Índia on la correlació sembla no existir.

Wichman [47] defineix el caràcter *densitat* com el nombre de llengües de la família dividit per l'edat estimada de la mateixa i calibrada com un valor entre 0 i 100. Després d'estudiar 36 famílies, arriba a la conclusió que sempre que el valor *densitat* es superior a 48, els pobles parlants de llengües de la família són principalment agricultors (12 casos). A les quatre famílies amb valors són inferiors a 8, es desconeix l'agricultura. Les 20 famílies amb valors entre 8 i 48 es reparteixen a parts iguals. Les 12 famílies agrícoles amb més densitat representen un 70% del total de llengües i un 90% dels parlants.

Yamamoto és un autor que ha treballat la distribució areal de l'ordre de constituents Objecte-Verb, tot comparant-lo amb d'altres per valorar la consistència i tractant d'esbrinar quin ha estat l'ordre ancestral per cada família. Va publicar al 1999 un primer article [48] limitats a la distribució i consistència i un segon [49] al 2005 en què hi afegia el factor històric. L'autor repassa, àrea per àrea, el que diu la bibliografia especialitzada sobre el tema i comparant amb les dades d'unes 3000 llengües.

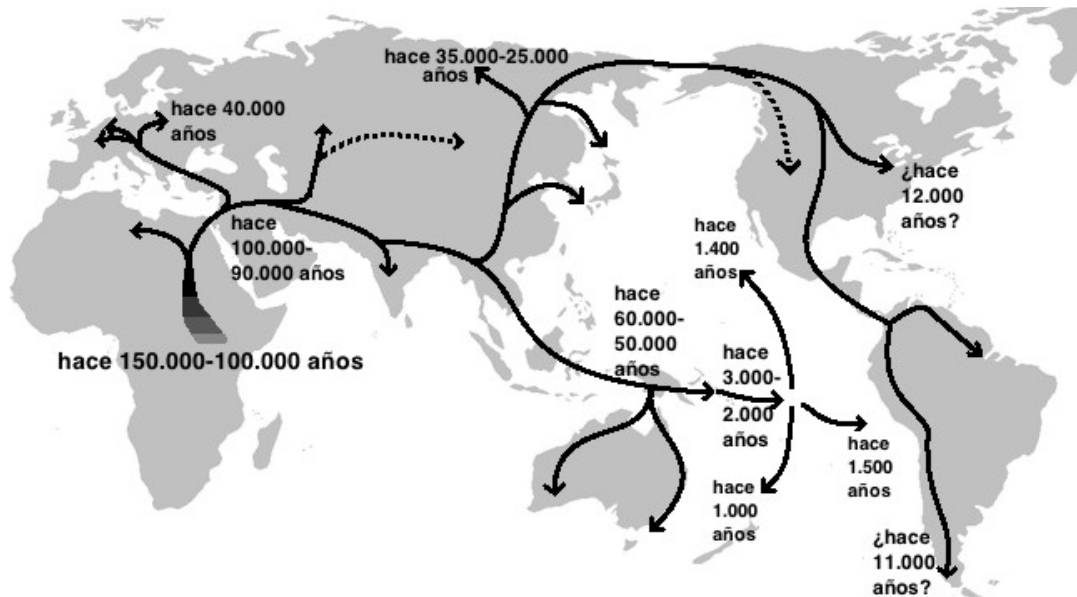
La seva conclusió general és que els ordres consistentment VO i els OV, es reparteixen ara a parts iguals la superfície terrestre, però al reexaminar-ho des d'un punt de vista diacrònic, la majoria d'àrees han tingut llengües amb l'ordre (S)OV en temps pretèrits i les de tipus VO tenien una extensió restringida abans d'expandir-se durant el Neolític.

Per últim, Merritt Ruhlen és l'autor d'una de les primeres classificacions globals de totes les llengües del món [63] i que a la darrera versió inclou un arbre únic amb arrel que correspondria a una hipotètica primera llengua humana (proto-world). Aquesta propietat, ha fet que sigui la preferida en aquesta tesi, malgrat haver-ni altres més modernes i possiblement acurades.

A Gell-Mann i Ruhlen 2011 [23], anomenat *The origin and evolution of word order*, l'autor, després d'analitzar 2135 llengües en la seva classificació arriba a les següents conclusions: (i) Si ha existit un protollengua comuna a la resta, tenia com ordre SOV. Aquí coincideix amb Givón [50] (ii) Excepte en casos de difusió, la direcció del canvi sempre és SOV > SVO > VSO/VOS, amb reversió ocasional a SVO, però mai a SOV. (iii) La difusió, encara que important, no és un procés dominant a l'evolució de l'ordre de constituents. (iv) Els ocasionals casos de OVS i OSV s'han originat a partir de SOV.

## 2.2 LES SOCIETAT HUMANES

### 2.2.1 DISPERSIÓ D'HOMO SAPIENS



El coneixement de l'expansió de l'home anatòmicament modern s'està aclarint de forma accelerada. Abans es basava només en proves arqueològiques i la comparació entre poblacions actuals, però als darrers 10 anys, les tècniques genètiques s'han sofisticat fins l'extrem de permetre analitzar l'ADN de fòssils de més de 100.000 anys, detectar el encreuament amb humans d'espècies properes (Neanderthals i Denisovans) o poder estimar l'antiguitat de l'anomenada Eva mitocondrial i el seu corresponent Adam.

Ara sabem que els *Homo sapiens* més antics coneguts daten de finals del plistocè (200.000 anys) a l'Àfrica [51]. Les dades genètiques i arqueològiques apunten que els humans que poblen la resta de continents deriven d'una única migració (o de dues segons alguns autors) que ara fa uns 60.000 anys va creuar l'estret d'Ogaden on una branca va seguir la línia de la costa per arribar fins a l'Índia, sud-est asiàtic, Nova Guinea i Austràlia. Els vedda de Sri Lanka, els andamanesos, els aeta filipins, els papús i els aborígens australians serien descendents moderns d'aquesta primer avançada.

Una segona branca va ocupar Europa, Pròxim Orient, Àsia central i després l'Índia i SE asiàtic (desplaçant o barrejant-se amb els humans de la primera branca) i les illes del Pacífic. També es va poblar Amèrica a partir de les poblacions siberianes, possiblement en 3 onades diferents. El dibuix d'aquesta expansió podria explicar en bona part la distribució de les principals famílies lingüístiques modernes.

Per contra a Africa, amb 200.000 anys més d'història, la diversitat genètica és molt més alta (fins a 10 vegades), presenta famílies lingüístiques força recents, de les poblacions més antigues, les pigmees han perdut la llengua original i només les khoisanides i algunes famílies menors resten per fer-se una idea del que va passar durant més de 200.000 anys. Els khoisanids semblen haver estat l'arrel que ha originat de la resta. Quasi un 99% de les llengües africanes pertanyen a només tres famílies molt esteses, però molt recents en comparació a la sortida d'Àfrica que va originar les de la resta del món. I d'aquestes, l'afroasiàtica provendria de poblacions asiàtiques reentrants.

## 2.2.2. DEMOGRAFIA HISTÒRICA

En aquesta tesi s'ha fet servir sovint el nombre de parlants com a variable, però el fet que el valor utilitzat s'ha tret d'una estimació del 2007 [52], molts segles després del moment de l'aparició de les llengües, i de mil·lennis en el cas del *genus*, obliga a tenir en compte algunes dades de demografia històrica.

Un exemple, potser exagerat, és el cas del Japó on hi vivien unes 100.000 persones al 400 A.C que s'han convertit en més de cent milions a hores d'ara.

Cal tenir en compte, però, que les afirmacions sobre aquest camp no poden deixar de ser especulacions i que poden haver discrepàncies entre diferents autors.

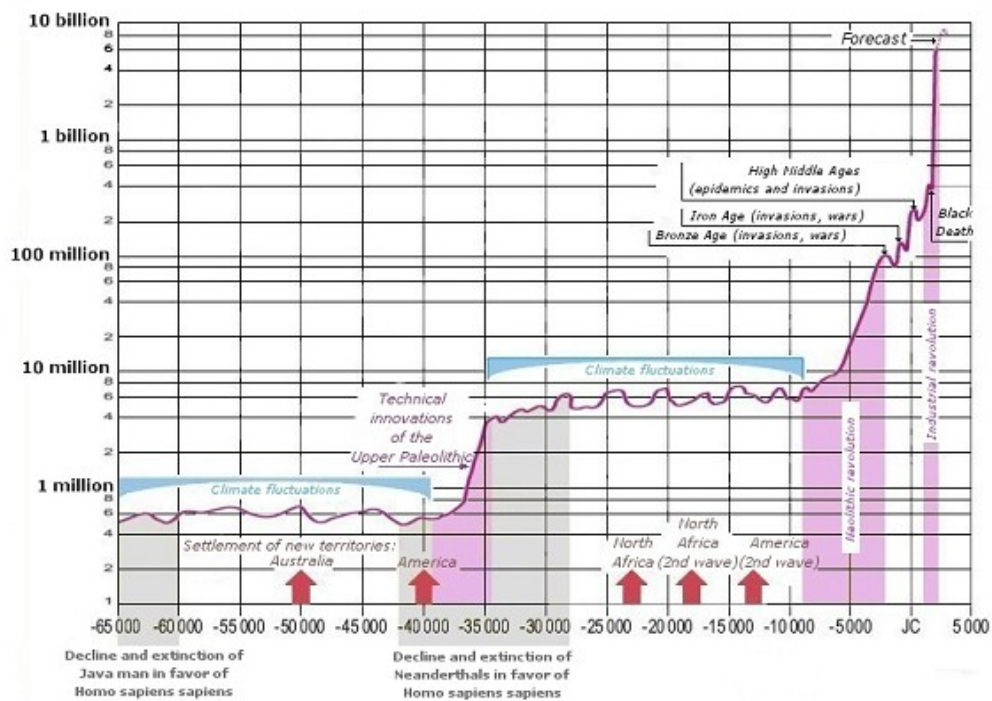


Fig. 18. Creixement de la població humana Font: [56]

Régions/Dates	-400	J.-C.	500	1000	1300	1400	1500	1700	1800	1900	2000
Chine (avec la Corée)	19	70	32	56	83	70	84	150	330	415	1273
Inde (avec le Pakistan et le Bangladesh)	30	46	33	40	100	74	95	175	190	290	1320
Sud-Ouest asiatique	42	47	45	33	21	19	23	30	28	38	259
Japon	0,1	0,3	2	7	7	8	8	28	30	44	126
Reste de l'Asie	3	5	8	19	29	29	33	53	68	115	653
Europe (avec la Russie)	32	43	41	43	86	65	84	125	195	422	782
Afrique du Nord	10	13	12	10	9	8	8	9	9	23	143
Reste de l'Afrique	7	12	20	30	60	60	78	97	92	95	657
Amérique du Nord	1	2	2	2	3	3	3	2	5	90	307
Amérique centrale et du Sud	7	10	13	16	29	36	39	10	19	75	512
Océanie	1	1	1	1	2	2	3	3	2	6	30
<b>Total mondial</b>	<b>152</b>	<b>250</b>	<b>205</b>	<b>257</b>	<b>429</b>	<b>374</b>	<b>458</b>	<b>682</b>	<b>968</b>	<b>1613</b>	<b>6062</b>

Taula 5 Creixement de la població per zones Font: [57]

La famosa Eva mitocondrial està datada fa més de 100.000 [53] anys a l'Àfrica i fins fa un 60.000 anys sembla que la població humana no va superar la xifra del milió ni tampoc havia sortit d'aquest continent. Entre fa uns 60.000 i 40.000 anys, coincidint amb la desaparició de l'*Homo erectus* es va estendre arreu del món fins arribar a Austràlia i Amèrica, però sense un augment significatiu de població.

Aquí s'acceptarà el model que sosté [54] que el conjunt de poblacions no africanes són descendents d'un únic grup d'humans, unes 1000-1500 persones com a màxim segons recents dades genètiques, que ara fa uns 50-60.000 anys van sortir d'Àfrica i van anar estenent-se i diversificant-se fins arribar a colonitzar la resta del món amb excepció de l'Antàrtida. Però la teoria, [55] que defensa que en realitat hi hauria hagut dues sortides d'Àfrica separades per uns 20.000 anys, no tindria tampoc cap efecte contradictori ni amb la metodologia ni amb les conclusions de la tesi.

Entre fa uns 40.000 i 30.000 anys [56], coincidint amb el declivi de l'home de Neanderthal, van aparèixer les innovacions tècniques del Paleolític superior i la població es va quadruplicar fins els 4 milions d'humans arreu del món i ja no va parar d'augmentar fins la revolució neolítica, començada fa un 10.000 anys que es comentarà a l'apartat 2.2.4.

Al neolític va començar un creixement exponencial que la va portar als 100 milions al 2000A.C i els 250 milions a l'any 0, xifra que es va mantenir fins el segle X (figura 18).

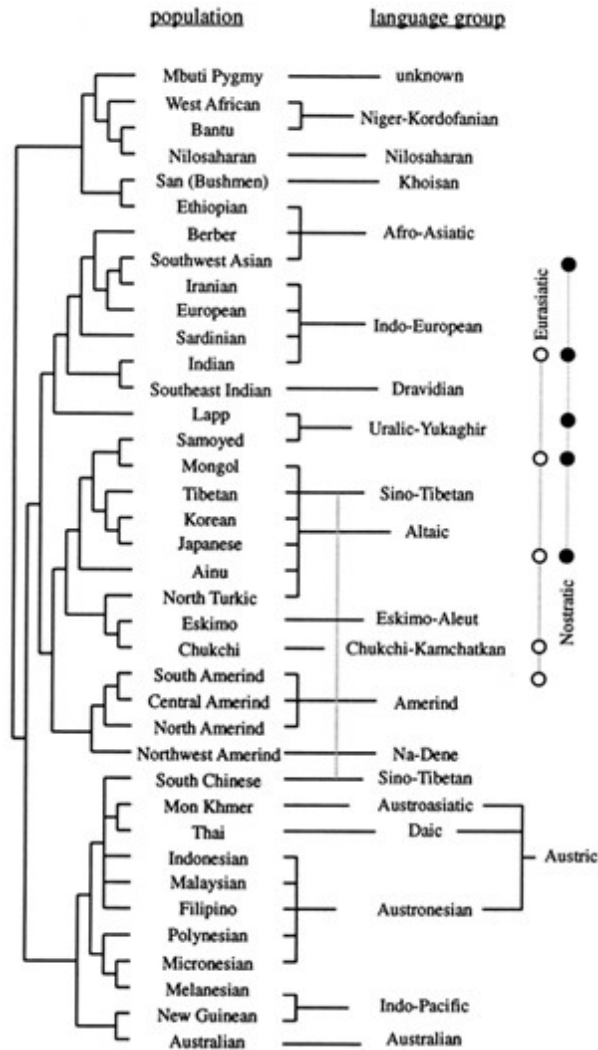
A l'any 0 la població estava distribuïda de manera molt irregular (taula 5) de manera que dels 250 milions de parlants del món, dos de cada tres vivien de l'arròs a la Xina, l'Índia i el sud-est asiàtic, Europa en tenia uns 43 i Àfrica 25. La resta del món se'n repartia 18 més. A l'Amèrica del nord, exclosa Mesoamèrica, només n'hi vivia un [57].

### 2.2.3 GENS I LLENGÜES

La idea d'associar pobles, nacions o "races" amb llengües és força intuïtiva i ha estat present al llarg de la història. La pràctica totalitat d'humans hem rebut tant l'herència genètica com la llengua dels pares i també les passem als fills. Això hauria de comportar l'existència d'una forta correlació sempre que no s'hagi vist alterada. De fet, la llista de causes de possibles excepcions és tan nombrosa que només té sentit mencionar les més freqüents: adulteris, adopcions, distinta llengua entre pare i mare, integració de la població immigrant, minorització, l'esclavatge, conquestes militars, etc.

Si adoptem una visió diacrònica, la majoria de famílies lingüístiques van originar-se dins de societats paleolítiques on cap dels factors esmentats eren rellevants. Fins el neolític, les poblacions humanes vivien en grups de fins unes 300 persones (numero de Fibonaci) dedicades a la caça i recol·lecció, en moviment constant. Dins cadascun d'aquests petits grups es donava tan l'aprenentatge i l'ús de la llengua com la reproducció. Quan un grup creixia fins més enllà del límit permès pels recursos, el grup s'escindia en dos de diferents que segueixen el seu propi camí.

El resultat, encara que amb excepcions, era una relació bijectiva entre llengua i gens per cada grA ls anys 80's, Cavalli-Sforza, [58] va ser pioner en reprendre el tema, però amb una base científica, deixant clar que hi ha un paral·lelisme evident entre les ramificacions de les poblacions humanes i les famílies lingüístiques.(Fig. 19).



**Fig. 19** Comparació entre poblacions genètiques i allunyades, demostrant que hi va haver contactes.

Només era una aproximació feta amb els medis de la època i va fer servir com a mesura “l'afinitat genètica” com un tot de la població. D'aquesta manera, es barrejaven de forma confusa les diferents herències genètiques que havien participat en cada població concreta, sense permetre reconstruir el pes de cada component i la història que havia portat a la situació actual.

La situació ha canviat totalment a partir de la generalització de l'ús de les tècniques basades en haplogrups. Aquestes tècniques han aportat dos tipus d'anàlisi impensables abans:

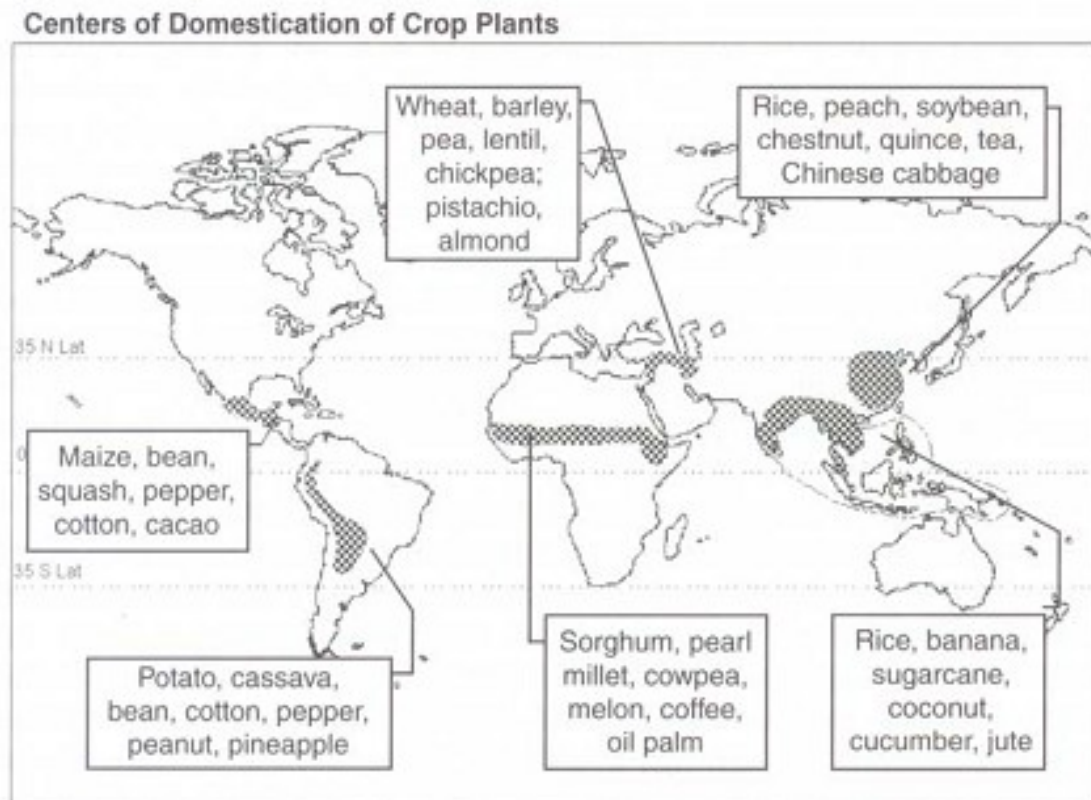
- 1) El que permet separar l'herència masculina de la femenina, no sempre coincidents i clau per reconèixer situacions excepcionals de barreja, com el cas d'Islàndia, on els homes son d'ascendència escandinava, però les dones són en bona part d'origen britànic.
- 2) Poder confirmar al 100% el parentiu entre individus concrets de comunitats avui

L'esquema de Cavalli-Sforza, malgrat els enormes avenços que s'han produït en aquest camp, continua com un clàssic sense refutar i és el que s'ha inclòs en aquesta pàgina.

## 2.2.4 NEOLÍTIC I DOMESTICACIÓ DE PLANTES

Possiblement el canvi més dràstic al mode de vida dels humans ha estat la revolució neolítica que va permetre a petites tribus itinerants de caçadors-recolectors convertir-se en pagesos sedentaris. L'acumulació de menjar, especialment en forma de farina, va permetre multiplicar la població i l'aparició de civilitzacions de creixent complexitat.

No hi ha cap dubte que el primer nucli neolític va ser el pròxim orient (8500 A.C) i que posteriorment el fenomen es va repetir de forma independent en altres zones del planeta, com la Xina i Mesoamerica (mapa 4).



**Mapa 4** Zones inicials de l'agricultura Font: [59]

El que ja resulta més complicat és saber el nombre exacte de nuclis independents on es va donar el fenomen, ja que la proximitat d'alguns d'ells no permet discernir la possible influència mútua, d'uns sobre els altres. Tampoc ha de ser és el mateix l'ordre d'aparició a cada focus neolític de l'agricultura, la ramaderia, la ceràmica, la metal·lúrgia o del nuclis urbans.

El Pròxim Orient és, com ja s'ha comentat, la zona amb una cronologia més antiga (8.500 A.C) i des d'on es va estendre pel Mediterrani, Europa i fins l'Índia. Les plantes estrella són el blat i l'ordi entre els animals la vaca, la cabra, l'ovella i el porc [59].

A l'Àfrica la situació és més confusa. Apart d'Egipte i altres parts properes a Àsia i Europa, on hi ha extensions del neolític del Pròxim Orient, hi ha 3 zones a tenir en compte: Sàhara, Nigèria i Etiòpia-sud del Sudan. La cultura sahariana s'originà cap el 8000 A.C quan el clima era més favorable i amb la progressiva desertificació els seus



protagonistes van haver de refugiar-se cap el nord o el sud. A Nigèria és domestica el nyam, planta que malgrat no ser un cereal, sembla haver sigut rellevant en l'expansió bantu a través de la selva fins i la sabana [60]. La zona d'Etiòpia-Sudan és la més dubtosa com a nucli independent ja que va rebre influència Sahariana i asiàtica.

A la Xina sembla que cap el 5770 A.C va començar el neolític, encara que es discuteix si en un focus o en dos independents [61], D'aquí el cultiu de l'arròs, va estendre's a la resta del sud-est asiàtic en una primera etapa, i posteriorment des de Taiwan per tota Oceania, tret d' Austràlia.

A Amèrica trobem dos nuclis principals, Mesoamèrica al nord i els Andes al sud. Els primers amb el triplet format per la mongeta, la carbassa i el blat de moro, i el segon amb la patata. La cronologia es força més recent pel que fa la domesticació de plantes, cap el 3000 A.C. A part hi ha dues zones secundàries més: l'est dels EUA (amb el tabac i gira-sol primer i el blat de moro arribat posteriorment desde Mesoamerica cap el 900 D.C) i l'Amazònia. Aquesta darrera es tractarà a l'apartat 7.5. A Nova Guinea hi ha una horticultura molt antiga (8000 A.C) entre els pobles papus [62].

### 2.2.5 NEOLÍTIC I I FAMILIES LINGÜÍSTIQUES

Renfrew, [29]], va ser el primer en proposar de forma seriosa que la distribució de les principals famílies lingüístiques va lligada en bona part a la difusió de l'agricultura. Encara que és una idea controvertida, el fet és que en molts casos sembla evident que la elevada demografia aconseguida amb el cultiu de cereals ha sigut un factor determinant.



Mapa 5 Famílies lingüístiques, segons Ruhlen

La genètica ha seguit a la paleoantropologia en assenyalar Àfrica com el lloc on té origen l'espècie humana. i on ha estat restringida durant almenys els primers 200.000 primers anys de la seva existència. Només entre els 70.000 AC i els 100,000 AC va començar l'expansió a la resta del món.

El mapa 5 mostra la distribució de famílies de Ruhlen [63] encara que avui està en part desfasada, permet una visió tradicional des de la que començar l'anàlisi.

Malgrat l'antiguitat de les poblacions africanes i la seva enorme complexitat genètica, el mapa de distribució de les famílies lingüístiques mostra un dibuix sorprenentment senzill i força més recent del que es podria pensar. Un dibuix també lligat com pocs continents a factors geogràfics i climàtics. Amb l'excepció algunes llengües no classificades o potser aïllades, el 99% de llengües africanes pertany a 4 úniques famílies.

El grup khoisanid, avui limitat, amb alguna excepció, al desert del Kalahari, és el més antic i són representants del que va ser la humanitat durant aquests primers 200.000 anys. Els seus parlants eren tots caçadors-recolectors, al menys abans de l'arribada dels bantus que els van anar recloent en aquest desert.

Dos grups més, Niger-Congo i Nilo-Sahariana, com el primer component del nom de cadascun ja permet entreveure, podrien haver-se originat en poblacions que viuen a les ribes dels rius Níger i Nil respectivament. La genètica ens diu que les poblacions que les parlen semblen haver-se originat només fa uns 50.000 anys, encara que les dues famílies semblen molt més recents, potser un pocs milers d'anys. Ambdós famílies estan conformades per pobles agrícoles i ramaders.

Per últim, hi ha la família afroasiàtica que sembla originada per poblacions no africanes que van retornar al continent i es van expandir per les muntanyes d'Etiòpia i les zones del desert del Sàhara i el Sahel. També són agrícoles i ramaders.

Euràsia a nivell lingüístic mostra regions molt diferenciades. D'Europa a l'Índia hi ha una família, la Indoeuropea, que a penes deixa excepcions al seu domini, com el basc, les llengües caucàsiques o les dravídiques als extrems. Sobre les raons de l'expansió d'aquesta família s'ha discutit molt, però ara els dubtes estan entre la hipòtesi que la lliga a l'agricultura [30] i els que defensen que seria la tecnologia relativa al cavall i al carro el que explicaria la conquesta a partir del segon mil·lenni en diferents onades [64]. Aquests pobles són també agricultors i ramaders.

L'àrea que correspon aproximadament amb l'antiga URSS se la reparteixen les famílies Uràlica, Altaica i Chukotko-Kamchatkan, d'oest a est. Els seus parlants ocupen zones poc apropiades per l'agricultura i bàsicament eren ramaders, amb densitats de població molt baixes. Potser emparentades amb les anteriors, hi quatre famílies més, ara mateix llengües aïllades, la Ainu, la Gilyak la Japonesa i la coreana. Les dues primeres de caçadors-recolectors i les dues darreres de pobles que viuen del cultiu de l'arròs.

Justament el cultiu de les dues subespècies d'arròs ha estat clau en la configuració del sud-est asiàtic a nivell lingüístic. A l'apartat 7.2 s'aprofundeix en les implicacions lingüístiques d'aquest cultiu. Domesticat a la Xina, els pobles que el cultivaven s'han anat desplaçant cap el sud, substituint a poblacions ancestrals. Bàsicament es trobaria la família Sino-Tibetana i el qüestionat tàxon Àustriac. Aquest en realitat és la suma de les

famílies Austroasiàtica, que ocupa l'antiga Indoxina i arriba a l'Índia, Tai des de Tailàndia fins el sud de la Xina i Austronèsica que és la família que ocupa un àrea més gran del planeta, per haver colonitzat tota Oceania menys Austràlia i la major part de Nova Guinea. També en forma part la família Miao-Yao que malgrat ser molt petita, té gran importància perquè es candidata a ser la primera en cultivar arròs [65].

La illa-continent d'Austràlia va ser poblada en una primera i única onada ara fa més de 50.000 anys i les llengües que hi ha pertanyen a més de 30 famílies diferents. Malgrat les seves diferències, els seus orígens han fet que sovint s'agrupin totes com una macrofamília, però mai s'ha pogut demostrar cap mena de parentesc entre elles. El fet més destacable és que una d'elles, la Pama-Nyunga, ocupa de forma sorprenent un 90-95% de la superfície, sense que hi hagi una explicació convincent d'aquesta expansió. Això ha portat a autors com Dixon [66] a negar l'existència mateixa de la família. Tots els pobles australians eren caçadors-recolectors.

A Nova Guinea es dona una situació similar a Austràlia, tret de l'arribada molt recent de llengües austronèsiques, el gruix de famílies lingüístiques són parlades per poblacions anomenades papús que van arribar fa uns 50.000 anys i per això s'ha proposat que totes haurien d'estar emparentades ni que sigui remotament, però sense cap confirmació. Ruhlen les va agrupar amb les llengües de les illes Andaman i les extintes de Tasmània en una controvertida família indopacífica. A Nova Guinea hi conviuen caçadors recol·lectors amb horticultors. Una família en concret, la Trans-NovaGuinea ocupa la major part del nord de l'illa gràcies a les tècniques agrícoles [67]

Austràlia i Nova Guinea representen un repte lingüístic pel fet que les dades genètiques deixen clar que van se poblades fa uns 50.000 anys i les famílies lingüístiques són prou diferents. Si afegim el fet que els australians només van poder arribar via nova Guinea, queda per resoldre el perquè no hi ha papús a Austràlia ni australians a Nova Guinea. Darrerament, però, s'ha proposat que els parlants llengües Sepik de Nova Guinea podrien ser romanents del pas dels australians barrejats després amb papús [68].

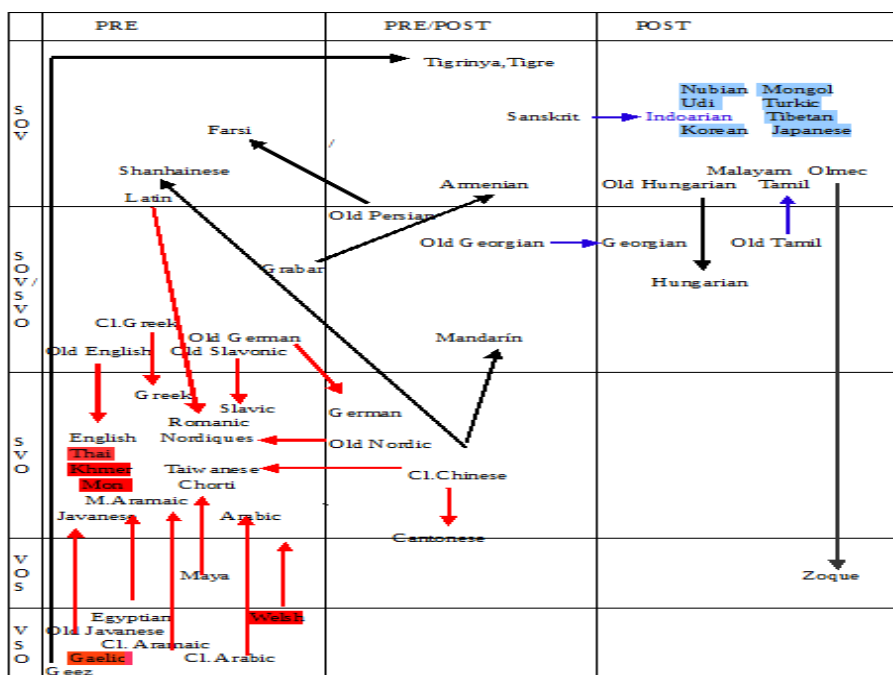
Després de Greenberg [69]), les llengües d'Amèrica s'acostumen a dividir en tres stocks que se suposa procedeixen de 3 onades arribades en diferents moments. L'esquimo-Aleutiana és la més recent i està formada per caçadors i pescadors circumpolars. Les altres serien la Na-Dene de l'Oest del Canadà i la que agruparia la resta, anomenada Ameríndia. Com en el cas d'Austràlia, la unitat d'aquest darrer grup no és més que una quimera, reforçada per les dades genètiques que mostren un possible origen comú a partir de la entrada d'un únic grup des de l'estret de Bering ara farà uns 15.000 o 20.000 anys [70]. No hi ha pobles ramaders, però sí una important agricultura a zones com els Andes, Mesoamèrica, l'Amazònia i l'est dels EUA.

## 2.2.6 LENGÜES DE LLARGA TRADICIÓ LITERÀRIA

L'escriptura ha aparegut de manera recurrent a les primeres societats urbanes com Mesopotàmia, Egipte, Mesoamèrica o la Xina. La societat mesopotàmica va ser la primera en escriure. Primer en sumeri cap el 3350 A.C que malgrat desaparèixer com a llengua parlada en uns 1500 anys, encara va sobreviure dos mil anys més com llengua literària i religiosa al costa de l'accadi de la família afroasiàtica que va substituir-lo com a llengua oral des del 2000 A.C, però que ja s'escrivia al 2800 A.C. L'egipci (2700 A.C-1000 D.C?), també afroasiàtic com l'accadi, és possiblement la llengua amb un recorregut documentat més llarg com a llengua parlada. De la mateixa família són les diferents llengües aramees, l'àrab, l'hebreu i el guez, totes també amb un llarg registre.

La família indoeuropea es l'altre gran font de coneixement de llengües antigues. Quasi totes les seves branques presenten llengües modernes descendents d'altres arcaïques amb una llarga literatura: cèltiques com el gaèlic, gal·lès i bretó, les germàniques com l'alemany o l'anglès o les nòrdiques. També l'armeni i sobre tot el grec. Dins el grup indoari de les indoeuropees tenim les emparentades amb el sànscrit, llengua sagrada de l'Índia, el persa i el iagnobi. També son indoeuropees el llatí i totes les seves descendents. Per últim l'antic búlgar malgrat no ser realment ascendent de les esclaves s'acostuma a usar com a tal degut al seu alt grau de conservadorisme De la resta de famílies tenim la uràlica hongaresa, altaïques com el mongol, el turc o potser també el japonès i el coreà. A les muntanyes del Caucas es troben el georgià i antic albanès. A l'Índia, les llengües dràvides es poden comparar amb l'antic tàmil donat que totes són força properes.

L'altre gran zona on abunda aquestes llengües es el SE Asiàtic i la Xina, amb el xinès, el tibetà, el thai, el monk o el khmer. A Indonèsia tenim el javanès. Ja fora d'Eurasia la llista acabaria amb el nubi a l'Àfrica i l'antic maia a Mesoamèrica.



Gràfic 2 Canvis documentats Font:elaboració pròpia

A la pàgina anterior hi ha el gràfic 3 que mostra canvis documentats a 7 variables per una mostra de 60 llengües modernes, principalment d'Euràsia, durant la seva història. La font és el treball de DEA que ha donat lloc a aquesta tesi. S'ha diferenciat entre canvis forts, de l'ordre primari, i canvis lleus d'opcions secundaries. El gràfic 22 mostra la variació d'ordre Subjecte-Verb-Objecte i adposicions. Les conclusions són:

a) Es pot veure un comportament diferent entre les llengües preposicionals i les postposicionals. Les primeres han patit molts més canvis d'ordre a tots els constituents. A les antigues hi ha molts ordres del subjecte, verb i objecte, però amb el temps han passat a SVO. Les llengües postposicionals són més estables i coherents.

b) Els diferents constituents presenten taxes de canvi molt diferents. S'han comptat els canvis de cada constituents i resulta evident que no canvien a igual ritme.

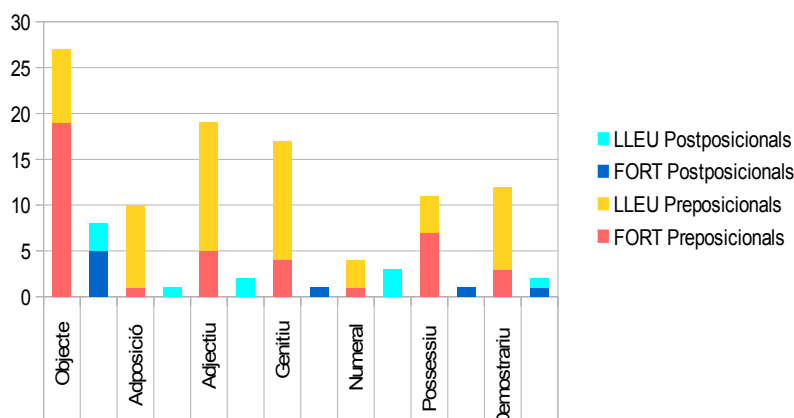
(S-O-V > Adjectiu, Genitiu > Demonstratiu, Possessiu > Adposició > Numeral)

El verb ha canviat de posició a la majoria de llengües, adjectius i genitius són els següents més freqüents, seguit de demostratius i possessius. Les preposicions són molt estables i els numerals constitueixen un fenomen apart. Només les llengües mon-khmer, thai, nuvi i tibetanes mostren preferència per la posició del numeral radera el substantiu.

c) Reducció d'ambivalència. A les llengües antigues de la mostra es troben molts més casos d'ordres mixtes que a les modernes. Això es particularment clar en el cas del verb, però pot ser efecte de l'elevat nombre de llengües indoeuropees.

d) Increment de coherència. La majoria de canvis semblen sotmesos a fenòmens d'analogia i canvien cap el sentit seguir de les adposicions.

Canvi d'ordre de constituents



Gràfic 3 Diagrama de les dades de la Taula 6

	Objecte	Adposició	Adjectiu	Genitiu	Numeral	Possessiu	Demonstratiu	
FORT Preposicionals	19	1	5	4	1	7	3	40
LLEU Preposicionals	8	9	14	13	3	4	9	60
FORT Postposicionals	5	0	0	1		1	1	8
LLEU Postposicionals	3	1	2	0	3	0	1	10
Total	35	11	21	18	7	12	14	

taula 6 Canvis documentats a set variables en una mostra de 60 llengües Font: elaboració personal.

## 2.3 CANVI LINGÜÍSTIC I FACTORS SOCIALS

Tradicionalment existeix una oposició de la majoria de lingüistes a acceptar explicacions extralingüístiques per fenòmens lingüístics. S'acostuma a buscar-ne les raons en el contacte amb altres llengües, especialment el substrat, o en qüestionades tendències a simplificar o acostar-se formes menys marcades.

Però hi ha alguns punts on la influència de factors socials extralingüístics és tan evident que s'ha anat obrint camí. McWorther [71] reconeix tres situacions en que la societat influeix a la gramàtica: el vocabulari particular, els usos particulars ( formes de respecte o rituals) i la simplificació gramatical a les llengües parlades massivament per població no nativa. Segons aquest autor, la complexitat seria la norma i l'excepció estaria en la dificultat dels nous parlants que aprenen la llengua passada l'adolescència.

Aquí cal ajuntar el debat amb un de paral·lel sobre si existeixen llengües més simples que d'altres o el nivell de complexitat és similar. Sense entrar-hi, el que ningú dubta és de la diferència de complexitat a nivell dels diferents nivells fonològic, morfològic i sintàctic. El dubte estaria en si es compensen de forma que la mitjana és constant o no.

El punt en comú entre el debat de la influència social i el de la diferència de complexitat està que les situacions proposades per unes són les mateixes en ambdós debats: pidgins, criolls, llengües en procés d'obsolescència i llengües franques. Tots aquests casos s'han proposat com exemple de simplificació, totes tenen a veure amb una part dels parlants no han après la llengua per transmissió normals dels seus pares i en tots ells sembla clar que els canvis socials hi tenen alguna cosa a veure. Als següents apartats es tractaran de manera més detallada.

### 2.3.1 PÍDGINS

Un *pidgin* o *jargon* és una forma simplificada de llengua que fan servir dues o més comunitats que no disposen de cap llengua en comú. Generalment s'utilitzen només en situacions concretes, especialment els negocis i intercanvis. No s'acostumen a fer servir a la vida diària ni tenen parlants nadius.

Bona part dels pídgins descrits han sorgit de l'encontre entre colonitzadors europeus i poblacions natives d'Àfrica i Amèrica i Oceania. En general la llengua base és la europea, però no sempre com en el fanagalo de Sud-àfrica de base zulu o el chinook jargon de la Colúmbia Britànica als seus inicis en que aparentment era una llengua de comunicació entre indígenes. .

La majoria de pídgins comparteixen trets lingüístics comuns que coincideixen en que comporten més simplicitat respecte a les llengües en que es basen. Segons McWhorter [72], serien l'absència de frases compostes, l'estructura sil·làbica bàsica Consonant-Vocal, en general només 4 o 5 vocals bàsiques i absència de variació morfofonemàtica i reduplicació per marcar els plurals ni tenen tons com a mínim en els monosíl·labs.

L'ordre sintàctic dels pídgins no mostra una clara preferència per l'ordre SVO, al contrari que als criolls, malgrat que acostumen a no tenir marques gramaticals per diferenciar l'objecte i el subjecte [73].

### 2.3.2 CRIOLLS

La idea bàsica de crioll és similar a la de pidgin, però amb un ús generalitzat a tots els aspectes de la vida, inclosa la transmissió als fills. Això implica que tot crioll té parlants nadius i que necessita estructures lingüístiques per permetre la comunicació que un pidgin pot estalviar-se. Com en el cas del pídgin, dels que molts potser en són descendents, la majoria es troben en zones de colonització o d'antiga societat esclavista on amos i esclaus parlaven llengües diferent. El vocabulari acostuma a basar-se a partir de la llengua colonial i la gramàtica manté molts trets de la llengua indígena, però al contrari dels pídgin, l'ordre pràcticament universal és SVO amb independència de l'ordre de les llengües inicials.

Encara que la posició tradicional és la que fa derivar cada crioll d'un pidgin anterior, hi ha autors], que sostenen que pídgin i criolls són fenòmens diferents sense cap relació [74], ja que en el cas dels criolls la llengua dominant és normalment compresa i utilitzada pels esclaus en la comunicació amb els amo després de la primera generació.

Igual que en el cas del pidgin, entre els criolls hi ha una gramàtica més simple que a la llengües de les que procedeixen. Tampoc tenen inflexió morfològica desenvolupada, ni tons en monosíl·labs i es formen paraules opaques semànticament. Tenen estructures sil·làbiques simples i la longitud de paraula és menor. La morfosintaxi és analògica.

En tot crioll s'ha donat el cas d'una primera generació en que els pares tenen una llengua materna diferent del crioll en que parlen als fills, possiblement parlants passius de la llengua dels pares. I també adults que decideixen aprendre la llengua amb totes les dificultats que això comporta. Així hi ha una doble interrupció en la transmissió de la llengua, la dels nens que no l'aprenen de pares nadius i dels adults que no la van aprendre dels pares.

### 2.3.3 LLENGÜES EN VIES D'OBSOLESCÈNCIA

Les llengües naturals disposen d'un munt de trets i matisos subtils que només els parlants fluents són capaços de fer servir amb naturalitat. A mesura que l'ús de la llengua va decaient es van perdent, tant en la pròpia comunitat nativa com en el cas d'emigrants que han marxat fora. El cas més greu seria el dels anomenats semiparlants, que mai han arribat a dominar-la, però són capaços de fer-se entendre.

Donat que la pèrdua d'ús d'una llengua comporta l'augment del d'una diferent, és natural que molts dels canvis que comporta l'obsolescència consisteixin en perdre trets que no es troben a la nova llengua dominant a l'hora que se'n manlleva alguns particulars d'aquesta. El que resulta rellevant és que també s'han detectat canvis recurrents, amb independència de les llengües amb que es tracti, que poden considerar-se predictibles. El vocabulari sofreix canvis similars als dels criolls.

A nivell fonològic, Andersen [75] defensa la disminució del nombre de distincions fonològiques, però amb persistència d'aquelles més funcionals. Campbell [76] proposa la tendència a reforçar les formes menys marcades o una generalització en l'ús d'aquestes.

A nivell morfològic Campbell [77] es veu una reducció morfològica (pèrdua d'al·lomorfs i un anivellament de paradigmes). A nivell sintàctic es donen canvis d'ordre sintàctic i la tendència de la llengua és a ser més analítica.

### 2.3.4 LENGÜES FRANQUES

No sempre que entren en contacte comunitats de diferents llengües es formen pídgin, en molts casos, senzillament una de les llengües acaba imposant-se com a *interlingua* pels parlants de la resta. Malgrat que en aquests casos és freqüent el nom de llengua franca, la verdadera llengua franca del Mediterrani a l'edat mitjana sembla que va ser en realitat un pidgin.

Alguns autors com Kusters [78] han defensat que el fet que un idioma esdevingui llengua franca comporta una simplificació, especialment a nivell morfològic. Ha comparat diferents varietats d'àrab, quítxua, suahili i llengües escandinaves per concloure que es dona una simplificació de la morfologia verbal a mesura que augmenta la complexitat de la societat. De manera que les varietats parlades per societats petites, tradicionals i amb majoria de parlants L1 han mantingut la complexitat original, mentre que aquesta ha disminuït paral·lelament a l'increment de població, l'apertura a l'exterior i l'existència d'un nombre rellevant de L2.

Segons les seves conclusions, per tal que es doni una simplificació de la morfologia verbal és necessari que la futura llengua franca tingui una expansió ràpida, un manteniment durant varies generacions i una certa desconexió entre unes varietats i altres per evitar l'anivellament.

### 2.3.5 COMPLEXITAT

El tema de la complexitat és un dels més polèmics de la lingüística. Va ser força popular al segle XIX al ser contemporània del primer darwinisme i del darwinisme social, que associaven complexitat a superioritat. Per això posteriorment, especialment després de Chomsky [79], es veu com un tema desfasat. Al final plana la idea que parlar una llengua menys complexa implica una ment més simple i comporta una ofensa als seus parlants que seria equivalent al mesurament de cranis de l'antropologia del XIX. En part és conseqüència de l'anomenada hipòtesi Whorf [80] que considera que la llengua és un reflex del pensament i obre la porta a pensar que una llengua simple suposa una ment simple.

El principal argument contra la diferència de complexitats és la creença que existeix un comportament balancejat de manera que allò que una llengua té de més en fonologia, ho tindrà de menys en morfologia o en sintaxis, i viceversa.

Encara que amb diferents noms, la majoria d'autors que han tractat el tema acostumen a reconèixer tres principis. El primer seria el d'economia, fer servir el mínim de marques morfològiques o categories. El segon seria el de transparència, és a dir, que hi hagi una relació un a un, entre forma i semàntica. El darrer és el d'somorfia que suposa que hi ha ordenacions més naturals que d'altres. La major part del canvi lingüístic es pot veure com la tendència a aplicar algun d'aquests principis allà on no s'acompleix.

Resulta evident que la manera d'aprendre una llengua dels que ho fan com a nens fins l'adolescència (L1) és força diferent dels que ho han hagut de fer com adults (L2). Els nens sempre aprenen la llengua fins l'estadi final, mentre que els adults no sempre, a l'hora que mostren grans dificultats per aprendre llistes de morfemes, conjugacions i declinacions complicades o fer servir tons i fonemes complexos no apresos de nens.



Això comporta que molts autors hagin coincidit en descriure les comunitats de parlants com a situades entre dos pols o extrems, en funció de la relació L1/L2. Kusters [81] proposa diferenciar com extrems del gradient entre les comunitats lingüístiques:

a) Comunitats Tipus 1. Son comunitats petites, on la majoria de gent es coneix. La gent només parla normalment amb gent de la pròpia comunitat i cal canviar de llengua per parlar amb comunitats veïnes. Normalment no reben visitants de fora i el cicle de la vida és predictable i poques les innovacions . La llengua s'usa per a les relacions socials. Pràcticament tots els parlants són L1 i tots els membres de la família la parlen a casa.

b) Comunitats Tipus 2. La llengua és l'únic element en comú perquè els membres no necessàriament viuen en un mateix espai ni comparteixen valors ni mode de vida. El nombre de parlants L2 és significatiu perquè adults de fora es mostren interessats en aprendre-la. Molts parlants només usen la llengua per negocis i intercanvis.

Evidentment, els dos tipus de comunitats lingüístiques són ideals i n'hi ha moltes que es situen entre mig i amb el pas del temps moltes poden passar d'una a l'altre. Curiosament, aquests dos tipus han estat proposats amb diferents noms per un munt de lingüistes diferents. Trudgill [82], [83] [84], que diferencia entre comunitats molt i poc cohesionades, Andersen [85] que parla de comunitats centrals i perifèriques, amb--- actituds exocèntriques i endocèntriques, Thurston [86],[87]que parla de comunitats esotèriques i exotèriques, Nichols [88] amb les comunitats d'àrees d'acreció i residuals i autors com Barth [89], Giles [90], LeVine & Campbell [91], Wallman [92] i Urciuoli [93] que han diferenciat entre comunitats lingüístiques compactes i expansives.

Darrera tot el que han escrit aquest autors, hi ha la convicció que les llengües es comporten diferent i no representen el mateix pels parlants d'una comunitat indígena de 2000 persones que pels d'una llengua amb 400 milions de parlants.



## 3 MATERIALS I MÈTODES

La primera decisió, fent de la necessitat virtut, és abandonar la llengua en favor del *genus* [14](conjunt de llengües amb un parentiu evident per qualsevol persona) com a unitat d'estudi principal. Apart d'un evident estalvi de feina, hi ha un seguit d'avantatges menys evidents, però igual d'importants, que ja es comentaran a l'apartat 3.1.2.

Una segona decisió és definir un nombre limitat de variables a partir de la informació recollida a diferents bases de dades tipològiques que poden consultar-se en Internet i a la bibliografia especialitzada per cobrir els buits d'informació. Tota la informació s'ha tabulat per tal de poder treballar amb ella posteriorment. S'utilitzaran tècniques estadístiques per obtenir correlacions i patrons de canvi a partir dels milers de dades recollides a les taules, per decidir si són significants.

Per últim, les variables es representaran en mapes areals i arbres filogenètics per veure tendències històriques que la simple aplicació de l'estadística no permetrien detectar. L'ús d'aquests arbres i mapes suposa un salt al buit ja que no existeix un consens sobre la classificació de les llengües i molt menys sobre la seva distribució històrica. Especialment dramàtic és el cas de les llengües de comunitats indígenes que han patit la colonització europea amb una dràstica reducció de la població o fins i tot l'extinció.

### 3.1 MOSTRES

#### 3.1.1 PROBLEMES DE MOSTRATGE

A l'hora d'aplicar l'estadística, la mostra esdevé un factor clau, però problemàtic. En el cas de la lingüística les dificultats són moltes i greus. Un requisit de bona part de les tècniques estadístiques exigeix que els diferents elements mostrejats siguin independents, de manera que el valor adoptat per un d'ells, no doni informació dels que tindrà el següent a mostrejar. Això no passa amb les mostres lingüístiques de llengües perquè a mesura que creixent, un major nombre de llengües mostrejades comparteixen proximitat filogenètica o areal.

La principal conseqüència d'això és que els resultats poden variar si en lloc d'agafar 10 llengües bantus i 10 romàniques, se n'agafen 5 de bantus i 15 de romàniques. I llengües aïllades com el basc, poden afectar de forma diferent a mesura que augmenta la mostra, perquè el seu pes sempre serà u, mentre que la resta de grups el van incrementant.

Els problemes de mostratge tampoc desapareix amb una mostra molt gran, ni tant sols agafant la totalitat de llengües del món. La raó és que hi ha branques amb una única llengua i d'altres amb 500. Agafar la totalitat de llengües suposarà infravalorar les llengües aïllades i les famílies poc ramificades i sobrevalorar-ne les més grans. Apart hi ha el tema de la variació dialectal. Allà on un autor veu una única llengua, altres en poden veure 5 o 10. Això farà variar el pes de cada llengua.

Per últim, no totes les llengües tenen la mateixa possibilitat de formar part d'una mostra perquè el coneixement que tenim varia molt d'una zona a una altre. Moltes mostres estan clarament esbiaixades a favor de llengües grans i accessibles als lingüistes.

La principal dificultat d'aquesta tesi és haver d'inferir canvis lingüístics no documentats històricament a partir de qualsevol tècnica que oferís una garantia de seriositat i rigor. Això ha implicat treballar amb un ventall de metodologies força diferents.

### 3.1.2 GENUS I LLEGÜES

La següent taula llista les principals diferències entre l'ús de *genus* i de llengües.

<i><b>Llengua</b></i>	<i><b>Genus</b></i>
<i>5000 a 7000 llengües</i>	<i>600 genus</i>
<i>80% no documentades</i>	<i>80% documentats</i>
<i>Problemes amb mostres</i>	<i>No cal mostrejar</i>
<i>Difícil delimitació.</i>	<i>Delimitació evident i universal</i>
<i>Sensible a influència dels veïns</i>	<i>Resistent a influència dels veïns</i>
<i>Fins 1000 anys d'edat?</i>	<i>Fins 3500-4000 any d'edat?</i>

**Taula 7** Principals diferències entre *llengua* i *genus*.

Encara que en alguns casos s'ha fet servir la llengua com unitat d'estudi, en la majoria de casos en aquesta tesi s'ha optat per usar el *genus*. La primera motivació va ser poder treballar de manera extensiva, sense haver de mostrejar. En lloc d'haver d'escollir unes poques llengües de cada tàxon, amb els problemes metodològics que això suposa, l'opció ha estat aprofitar qualsevol dada disponible.

El nombre total de *genus* és aproximadament d'un 600, una xifra més de 10 vegades inferior a les prop de 7000 llengües amb codi ISO llistades a la 16a edició de l'Ethnologue. Al mateix temps, encara que la immensa majoria de llengües no està prou documentada per formar part d'un estudi comparatiu (de moltes a penes se'n coneix més que el nom), en el cas dels *genus*, si que és possible trobar dades suficients com a mínim d'alguna llengua que en formi part.

Però a part dels avantatges a l'hora de confeccionar una taula més complerta, n'hi ha d'altres que justifiquen l'ús dels *genus* en lloc de llengües. En primer lloc, els límits que separen una llengua d'un altra són força nebulosos. Allà on un autor veu una sola llengua, altres en veuran 5 o 10. En canvi, d'acord a la seva definició, no ha d'haver-hi pràcticament cap dificultat en assignar una varietat lingüística a un *genus* o un altre: *A genus is a group of languages whose relatedness is fairly obvious without systematic comparative analysis, and which even the most conservative "splitter" would accept.* (Dryer 2005:584) [2]

En segon lloc les llengües són sensibles a la influència de les seves veïnes que pot suposar canvis tipològics que poden confondre a l'hora de buscar patrons generals. L'ús del *genus* com unitat d'estudi permet fer una comparació de les llengües que el formen i deduir la tipologia original ignorant aquestes influències si només han afectat a unes poques llengües que suposin excepcions clares.

Per últim, els *genus* tenen mil·lennis d'antiguitat, més de 4.000 anys en el cas de semític, davant dels segles per part de les llengües. Mentre que una llengua pot revertir un ordre sintàctic en 200 anys, les possibilitats que aquest canvi s'hagi donat a totes les llengües del mateix *genus* és ínfima i disminueix quant més gran sigui aquest.

### 3.1.3 RECOLLIDA DE DADES.

Per realitzar aquesta tesi s'ha treballat amb desenes de milers de dades corresponents a milers de llengües. Evidentment si això ha estat possible és perquè prèviament aquesta informació ha estat recollida i tabulada per altres autors en bases de dades d'accés públic o publicades a la bibliografia especialitzada.

Apart de bibliografia específica per anar omplint forats, el gruix de les dades s'ha obtingut de les següents bases de dades disponibles com annexos a la tesi:

– **WALS** (Dryer, Matthew S. & Haspelmath, Martin (eds.). 2011. *The World Atlas of Language Structures Online.*) [95] (disponible en línia a <http://wals.info/export>, consultat el 20.11.2015) La base conté dades de 2676 llengües i dialectes dels quals es mostra la distribució de més de 200 trets estructurals que pertanyen a dominis com la fonologia, la morfologia, la sintaxi. Conté informació sobre la classificació com la família, la subfamília i el *genus*. També d'extralingüística com el nombre de parlants. Es dona també la posició geogràfica (latitud i longitud) de cada llengua, representada per un únic punt que es pot visualitzar en un mapamundi on es poden combinar trets i llengües de forma interactiva. Cadascuna de les varietats lingüístiques està associada al codi ISO i un altre codi de 3 lletres anomenat codi *Wals*.

– **Matsumoto (1987)** [96]. Disponible en línia només en japonès, però amb un resum en anglès. <http://www.tulips.tsukuba.ac.jp/limedio/dlam/M17/M171781/2.pdf>. (consultat el 11.11.2015). Conté les dades de 1430 llengües, a més del nom es cita el continent, la classificació segons Greenberg [97], els països on es parla i 4 variables d'ordre de constituents (els 4 principals de Greenberg). També hi ha taules resum dels diferents tipus segons família, regió geogràfica i tipus d'ordre de constituents.

– **Hawkins (1983)** [10] En un apèndix es llisten 350 llengües amb la seva classificació segons Ruhlen (1978), els països on es parlen, el tipus dins la classificació tipològica de Greenberg, algunes variables sintàctiques i la font bibliogràfica.

– **A global linguistic database (2008)** compilada per M.Ruhlen. Està disponible en línia a <http://starling.rinet.ru/typology.xls> (consultat el 11.11.2015). Conté informació sobre 5736 llengües i dialectes, amb un total de 43 variables com la distribució geogràfica, classificació segons Ruhlen (1995) [63], noms alternatius i cites bibliogràfiques de descripcions de la llengua. La informació lingüística és centra a l'inventari fonològic, però hi ha dades sobre algunes variables sintàctiques i morfològiques.

– **Gell-Mann, M., & Ruhlen, M. (2011)** [23]

Article en que els autors llisten 2135 llengües agrupats segons l'ordre del subjecte, del verb i de l'objecte.

– **Siewierska (1997)** [94]. El títol de l'obra, *Constituent Order in the Languages of Europa*, deixa clar el seu contingut. És un volum d'una col·lecció de nou que són fruit del projecte de recerca europea *Typology of Language in Europe* (EUROTYP). A l'apèndix de les pàgines 783-812 es llisten 137 llengües d'Europa amb la seva classificació i 12 variables d'ordre de constituents. La informació sobre cada parella nucli-complement es molt completa i té en compte a més de l'ordre, informació sobre

l'estatus morfològic: lliure, clíctic, afix o quasi compost. També s'indica si hi ha un ordre secundari. Les primeres quatre són de constituents a nivell oracional i les altres vuit a nivell del sintagma nominal. Malauradament està limitada a Europa.

– **The ASJP Database (versió 16).** Automated Similarity Judgment Program (ASJP) és el nom d'un equip de múltiples lingüistes que dirigits per Søren Wichmann han desenvolupat un programa informàtic per la classificació automàtica de llengües. Per tal d'implementar-lo han creat una immensa base de dades que recull una versió reduïda (40 paraules) de la llista de Swadesh transcrita en un alfabet fonètic simplificat per 6895 varietats lingüístiques corresponents a 4401 llengües amb codi ISO diferents. Cada registre aporta dades com la classificació i el nombre de parlants segons Ethnologue 16, el *genus* i dades geogràfiques. Està disponible en línia a <http://asjp.clld.org> (consultat el 11.1.2015).

– **Phoible.** És una base de dades que recull els inventaris fonològics de 1672 llengües juntament amb el codi ISO, el *genus* i les coordenades geogràfiques que es poden visualitzar agrupades en un mapa segons les diferents variables. Resulta força útil que de cadascun dels 2160 fonemes diferents que es recullen, s'hi dona els valors en un matriu de trets distintius. Pot consultar-se online a <http://phoible.org>. La base actual s'ha format de la suma de varies bases anteriors.

– **Ethnographic atlas.** És una base de dades antropològiques codificada per Gerges R. Murdok [98] i publicada en 29 números consecutius de la revista ETHNOLOGY en el període 1962-1980. De cada societat descrita es donen més de 200 variables d'interès antropològic i lingüístiques com el codi ISO o la classificació.

### 3.1.4 TRACTAMENT DE DADES.

L'heterogeneïtat de les fonts representa un repte a l'hora d'elaborar una llista de variables que pugui aprofitar el màxim d'informació disponible. Les dades originals provenen de milers de referències bibliogràfiques d'autors diferents de diverses èpoques que han treballat amb metodologies completament diferents. Per una banda hi ha llengües sense cap dada i per l'altra llengües amb molta informació, però a cops contradictòria. En alguns casos, els autors no es posen d'acord ni en quina és la llengua que han descrit ni en si és la mateixa que la del poble del costat.

Malgrat que bona part de les anàlisis s'han fet al nivell de *genus*, cal tenir present que totes les dades primàries s'han obtingut al nivell de llengua. En el cas de variables quantitatives com el nombre de parlants d'una llengua o el total de fonemes, la qüestió es centra en escollir entre les diferents fonts quan no coincideixen. Només és qüestió de guiar-se per criteris de qualitat, prestigi o antiguitat. La casuística és tan extensa que no és possible justificar la decisió presa en cada cas.

Pel que fa a les variables qualitatives la cosa és molt més complicada ja que es necessari reduir-les a paràmetres comuns per tal que pugin ser tractades conjuntament. Els detalls s'expliquen al següent apartat.

La part més feixuga de la tesi ha sigut justament l'elaboració d'unes poques taules que recullen la major part de la informació consultada, permetent el seu posterior tractament i comparació.

### 3.1.2 3.1.5 ORDRE DE CONSTITUENTS.

A la majoria de llengües, l'ordre en que les paraules són produïdes aporta una quantitat substancial del significat de l'oració. Justament el fet que alguns d'aquests ordres siguin molt més freqüents en unes llengües que en altres, ens permet aprofitar-los com a variables útils per fer-ne comparacions.

S'han escollit 9 constituents que són els més àmpliament usats a la bibliografia. I s'han creat 5 taules anomenades MATSUM, HAWKINS, TYPOL, RUHLEN i WALS a partir de les 5 primeres fonts llistades prèviament. A més, dues taules més (EUROP i DEA) s'han fet servir només per l'Anàlisi de Components Principals (veure apartat 3.3). La primera a partir de Siewierska (1997) [94] i la segona, es comenta a punt 2.2.6.

Donat que els autors de cadascuna de les bases de dades consultades ha fet servir criteris diferents per caracteritzar les relacions d'ordre Nucli-Complement, s'ha procedit a unificar-les de la mateixa manera:

Només s'han distingit 5 categories diferents: S'ha assignat un valor +10 en els casos en que el nucli segueix al complement i -10 en cas contrari. Si els dos ordres coexisteixen sense dominància el valor assignat és 0. Els valors +5 i -5 s'apliquen si es troben dos ordres, però un és clarament dominant. La taula següent mostra les variables escollides, els valors extrems i les fonts utilitzades.

<b>VARIABLE</b>	<b>DESCRIPCIÓ</b>	<b>RUHLEN</b>	<b>WALS</b>	<b>HAWKIINS</b>	<b>TYPOL</b>	<b>MATSUM</b>	<b>+ 10</b>	<b>-10</b>
<i>SUBJECTE</i>	<i>Ordre del subjecte respecte el verb.</i>	x	x	x	x	x	SV	VS
<i>OBJECTE</i>	<i>Ordre de l'objecte respecte el verb</i>	x	x	x	x	x	OV	VO
<i>ADPOSICIÓ</i>	<i>Ordre de l'adposició respecte el substantiu.</i>	-	x	x	x	x	POST	PREP
<i>ADJECTIU</i>	<i>Ordre de l'adjectiu respecte el substantiu.</i>	-	x	x	x	x	AN	NA
<i>GENITIU</i>	<i>Ordre del genitiu respecte el substantiu.</i>	-	x	x	x	x	GN	NG
<i>NUMERAL</i>	<i>Ordre del numeral respecte el substantiu.</i>	-	x	-	x	-	NumN	NNum
<i>RELATIU</i>	<i>Ordre del relatiu respecte el substantiu.</i>	-	x	-	-	-	RN	NR
<i>DEMOSTRATIU</i>	<i>Ordre del demostratiu respecte el substantiu.</i>	-	x	-	x	-	DN	ND
<i>POSSESIU</i>	<i>Ordre del possessiu respecte el substantiu.</i>	-	-	-	x	-	PossN	NPoss

**Taula 8** Variables fetes servir a cada base de dades i valors màxims i mínims.

Un cop tabulats els valors per cada llengua, s'ha fet el càlcul per cada *genus*. A les variables quantitatives com la població, només ha calgut fer la suma, mentre que en d'altres com el nombre de fonemes, el que s'ha fet es recollir-les i calcular la mitjana.

A les variables qualitatives, com en el cas dels ordres de constituents, els resultats s'ha passat a un percentatge sobre el total del recorregut. S'han creat 5 categories discretes a les quals se'ls hi ha assignat un codi de colors, d'acord a la següent taula per tal de

representar-se en mapes i arbres filogenètics. Per càlculs posteriors, a cadascuna se li ha assignat un valor numèric de 0 al 4, d'acord a la Taula 9.

<i>Valor genus</i>	<i>Genus</i>	<i>color</i>
0-10%	<i>Nucli final</i>	0
10%-30%	<i>Nucli final dominant</i>	1
30%-70%	<i>Sense ordre dominant</i>	2
70%-90%	<i>Nucli inicial dominant</i>	3
90%-100%	<i>Nucli inicial</i>	4
<i>Sense dade</i>	<i>Sense informació de cap llengua</i>	-

**Taula 9** Color i valor de les categories definides segons els valors.

### 3.1.6 POBLACIÓ DE *GENUS*.

El càlcul de parlants de cada *genus* s'ha fet a partir de les xifres de l'**ASJP Database (versió 16)** que semblen extretes de l'**Ethnologue, 15 edició** [52]. En els casos de *genus* que no hi constaven, s'ha anat a buscar la xifra directament a la web. A totes les llengües amb menys de 500 parlants, se'ls hi ha assignat aquest nombre al considerar que és el mínim de parlants d'una llengua natural que no estigui sotmesa a processos de minorització.

L'estimació de la població de parlants de cada *genus* presenta com a gran dificultat el fet que en molts casos la colonització europea ha representat una reducció dramàtica del nombre de parlants. La importància d'aquest fet s'ha reduït força pel fet de fer servir com a variable el logaritme de la població en lloc de la xifra concreta.

### 3.1.7 PROGRAMARIS ESPECÍFICS.

A) Mesquite és un programari pensat per a la biologia evolutiva, dissenyat per ajudar els biòlegs a organitzar i analitzar dades comparatives sobre els organismes. El seu èmfasi està en l'anàlisi filogenètic, però alguns dels seus mòduls es refereixen a la genètica de poblacions, mentre que altres fan anàlisi multivariable no filogenètica. És gratuït i pot descarregar-se a <http://mesquiteproject.org/> (consultat el 24.11.2015).

Del conjunt de possibilitats que ofereix, només s'ha fet servir l'opció que permet a partir d'un arbre filogenètic i una distribució d'estats de caràcters en els tàxons observats, intentar reconstruir la història dels canvis en els nodes ancestrals. S'obté una proposta d'història d'acord al principi de màxima parsimònia.

El programa també permet pintar de colors les branques d'un arbre filogenètic segons els valors de les diferents variables i deduir-ne les tendències històriques de canvi tipològic d'un cop de vista.

B) XLSTAT és un paquet estadístic comercial que té per principal característica el fet d'integrar-se en el full de càlcul Excel d'Office. Això permet aprofitar els fulls d'*excel*



per tabular, tractar les dades, aplicar un munt de funcions estadístiques i crear gràfics de gran qualitat, sense necessitat d'anar fent canvis d'un programa a un altre.

S'ha fet servir bàsicament per buscar correlacions entre variables i per l'aplicació de l'anàlisi de components principals (ACP), una metodologia especialment indicada per simplificar el nombre de variables estudiades.

C) Photoshop CS4. Aquest programa s'ha fet servir bàsicament per elaborar els mapes utilitzats de la tesi. Molts mapes s'han creat a base de superposar-ne altres i fusionar-los amb l'ajuda de Photoshop. També s'ha fet servir naturalment per posar noms i pintar-ho d'acord amb els valors de cada categoria i variable.

## 3.2 CORRELACIONS I DIAGRAMES DE BARRES

En molts casos, la manera més clara per fer evident la relació entre varies variables és fer una senzilla correlació entre elles o fer un diagrama de barres per obtenir-ne una representació visual. Són tècniques senzilles que s'han portat a terme gràcies al paquet estadístic Xlstat.

## 3.3 ANÀLISI DE COMPONENTS PRINCIPALS

### 3.3.1 PRESENTACIÓ DEL MÈTODE.

Un cop seleccionades les variables que seran objecte d'estudi i construïdes les taules pertinents, s'ha procedit a fer-ne l'anàlisi estadística. En concret, en aquest capítol, s'ha utilitzat la tècnica de l'Anàlisi de Components Principals (ACP) que facilita enormement la visió global.

És una tècnica estadística que persegueix la reducció del nombre de les variables, però conservant el màxim d'informació. La tècnica permet substituir les variables inicials per uns factors que són combinació lineal de les primeres i independents entre sí. L'elecció dels factors es fa de tal manera que el primer recull la màxima proporció possible de la variabilitat original; el segon factor ha de recollir tota la possible que no ha recollit el primer, i així successivament. Aquests factors s'anomenen components principals. Com que existeixen diferents variants d'ACP, en el nostre cas, s'ha fet servir la variant francesa tal com apareix al programa XLSTAT amb les opcions per defecte.

Un aspecte cabdal de la tècnica és que els factors principals definits no es corresponen obligatòriament amb realitats evidents, sinó més aviat amb variables abstractes. A la majoria de casos cal estudiar-les minuciosament abans d'aventurar possibles interpretacions. Un exemple que s'usa amb freqüència és el de les notes de totes les assignatures d'un escolar que es poden explicar amb factors com memòria, expressivitat oral o facilitat per les matemàtiques.

### 3.3.2 APLICACIÓ.

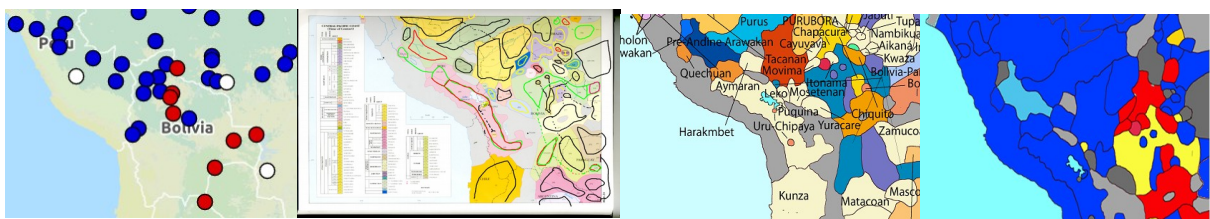
L'ACP s'ha aplicat a les varies taules sobre fonemes i ordre de constituents que es descriu a l'apartat 3.1.4. Cal destacar que les taules **DEA, EUROP, WALS, HAWKINS, MATSUM, i TYPOL** contenen exclusivament dades sobre ordre de constituents i per això la seva informació és clarament redundant. Es diferencien només pel que fa a la metodologia de l'autor, per les llengües que les formen i pel nombre de constituents estudiats.

Aquesta insistència en repetir els càlculs sobre bases de dades similars té com objecte comprovar fins a quin punt els resultats obtinguts estan esbiaixats per factors metodològics com els criteris dels autors o la mida de la mostra.

A cada taula citada s'ha realitzat la prova per cinc variables (Subjecte, OBBjecte, Adposició, Genitiu i Adjectiu. A les 3 primeres també s'ha repetit per vuit variables (les anteriors més Demonstratiu, Numeral i Possessiu. Finalment per EUROP s'ha pogut fer per les 11 variables. La unitat d'anàlisi ha sigut la llengua i no el *genus*. Un dígit radera el nom de cada taula indica el nombre de variables que s'ha fet servir. Per exemple, WALS5 indica la taula WALS limitada a les cinc primeres variables. Donada la gran quantitat de comparacions fetes i de resultats obtinguts (a més de per coherència amb altres capítols de la tesi), s'ha optat per donar preferència a la base WALS, els resultats de la qual seran presos com referència a les comparacions i els gràfics de la qual es mostren al cos principal de la tesi. Els resultats de la resta de bases es compararan amb aquesta..

### 3.4 MAPES

Una part fonamental de la tesi, i possiblement la que ha presentat més dificultats, ha estat la confecció de mapes per representar la distribució dels principals trets estudiats. Al contrari que als mapes del WALS, que presenten les llengües com punts de colors diferents segons el valor de les variables (mapa 6), s'ha optat per fer servir l'àrea de cada *genus* com unitat.



Mapa 6

Mapa 7

Mapa 8

Mapa 9

S'ha fet principalment perquè l'ús de punts no permet la funció bàsica del mapa d'estudiar la distribució sobre el terreny. Representar una llengua com el rus amb un punt, no permet fer-se idea de quines són les veïnes amb les que ha pogut interaccionar.

Els mapes 6 permet veure la variable Adposició si fem servir un punt per cada llengua de diferent color segons se seu valor. Per treballar amb *genus* el que s'ha fet és reunir totes les llengües de cada genus (mapa 7), unir les seves superfícies en un mapa de *genus* (mapa 8), per més tard pintar-les segons els valors mitjans de la variable per les llengües documentades que en formen part (mapa 9). Els avantatges que pot aportar l'ús de *genus* ja s'ha comentat a 3.1.5, però les dificultats són també considerables.

La distribució original s'ha vist alterada per les colonitzacions europees a partir del segle XV i altres expansions anteriors com la bantu, la polinèsia o la de llengües d'àrabs i berbers. Això fa que la distribució mostrada als mapes coincideix aproximadament amb l'actual (o potser més exactament la de principis del segle XX) en el cas d'Àfrica i Euràsia i Nova Guinea, amb la del segle XVIII per les llengües australianes i d'Amèrica del nord i amb la del segle XV pel cas de d'Amèrica del sud.

Els mapes s'han creat amb Photoshop a partir dels mapes de *Moseley 1993* [99], excepte per Àfrica i Amèrica del nord en què s'ha donat prioritat als de Greenberg [100] i Mithun [101]. També s'han consultat un munt de mapes per zones concretes citats a la bibliografia. En els mapes de distribució d'ordres de constituents s'ha fet servir e codi de colors de la Taula 9. En principi, s'han fet 6 mapes individuals de 6 macro-àrees: Amèrica del nord, Sudamèrica, Austràlia, Nova Guinea, Euràsia i Àfrica. Després s'han ajustat i fusionat en un mapamundi. Cada mapa consta de diverses capes, de manera que en cada una s'hi ha representat una variable, a més de les generals amb noms, fletxes, fronteres, etc.

A partir dels mapes de macro-àrees, s'han retallat altres de zones d'especial interès per l'anàlisi de les dades: Columbia Britànica, Mesoamèrica, Amazonia Boliviana, Cinturó africà de fragmentació, Regió Euro-mediterrània i Sud-est d'Àsia.

En total, s'han representat 590 *genus* i malgrat això han quedat amplies zones sense dades que s'han pintat de color gris. La pertinença d'una llengua a un *genus* determinat s'ha decidit en primer lloc d'acord al WALS quan ha estat possible i secundàriament segons AJSP i Moseley [99]. La ubicació de cada llengua i els conjunt de mapes es poden consultar als l'annexos.

### 3.5 ARBRES

Els arbres s'han construït amb el programa Mesquite, descrit a 3.1.6. Encara que pot fer-se servir en lingüística, el programa està pensat per la biologia. Els arbres filogenètics són representacions gràfiques de la història de les relacions entre organismes (Page and Holmes, 1998). [102] Un cop construït l'arbre, qualsevol característica dels organismes relacionats es pot analitzar en el marc especificat per l'arbre. Els canvis en els caràcters es poden deduir observant les diferències entre els organismes descendents i els seus ancestres, fins i tot si aquests mai van ser directament observats (Maddison and Maddison, 2000). [103]

Al contrari que altres programes similars, Mesquite és incapaç de proposar cap arbre filogenètic a partir dels valors de les branques terminals. Així que cal prèviament introduir-li l'arbre filogenètic complet sobre el que treballar. L'arbre escollit ha sigut el de Ruhlen del 1983 [63]. Al cos de la tesi, els arbres apareixen en una escala que només permet fer anàlisis a nivell global, però a l'annex 1 es pot visualitzar els valors de cada llengua.

En el nostre cas, les unitats a classificar han estat els *genus* lingüístics descrits 3.1.1 i el criteri de reconstrucció dels estadis previs a partir dels terminals ha estat el de màxima parsimònia entesa com el principi que minimitza el nombre de canvis necessaris. Encara que cada arbre és del total de *genus* estudiat, en alguns casos s'han creat imatges d'algunes branques en particular per tal de mostrar els tàxons implicats en a les àrees d'especial atenció llistades a 3.4.

Un cop fet els arbres, s'han pintat les branques terminals d'acord amb el codi de colors mostrat a la Taula 9. Al cos de la tesi, els gràfics que mostren els arbres porten incorporats rètols a les zones on les branques es corresponen amb alguna de les zones d'especial interès ja mencionades.

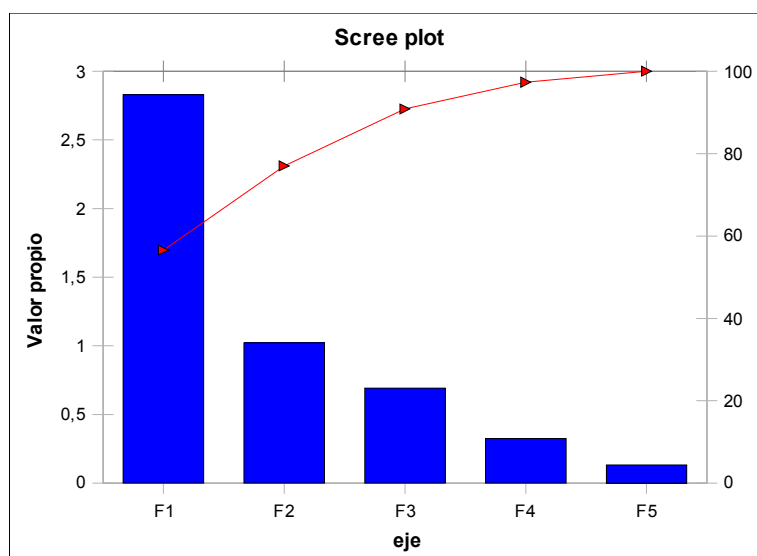
L'opció **Trace Character History** permet reconstruir els nodes intermedis, a més de l'arrel. Això permet proposar una història diacrònica dels canvis de cada tàxon.

## 4 RESULTATS ACP

Donat que aquesta prova estadística s'ha aplicat a un total de 10 taules (5 bases de dades i diferent nombre de variables per cadascuna), s'ha començat fent una presentació amb WALS5 (taula WALS amb 5 variables), per després comparar els resultats obtinguts amb la taula ampliada (WALS8) i després amb la resta de bases de dades.

### 4.1 RESULTATS ACP DE WALS 5 I INTERPRETACIÓ

L'Anàlisi de Components Principals (ACP) ens ha permès transformar les variables estudiades en un nou conjunt de variables independents que ens ajudarà a explicar la variabilitat.

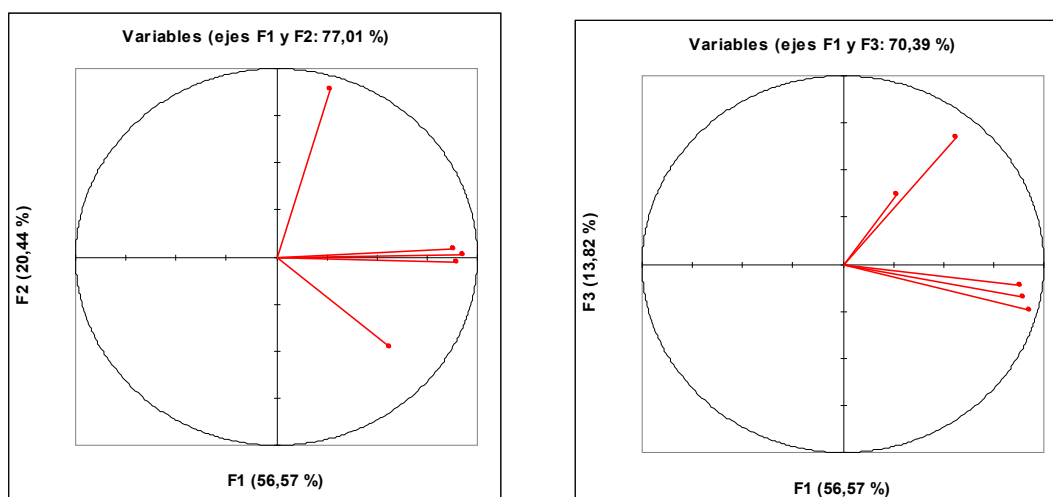


**Gràfic 4** Diagrama de barres corresponent a la taula 10, corresponent als factors de WALS5. El percentatge de la dreta correspon al percentatge explicat de la variància.

	F1	F2	F3	F4	F5
Valor propio	2,829	1,022	0,691	0,325	0,133
Variabilidad (%)	56,571	20,443	13,824	6,506	2,656
% acumuladc	56,571	77,014	90,838	97,344	100,000

**Taula 10** Valors dels factors principals de WALS5

Si ens fixem en els resultats obtinguts a la base de dades WALS amb 5 variables (gràfic 4) i la taula corresponent, es pot veure que un únic factor explica la major part de la variabilitat. (56,57%). El segon acumula un 20%, que representa la variabilitat esperada sense necessitat de cap transformació i la resta dona valors molt inferiors als que caldria esperar. Els resultats són coincidents amb els d'Ehara [15] comentat a 1.1.1.



**Gràfic 5** Diagrames circulars de F1 amb F2 i F1 amb F3, per WALSS

Correlaciones entre las variables y los factores:

	F1	F2	F3	F4	F5
OV	0,883	-0,050	-0,229	-0,348	-0,208
SV	0,552	-0,545	0,629	-0,049	0,028
POST	0,917	0,003	-0,266	-0,053	0,294
AN	0,420	0,807	0,412	-0,058	0,006
GN	0,872	0,005	-0,084	0,467	-0,118

**Taula 11** Correlacions entre variables i factors, per WALSS

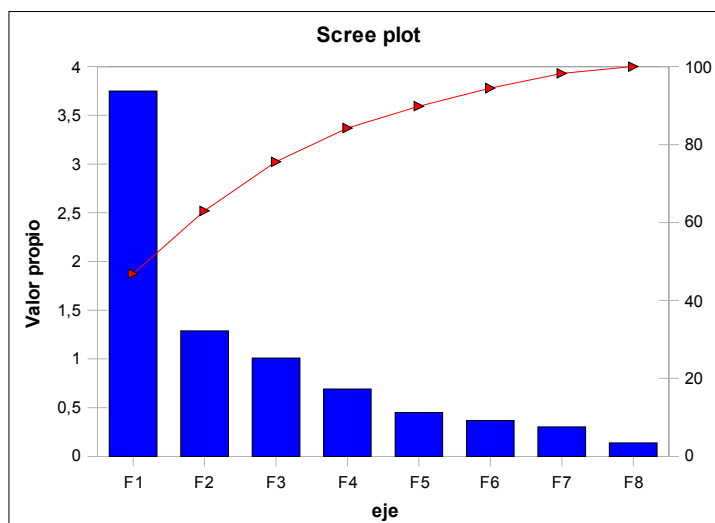
La taula 11 permet fer-se una idea del que hi ha darrera de cadascun dels factors principals. El primer factor o F1 està correlacionat en un 88,3% amb la posició de l'objecte respecte al verb, en un 91,7% amb la posició ocupada per les adposicions i en un 87% amb l'ordre dels genitius.

En aquesta taula, cada variable s'anomena per les inicials dels dos constituents que la formen, així OV fa referència a la variable Objecte, i no el valor concret OV d'aquesta contraposat al valor VO. I així amb la resta.

En altres paraules, el genitiu, l'objecte i l'adposició tendeixen a variar a la mostra estudiada de forma conjunta, sent l'última la que sembla dominant al tenir un percentatge una mica més gran. Les gràfiques circulars, on es representa cadascuna de les cinc variables segons els components F1 i F2 a l'esquerra o F1 i F3 a la dreta, confirmen aquesta similitud de comportament, ja que les tres pràcticament es superposen.

Per contra, les variables Subjecte-Verb i Adjectiu-Nom es comporten de forma similar pel que fa al tercer factor, però difereixen clarament en el segon. Aquestes dues variables semblen intervenir en bona part en la composició del tercer factor principal, tal com es veu en els gràfics circulars (gràfic 5).

## 4.2 COMPARACIÓ AMB WAL58



**Gràfic 6** Diagrama de barres corresponent a la taula 12, corresponent als factors de WAL58

Valores propios:

	F1	F2	F3	F4	F5	F6	F7	F8
Valor propio	3,750	1,288	1,008	0,691	0,451	0,370	0,302	0,148
Variabilidad (	46,869	16,101	12,604	8,642	5,635	4,620	3,781	1,748
% acumulado	46,869	62,971	75,574	84,216	89,852	94,471	98,252	100,000

**Taula 12** Valors dels factors principals de WAL58

Correlaciones entre las variables y los factores:

	F1	F2	F3	F4	F5	F6	F7	F8
OV	0,862	0,209	0,068	-0,218	0,064	-0,192	-0,270	-0,220
SV	0,430	0,643	-0,005	0,604	-0,086	-0,164	-0,031	0,036
POST	0,881	0,202	0,031	-0,254	-0,047	0,091	-0,184	0,272
AN	0,538	-0,613	-0,176	0,443	0,095	0,221	-0,220	-0,020
GN	0,833	0,194	-0,009	-0,076	-0,240	0,371	0,231	-0,119
NumN	0,004	-0,119	0,986	0,105	0,009	0,056	-0,017	0,001
RN	0,813	-0,111	0,011	-0,021	0,491	-0,089	0,275	0,030
DN	0,635	-0,591	0,009	-0,009	-0,359	-0,317	0,131	0,031

**Taula 13**

Al repetir l'anàlisi amb WAL 8 afegint-hi tres variables extres no es veu un comportament diferent del que s'ha descrit a l'apartat anterior. Excepte pel fet que els percentatges baixen lleugerament. Efectivament, tal com es pot veure als gràfics i taules de WAL58 immediatament superiors, segueix havent-hi un primer factor que explica per ell mateix quasi la meitat del total de variabilitat.

Igualment, la taula 13 permet veure que genitiu, objecte i adposicions es comporten de forma similar, i ara s'hi afegeix el relatiu i en menor part el demostratiu. I el segon factor sembla connectat amb les variables *Subjecte* i *Adjectiu*.

Aparentment, el tercer factor de WAL55 ha passat al quart lloc i ha aparegut un nou F3 que coincideix en un 98,6% amb els numerals, que semblen anar per lliure.

### 4.3 COMPARACIÓ AMB LA RESTA DE BASES DE DADES

	F1		F2		F3		F4		F5	
	Variab. (%)	% acumul.	Variab. (%)	% acumul.	Variab. (%)	% acumul.	Variab. (%)	% acumul.	Variab. (%)	% acumul.
TYPOL5	57,924	57,924	0,990	77,733	0,692	91,567	0,280	97,169	0,142	100
DEA5	66,085	66,085	16,015	82,100	11,671	93,771	4,182	97,952	2,048	100
WALS5	56,571	56,571	20,443	77,014	13,824	90,838	6,506	97,344	2,656	100
EUROP5	59,441	59,441	18,523	77,964	13,511	91,475	4,552	96,028	3,972	100
MATSUM5	60,656	60,656	18,410	79,066	14,585	93,651	4,806	98,458	1,542	100
HAWKINS5	59,612	59,612	18,092	77,704	12,607	90,311	6,834	97,145	2,855	100
WALS8	46,869	46,869	16,101	62,971	12,604	75,574	8,642	84,216	5,635	89,852
DEA8	49,855	49,855	18,499	68,354	12,414	80,769	8,495	89,264	3,787	93,050
EUROP8	42,706	42,706	21,491	64,197	11,622	75,818	11,043	86,861	6,460	93,322
EUROP11	51,247	51,247	17,993	69,240	8,646	77,886	7,783	85,669	6,212	91,882

**Taula 14** Comparació entre els factors principals de totes les mostres.

La taula 14 mostra com malgrat que s'han fet servir bases de dades compilades per diferents autors que fan servir els seus propis criteris i que la bibliografia que les nodreix ha estat escrita per milers d'autors d'èpoques i concepcions diferents, els resultats finals són sorprenentment homogenis.

Sempre apareix un factor principal que explica entre una meitat i dos terços de la variabilitat total (només en EUROP baixa al 43%). Després hi ha un parell de factors mitjans (entre 10% i 20%) i la resta presenten percentatges inferiors al 10%. EUROP és la base de dades que més s'aparta del model, però es pot justificar en el fet que només conté llengües europees.

La correlació entre els factors principals i les variables de cadascuna de les Anàlisis de Components Principals realitzats es pot veure, pel seu gran volum, al annexos. De tota manera comentarem les principals conclusions.

El primer factor sempre està correlacionat amb el genitiu, l'objecte i l'adposició. Un segon factor ho està, encara que de manera menys clara amb l'adjectiu mentre que els numerals sembla que no es correlacionen amb cap altre. La resta de constituents sembla ordenar-se a les frases seguint els patrons que determinen aquests

Donat que la lingüística històrica ens mostra com, en els casos documentats, cada constituent va canviant en moments històrics particulars, aquesta correlació només s'explicaria si en cada llengua, el canvi d'un constituent arrossega a la resta.

A les quatre bases de dades més grans sempre és l'adposició la variable que més fortament es correlaciona amb el primer factor, encara que per poc amb l'objecte.

### 4.4 CONCLUSIONS DE LA PROVA DE L'ACP

- Els constituents sintàctics es comporten jeràrquicament de tal manera que n'hi ha que arrosseguen a la resta a l'hora d'establir-ne l'ordre.
- Aproximadament la meitat de la variabilitat que presenten les llengües a nivell d'ordre de constituents està determinada per un únic factor.
- L'adposició o l'objecte són els candidats millor situats per ser al darrera del principal factor de variabilitat.



## 5 RESULTATS: CORRELACIONS

Al capítol anterior s'ha definit un factor principal F1 que explica per si mateix prop de la meitat de la diversitat de valors observada a les variables estudiades.

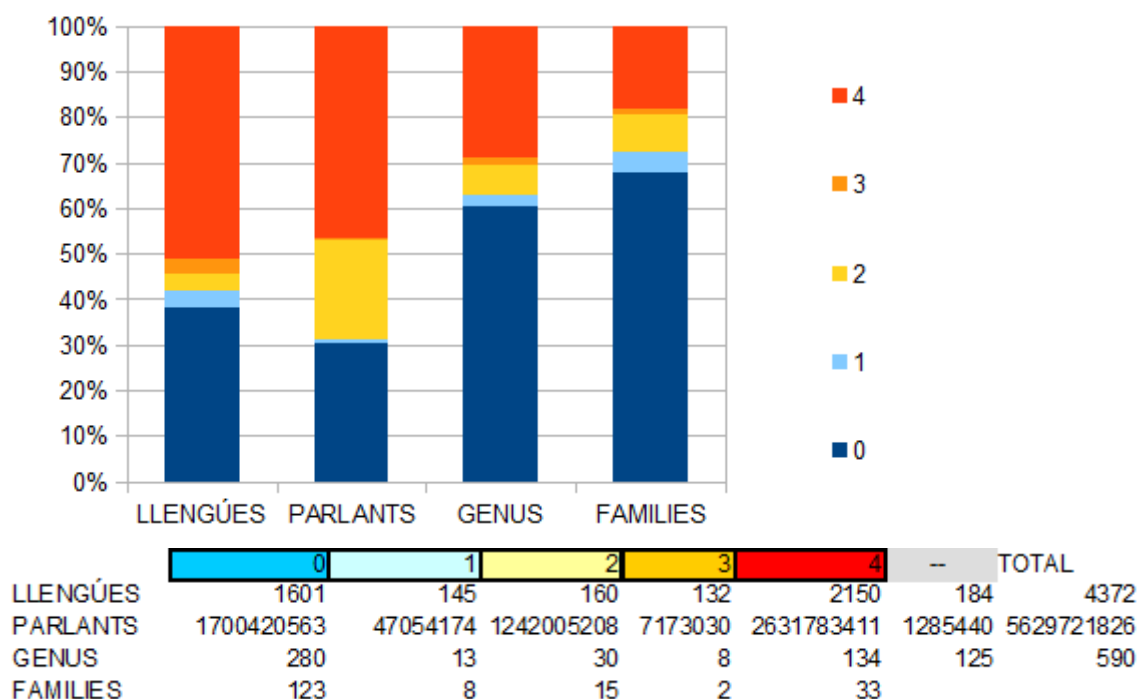
En aquest capítol s'hauria d'analitzar com varia aquest F1 als diferents *genus* segons diferents factors a estudiar. Donada la dificultat de calcular el factor F1, per raons pràctiques, s'ha optat per substituir-lo per l'ordre de l'adposició, molt més senzill d'obtenir per qualsevol llengua i que correlaciona amb ell en un 90%.

En aquests gràfics, el valor 4 s'aplica a l'ordre de Nucli inicial, el 3 a un ordre amb predomini del Nucli inicial, el 2 quan no hi ha ordre dominant. Els valors 1 i 2 quan l'ordre és total o predominantment del Nucli final. El color gris implica falta de dades.

### 5.1 ADPOSICIÓ I VARIABLES SOCIALS

El nombre de llengües preposicionals coincideix amb el de postposicionals fins al punt que l'explicació més evident pot ser considerar aquesta proporció 50-50 com el resultat de la igualtat d'eficiència comunicativa entre ambdós ordres possibles. Un punt d'equilibri.

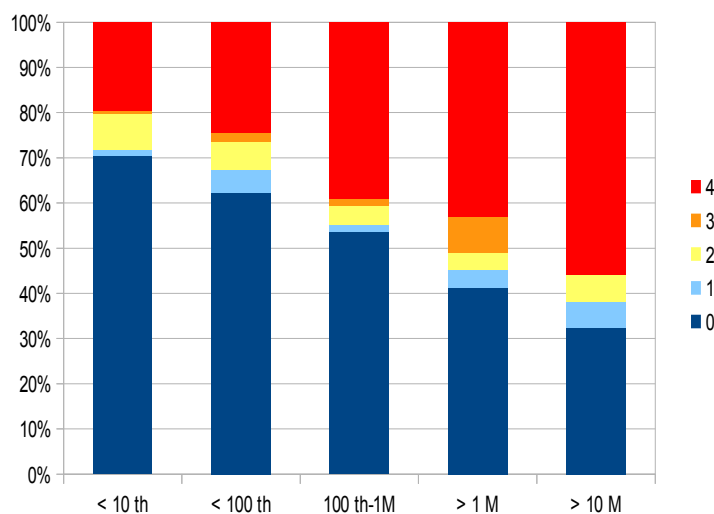
Curiosament aquesta proporció global desapareix si s'analitza per famílies lingüístiques, els *genus*, el nombre de parlants o diferents àrees geogràfiques (gràfic 23).



Taula 15 i Gràfic 7 La variable Adposició segons diversos factors

El 50% de llengües postposicionals de la teoria només representen un 37% del total de llengües dels *genus* definits com postposicionals de la mostra, en aquest (possiblement perquè són llengües més inaccessibles als lingüistes) i no arriba a acumular ni un 30% dels parlants. En canvi representen un 60% dels *genus* i un 67% de les famílies estudiades, més que triplicant en nombre a les famílies preposicionals. Justament aquestes asimetries és el que s'examina en els següents apartats.

## 5.2 ADPOSICIÓ I POBLACIÓ



*ADPOSICIÓ*

	0	1	2	3	4	-	TOTAL
< 10 th	150	3	17	1	42	106	319
< 100 th	61	5	6	2	24	16	114
100 th-1M	37	1	3	1	27	3	72
> 1 M	21	2	2	4	22	0	51
> 10 M	11	2	2	0	19	0	34
	<b>280</b>	<b>13</b>	<b>30</b>	<b>8</b>	<b>134</b>	<b>125</b>	<b>590</b>

**Taula 16 i Gràfic 8** La variable Adposició segons població

Al gràfic 16 es mostra el percentatge dels diferents valors de la variable *Adposició* ja explicats més amunt. Els *genus* s'han classificat en cinc grups d'acord amb el nombre de parlants: menys de 10.000 parlants, entre 10.000 i 100.000 parlants, entre 100.000 i un milió, entre un milió i 10 i més de 10 milions.

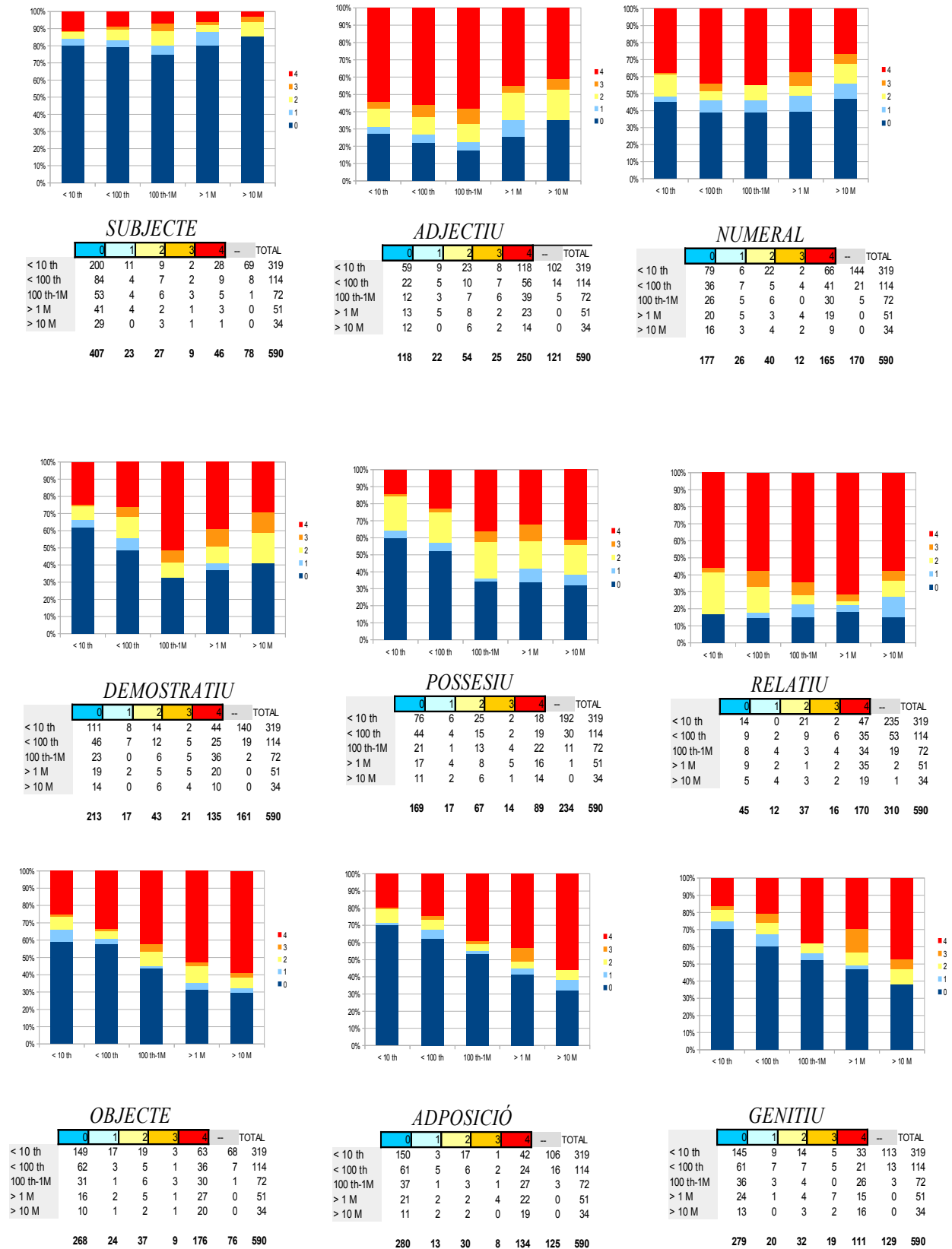
Una primera mirada al gràfic permet extreure dues conclusions evidents. La primera que la immensa majoria de *genus* presenta un únic ordre de constituents dominant, mentre que els que mostren ambivalència o convivència d'ordres diferents representen aproximadament un 10%.

La segona, abans de buscar la confirmació estadística, és que a major nombre de parlants, menor és la proporció de postposicions.

La baixada és força regular, del 70% de *genus* postposicionals que trobem entre les llengües de menys de 10.000 parlants fins al 30% dels *genus* amb més de 10 milions.

En contrapartida, els *genus* preposicionals passen d'un minoritari 30% fins a esdevenir clarament majoritaris quan el nombre de parlants supera els 10 milions.

## 5.3 ORDRE DE CONSTUENTS I POBLACIÓ



Gràfic 9 Multigràfic que mostra el comportament de nou variables respecte població, amb la taula corresponent.

En aquest apartat s'ha fet una comparació per veure si la resta de variables dels ordres de constituents es veuen afectades pel nombre de parlants, tal com succeeix en el cas de l'adposició, ja examinat a l'apartat anterior. Les taules i colors s'han d'entendre igual que en gràfics de les adposicions.

L'anàlisi dels gràfics de la pàgina precedent permet extreure varies conclusions evidents abans de fer la necessària anàlisi estadística.

El patró observat a l'adposició, de descens regular de la proporció de *genus* de nucli final a mesura que s'incrementa el nombre de parlants, també és evident als gràfics de l'objecte i al genitiu. Tots tres es poden confondre de tan semblants com són. Això és el que era esperable i corroborà el fet que totes tres variables mostren un grau de proporcionalitat a F1 proper al 90%.

Encara que de forma molt menys evident, també als demostratius i possessius sembla que es dona una relació amb el nombre de parlants, però només sembla afectar a les tres primeres columnes. A partir d'un milió de parlants, desapareix qualsevol proporcionalitat i l'anàlisi esdevé confusa.

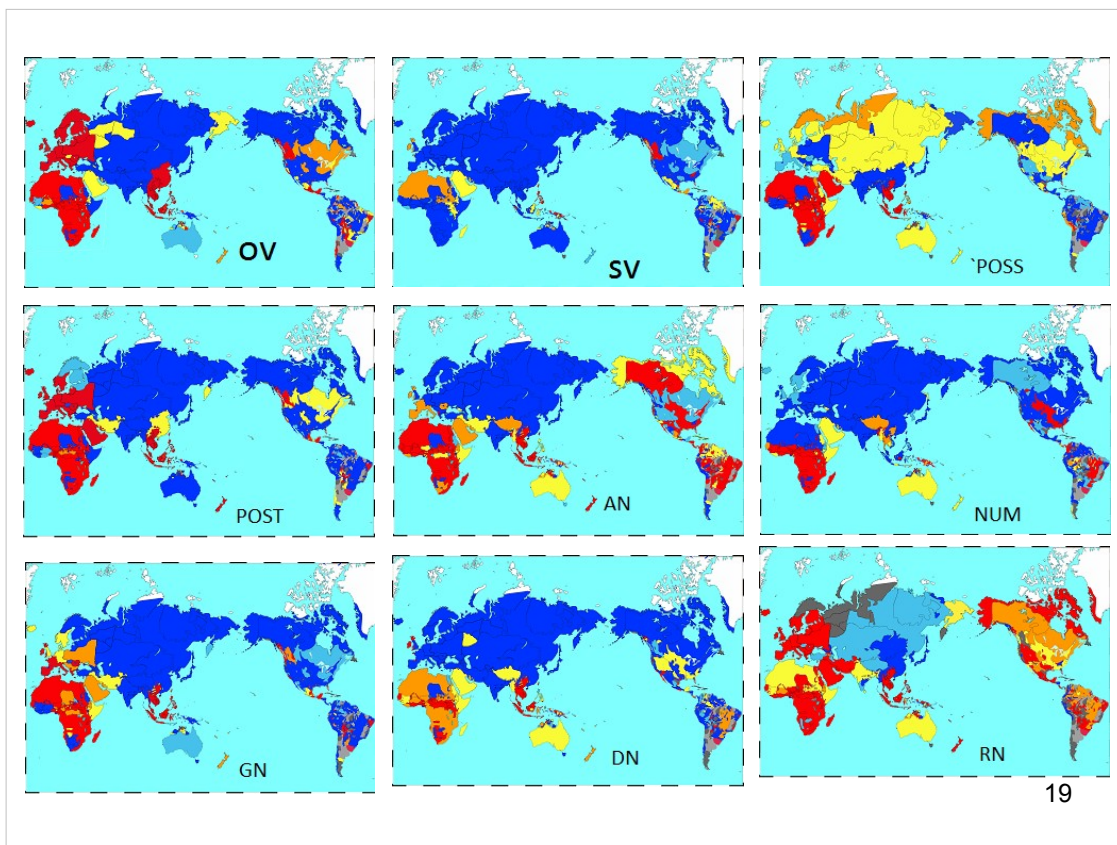
La resta de variables (subjecte, relatiu, adjectiu i numeral) semblen no veure's afectades pel nombre de parlants. En el cas del subjecte, hi ha un claríssim biaix cap a l'ordre Subjecte davant del Verb que domina en percentatges aclaparadors a tots els grups poblacionals. Per contra, les altres tres tampoc sembla que es distribueixen de forma més repartida, encara que en general predomina l'ordre Nucli-Complement.

També crida l'atenció, en comparació amb les adposicions analitzades a l'apartat anterior que no a totes les variables es dona una proporció d'ordres intermedis tant baixa. En el cas dels demostratius i possessius les franges de colors groc, taronja o blau cel sumen un nombre significatiu de *genus*. De forma menys segura es pot dir el mateix de relatiu, adjectiu i numerals.

Intuïtivament els dos fets discutits anteriorment semblen lligats. Si per una llengua es important, per la raó que sigui, tenir un ordre o el contrari en funció del de parlants, és normal que això impliqui que els ordres intermedis acabin confluint en un dels extrems. I si no hi ha cap tendència especial, també és normal que el ordres no dominants es mantinguin sense problema en percentatges més alts.

## 5.4 ORDRE DE CONSTITUENTS I GEOGRAFIA

Un cop examinat com afecta la població a la distribució de les variables, sembla interessant veure com es distribueixen a nivell geogràfic, amb tota la complicació que suposa, ja comentada a l'apartat 3.4 i que es pot resumir en el fet que les colonitzacions han esborrat qualsevol vestigi de les antigues distribucions a la major part del planeta. colonitzacions han esborrat qualsevol vestigi de les antigues distribucions a la major part del planeta.



**Mapa 10** Multimapa que mostra el comportament de nou a nivell geogràfic.

Els mapes mostren que l'ordre de distribució de les variables presenta patrons geogràfics comuns al llarg de desenes de milers de kilòmetres, la qual cosa no deixa de ser sorprenent donada la facilitat com moltes llengües han canviat l'ordre de constituents a Europa en els darrers dos mil anys.

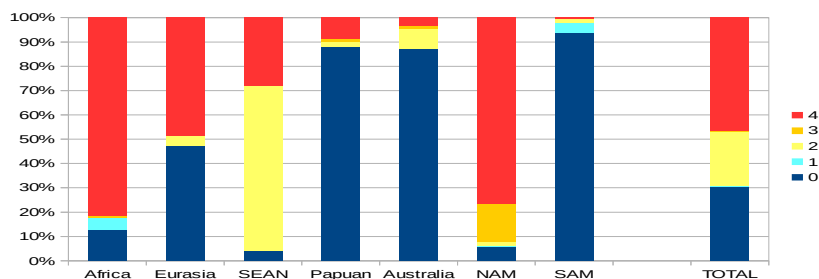
Igual que a l'apartat anterior, hi ha una sintonia entre adposicions, genitiu i objecte. A cop d'ull es veu als mapes d'aquestes tres variables com el color vermell predomina a l'Àfrica, Europa, Sud-est asiàtic, la costa pacífica canadenca i Meso-Amèrica. Per contra, el blau ocupa Sibèria, l'Índia, Austràlia i la major part d'Amèrica. Aquesta distribució serà discutida més endavant en constituir una aspecte central en aquesta tesi.

A la resta de mapes resulta difícil trobar patrons clars coincidents entre dues o més variables. Encara que algunes zones com l'Índia i Sibèria aposten pel blau a pràcticament tots els mapes i el Sud-est asiàtic i Àfrica es mostren en general de color vermell.

Austràlia alterna el groc amb el blau segons la variable i Amèrica mostra una varietat de colors que no permet suggerir cap patró.

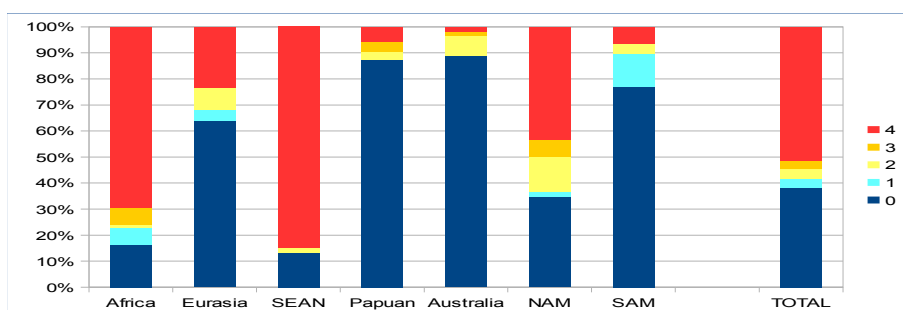
Un factor a tenir en compte és la diferent distribució areal de cada *genus*. Les llengües bantus i les varietats d'àrab i berber, malgrat que només suposen 3 *genus* de la vuitantena que hi ha a l'Àfrica es reparteixen uns dos terços del territori d'aquest continent. Igualment alguns *genus* com l'algonquí a Nord-Amèrica, les llengües Pama-Nyunga a Austràlia o unes poques a Sibèria i l'Índia, poden, degut a la seva ampla distribució, donar una idea enganyosa de la distribució geogràfica de cada variable tipològica.

## 5.5 ADPOSICIÓ PER CONTINENTS



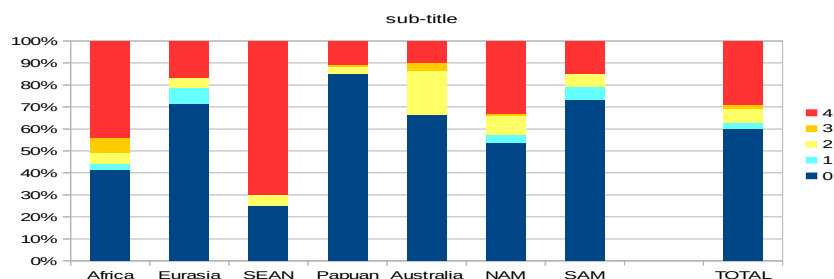
Població						TOTAL
	0	1	2	3	4	
Africa	91784612	34270210	915251	5518180	584690535	70000
Eurasia	1518317245	12062370	123294452	1573014733		3226688800
SEAN	69254736		1117185700	465725113	938150	1653103699
Papuan	3246343		76340	44700	319330	165100
Australia	95420		9050	1500	3530	5000
NAM	608860	33044	187395	1608650	7950420	7360
SAM	17113347	688550	337020		79750	99830
<b>TOTAL</b>	<b>1700420563</b>	<b>47054174</b>	<b>1242005208</b>	<b>7173030</b>	<b>2631783411</b>	<b>1285440</b>

Taula 17 i Gràfic 10 La variable Adposició segons població i continent.



Llengües						TOTAL
	0	1	2	3	4	
Africa	206	82	16	83	878	1265
Eurasia	226	14	31		82	353
SEAN	155		17		971	1143
Papuan	466		17	21	30	534
Australia	161		14	3	3	181
NAM	133	7	52	25	165	382
SAM	254	42	13		21	330
<b>TOTAL</b>	<b>1601</b>	<b>145</b>	<b>160</b>	<b>132</b>	<b>2150</b>	<b>4188</b>

Taula 18 i Gràfic 11 La variable Adposició segons nombre de llengües i continent.



Genus						TOTAL
	0	1	2	3	4	
Africa	31	2	4	5	33	7
Eurasia	30	3	2		7	42
SEAN	15		3		42	10
Papuan	80		3	1	10	46
Australia	20		6	1	3	8
NAM	44	3	7	1	27	13
SAM	60	5	5		12	41
<b>TOTAL</b>	<b>280</b>	<b>13</b>	<b>30</b>	<b>8</b>	<b>134</b>	<b>125</b>

Taula 19 i Gràfic 12 La variable Adposició segons nombre de *genus* i continent.

El fet que a nivell areal hi ha uns pocs *genus* d'amplíssima distribució geogràfica i molts altres reclosos en espais minúsculs pot afavorir que es treguin conclusions equivocades si només es tenen en compte els mapes i prou. El cas més evident és el de les llengües de Nova Guinea que malgrat superar el centenar de *genus*, pràcticament no es distingeixen a nivell d'un mapamundi. Per això s'han analitzat les xifres del nombre de parlants, del nombre de llengües i del de *genus* per set àrees geogràfiques: Euràsia, SE asiàtic, Àfrica, Amèrica del nord i del sud, Austràlia i Nova Guinea.

La comparació dels tres gràfics corresponents a parlants, *genus* i llengües permet extreure una conclusió bàsica, si transcendim el fet que hi ha informació suficient per fer anàlisis molt més exhaustives que ocupin un munt de pàgines, si es comentés cada xifra. Estadísticament es confirma que cada cop que una llengua s'ha expandit, ja sigui en nombre de parlants o perquè s'ha ramificat en varies llengües filles que alhora han pogut donat lloc a altres *genus* germans, la probabilitat de ser o mantenir l'ús de postposicions disminueix. De manera que els *genus* amb més llengües i amb més parlants mostren clarament més proporció de color vermell i les més aïllades i petites tendeixen al color blau.

El color groc de la barra de població SEAN( sud-est asiàtic) és degut a una única llengua, el mandarí, que distorsiona el gràfic.

La tendència general és que el percentatge de postposicions és major als *genus* i disminueix a la barra equivalent del gràfic de llengües i encara més al de parlants. Aquesta tendència es compleix a la majoria d'àrees, excepte a Austràlia i Sudamèrica, i per poc. En el cas sud-americà, la raó està en un parell de *genus* andins, el quítxua i l'aimara, que representen a la majoria de parlants del continent.

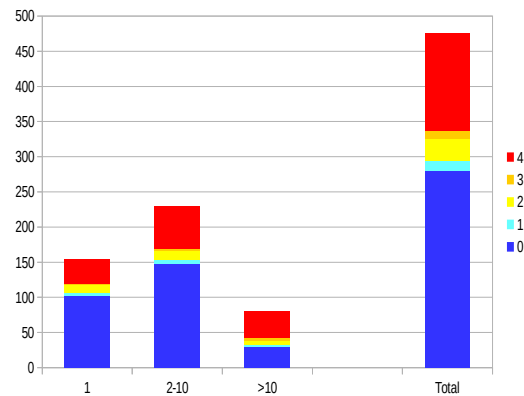
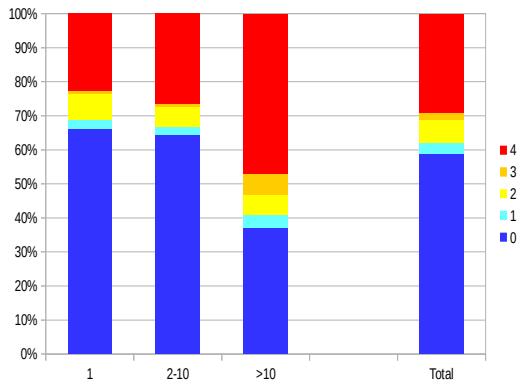
A nivell global hi ha una paritat entre llengües preposicionals i postposicionals, però al mirar per continents queda clar que només és casualitat, perquè les proporcions en cadascun són molt diferents.

En particular, el fet que a Austràlia i Nova Guinea l'ús de postposicions rondi el valor de 80% a tots tres gràfics, després de més de 50.000 anys d'aïllament mereix atenció i reflexió. És una evidència que no n'hi ha prou amb sostenir la igualtat de capacitat comunicativa entre els dos tipus d'adposició i el pas del temps, per pensar que s'arribaran a equiparar els percentatges de preposicions i postposicions.

## 5.6 ADPOSICIÓ I RAMIFICACIÓ

Existeix un consens en que la totalitat de *genus* i famílies lingüístiques que es parlen són descendents d'una protollengua que amb el pas del temps s'ha anat ramificant. En aquest apartat s'estudia com aquest procés ha alterat la distribució de la variable adposició

### 5.6.1 RAMIFICACIÓ DE *GENUS*



**Gràfic 13** Diagrama de barres de la taula 20 en percentatge nombres absoluts

**Gràfic 14** Diagrama de barres de la taula 20 en nombres absoluts

	0	1	2	3	4
1	102	4	12	1	35
2-10	148	6	13	2	61
>10	30	3	5	5	38
<b>Total</b>	<b>280</b>	<b>14</b>	<b>32</b>	<b>11</b>	<b>138</b>

**Taula 20** Nombre de *genus* segons llengües/*genus* i valor de la variable Adposició

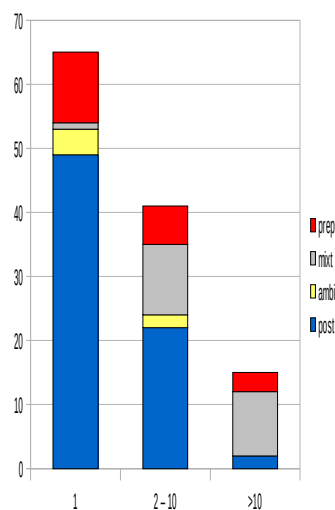
Els dos gràfics superiors fan referència al nombre de llengües per *genus*. El de l'esquerra en percentatge i el de la dreta en valors absoluts.

Les columnes que representen els *genus* amb un màxim de 10 llengües són majoritàriament postposicionals, mentre que en la columna dels que han assolit la xifra de més de 10 llengües, el percentatge de *genus* amb postposicions es redueix a la meitat i els preposicionals esdevenen majoritaris.

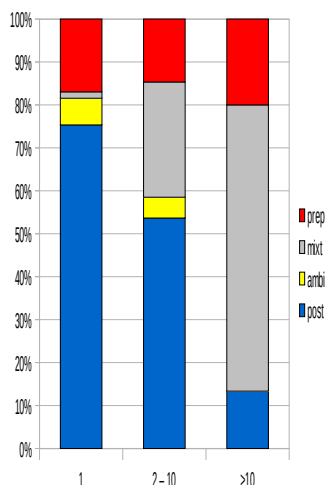
La conclusió és que una bona part dels *genus* que s'han ramificat de forma significativa, han adoptat de forma paral·lela l'ús de preposicions.



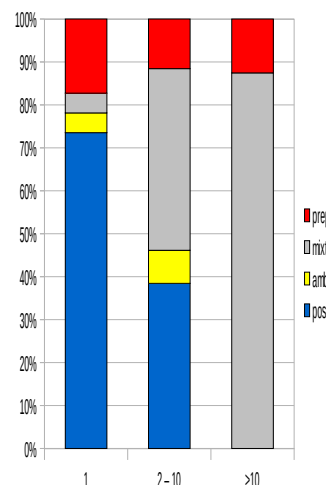
## 5.6.2 RAMIFICACIÓ DE FAMÍLIES



Gràfic 15



Gràfic 16



Gràfic 17

LLENGÜES/FAMÍLIA	post	ambi	mixt	prep	total
1	49	4	1	11	65
2-10	22	2	11	6	41
>10	2	0	10	3	15
total	73	6	22	20	121

Taula 21

genus/FAMÍLIA	post	ambi	mixt	prep	total
1	64	4	4	15	87
2-10	10	2	11	3	26
>10			7	1	8
total	74	6	22	19	121

Taula 22

Els tres gràfics de dalt fan referència a les famílies lingüístiques. Els dos de l'esquerra són de llengües/família (en percentatge i xifres absolutes) i el de la dreta de *genus*/família en percentatge. Cada gràfic consta de 3 columnes. Una per llengües, o *genus* segons el cas, aïllades. Una altre per aquelles amb un nombre limitat d'elements, de dos a deu, i un per les famílies expansives amb més de deu llengües o *genus*.

Els resultats mostren unes variacions extraordinàries. En el cas de *genus* o llengües aïllades, el percentatge d'ús de postposicions domina de forma aclaparadora fins assolir percentatges superiors al 70% i pràcticament no hi ha casos de mixtes o neutres.

En el cas de famílies molt ramificades la situació és la inversa. No hi ha cap família de llengües postposicionals que arribi als deu *genus*, i les que han arribat a la meta molt més modesta de les deu llengües són una minoria que apenes supera el 10%. Entre les famílies molt ramificades el predomini de *genus* mixtes és molt predominant.

Això també és important. Les famílies lingüístiques identificades fins ara no tenen molt més de 10.000 anys i totes se suposen descendents d'una protollengua. Això implica que en elles, la ramificació ha anat paral·lela a un progressiu canvi en l'ordre de adposició i el Nucli, com a mínim d'algunes llengües.

La conclusió final és que a mesura que un *genus* o una família s'expandeix, tant si és en nombre, de llengües o de *genus*, una part d'aquests passa de fer servir postposicions a preposicions, fins a l'extrem que a les famílies més ramificades ja no queda cap *genus* on no s'hagi començat a produir aquest procés de canvi.

## 5.7 ALTRES FACTORS

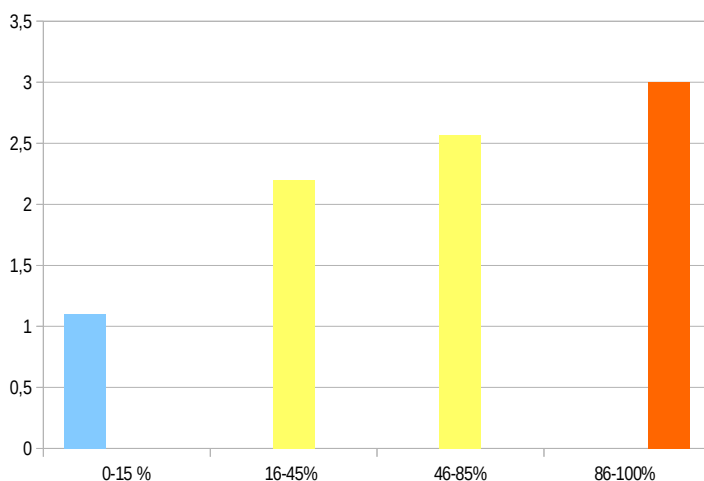
En la realització d'aquesta tesi s'han elaborat moltes taules necessàries per poder quantificar relacions entre diverses variables lingüístiques i socials. Un cop acabada, ha resultat interessant aprofitar-ne algunes per comprovar algunes hipòtesis proposades per diversos autors com els comentats a l'apartat 2.1.3

### 5.7.1 ADPOSICIÓ I DEPENDÈNCIA DE L'AGRICULTURA

En aquest apartat es comenta el fet que hi ha una relació entre el percentatge d'aliment que s'obté de l'agricultura i la tendència a usar un tipus o altres d'adposicions.

El gràfic d'aquesta pàgina mostra que entre les comunitats amb poca dependència de l'agricultura (menys d'un 15% de), la mitjana de la variable adposició del *genus* al que pertany la seva llengua està al voltant d'1 (majoritàriament postposicional). Mentre que arriba a 3 (majoritàriament preposicional) en el cas d'aquelles amb una dependència superior al 85%. Els casos intermedis mostren valors intermedis.

Tal com s'ha explicat a l'apartat 3.1.4, un valor 1 implicava un ús de postposicions entre el 70% i el 90% dels casos, mentre que el 2 anava del 30% al 70% i el 3 entre 10% i 30%, d'acord a la taula 9. En altres paraules, els pobles agrícoles fan servir llengües predominantment preposicionals i els caçadors-recolectors llengües predominantment postposicionals.

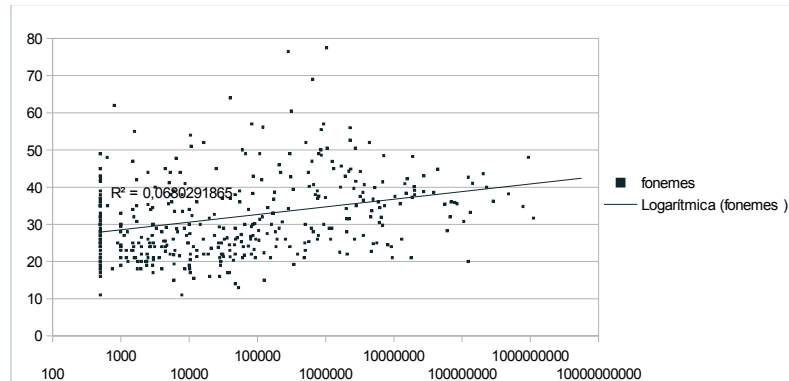


**Gràfic 18.** Percentatge de dependència de l'agricultura i tipus d'adposició.

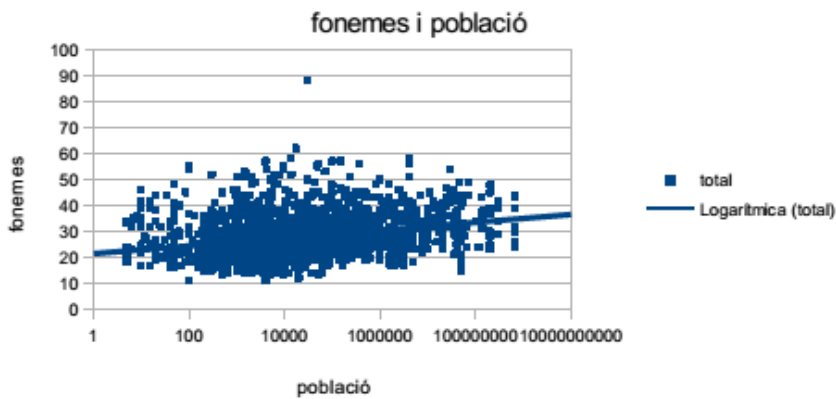
Cal tenir en compte que el gràfic només té en compte l'agricultura i no altres activitats com la ramaderia o la pesca que en moltes societats neolítiques o urbanes poden tenir un pes similar a l'agricultura.

## 5.7.2 POBLACIÓ, FONEMES I ORDRE DE CONSTITUENTS

A l'apartat 1.1.3 s'han comentat diferents autors que han proposat l'existència d'una correlació entre la població i el nombre de fonemes. Donat que a la tesi s'ha tractat justament de la relació entre el nombre de parlants i l'ordre de constituents, s'ha aprofitat la informació de les taules per fer una ràpida ullada i veure que en surt.

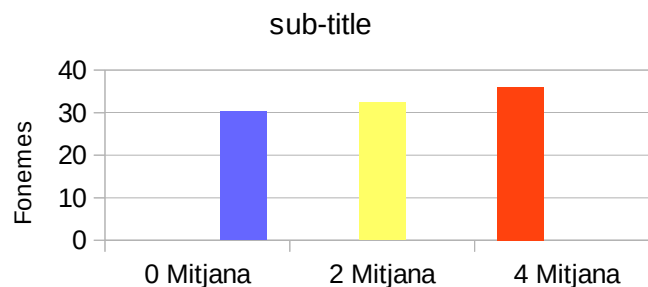


Gràfic 19 Nombre de fonemes i població, per genus. Dades personals.



Gràfic 20 Nombre de fonemes i població, per llengua. Dades personals.

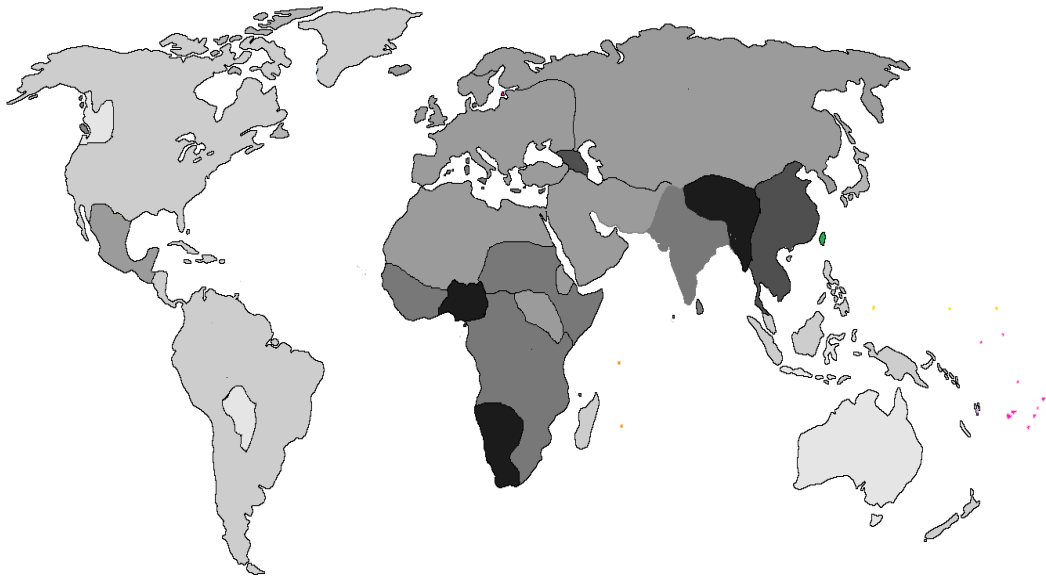
Als gràfics superiors es correlaciona el nombre de fonemes amb el nombre de parlants. Al superior, a nivell de *genus*, i al de sota per llengua. A tots dos es repeteix l'increment de fonemes a mesura que s'incrementa la població, tal com han dit els autors citats. La correlació és baixa i la distribució adopta una forma de nuvol que pot fer dubtar que sigui altre cosa que un artefacte sense significació.



Gràfic 21 Nombre de fonemes segons pertinença a valor d'Objecte verb, per *genus*.

El gràfic superior mostra com els *genus* amb ordre VO mostren un nombre de fonemes inferior als d'ordre no dominant i encara més baix que als d'ordre VO. La diferència representa casi 6 fonemes més a les VO dels 30 de mitjana de les OV, un 20%.

### 5.7.3 GEOGRAFIA DEL NOMBRE DE FONEMES, WL, I DISTÀNCIA A L'ÀFRICA.



**Mapa 11** Mapa de nombre de fonemes per àrea

Atkinson va sorprendre amb la hipòtesi que el nombre de fonemes de les llengües anava minvant a mesura que augmenta la distància al sud-oest Africa com s'ha comentat a 2.1.3.

Al mapa de fonemes, s'ha representant la mitjana de fonemes dels *genus* dins unes àrees definides en funció del seu valor. Curiosament el resultat és compatible amb el d'Atkinson, però amb una segona tendència més.

Es pot veure que el màxim de fonemes és al territori de les llengües khoisanides, amb més de 80. Després la xifra baixa fins poc més de 60 a la resta d'Àfrica i als 49 a Europa fins l'Índia. Sibèria i Nord-Amèrica. Per últim a Oceania i Amèrica del sud la mitjana de fonemes es queda en poc més d'una vintena.

Al costat d'aquesta tendència relacionada amb les migracions humanes que van poblar el planeta en temps prehistòrics hi ha una segona tendència superposada. Les regions com Nigèria, Sud-est asiàtic, Mesoamèrica o Colúmbia Britànica i potser la regió Euromediterrània, hi ha un increment de fonemes respecte les zones del voltant. Aquestes zones, com es veurà posteriorment, són zones on possiblement s'ha produït un canvi tipològic de llengües postposicionals a preposicionals.

Cal remarcar que aquest mapa corrobora els resultats d'Atkinson [41] i Wichmann et al. [44], malgrat que totes tres aproximacions s'han fet amb metodologies totalment diferents i ni tant sols sobre es mateixes variables. Perquè mentre el primer estudia nombre de fonemes per llengua i el segon els fonemes per llengua que apareixen en una transcripció simplificada d'un llistat de paraules, en el meu cas he calculat la mitjana de fonemes per *genus* i àrea geogràfica.



**Mapa 12** Mapa de WL per àrea

El mapa de WL mostra la longitud mitjana en fonemes d'una mostra escollida de la llista de Swadesh de l'AJSP. Amb alguna excepció, sembla concordar força amb el nombre de fonemes, de manera que es confirmaria que a major nombre de fonemes, les paraules són més curtes.

Regions allunyades d'Àfrica com Austràlia, Nova Guinea, Amèrica o Sibèria mostren el màxim de longitud de paraula, mentre que les regions com el SE asiàtic, Nigèria, Mesoamèrica o la Columbia Britànica mostren una reducció respecte les seves regions veïnes.



## 6 DISTRIBUCIÓ DE LES VARIABLES

En aquesta tesi s'ha fet servir informació procedent de diverses bases de dades, però la falta d'informació d'aquestes sobre alguns *genus*, ha portat a consultar directament en gramàtiques i bibliografia especialitzada sobre llengües particulars, per tal d'omplir algunes cel·les en blanc. La gran varietat de fonts comporta que no hi hagi un criteri únic a l'hora d'atribuir valors a cadascuna de les llengües per cada variable, ja que els diferents compiladors han seguit els seus propis criteris, no sempre coincidents.

En WALS [95] hi ha un capítol sencer dedicat a aquesta qüestió que pot servir per tenir una visió general del problema. Potser la principal idea és que sota una determinada opció, com per exemple l'ordre Adjectiu.-Nom, s'agrupen dos tipus de possibilitats diferents. La de les llengües que la presenten de forma exclusiva i la d'aquelles que malgrat que puguin tenir aquest ordre la major part de les vegadesdo, no exclouen altres ordres alternatius.

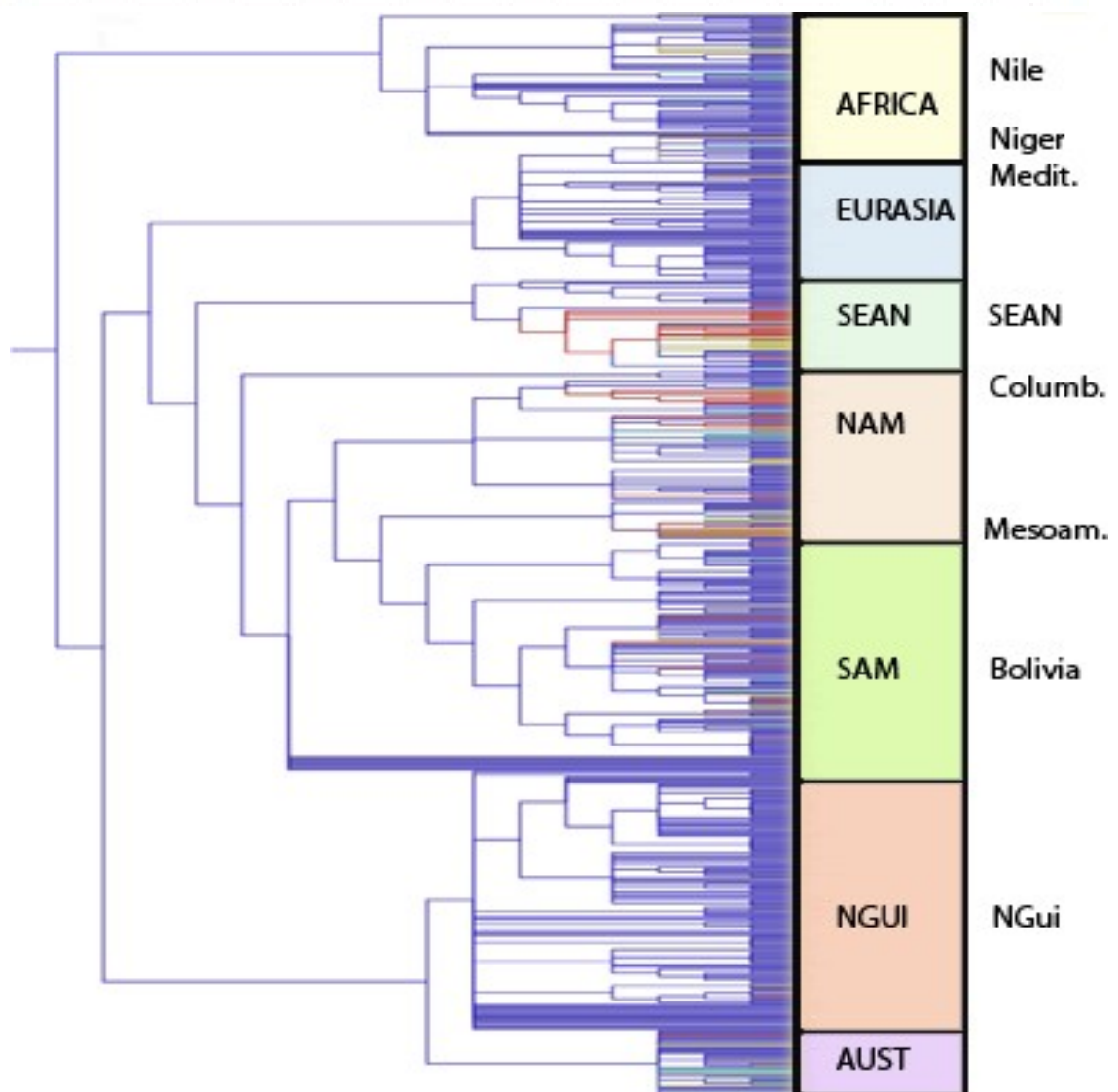
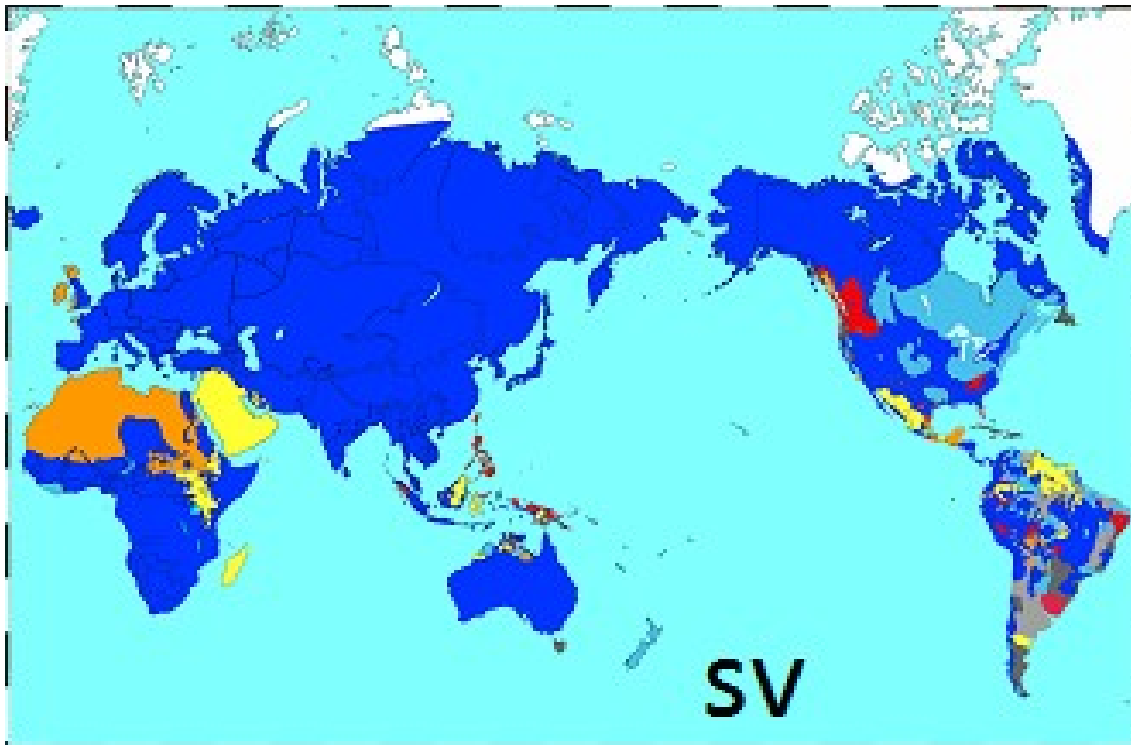
A moltes llengües, com al català mateix, la majoria d'adjectius segueix el nom, però hi ha petit grup que poden precedir-lo (com pot ser el cas de *gran*). Això pot provocar dubtes a l'hora de valorar-ho que es resolen de forma no coincident pels diferents autors. El mateix es pot donar en altres variables com adposicions, demostratius, possessius, etc. També resulta conflictiu el cas dels numerals ja que algunes llengües canvien l'ordre a partir d'una determinada xifra.

Aquesta diferència de criteris i metodologies usades s'ha solucionat establint unes normes d'homogeneïtzació per tal de convertir cada valor individual en una taula en un dels 5 resultats possibles (+10,+5,0,-5 i-10) que s'han comentat a l'apartat 3.1.4.

Els colors dels mapes i arbres sempre són els de la taula 9.

La diferent valoració dels diferents autors s'ha tractat de solucionar fent la mitjana entre ells, però el sistema no funciona quan són tan contradictoris que impliquen un clar error en alguns d'ells. En aquest cas s'ha optat per rebutjar els més antic o d'autors i publicacions que semblaven menys prestigioses.

## 6.1 SUBJECTE





L'ordre entre subjecte i verb és la variable que mostra menys equilibri. La distribució de l'ordre VS és força marginal al llarg del món i només és totalment dominant a la regió de la Colúmbia britànica. En altres zones sembla que l'ordre està en reflux a partir d'una distribució passada més extensa, per exemple entre les llengües semítiques i austronèsiques. Sovint s'ha proposat una tendència universal del subjectes d'ocupar la posició del Tòpic precedint el verb [13].

A l'Àfrica no hi ha cap *genus* amb ordre dominant Verb-Subjecte, encara que aquest ordre és fàcil de trobar entre les llengües d'alguns grups. Per una banda entre les llengües sudàniques orientals La llengua hadza, encara que allunyada filogenèticament d'aquestes, n'és veïna i mostra una ambivalència possiblement resultat del contacte.

Però l'ordre VS també es troba entre les afroasiàtiques més properes al Mediterrani (semítiques, berbers i antic egipci). Aquí és molt important destacar que històricament la tendència ha estat cap un canvi a l'ordre majoritari SV, tant a la branca egípcia com a la semítica i possiblement a la berber.

A Euràsia trobem llengües amb ordre VS als dos extrems oposats. A l'oest, entre les llengües cèltiques i a l'est a la família austronèsica amb més de la meitat dels *genus* VS (entre ells la totalitat dels de les illes Filipines) i la resta repartits entre SV i mixtes. En aquesta família també sembla clar que s'ha donat una regressió des de l'ordre original amb subjecte inicial, reconstruït com el més parsimoniós, cap el que domina a nivell global. Madagascar, com territori africà on es parlen llengües austronèsiques presenta llengües VS.

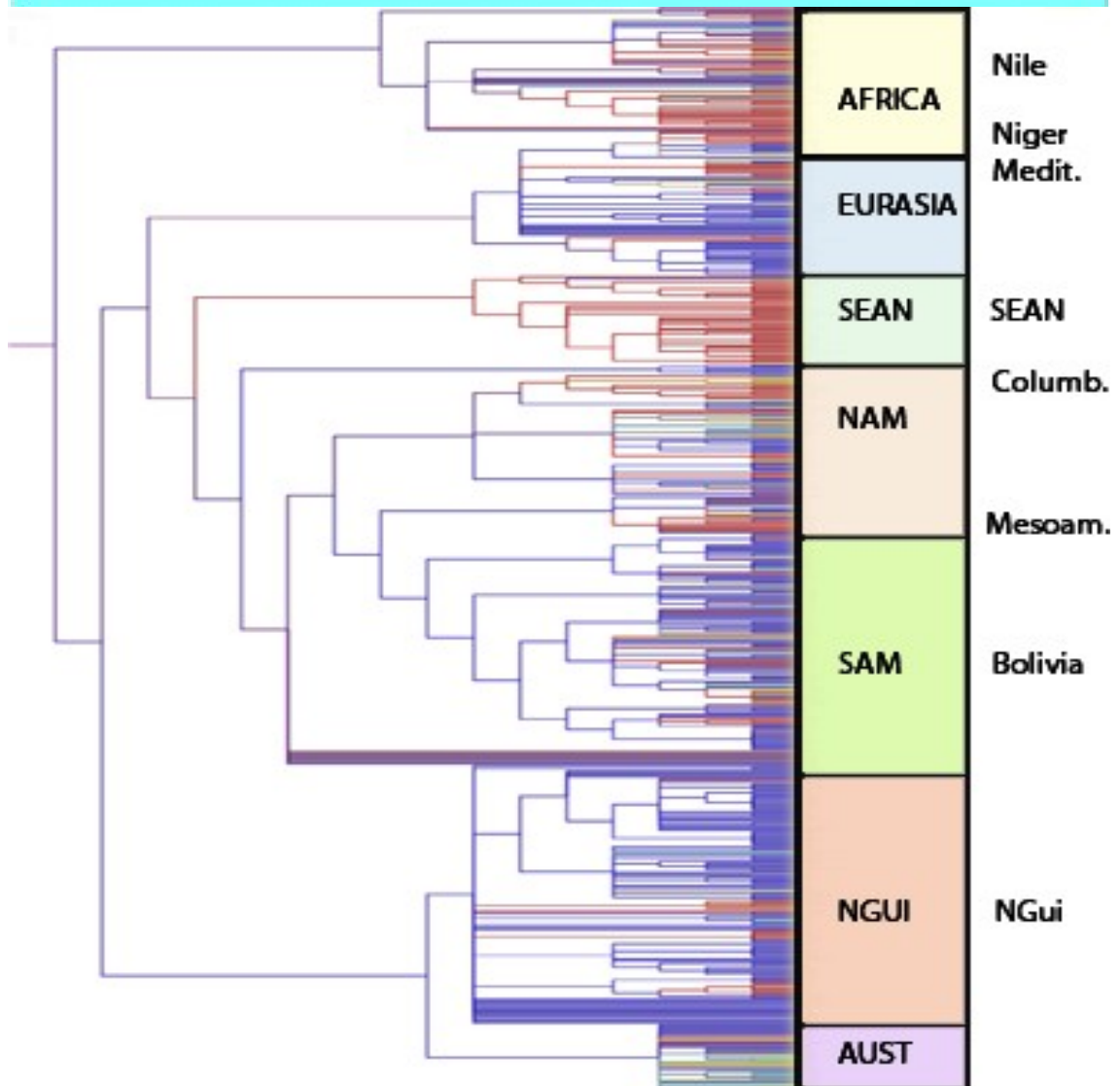
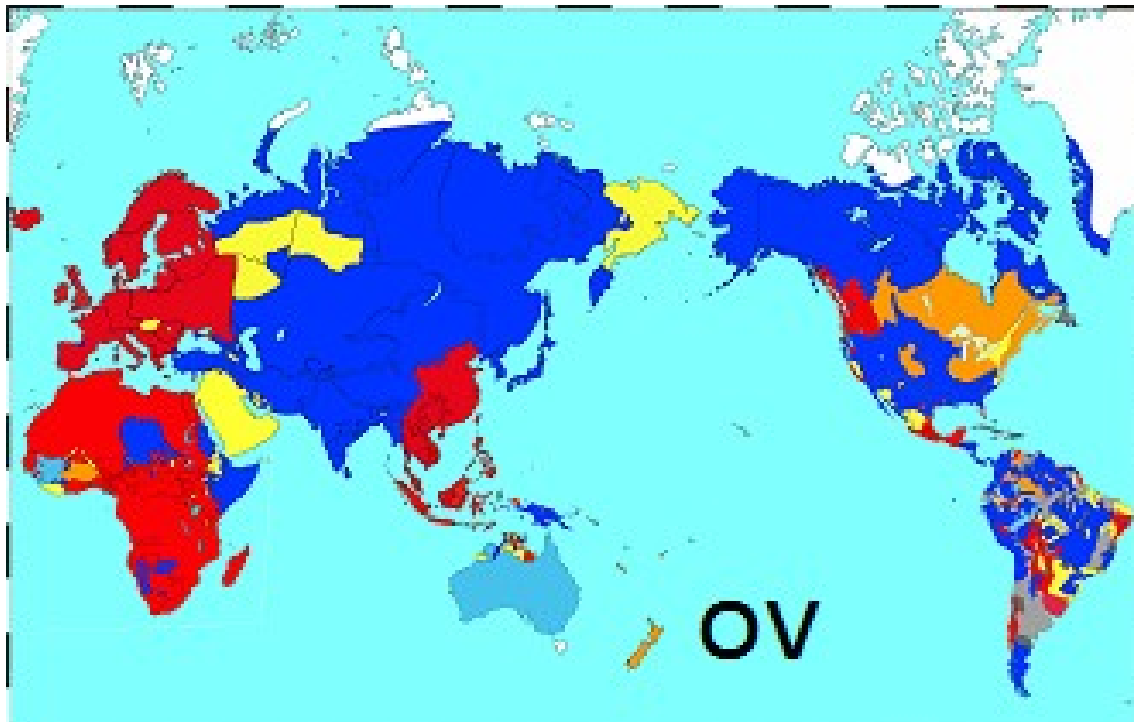
Amèrica és el continent amb més presència de *genus* amb ordre Verb-Subjecte, però bàsicament a tres àrees concretes, com es pot veure a mapa de la pàgina precedent. La regió que anomenarem de la Columbia Britànica, encara que els seus límits no coincideixen exactament, l'Amazònia boliviana i a Mesoamèrica, en ordre de major a menor consistència. A la primera pràcticament només trobem *genus* amb ordre VS, a la segona hi ha un cert predomini, però amb nombroses excepcions. A la regió mesoamericana hi ha una situació molt més equilibrada sense un ordre dominant. A part, d'aquestes zones, alguns altres *genus* amb llengües VS és troben de forma dispersa per la resta del continent.

A la regió de Nova Guinea no hi ha cap llengua VS tret del Kuot que es troba en una illa rodejat de llengües austronèsiques i potser l'ha canviat per aquest contacte.

S'ha de tenir en compte que el fet que es digui que un *genus* té un ordre determinat no vol dir que no hi puguin haver llengües amb ordres diferents, sinó que aquestes suposen una proporció tan baixa que no impedeix l'assignació de la categoria. Així és troben llengües sense un ordre dominant a diversos *genus* d'Europa i Nord-Amèrica, però sense que afecti a la mitjana del conjunt.

Possiblement el fenomen més interessant que s'observa en aquesta variable és el fet que les zones on abunden més les llengües VS siguin subàrees d'altres que es troben en altres variables, juntament amb el possible reflux.

## 6.2 OBJECTE



L'ordre entre l'objecte i el verb és una variable amb una distribució geogràfica força clara i les llengües OV i les VO semblen repartir-se el món de forma equilibrada.

El mapa d'Àfrica ofereix una primera impressió de domini pràcticament total dels *genus* amb ordre VO, amb l'ordre OV reclòs a la banya d'Àfrica, voltants de Guinea i al centre dels deserts del Sàhara i el Kalahari. Però aquesta visió és deguda a l'enorme superfície ocupada pel *genus* bantu, el berber i la llengua àrab que ocupen junts més del 80% de la superfície africana.

En realitat ambdós ordres es reparteixen Àfrica d'una manera molt més equilibrada, encara que els *genus* VO segueixen més que doblant en nombre als OV, però quedant lligats de forma clara a dos territoris concrets al voltant dels trams dels rius Níger i Nil més allunyades dels rigors del Sàhara. Les llengües khoisanides del nord i sud són d'ordre VO, com les bantus que les envolten, però possiblement és degut a la influència d'aquestes, ja que les del centre que viuen a zones més aïllades, es mantenen l'ordre OV.

Les llengües berbers, semítiques i l'antic egipci constitueixen també un àrea VO a l'Àfrica, però possiblement la seva tipologia té a veure amb el Pròxim Orient igual que passa amb Madagascar respecte el sud-est asiàtic.

A Euràsia, igual que passava a la variable subjecte, també hi ha llengües de Nucli inicial als dos extrems del continent, però en àrees molt més esteses. Ara no són només les llengües cèltiques sinó pràcticament tot Europa, amb alguna excepció com el basc. A l'extrem contrari, ja no són les llengües austronèsiques les úniques, sinó que l'ordre VO s'estén al conjunt del sud-est d'Àsia, tret de Birmània i també al *genus* xinès.

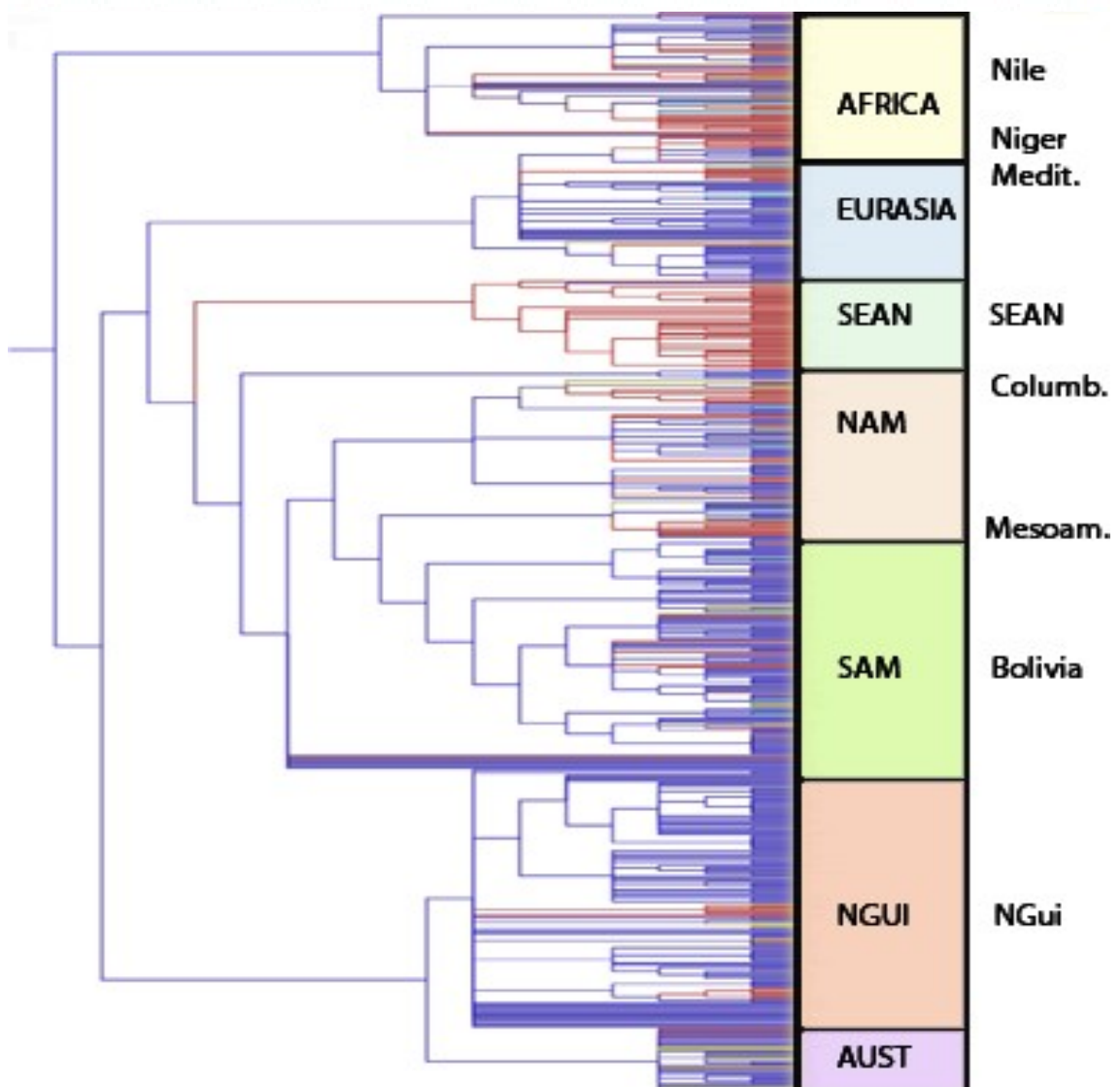
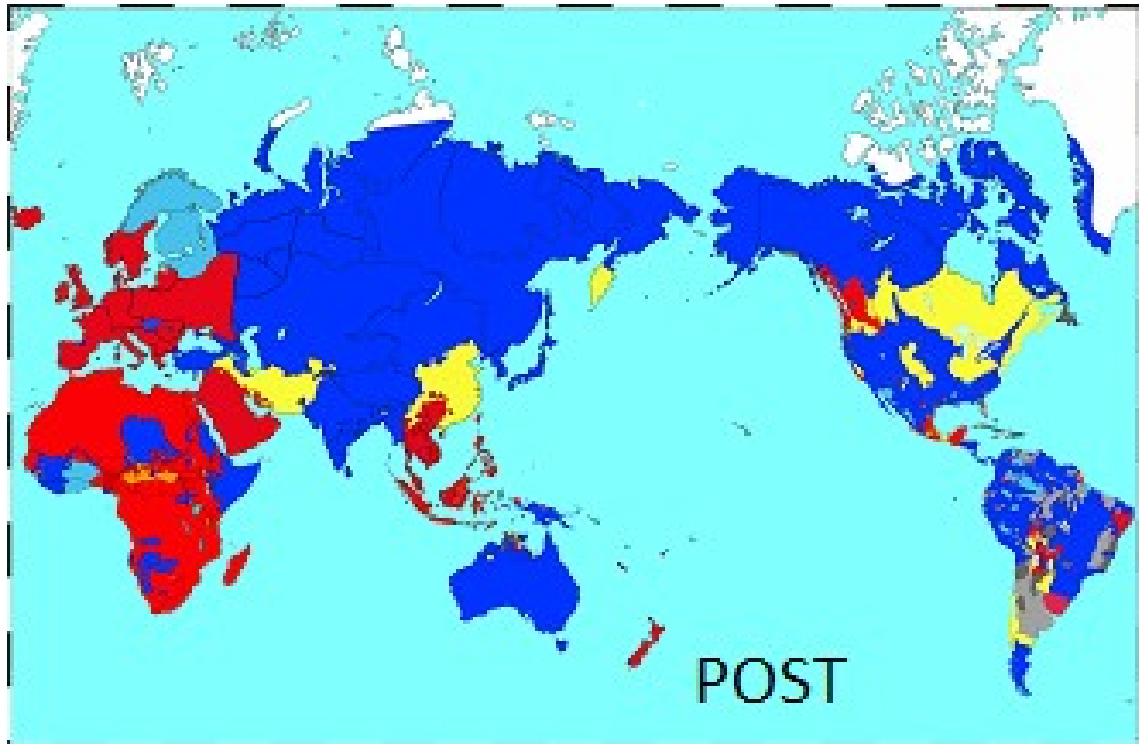
La resta del continent, des de Birmània a Anatòlia passant pel Caucas, Sibèria, el Tibet o el Japó són inequívocament OV. Les poques excepcions són alguns *genus* on la influència de llengües eslaves sembla haver portat a un canvi. Les llengües semítiques han suposat un dubte l'hora de fer el mapa ja que malgrat que són un *genus* mixt, això és perquè a Euràsia tendeixen a VO i a l'Àfrica a OV. Tal com s'ha comentat anteriorment, el cas de les llengües europees i mediterrànies és força interessant perquè el registre escrit ha documentat que l'ordre dominant avui ha estat el resultat d'un canvi total que s'ha produït en els darrers dos mil·lennis.

A les Amèriques trobem les mateixes 3 àrees divergents de nucli inicial comentades a la variable Subjecte, destacades sobre un fons blau de llengües OV. A l'Amèrica del Nord, l'àrea de la Colúmbia britànica és veu perllongada cap a l'est per la forta expansió de les llengües algonquines, *genus* de categoria 3, mentre que a Mesoamèrica i l'Amazonia boliviana molts *genus* sense un ordre dominant pel que fa al subjecte i el verb, es decanten clarament per l'opció de nucli inicial en el cas de l'objecte.

A Nova Guinea, els *genus* d'origen papú són molt majoritàriament OV i les excepcions s'acostumen a atribuir a la presència de llengües austronèsiques, tal com es comenta de forma més ampla a l'apartat 7.7

Per acabar, Austràlia presenta un ordre bàsic OV [66], però les gramàtiques de les seves llengües acostumen a manifestar una gran llibertat i això fa que es puguin trobar més *genus* amb ordres distints dels que caldria esperar a priori. Factors com l'animacitat poden explicar algunes d'aquestes excepcions.

### 6.3 ADPOSIÓ



Tal com ja s'ha comentat abans, l'adposició coincideix en gran manera amb l'objecte i els mapes i arbres de les dues variables es corresponen a grans trets. Per això bona part del que aquí s'explica és una reiteració de l'apartat anterior.

La diferència principal és que a la majoria de regions que s'han identificat per agrupar *genus* amb nucli inicial en el cas de l'objecte, quan es comparen amb l'adposició es desdibuixen una mica i mostren una superfície menor.

Així a l'Àfrica hi ha un predomini geogràfic dels *genus* amb ordre preposicional, degut en bona part a l'àrea ocupada per llengües bantus, berbers i l'àrab, però diversos *genus* de les àrees del Níger i del Nil mostren una relaxació en la tendència a l'ordre de nucli inicial que es veia a l'Objecte, com passa amb els *genus* Gur i Kwa al golf de Guinea o la major part de les llengües sudàniques centrals. En el cas de les llengües khoisanides, la suposada influència de les llengües bantus no hauria arribat al *genus* del nord. En canvi, les llengües semítiques són preposicionals a les dues ribes del mar Roig.

A Euràsia també trobem les dues àrees discordants a Europa i sud-est asiàtic de llengües amb preposicions, però alguns *genus* varien d'ordre en comparació a l'ordre de l'objecte i verb. Els casos més evidents són alguns *genus* de la família Uràlica que són postposicionals de forma clara i el *genus* xinès que no presenta cap ordre dominant.

En el cas d'Europa i del Pròxim Orient, la literatura d'aquestes llengües els darrers mil·lennis permet veure que no s'ha donat a les adposicions uns canvis tan nombrosos com els que semblen haver-se donat a la variable objecte.

A l'àrea de la Colúmbia Britànica hi ha alguns *genus* que malgrat ser VO a la variable objecte, pel que fa a les adposicions són total o parcialment postposicionals: Kutenai, Sahaptian, Costanoan, Palaihninan. Altres *genus* en canvi, malgrat no tenir un ordre a la variable objecte, són postposicionals com en el cas dels *genus* Coosan, Miwok i Karok. El fet que moltes llengües algonquines no tinguin adposicions ajuda a definir millor aquesta àrea.

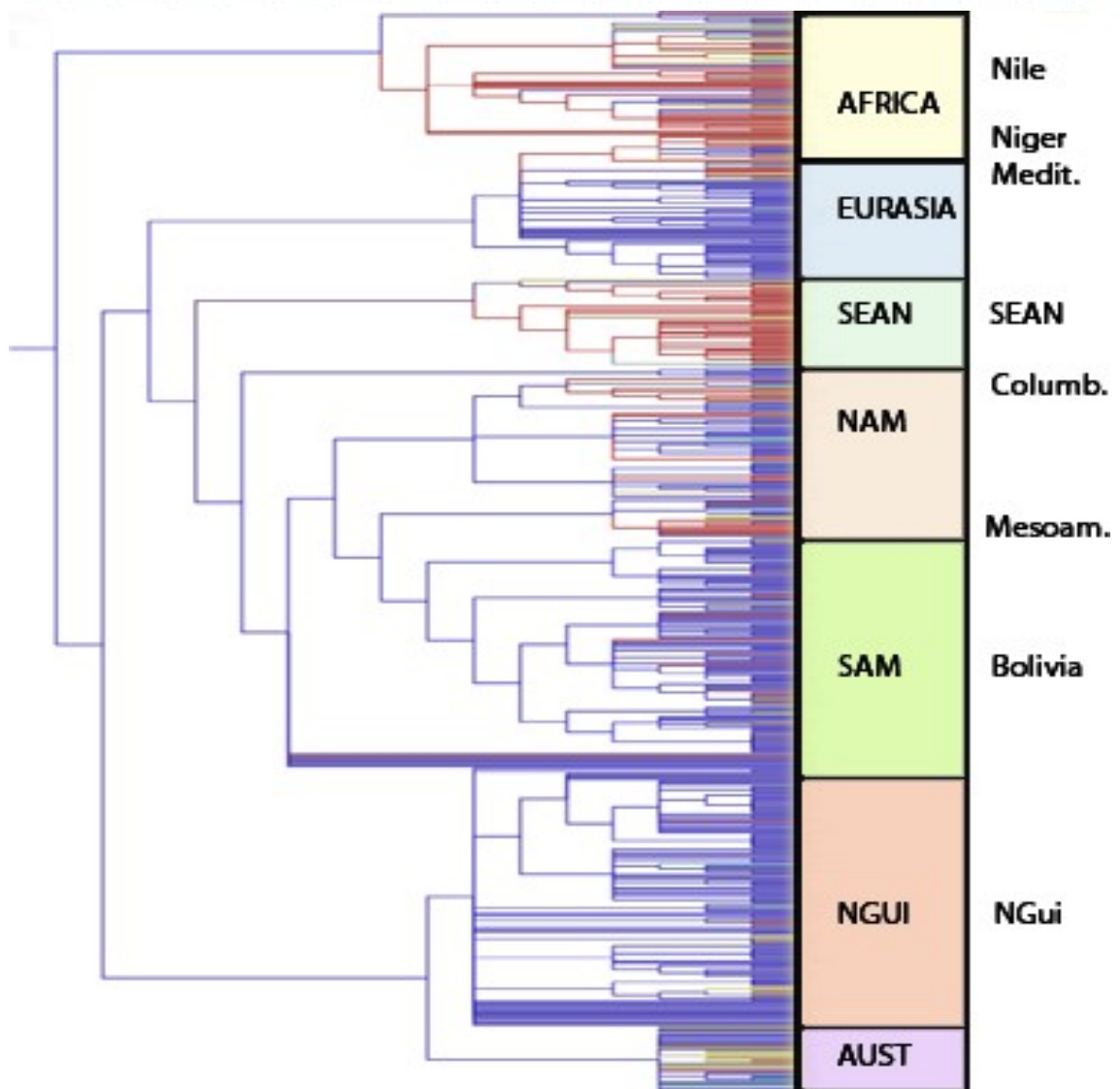
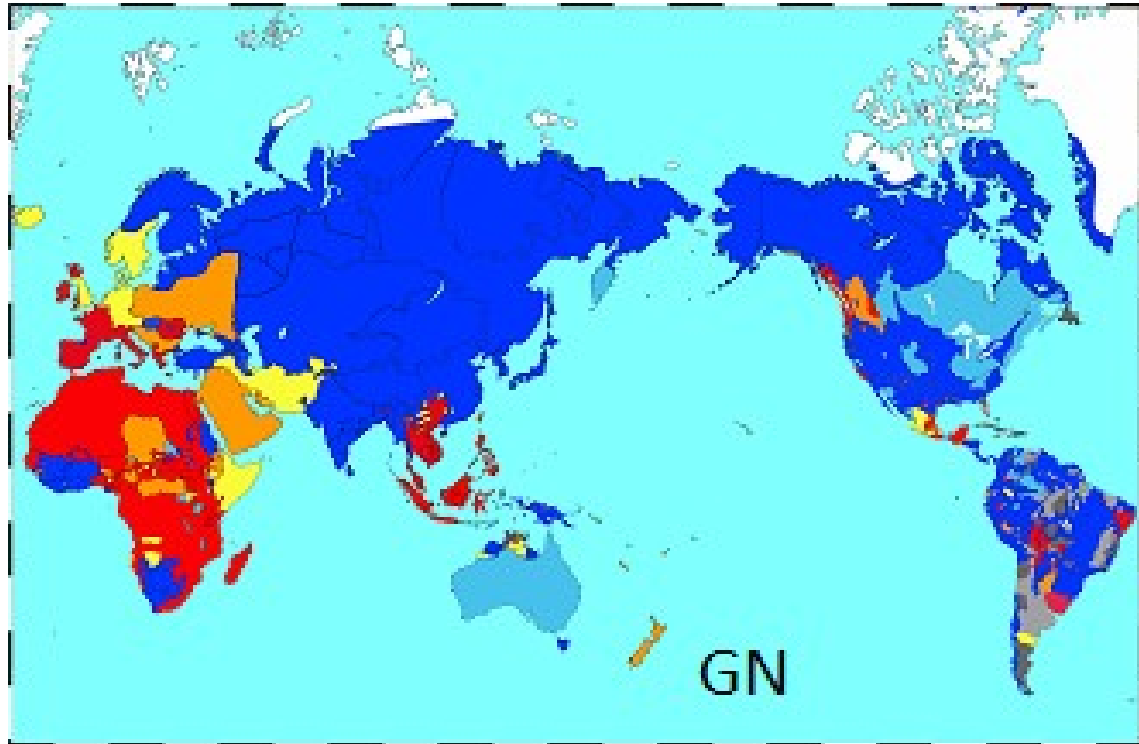
A Mesoamèrica la tendència es repeteix, però suavitzada. Només 2 *genus* amb ordre VO (Mixe-Zoque i Tarascan) no són preposicionals i dos més que no mostren dominància en l'objecte (Tepiman i Corachol) si que la mostren en l'adposició i decantant-se per les postposicions.

A l'Amazonia boliviana trobem uns 5 *genus* amb un ordre dominant VO que són postposicionals (Mosetenan, Arawak del nord i Arawak pre-Andí) o sense un ordre adposicional dominant (Guaicuruan i Bolívia-Parana)

A Nova Guinea, només *genus* Bilua és VO i postposicional. A la resta dels que hi ha informació, l'opció nucli final o inicial és compartida tant per objectes com per adposicions.

Per acabar, Austràlia presenta a les adposicions la mateixa llibertat d'ordenació que ja s'ha comentat per l'objecte, però en un 80% dels *genus* la tendència al nucli final es veu accentuada per 2 en contra (Murrinh-Patha i Anindilyakwa). En canvi només dos *genus* són alhora VO i preposicionals (Mangarrayi i Tiwian), potser per criteris d'animacitat.

## 6.4 GENITIU



La variable genitiu està fortament correlacionada amb l'objecte i l'adposició tal com ja s'ha comentat, per això arbre i mapes són força similars a primer cop d'ull. La diferència principal és que a la majoria de regions caracteritzades com de nucli inicial a les variables objecte i adposició, quan es comparen amb el de la variable genitiu, es mostren més reduïdes i menys consistents.

A l'Àfrica ens trobem igualment amb el predomini geogràfic dels *genus* amb nucli inicial, degut a l'àrea ocupada per llengües bantus, berbers i de l'àrab. Els *genus* que estan més enllà de Nigèria cap a l'oest, com el Kru, el Kwa, el Gur o el Mande de l'est, són de nucli final, a diferència dels casos d'adposicions i objecte. A la banya d'Àfrica no hi ha ordre dominant. En canvi a l'àrea del Nil, el color vermell ocupa un àrea similar al parell de variables mencionades. A les llengües khoisanides, l'absència de *genus* amb nucli final indica que no hi ha cap influència de llengües bantus.

A Euràsia també trobem també dues àrees distals, a Europa i sud-est asiàtic. A la regió Euromediterrània hi ha una barreja pel que fa als ordres que trobem. Els colors coincideixen força amb el mapa de la variable adposició, però alguns *genus* varien cap a una menor rigidesa de les opcions amb nucli inicial. Al sud i oest, els genitius segueixen als substantius mentre que a l'est passa el contrari, i entremig hi ha les llengües germàniques sense un ordre dominant. Al sud-est asiàtic, el color vermell ocupa un àrea similar al de les variables adposició i objecte, però el *genus* xinès és de nucli final.

En el cas d'Europa i del Pròxim Orient, la literatura d'aquestes llengües mostra canvis a la majoria de llengües indoeuropees, però no sempre en un mateix sentit (apartat 2.2.6).

A l'àrea de la Colúmbia Britànica també hi ha alguns *genus* que malgrat ser VO, són de nucli final pel genitiu, com el cas del *genus* Sahaptian i algunes de Califòrnia, com els *genus* Palaihnan i Costanoan, la qual cosa perfila millor l'àrea.

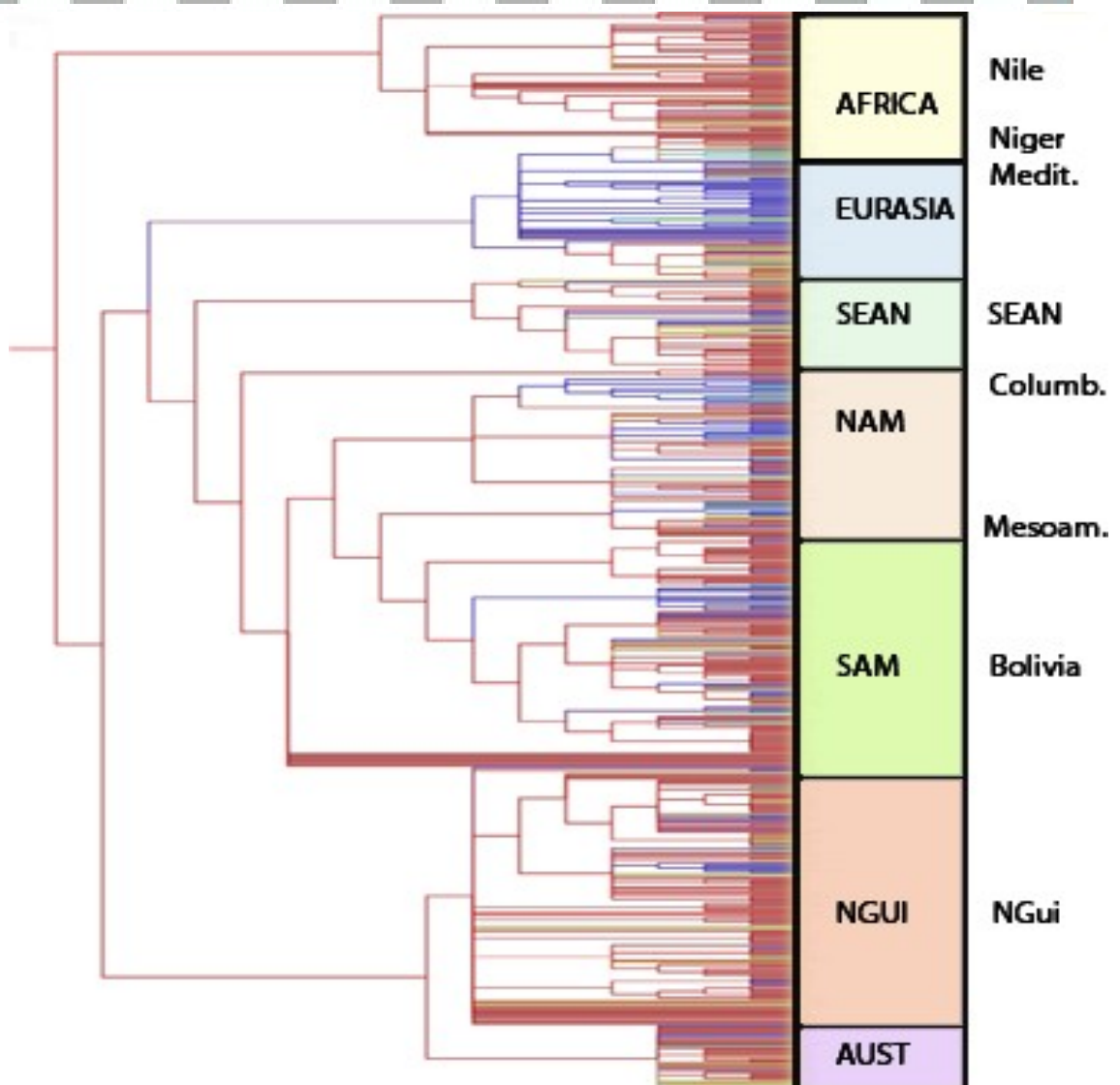
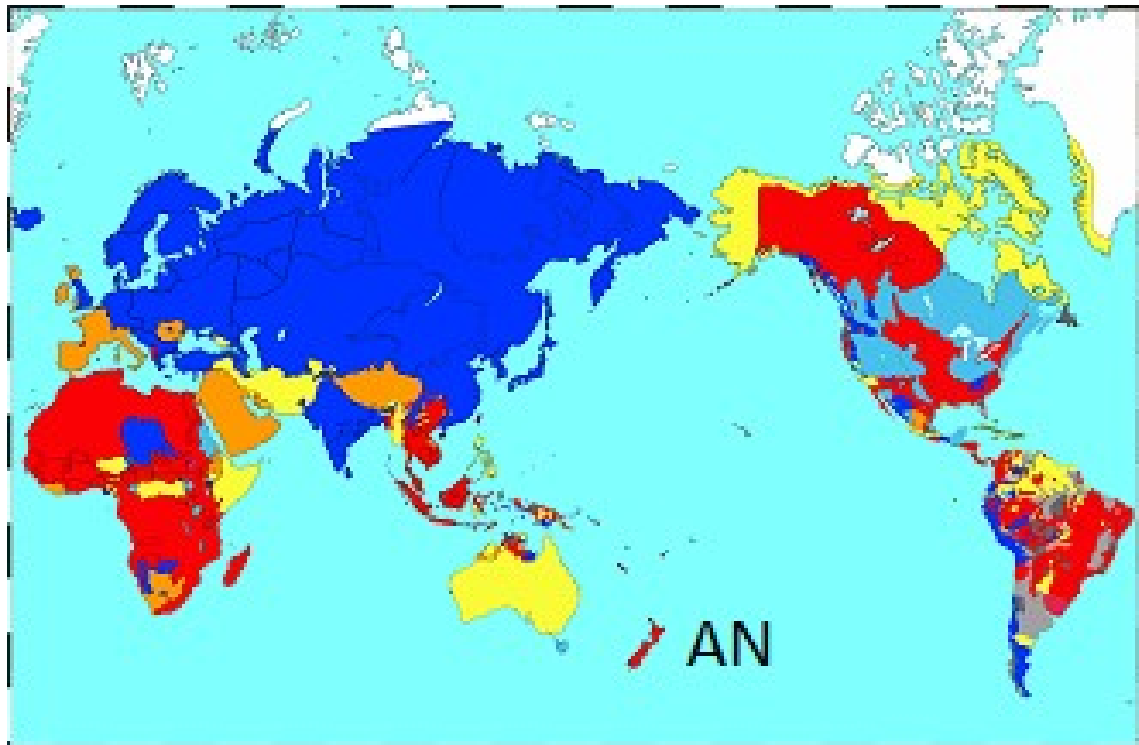
A Mesoamèrica l'àrea amb ordre Nom-Genitiu està fragmentada entre el territori maya, Oaxaca i una tercera més al nord. *Genus* com l'asteca, el tarasc i Corachol no mostren un ordre dominant mentre que les llengües Mixe-zoque són de nucli final.

A l'Amazonia boliviana trobem una distribució dels *genus* de nucli inicial similar a la de la variable objecte, excepte en el cas d'alguns *genus* amb el genitiu precedint al substantiu com en el cas dels Zamucoan, Mataco i Moseten.

A Nova Guinea, la pràctica totalitat de *genus* papús mantenen un ordre Genitiu-Nom, amb algunes excepcions com el *genus* Kayagar, Krisa, Kuot i Warapu que són d'ordre Nom-Genitiu. Alguns altres més no mostren cap ordre dominant, però no modifiquen la visió global que ofereixen.

Austràlia té uns dos terços de *genus* amb genitius que precedeixen al nom, incloent-hi el Pama-Nynga que pertany a a categoria 1 (entre 10% i 30% de llengües amb nucli final pel que fa a la variable), la qual cosa tenyeix de blau clar el mapa. El terç restant no presenta ordre dominant excepte tres *genus* amb ordre Nom-Genitiu: Wagiman, Laragiya i Murinh-Patha.

## 6.5 ADJECTIU





Tal com s'ha vist als mapes de l'apartat 5.4, l'ordre dels adjectius no sembla compartir patrons amb altres constituents i la majoria de les àrees que presenten tendències compartides per les variables objecte i adposició queden partides al mirar els adjectius.

El mapa d'Àfrica ofereix un aspecte en vermell que acredita una ampla distribució dels *genus* amb adjectius que segueixen al substantiu. Les excepcions a la regla es troben al centre del Sahara, la Banya d'Àfrica i entre llengües khoisanides, però al contrari que en variables com les adposicions, aquest gran predomini al mapa es veu confirmat al mirar l'arbre filogenètic. Els *genus* africans on l'adjectiu precedeix el substantiu són una mitja dotzena.

A Euràsia els *genus* on els adjectius sempre van davant del nom només es troben al sud-est d'Àsia (tret d'Albània) i encara allà es troben força excepcions, especialment a la zona de les Filipines). A banda d'aquestes, hi ha un seguit de *genus* entre el Tibet i Irlanda de categoria 3, això és amb un predomini Nom-Adjectiu, però amb excepcions abundants. Serien els *genus* Bodic, Semític, Romance i Cèltic. Les llengües iràniques no mostren un ordre dominant.

La lingüística històrica ha documentat molts dels canvis en aquesta variable a Euràsia, tal com s'ha comentat a l'apartat 2.2.6. I queda clar que la direcció del canvi sempre ha sigut lliure i sense cap tendència a convergir amb l'ordre de les adposicions o del verb i l'objecte. El ritme dels canvis és superior al de la majoria de variables estudiades.

A l'Amèrica del Nord, Mesoamèrica inclosa, els colors blau, groc i vermell es combinen sense cap ordre aparent, fent impossible qualsevol intent de descriure-ho amb paraules.

A l'Amèrica del sud, encara que hi ha un desordre similar si que es pot veure un predomini de l'ordre Adjectiu-Nom a la zona Andina i del Nom-Adjectiu a la vessant atlàntica mentre que a la zona caribenya predomina el groc dels *genus* sense ordre dominant. Tot i això les excepcions són tan nombroses que no val la pena llistar-les.

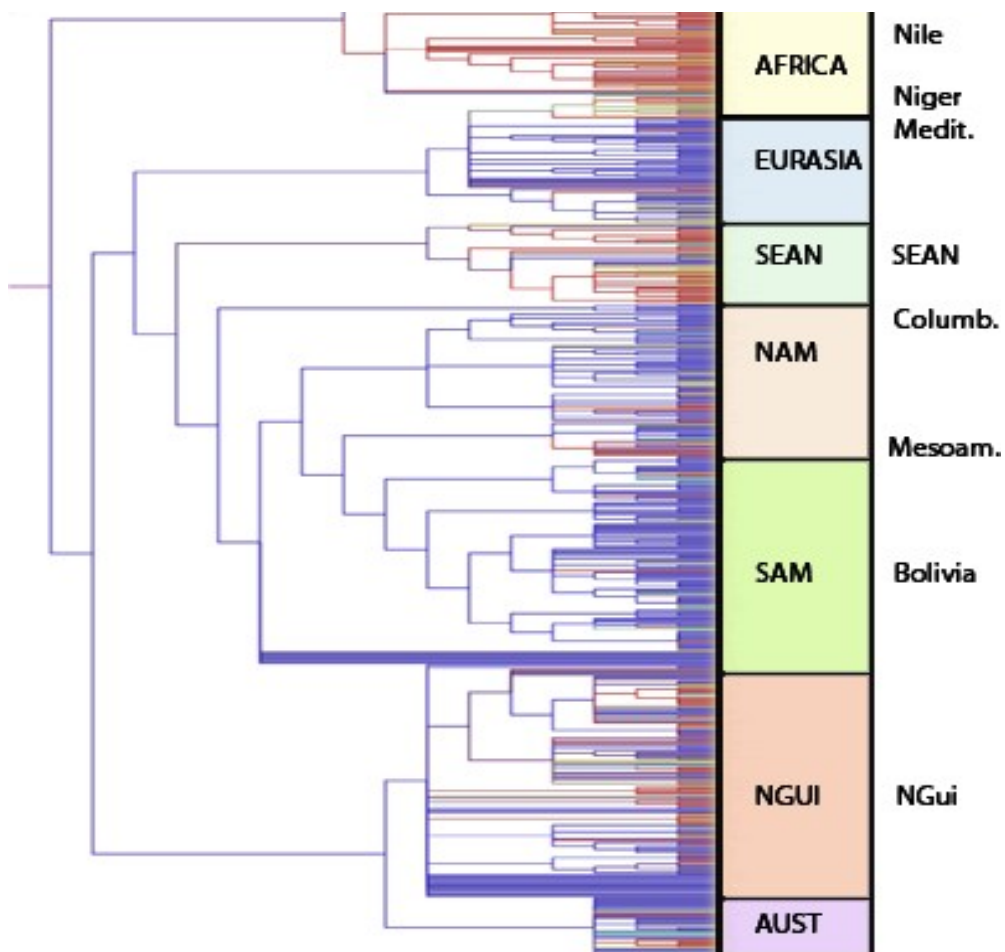
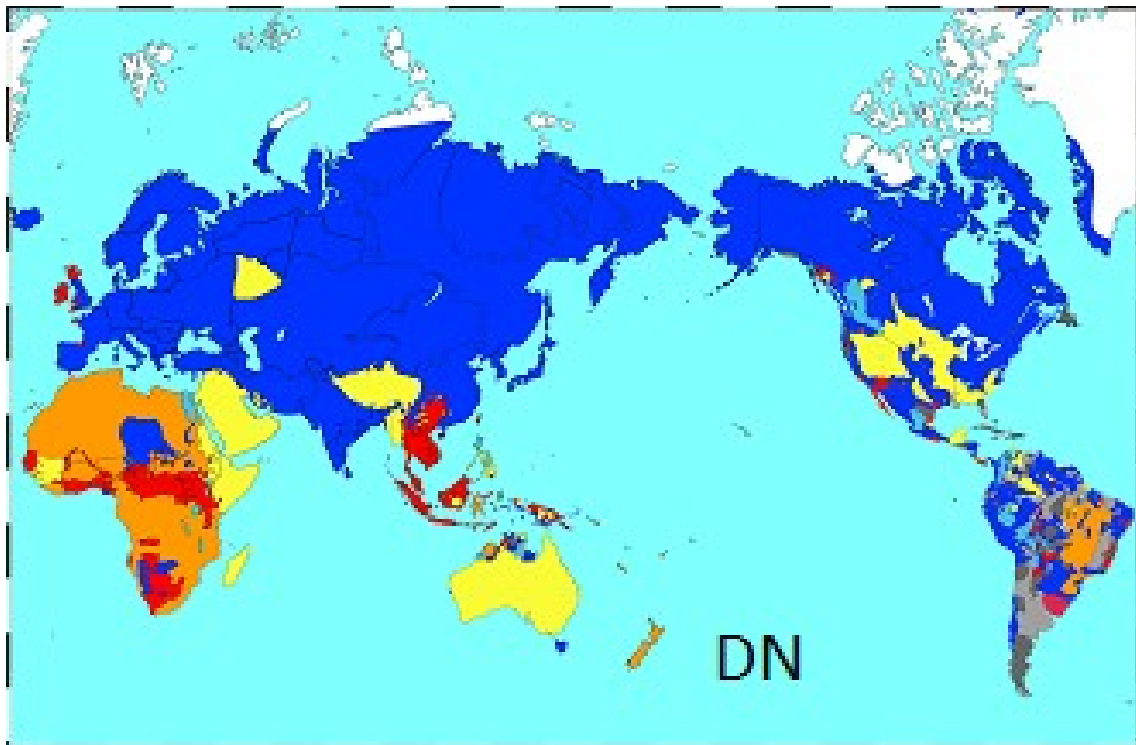
A Nova Guinea, encara que hi ha molta barreja, també s'hi poden veure certs patrons. Els *genus* sense ordre dominant es troben al centre de l'illa, mentre que aquells de nucli final són més freqüents al sud i oest, deixant la resta a *genus* d'ordre Nom-Adjectiu.

A Austràlia l'ordre Nom-Adjectiu és clarament dominant pel que fa al nombre de *genus*, concentrats tots a la part central de la costa nord, però el fet que les llengües pama-nyunganes, que ocupen el 95% de la superfície australiana no tinguin un ordre dominant converteix el color groc en pràcticament el del continent sencer.

Queda per explicar com pot ser que aquesta variable es comporti d'una manera tan imprevisible en alguns continents, mentre que a Euràsia i Àfrica i Austràlia sembla que els ordres han confluït en alternatives oposades, però força consistents en cada continent.

Una mirada a l'arbre permet veure que en general el color vermell de l'ordre NA domina totalment sobre el blau del AN.

## 6.6 DEMOSTRATIUI



Aquesta variable és la que presenta menys dominància d'ordres absoluts. Els colors taronja i groc ocupen una superfície considerable especialment a l'hemisferi sud. Una possible raó és que els demostratius a moltes llengües adopten la forma d'afixos i en alguns casos poden anar simultàniament davant i darrera del substantiu.

A l'Àfrica hi ha un predomini casi total dels *genus* amb demostratius que segueixen al nom. Malgrat això, el fet que les llengües bantu i berber coincideixin en el color taronja fa que aquesta tonalitat ocupi la pràctica totalitat de la superfície. Les excepcions principals es troben a el centre del Sahara i del Kalahari on els *genus* són de nucli final. A la Banya d'Àfrica no hi ha cap ordre dominant, com a les llengües semítiques i les de Madagascar.

A Euràsia es repeteix el patró ja vist abans. Tot el continent és de nucli final, des d'Europa al Japó passant per la Índia, Xina, Sibèria i el pròxim Orient. Per contra, als dos extrems trobem el color vermell que implica que els demostratius segueixen al nom, com passa a les llengües cèltiques, al basc i la majoria de les del sud-est asiàtic tret de part de les austronèsiques, especialment a les Filipines i Taiwan. Aquestes mostren una gran variabilitat d'ordres possibles. El Tibet i Birmània com a zones de transició no mostren tampoc un ordre dominant.

A les dues Amèriques hi ha un predomini general dels *genus* de nucli final, però les excepcions són nombroses i ocupen amplies àrees. Entre els *genus* sense ordre dominant es troben les llengües siuanes i altres de les grans planes americanes, les Maies i les Arawak del nord. Els demostratius darrera dels substantius són l'ordre dominant a zones com la Península de Califòrnia, Oaxaca (al sud de Mèxic) i a diversos *genus* de l'Amazonia.

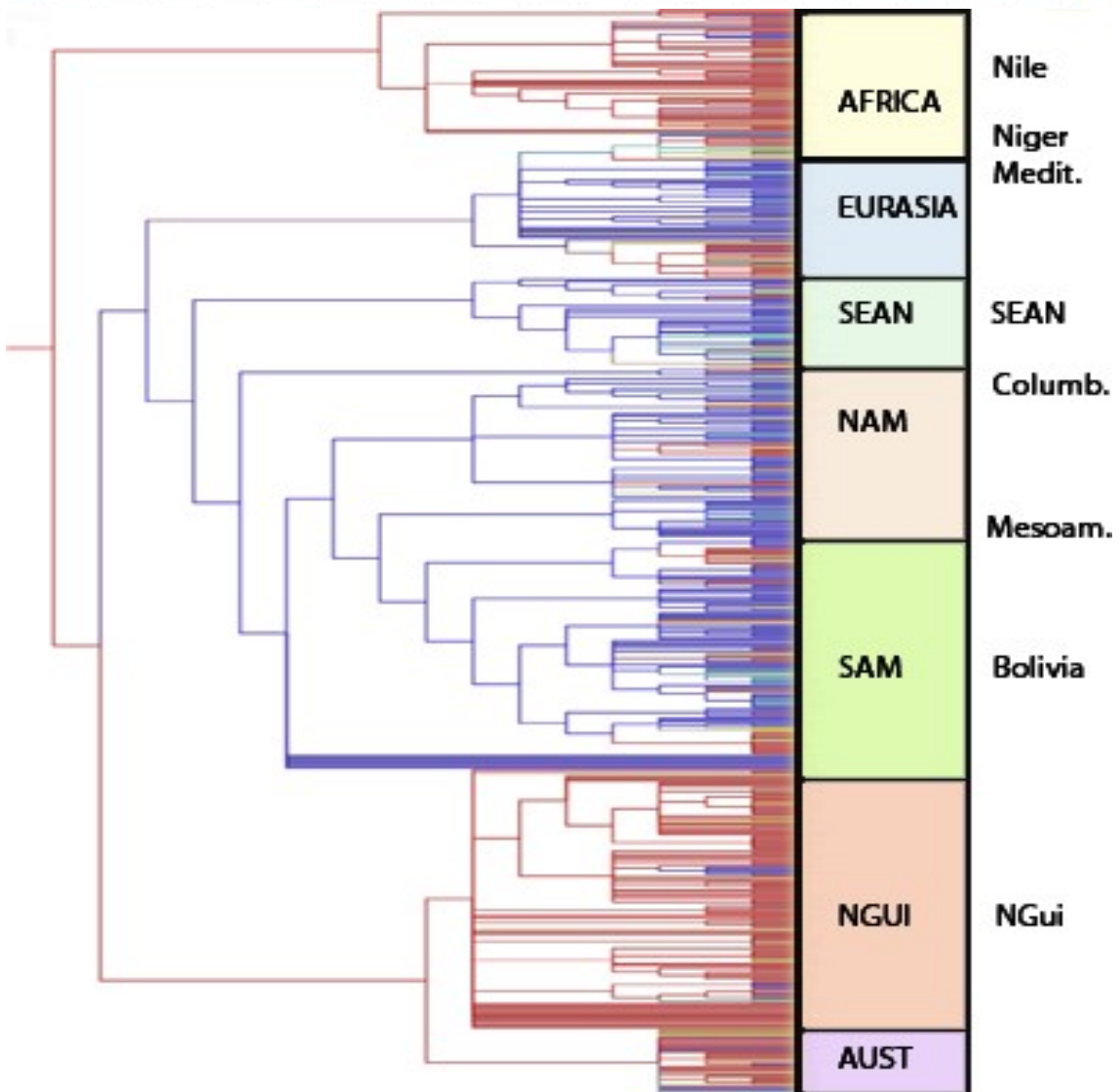
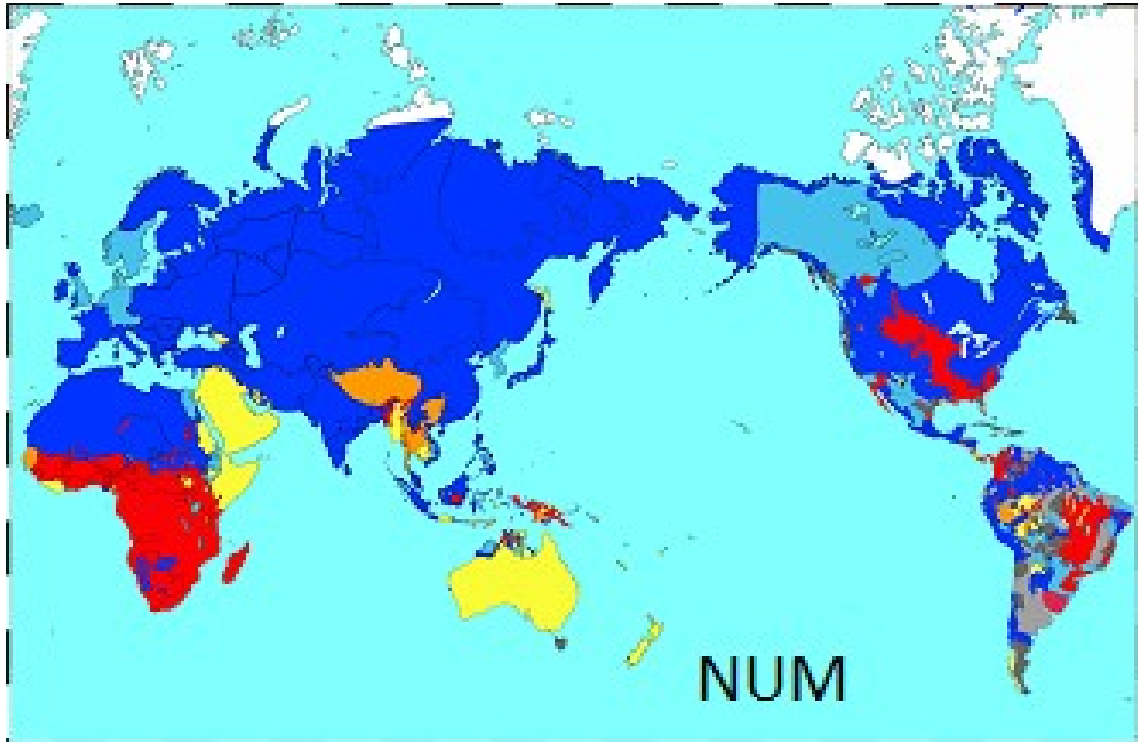
A Nova Guinea hi ha una barreja de tots els tipus d'ordres, però amb una tendència dels *genus* amb nucli final a ocupar la part sud de l'illa, mentre que els d'ordre inicial ocupen el nord i els d'ordre no dominant es troben a les parts centrals.

Els *genus* australians són majoritàriament d'ordre final, excepte la família Daly i les dels *genus* Murrin-Patha i Wororan on predomina l'ordre de nucli inicial i algunes sense ordre dominant. Aquest darrer és el dels *genus* Limilngan, Jaminijungan, Gunwinygic i les PamaNyunga que atès que ocupen un 95% de la superfície d'Austràlia tenyeixen el mapa de color groc, excepte l'extrem nord on es concentra la resta..

Com en altres variables, crida l'atenció que a zones com les Filipines s'hagin diversificat els ordres en uns pocs mil·lennis mentre que que hi hagi un continu sense variacions a l'hemisferi nord des de Portugal fins el Quebec i Groenlàndia.

A nivell històric s'ha observat que a Euràsia és una variable amb poca tendència al canvi, però quan s'ha produït ha estat des de llengües ancestrals amb un ordre no dominant fins a les descendents que han optat per la disposició majoritària al continent, com ha estat en el cas de les llengües germàniques nòrdiques, les esclaves i l'armeni.

## 6.7 NUMERAL



L'ordre dels numerals mostra patrons clars en la seva distribució geogràfica i permet, en general, fer una descripció força acurada.

La distribució a l'Àfrica subsahariana és pràcticament tota d'ordre Nom-Numeral, excepte a les llengües khoisanides del centre i a mitja dotzena de *genus* principalment nilo-saharians: Maban, Nara, Shabo, Eastern Jebel, Lendu i Kresh, Ijoid i Kru. També a la Banya d'Àfrica predominen els *genus* sense un ordre dominant, com en altres variables ja comentades i les llengües berbers són d'ordre Numeral-Nom com pràcticament tota la riba del Mediterrani.

A Euràsia, fora del sud-est asiàtic, els numerals precedeixen als substantius pràcticament arreu, excepte les llengües semítiques, caucàsiques del nord-oest i el Nivkh que no tenen un ordre determinat. Al Tibet i a diversos *genus* de Birmània i Tailàndia, en canvi, hi ha força casos de numerals que segueixen al substantiu. El nombre de *genus* al sud-est asiàtic amb un ordre o altre és força similar i és de les poques zones, on costa trobar patrons de distribució clars.

A l'Amèrica del nord, com a la resta de l'hemisferi nord, hi ha un predomini casi absolut dels *genus* amb numerals precedint els substantius, excepte a les planes del centre dels EEUU i a la península de la Baixa Califòrnia.

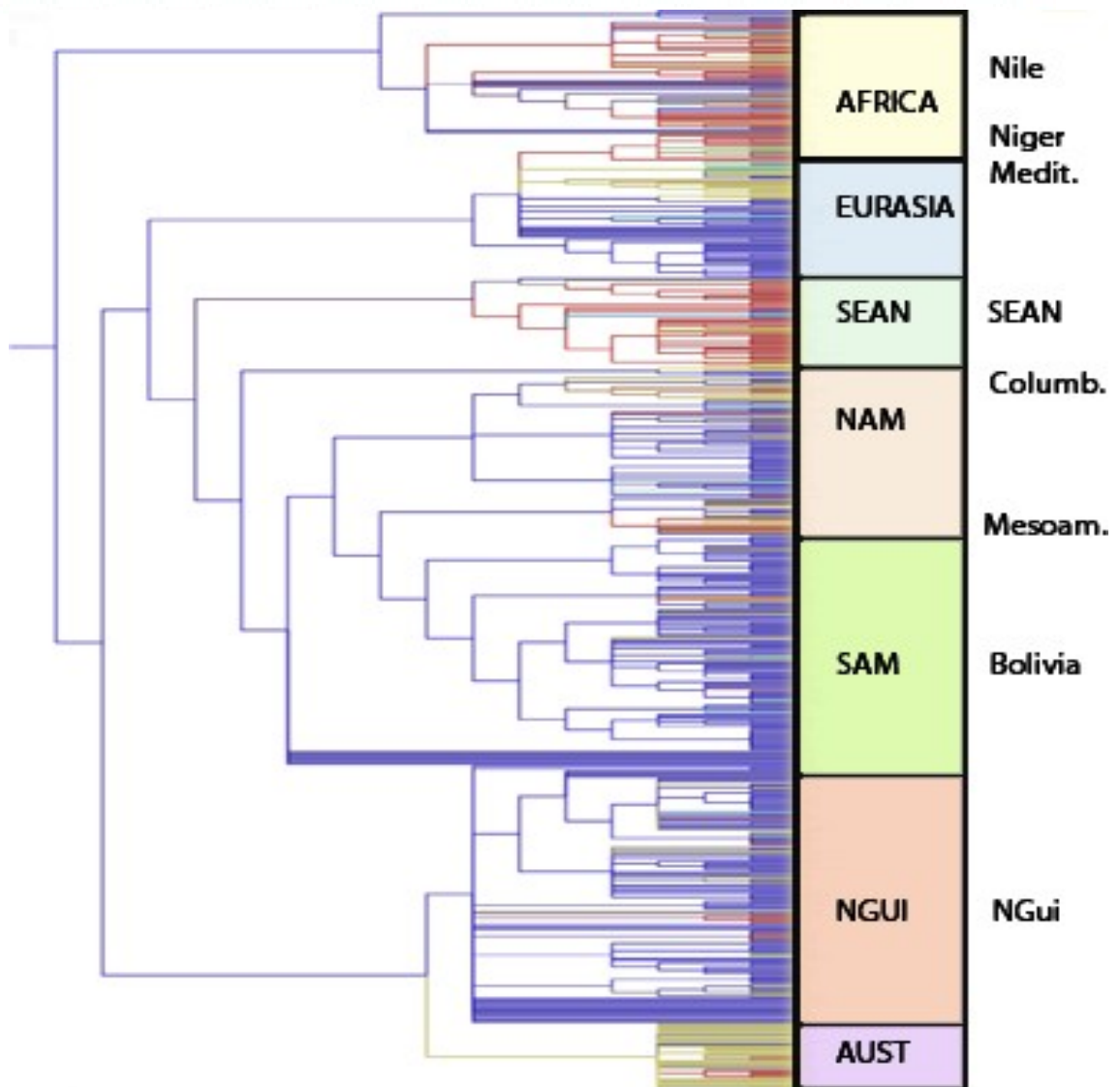
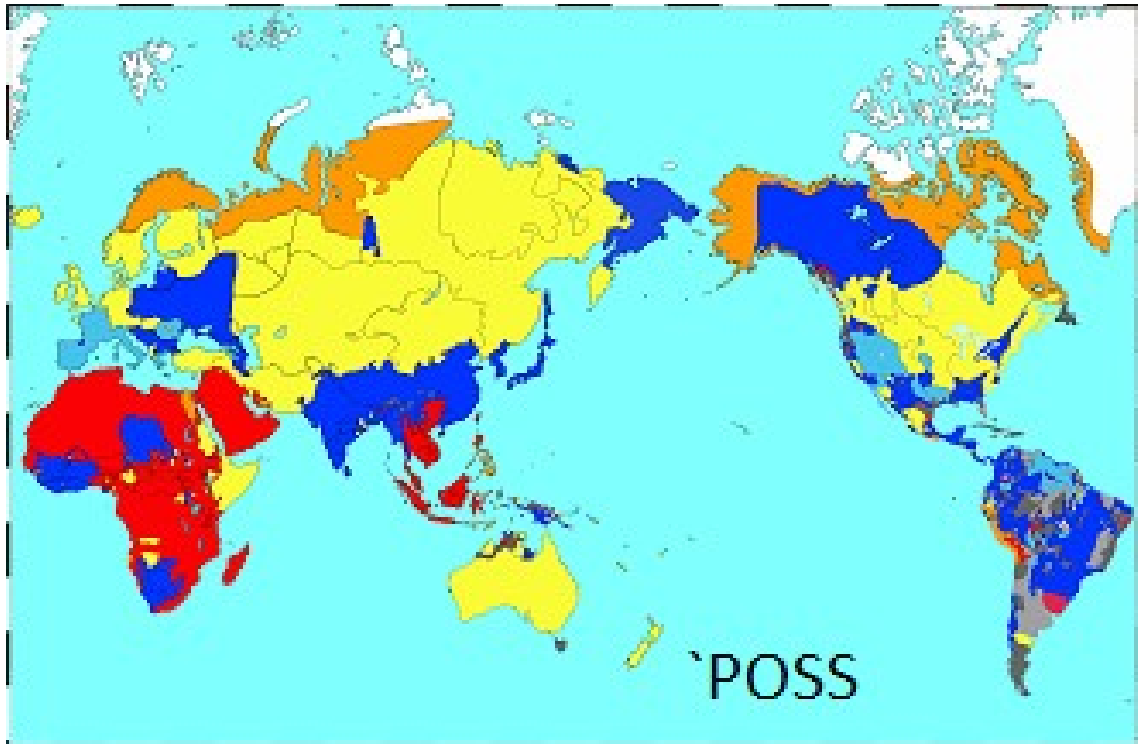
Sudamèrica mostra molta més diversitat. L'ordre Numeral-Nom és dominant, però a bona part de Colòmbia, Centramèrica i del Brasil es dona l'ordre contrari amb clapes d'ordres mixtes a l'oest de l'Amazonia .

A Nova Guinea pràcticament tots els *genus* de llengües papús són de nucli inicial pel que fa a aquesta variable, amb algunes excepcions amb ordre Numeral-Nom, especialment entre les del grup filogenètic anomenat Trans-Fly i alguns d'illes on conviuen amb forta presència de llengües austronèsiques. Aquest seria el cas del Bilua, el Savosavo i el Kuot. Alguns pocs no presenten cap ordre dominant com als *genus* Middle Sepik, Senagi i Suki.

Pel que fa a Austràlia, es repeteix el patró ja vist a d'altres variables. Hi ha una convivència de molts ordres diferents entre la multitud de *genus* del nord, però la falta d'un ordre dominant entre les llengües Pama-Nyunganes i això fa que el mapa d'aquesta illa-continent sigui pràcticament tot de color groc.

A nivell diacrònic, entre les llengües d'Europa que ens han deixat documentades històricament no hi ha pràcticament cap canvi documentat, i quan n'hi ha es tracta només de la desaparició d'ordres secundaris, cosa que ha passat a llengües germàniques com l'anglès o l'alemany, i també amb l'armeni. A l'Índia, el cingalès ha adoptat un ordre Nom-Numeral contrari al general del seu *genus*, mentre que l'antic Tàmil, que no tenia un ordre dominant, ha donat lloc a llengües descendents amb numerals que precedeixen als substantius.

## 6.8 POSSESIU



L'aspecte més remarcable de la distribució geogràfica d'aquest constituent és el fet que en amplíssimes zones del planeta no mostra un ordre dominant. Això passa a la major part dels Estats Units i a bona part de Sibèria i Austràlia.

Àfrica és l'únic continent on geogràficament predominen els *genus* on els possessius segueixen als noms i per això el mapa apareix pràcticament vermell. Al mirar en detall es veu que les excepcions són més nombroses del que sembla a primera vista, com en altres constituents ja comentats. Les excepcions poden ser *genus* amb nucli final que es troben principalment al centre del Sàhara, a les llengües khoisanides del centre i del sud, i a la franja de territori entre Nigèria i el Senegal amb les llengües Kru, Kwa, Mandé, Gur i Songhai. Per contra els ordres no dominants es troben a la Banya d'Àfrica, al Beja i a poques més com les Khoisànides del nord.

A Euràsia es troben *genus* amb tots els tipus d'ordenacions, però la seva distribució no és a l'atzar i al mapa apareixen zones de colors amb una distribució molt consistent. El sud-est asiàtic amb les illes d'Indonèsia i Filipines són d'ordre Nom-Possessiu. Des de l'Índia fins el Japó, l'ordre és el contrari. Aquest ordre també el trobem a la major part d'Europa. En canvi, a la major part del continent domina el color groc de l'absència d'un ordre dominant, des de les llengües cèltiques, passant per les germàniques i d'aquí fins les fineses i turques i la resta fins Sibèria. La majoria de llengües uràliques circumpolars comparteixen ordre amb les dels pobles esquimals (el taronja de la categoria 3).

La lingüística històrica ha documentat canvis a l'ordre dels possessius a força llengües de la regió Euromediterrània com és el cas de l'antic egipci, l'islandès, llengües esclaves, grec, armeni o persa. Malgrat aquests canvis, és una variable que canvia un ritme inferior a moltes altres variables, som s'ha vist a l'apartat 2.2.6.

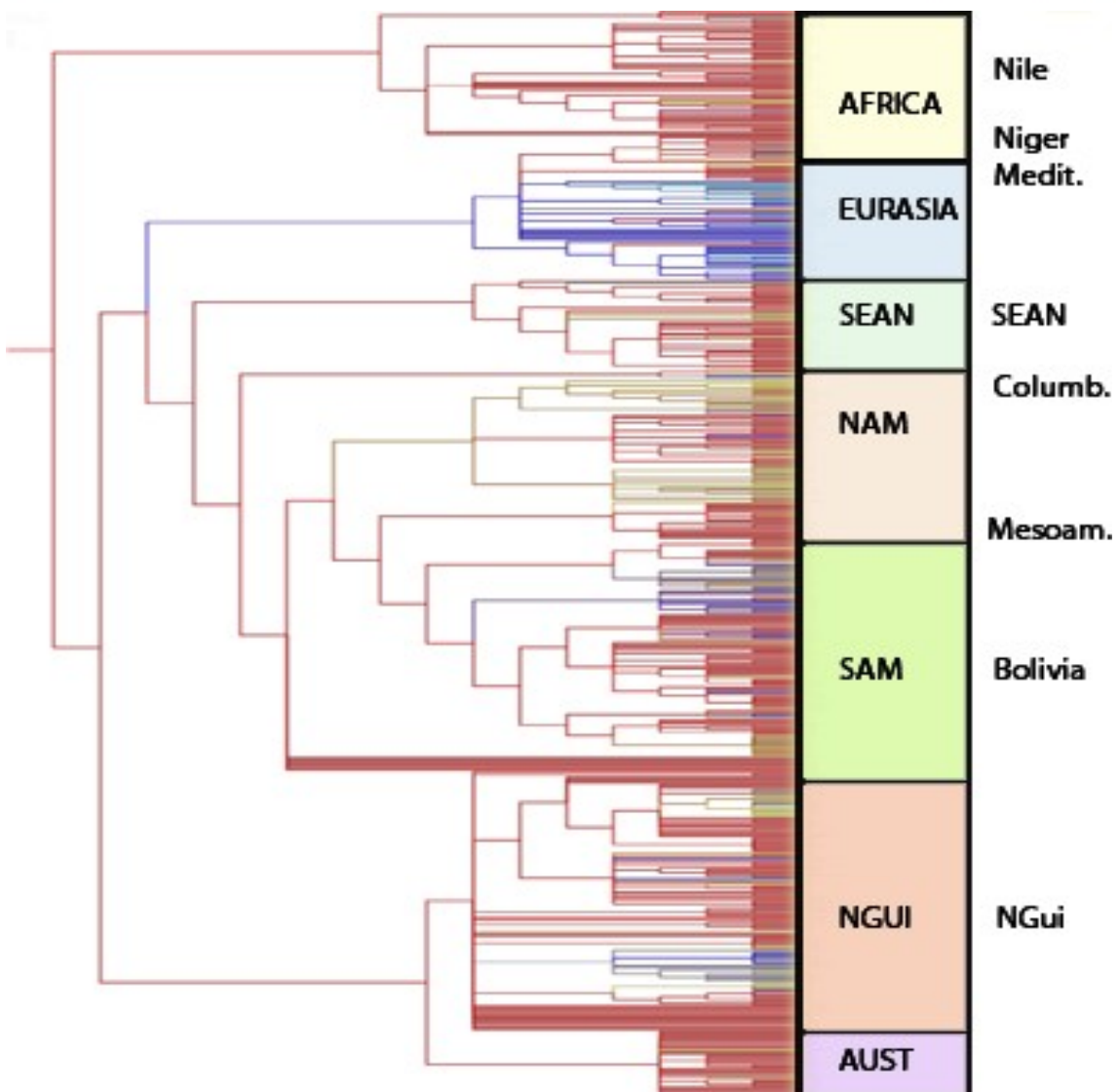
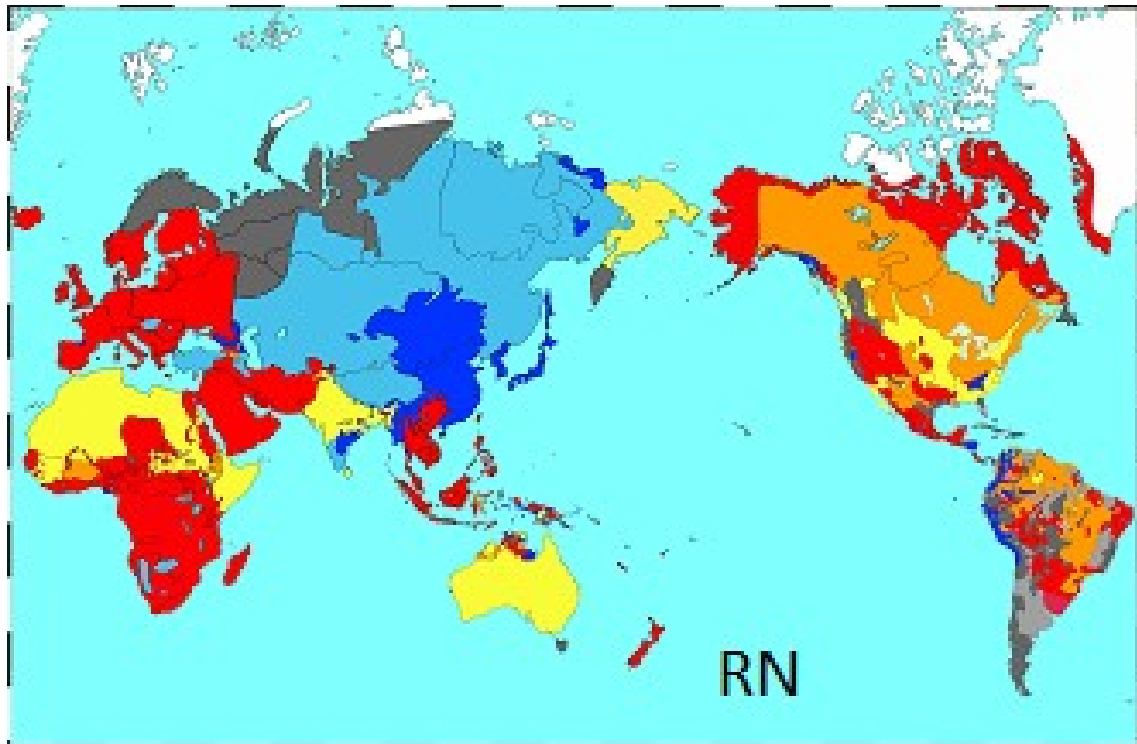
A l'Amèrica del Nord, Mesoamèrica inclosa, hi ha una clara majoria de *genus* amb el possessiu davant del nom. L'ordenació contrària es dona parcialment a les llengües esquimals i en alguns *genus* de la Colúmbia Britànica i Mesoamèrica. També en aquestes dues zones, justament amb les plans planes dels Estats Units on es troben *genus* sense ordre dominant.

A l'Amèrica del sud hi ha un domini pràcticament total de l'ús dels possessius al davant dels substantius. Com excepcions amb nucli final només hi ha els *genus* Quechua, Aimara i Chapacura-Wanham. I com excepcions sense ordre dominant hi ha el Puelche, el Movima, el Cayuvava i les llengües Jivaroan.

A Nova Guinea hi predomina l'opció amb possessius davant dels substantius de forma general entre les llengües papús. L'ordre alternatiu només és preferit per les llengües de la família Skou, el Kayagar, el Dagan i part de la família Torriceli. Uns pocs *genus* sense ordre dominant completen el panorama.

A Austràlia, tret dels territoris ocupats pels *genus* de nucli final com la família Daly i el *genus* Wagiman, per una banda i per l'altre els *genus* Tiwian, Iwaidjan i Garrwan, de nucli inicial, la resta del continent pertany a *genus* que no tenen un ordre dominant pel que fa aquesta variable.

## 6.9 RELATIU





Aquesta variable presenta una fortíssima tendència a tots els continents, tret de la major part d'Àsia, a situar el relatiu darrera el nom.

A l'Àfrica, els *genus* amb relatius situats davant del nom són només l'Ijoid i el cuixític central, de categoria 1, i el khoisanid del centre del Kalahari i Omotic del nord de categoria 2. La falta d'ordre dominant es pròpia de quatre *genus* més: Lendu, Dogon, cuixític de l'est i mande de l'oest. La resta tendeix a posar el relatiu darrera el nom.

Euràsia és el continent amb més diversitat d'opcions. Als dos extrems que són la regió Euromediterrània i el sud-est asiàtic amb les illes pròximes presenta un ordre Nucli-Relatiu. Aquestes dues àrees pràcticament coincideix amb la situació que ja s'ha vist a l'apartat de les adposicions o de l'objecte. Hi ha diferències com pot ser que les llengües índiques no tenen ordre dominant o que les iràniques segueixen el patró europeu.

Per contra, la resta del continent, des de Turquia fins el Japó, passant per Sibèria, Sud de l'Índia, Tibet, Birmània, la Xina i fins i tot l'eusquera i l'hongarès mostren el color blau dels ordres de nucli final, encara que repartits entre *genus* de categoria 0 i 1.

L'Amèrica del nord pràcticament no té cap *genus* amb relatius davant del nom, tret del Tlingit, l'Iroquès del sud i el Maiduan. La resta de *genus* es reparteixen molt majoritàriament entre les d'ordre estricte Nom-Relatiu, majoritàriament a les costes i que coincideixen en bona part amb les llengües Hoka de Ruhlen i els *genus* mesoamericans, per una banda i per l'altre les d'ordre NR de categoria 3 (amb entre un 10% i un 30% d'excepcions) que es troben més a l'interior. La zona de les grans planes està ocupada per *genus* que no tenen un ordre dominant.

En el cas de l'Amèrica del sud es pot dir que, en general, els *genus* amb nucli final ocupen la costa pacífica i els Andes, mentre que a la resta predominen amb diferent intensitat les llengües amb ordre Nom-Relatiu. Fins i tot es troben alguns *genus* sense ordre dominant com el Waorani, el Tupari, les llengües Yanomamis i l'Aruak.

A Nova Guinea es barregen totes els ordres possibles, però aproximadament es pot dir que l'ordre Relatiu-Nom ocupa la part sud de l'illa, l'ordre Nom-Relatiu la part superior i al centre hi ha zones d'ordre no dominant. Al mirar els arbres, però, es pot veure que la majoria de *genus* amb nucli final pertanyen a una mateixa família (Sepik).

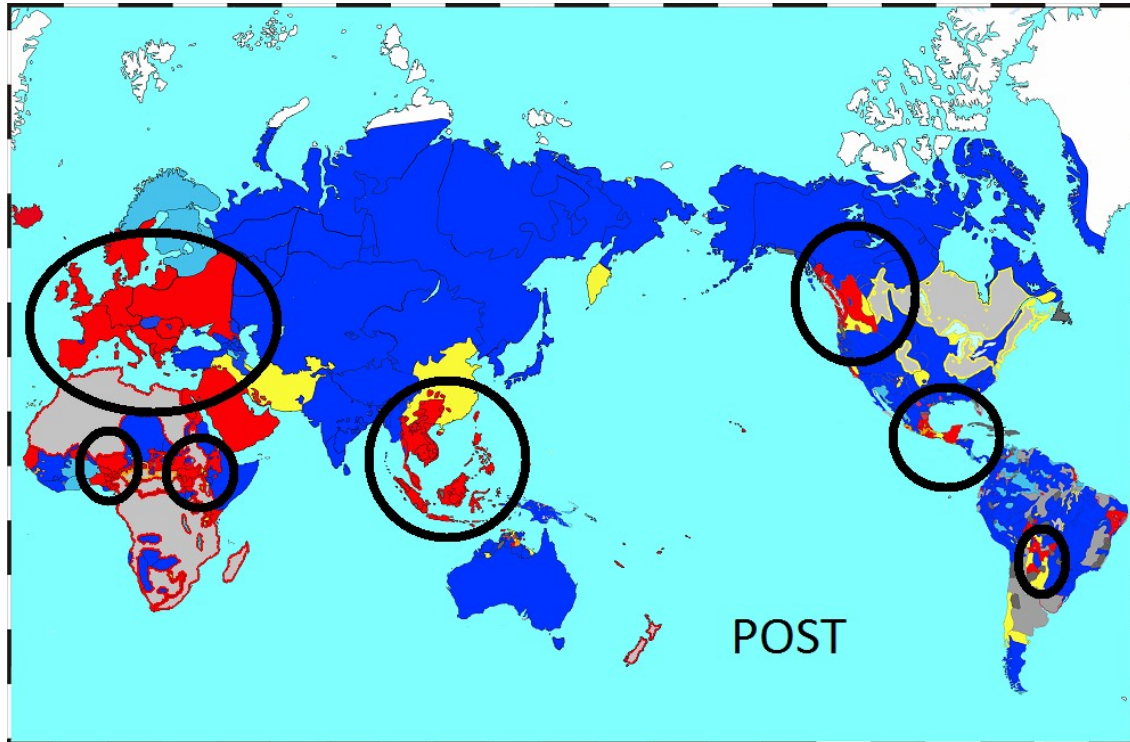
En el cas d'Austràlia, el predomini de *genus* amb relatiu posterior al substantiu és total i només els *genus* Murrinh-Patha i Daly del sud escapen a la norma al no tenir un ordre dominant. Com en altres casos, el fet que la família Pama-Nyunga pertanyi a una categoria diferent a la majoritària al continent fa que el color d'Austràlia sigui al final el taronja i no el vermell.

Per últim cal destacar que la distribució de l'ordre Nom-Relatiu tan extensa que podria ser considerada casi com universal, es compatible amb les àrees detectades a les variables adposició, objecte i genitiu en el Vell Món que s'estudiaran de forma detallada al capítol 7.

És de les poques variables en que l'ordre nucli-complement sembla ancestral i és majoritari a l'arrel de l'arbre.



## 7 ÀREES ESPECIALS



**Mapa 22** Àrees especials de genus amb ordre d'adposicions de nucli inicial

Després d'examinar els mapes de distribució de les variables al capítol anterior, ha quedat clar que en algunes d'elles, els *genus* de nucli inicial es concentren a unes poques regions que destaquen en vermell sobre un fons general en blau que correspon a l'ordre de nucli final. Aquestes serien les variables Objecte, Adposició, Genitiu i Subjecte, però no coincideixen ni en el nombre de *genus* implicats ni en àrea ocupada ni tampoc en consistència.

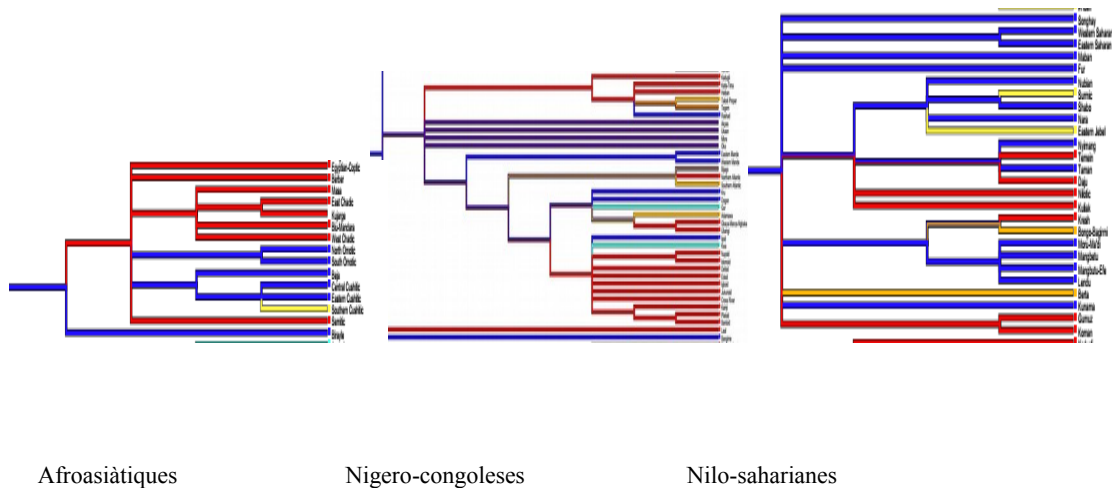
També sembla que en aquestes regions majoritàriament de nucli inicial, la variable Objecte ocupa sempre l'àrea més gran, mentre que les variables Adposició i Genitiu són progressivament més petites i la variable Subjecte ocupa un espai molt reduït.

També ha quedat clar que aquestes àrees estan força fragmentades de manera que encara que representen una part mínima de la superfície de cada continent (5%-10%), acumulen aproximadament una tercera part dels *genus* del món (exclosa la prolífica Nova Guinea on conviuen llengües de més d'un centenar llarg de *genus*).

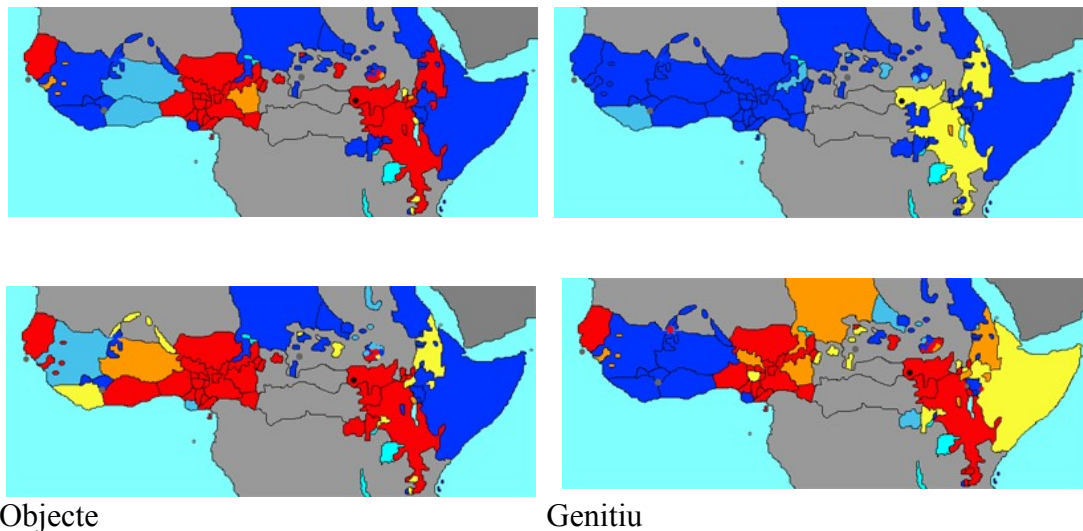
Des d'aquestes àrees d'alta fragmentació s'han produït expansions per la sortida d'alguns *genus* que ha estès la seva tipologia. Exemple en serien els *genus* bantu, berber, semític, arawak del nord de forma clara i de forma més dubtosa i/o limitada es podrien considerar així l'algonquí, l'asteca, el xinès, el nilòtic, l'ubangi o alguns d'indoeuropeus.

Aquestes àrees es comenten de forma més detallada en aquest capítol a partir de mapes areals de les quatre variables implicades.

## 7.1 ÀFRICA



**Arbre 10** Multiarbre dels valors dels de quatre variables a les zones africanes



**Mapa 23** Multimapa dels valors dels de quatre variables a les zones africanes.

Per trobar els patrons dels possibles canvis en l'ordre de constituents a l'Àfrica, s'ha eliminat de l'anàlisi a quatre *genus* (Ubangi, Bongo-Bagirmi, Bantu i Berber) dels mapes. La raó és que les dues primeres ocupen l'espai geogràfic situat entre les dues àrees on històricament s'han donat els canvis a estudiar i les dues últimes perquè la seva recent expansió desdibuixa la situació prèvia ja que junts representen, juntament amb l'àrab, un 80% de parlants en prop d'un 80% de la superfície, deixant el que queda a la resta dels més de 90 *genus* africans.

Bàsicament els canvis s'han produït en dues zones molt determinades: El delta del Niger i la zona subsahariana del Nil. En cadascuna conviuen dotzenes de *genus*. Amb alguna excepció, les del Niger pertanyen a la família Nigerocongolesa i les del Nil són Afroasiàtiques i Nilosaharianes, amb algunes d'aïllada o potser khoisanides. Els arbres filogenètics permeten constatar que els canvis a ambdues zones són fenòmens areals ja que els *genus* implicats no són monofilètics i que els canvis s'han produït un cop les famílies implicades ja estaven en procés de diversificació.

L'ordre SV domina a tota l'àrea del Níger i en casi tota la del Nil. En aquesta darrera el *genus* Kuliak és VS i en el *genus* Nilòtic no hi ha ordre dominant. Les llengües semítiques apareix com d'ordre no dominant, però cal tenir en compte que aquestes també es troben a l'Àsia i el valor respon a la mitjana d'ambdós continents.

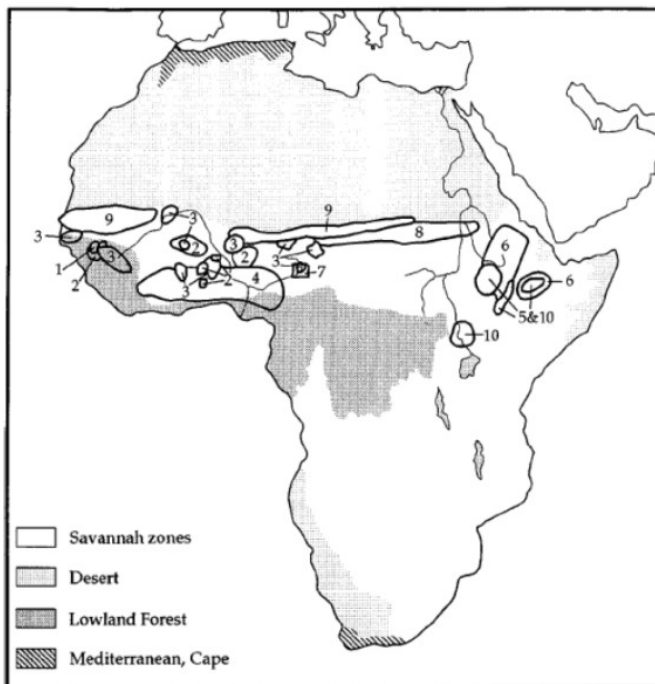
L'ordre VO està present en el mapa en tres zones. Una en el que seria Senegal que correspon a les llengües atlàntiques i que no s'ha tingut en compte al trobar-se massa lluny de l'àrea de Nigèria per tractar-la conjuntament. Aquesta última esta formada per una quinzena de *genus* i es perllonga cap a l'oest amb els *genus* kwa i gur i cap a l'est amb les llengües bantus. Per últim un àrea amb mes de 25 *genus* a la zona del Nil.

La distribució de la variable adposició coincideix bàsicament amb la de l'objecte però amb un àrea més restringida, al faltar les llengües kwa i gur a l'àrea de Nigèria i els *genus* Manbutu-Efe i Mangbetu a la del Nil. Un parell de *genus* més (Talodi proper i Daju) són posposicional malgrat que respecte l'objecte no tenen ordre dominant, justament el contrari que el *genus* Easter Jebel.

La variable Genitiu mostra al mapes i arbres un patró molt similar a l'Objecte, però les àrees definides són lleugerament menors i no mereixen massa més comentaris, per exemple, els *genus* Nupoid i Berta no tenen ordre dominant malgrat ser preposicionals.

Al mapa 24 es constata que la zona del Níger és la de domesticació d'un grapat de plantes, amb especial importància el nyam, base de la dieta a Nigèria. La zona del Nil també ha protagonitzat algunes domesticacions de plantes i d'animals ja que molts dels pobles són ramaders.

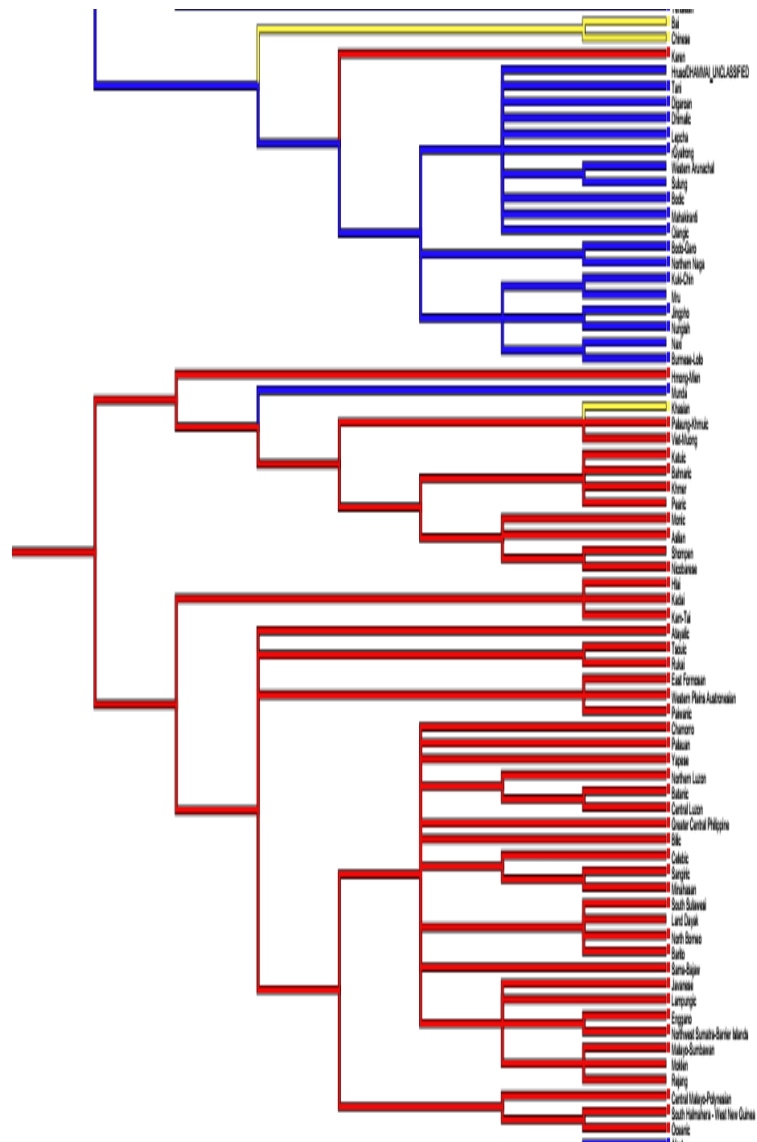
En total hi ha una desena de plantes domesticades que han desenvolupat una agricultura mil·lenària a les zones estudiades i més enllà. També a la zona on el parell de *genus* de les llengües atlàntiques semblen haver patit canvis similars.



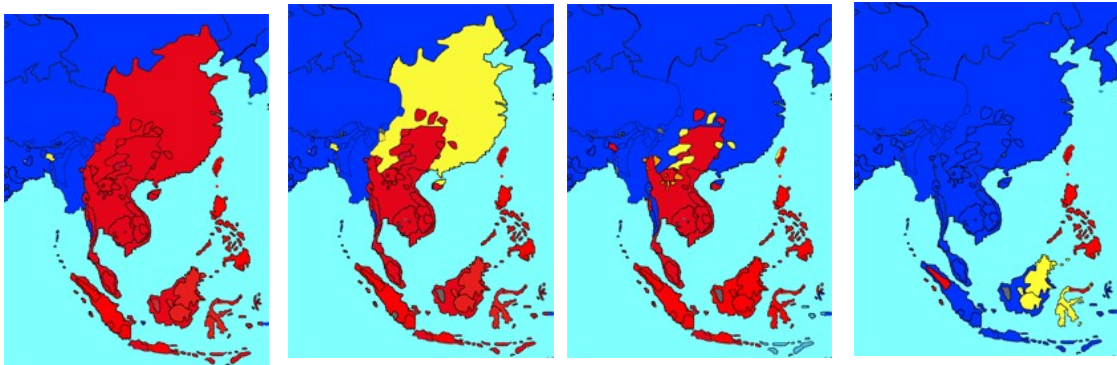
**Fig. 3.** Proposed areas of domestication of African plants (after Harlan, 1971). 1. Guinea millet; 2. Fiono and black fonio; 3. African rice; 4. Yam (*Dioscorea cayenensis* complex); 5. Enset; 6. Tef; 7. Groundnuts (*Kerstingiella* and *Vooandzeia*); 8. Sorghum; 9. Bullrush/pearl millet; 10. Finger millet.

**Mapa 24** Mapa de zones inicials agrícoles africanes

## 7.2 SUD-EST D'ÀSIA



**Arbre 11** Multiarbre dels valors dels de quatre variables al sud-est asiàtic.



Objecte

Adposició

Genitiu

Subjecte

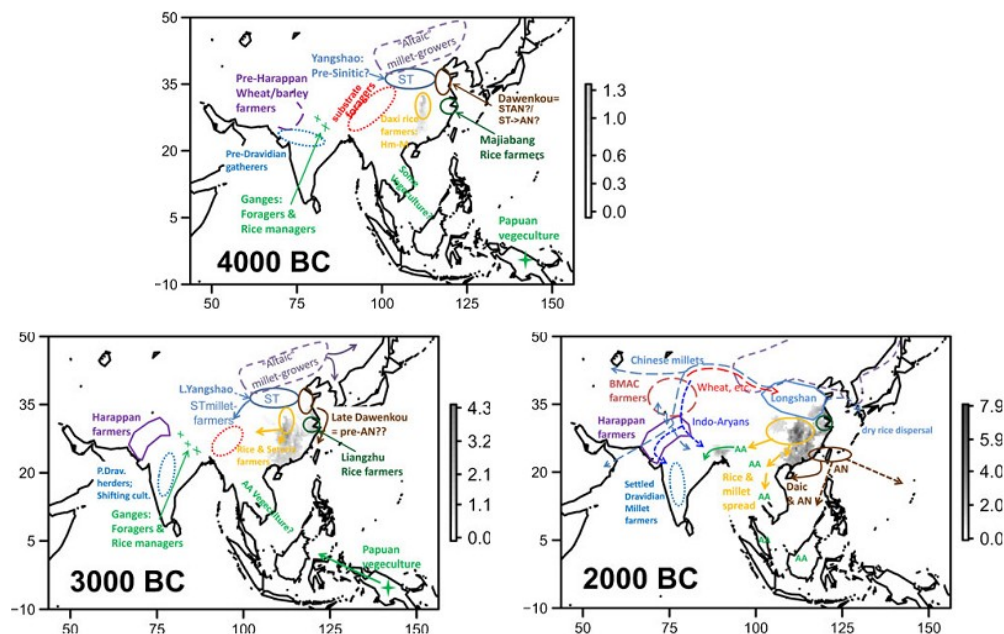
**Mapa 25** Multimapa dels valors dels de quatre variables al sud-est asiàtic.

El sud-est asiàtic és una zona amb una densitat de població elevada i també on es barregen un munt de *genus* i famílies lingüístiques amb una història lligada al desenvolupament de l'agricultura, especialment de l'arròs.

En la regió hi participen principalment *genus* de famílies diferents que Ruhlen agrupava en la família Austrica, però que darrerament s'han tornat a separar: Austro-Asiatic, Tai-Kadai i Hmong-Mien, entre tots sumen una cinquantena de *genus* diferents. A la regió se li pot sumar els *genus* xinès, Karen i Bai entre les Sino-Tibetanes i el Shompen, una llengua aïllada de les illes Nicobar que possiblement ha patit la influència de les veïnes Austro-Asiàtiques.

L'ordre VS només és present avui en uns pocs *genus* austronèsics, especialment a les Filipines, encara que sembla que en el passat era una opció molt més estesa [23]. Aquesta regió és molt coherent i a les variables objecte, adposició i genitiu, el mapa és pràcticament idèntic en el predomini del color vermell, propi d'ordres amb nucli inicial. La diferència està als *genus* Xinès i Bai que són de nucli inicial a la primera variable citada, no tenen un ordre dominat a la segona i finalment són de nucli final a la tercera. El genitiu d'algunes *genus* austronèsics hi ha ordre no dominant o nuclis finals.

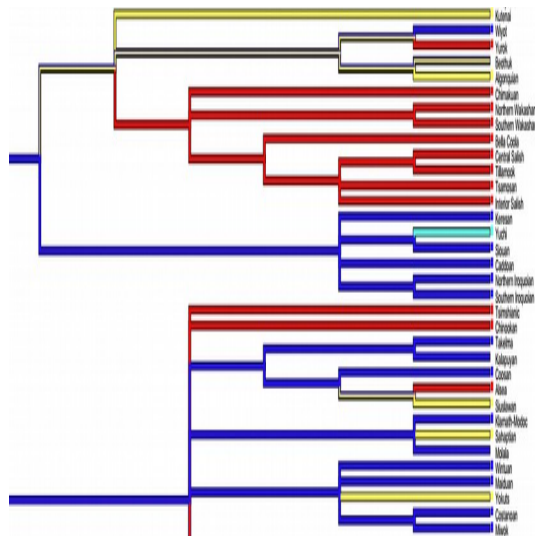
Resulta interessant el cas del *genus* Munda, l'únic entre les llengües Austroasiàtiques que presenta una tipologia divergent a les de la resta de la família i la té igual a la de les llengües de la Índia. Després de moltes discussions, ara s'ha demostrat que són una població genèticament barrejada d'homes d'origen Austroasiàtiques amb dones de l'Índia [105].



Mapa 26 Multimapa de l'expansió des pobles lligats al cultiu de l'arròs. Font: [106].

La paleontologia lingüística apunta a les llengües Hmong-Mien (i potser les austroasiàtiques de forma paral·lela) com a candidates a primeres poblacions en domesticar de l'arròs al Yangze [106]. El mapa 26 mostra com juntament amb l'agricultura, hi ha hagut un gran moviment de pobles, que han anat baixant i colonitzant nous territoris, primer cap a Indoxina i posteriorment cap a les Illes d'Oceania. A l'Índia, Corea i el Japó hi ha arribat el cultiu de l'arròs sense que hi hagi hagut cap substitució lingüística en favor de les llengües abans anomenades àustriques.

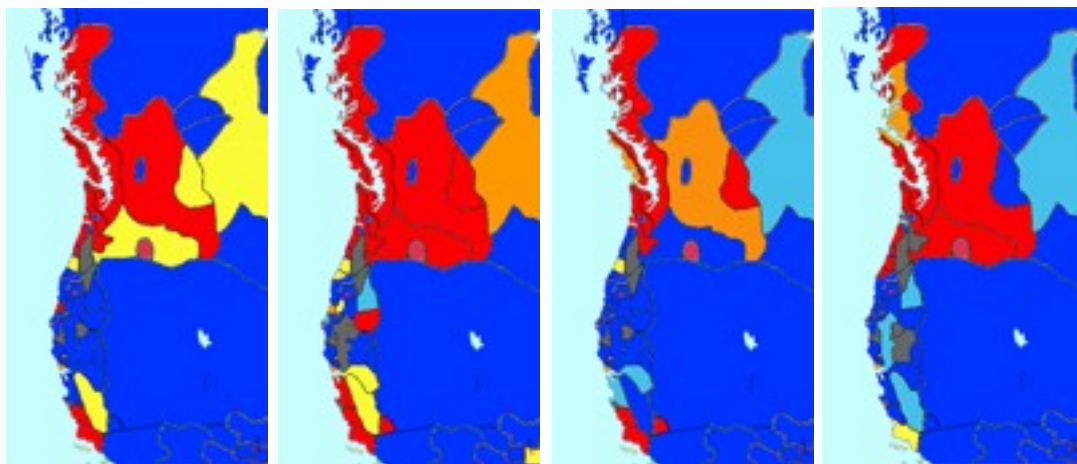
### 7.3 COLÚMBIA BRITÀNICA



**Arbre 12** Multiarbre dels valors dels de quatre variables a la Columbia Britànica.



**Mapa 27** Àrees culturals d'Amèrica del nord. Font[108]



**Mapa 28** Multimapa dels valors dels de quatre variables a la Columbia Britànica.

A aquesta regió és difícil trobar-hi uns límits clars. Coincideix la desaparició de llengües a conseqüència de la conquesta i una gran fragmentació lingüística amb una història de desplaçaments i intrusions impossibles de reconstruir de forma segura.

Hi ha una zona nord que no ofereix dubtes, formada per tots els *genus* de la família Salishan, Wakashan, Oregon Coastal, uns pocs de la família Petunia i alguns aïllats com el chimakuan i el kutenai, uns quinze *genus* en total. Baixant per la costa, es troben tres *genus* més (Karak, Palaihnihan i Beothuk) que semblen participar de la tipologia de la regió, però que n'estan separats per illots de parlants athapascans.

Encara més avall hi ha un altre grup similar de sis *genus*, tres de la família Petunian, els *genus* Salinan, Chumash i un, el Tulatulabal, que pertany a la família utoasteca. Aquest darrer grup, separat dels altres per uns centenars de kilòmetres de costa, és el que genera dubtes de si podria haver adquirit una tipologia similar per contacte o hi va arribar de forma paral·lela.



La variable subjecte presenta l'ordre SV al *genus* Kutenai i a tots els de les altres dues zones comentades, tret del Chumash que no té ordre dominant. La resta de *genus*, tots a la zona nord, és d'ordre VS sense excepció.

El mapa de la variable objecte coincideix pràcticament amb el del subjecte pel que fa a la zona nord, encara que en aquest cas els *genus* Coosan i Siuslawan no tenen ordre dominant. La diferència és a les zones meridionals on el color vermell tenyeix tots els *genus*, tret dels Karok, Miwok i Yokuts que no tenen ordre dominant.

En el cas de les adposicions, el mapa també és bàsicament el mateix que el de la variable Objecte, però més restringida. A la zona nord només hi ha tres *genus* amb ordres diferents: El Coosan és de nucli final i el Kutenai i Sahaptian no tenen ordre dominant.. En canvi, a les dues zones meridionals la situació és força diferent i només tres *genus* són preposicionals (Yurok, Chumash i Salinan) són preposicionals i no hi ha ordre dominant a la llengua Yokuts, la resta són posposicionals.

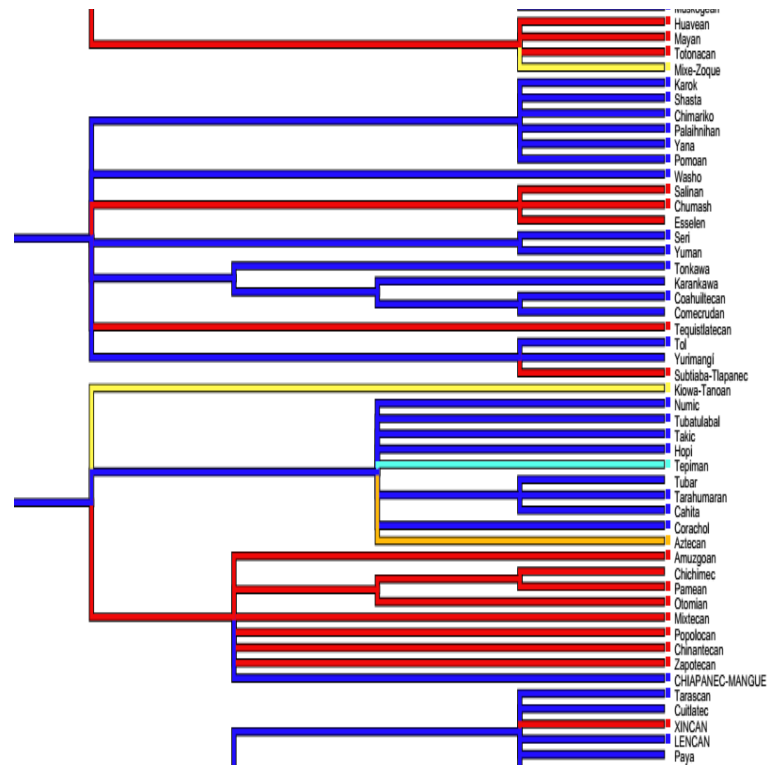
A la variable Genitiu, l'àrea ocupada per tons vermells és encara més petita. A les zones meridionals els *genus* amb ordre NG són només tres: Tubatulabal, Chumash i Salinan, a la resta el genitiu precedeix al nom. A la zona nord els ordres coincideixen amb el de l'objecte excepte en els *genus* Sahaptian, Alsea, Tillamok i Coosan que són de nucli final per la qual cosa estan al mapa de color blau.

Les llengües algonquines formen un *genus* que s'estén fins la costa atlàntica. El genitiu i l'objecte presenten un ordre bàsic de nucli inicial i per això estan de taronja al mapa i no tenen un ordre dominant a les adposicions. Com que aquest *genus* darrerament sembla que està fortament emparentat amb les llengües Wiyot Yurok de Califòrnia, i aquesta darrera presenta la tipologia de l'àrea, podria tractar-se d'una expansió cap a l'est des de l'oest [107].

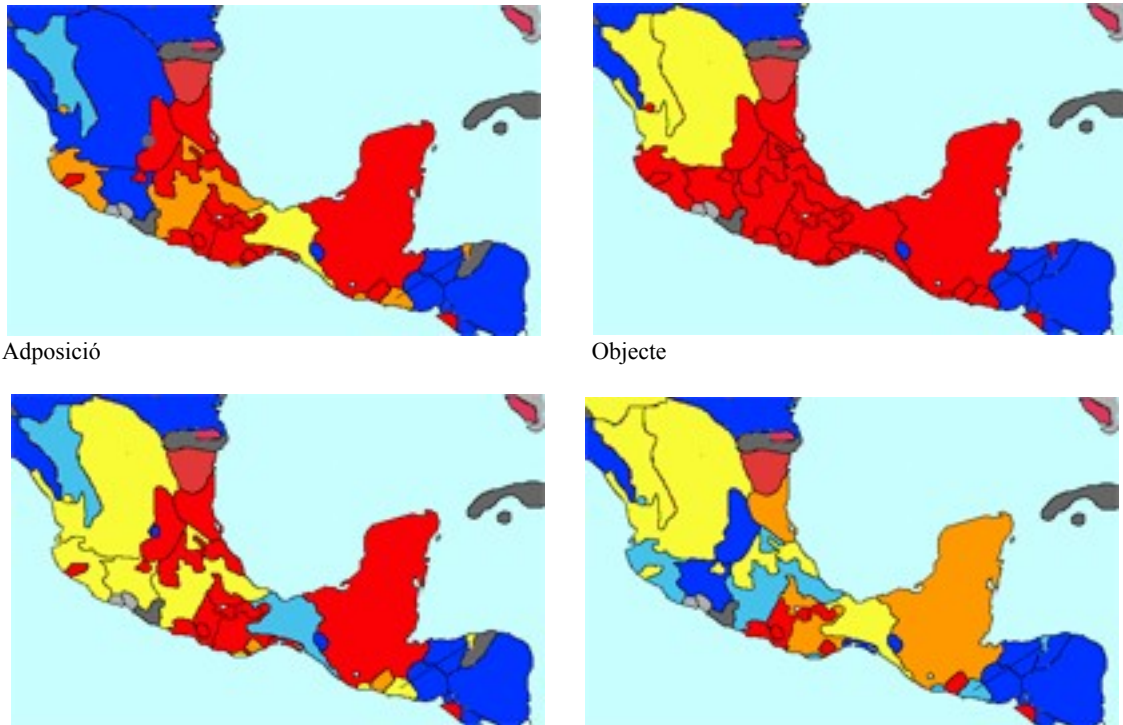
Aquesta regió coincideix, excepte els dubtosos de la costa californiana, amb la suma de les regions Northwest Coast i Plateau les regions culturals dels indis nord-americans definides per antropòleg Kroeber [108] i que coincideix amb àrea de pesca del salmó. És una zona que suposadament va estar coberta de gel a la glaciació i un cop lliure de gel va ser colonitzada per poblacions de diversos grups lingüístics. Malgrat que aquestes poblacions són considerades de cultura paleolítica van arribar a densitats de població molt altes gràcies a una explotació intensiva del salmó i en alguns casos a l'elaboració de farina de peix que ocupava un rol similar a la de cereals en zones neolítiques [109].

Encara que la majoria de *genus* implicats pertanyen un mateix grup lingüístic segons Ruhlen, el fet que també hi ha implicades altres famílies, fa pensar en un fenomen areal, segons es veu als arbres filogenètics.

## 7.4 MESOAMÈRICA



**Arbre 13** Multiarbre dels valors dels de quatre variables a Mesoameèrica.



**Mapa 29** Multimapa dels valors dels de quatre variables a Mesoameèrica.

Mesoamèrica és una regió cultural prehistòrica ben delimitada on conviuen pobles que compartien un conjunt de trets culturals com el sedentarisme, alimentació a base de tres plantes domesticades (blat de moro, carbassa i fesol) i diversos elements religiosos comuns com la construcció de temples en forma de piràmide, el joc de pilota, sacrificis humans o el calendari. A nivell lingüístic es pot mencionar l'ús de sistemes de numeració vigesimal i sistemes d'escriptura pictogràfica. La seva definició és de principis dels anys quarantes per part de Paul Kirchoff [110].

Tal com es veu als arbres, malgrat aquesta homogeneïtat cultural, els seus pobladors procedeixen de grups lingüístics sense cap relació i que van arribar-hi en diferents moments històrics de diverses procedències. Dels dinou *genus* que compondrien la regió, onze pertanyen a un parell de famílies (Oto-Mangue i Utoasteca) i la resta són alguns *genus* importants en nombre de parlants amb parentius discutits com el Maya, el Totonaca o el Mixe-Zoque, o bé *genus* considerats tradicionalment com llengües aïllades com el Huave, el Xinca, el Tarasco i el Cuiclatec.

No està clar el lloc ni el grup lingüístic que va començar a domesticar el blat de moro ara fa uns 8000 anys, però els candidats més destacats serien l'Oto-Mangue i el Mixe-Zoque [111], mentre que l'utoasteca sembla dels darrers en arribar des del nord en temps molt posteriors.

Com en altres casos, l'ordre amb nucli inicial representa per la variable Subjecte una menor freqüència que per la resta de variables. En concret, els *Genus* VS només predomina, i amb les excepcions que suposa la categoria tres de color taronja, a tres zones: Oaxaca, la Península del Iucatan i l'antic Totoncapán al nord. Entre mig predomina el blau propi de l'ordre SV i el groc dels ordres no dominants. Per això és una regió on conviuen equilibradament *genus* entre els diferents ordres possibles.

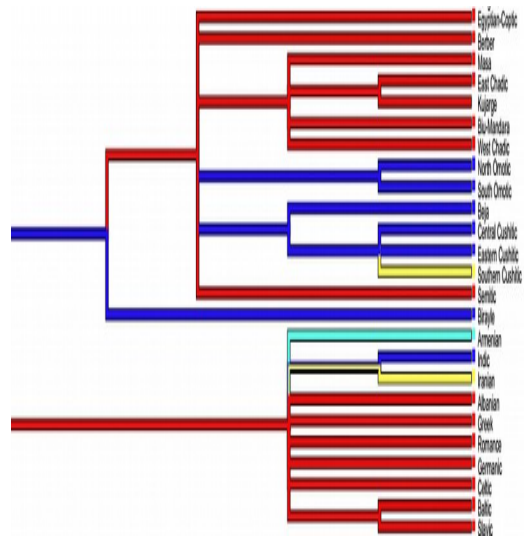
El mapa de la variable Objecte mostra que aquesta variable és la que més ha canviat si suposem un ordre ancestral Objecte-Verb. I la que presenta menys excepcions. Els *genus* Tepiman i Corachol presenten ordre no dominant en aquesta variable igualment que a les variables genitiu i subjecte, malgrat que no formen part de la Mesoamèrica cultural tal com s'ha definit més amunt.

La distribució de la postposició coincideix força amb l'objecte, però alguns *genus* perifèrics com els Tepiman, Corachol i Tarascan són de nucli final. L'únic *genus* sense ordre dominant seria el Mixe-Zoque. El *genus* Asteca està format majoritàriament per llengües preposicionals, contràriament al nàhuatl clàssic que en temps de la conquesta tenia bàsicament posposicions.

En el cas del genitiu, el mapa s'assembla força al de l'objecte, però amb algunes diferències, com en altres zones, en sentit de menor consistència. Respecte a l'objecte, el Nahuatl i el tarasc no mostren ordres dominants i entre les llengües Mixe-Zoque predomina opció de genitiu davant del nom.

En aquesta zona, especialment al nord, moltes llengües han desaparegut i els límits són en realitat menys clars del que el mapa suggereix.

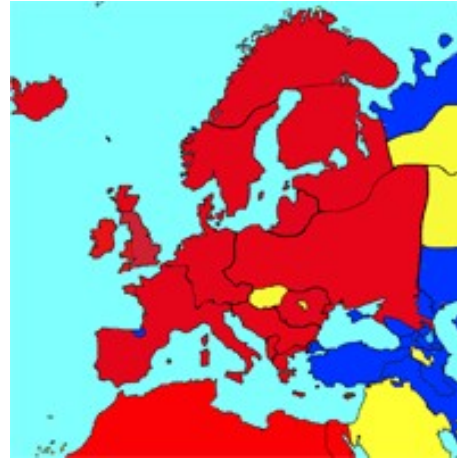
## 7.5 REGIÓ EUROMEDITERRÀNIA



**Arbre 14** Multiarbre dels valors dels de quatre variables a l'àrea Euromediterrània.



Adposició



Objecte



Genitiu



Subjecte

**Mapa 30** Multimapa dels valors dels de quatre variables a l'àrea Euromediterrània.

Com és prou conegut, aquesta regió concentra un gran quantitat de textos molt antics de dotzenes de llengües diferents. En molts casos hi ha hagut continuïtat entre les llengües ancestrals i les seves nombroses descendents contemporànies, la qual cosa permet fer-se una idea més que suficient dels canvis en l'orde de constituents que s'han produït.

Europa està pràcticament ocupada per llengües Indoeuropees excepte al territori basc, el nord d'Escandinàvia i el Caucas on sembla que s'hi mantenen *genus* més antics. Però aquest dibuix simple oculta l'expansió de llengües romàniques, eslaves i germàniques sobre cèltiques, etrusc, l'ibèric i moltes altres menys conegudes. La irrupció de l'uràlic magiar data només del segle XIII.

La riba sud del Mediterrani ha sigut històricament sempre afroasiàtica. Els únics canvis rellevants ha estat l'arraconament de les llengües berbers i la desaparició del copte per la pressió de l'àrab. L'ordre VS no es dona de forma absoluta en cap dels *genus* ara, però entre les llengües afroasiàtiques i cèltiques sembla clar que l'ordre bàsic en temps històrics i encara ho és en moltes de les seves llengües.

L'ordre verb-objecte ocupa el territori de forma pràcticament total a Europa fins els Urals i la costa mediterrània meridional del Marroc fins a Síria. El basc és la única excepció antiga i les llengües caucàsiques que ja quedarien fora de la zona. Les semítiques surten al mapa de color groc, però és perquè aquest *genus* es troba també a Etiòpia amb diferents ordres que afecten a la mitjana. Alguns *genus* de famílies típicament OV com l'Ugric, el Mordvin i el Pèrmic, a més de l'armeni han adoptat l'ordre no dominant possiblement per la influència del rus i altres llengües europees.

El que resulta més sorprenent és el fet que malgrat que a Europa l'ordre VO és pràcticament general, només fa uns 2000 anys, en temps dels romans, la situació era totalment diferent i les llengües de nucli final eren majoria. A molts *genus* s'han donat canvis, com s'ha vist a l'apartat 1-2-6 i també cal tenir en compte que hi ha hagut importants moviments de pobles i conquestes.

Les distribució de les llengües preposicionals coincideix amb les d'ordre VO amb poques excepcions. Per una banda les llengües uràliques d'Escandinàvia i voltants, i per l'altre les llengües que havien adoptat ordres no dominants però que aquí no han canviat. A la riba sud de la mediterrània només trobem preposicions.

En el cas del Genitiu, històricament ens trobem que moltes llengües han canviat de GN a NG adoptant l'ordre de nucli inicial de les adposicions i/o l'objecte. El resultat final és una barreja de colors. Les llengües romàniques, cèltiques, el grec, l'albanès, berbers i l'antic egipci mostren l'ordre NG sempre i una majoria de eslaves i semítiques. Les altaiques, uràliques, el basc i en el Caucas, l'ordre és GN i entre les germàniques i les iràniques no hi ha ordre dominant.

En aquesta regió no cal fer referència a l'agricultura perquè aquesta està present des de fa mil·lennis, juntament amb la ramaderia, i sembla indestruïble de fenòmens com l'increment demogràfic i la urbanització.



L'Amazònia boliviana és un àrea difusa on diversos *genus* pròxims geogràficament, però sense parentiu filogenètic semblen haver canviat pel que fa a les quatre variables estudiades, des d'un ordre bàsic Complement-Nucli a un de Nucli-Complement, encara que a ritmes diferents.

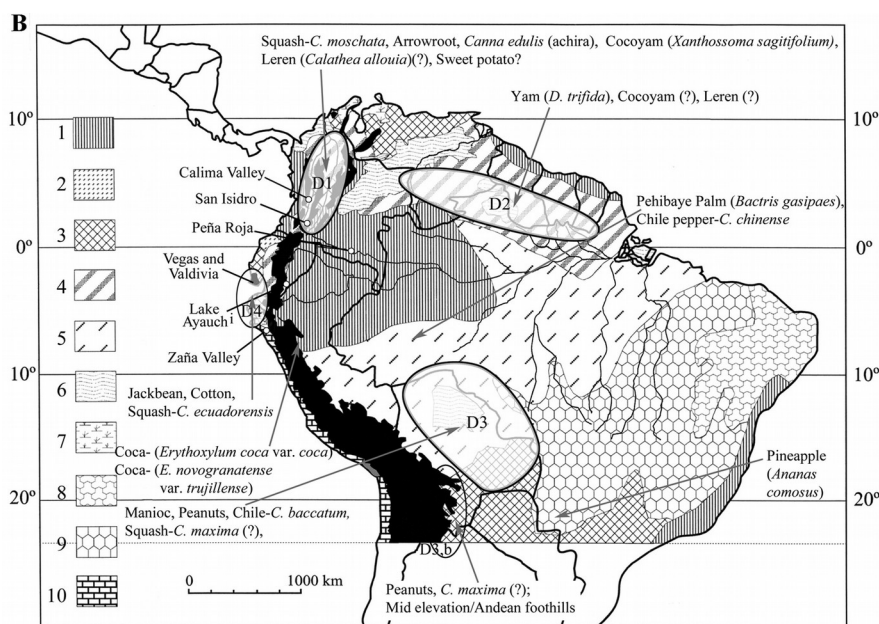
Aquesta regió és potser la que presenta més incertesa a l'hora de fer un mapa. Els 500 anys que han passat des de la conquesta ha fet desaparèixer molts pobles indígenes i qualsevol intenten de reconstruir els límits exactes o passats contactes és pura especulació. Potser per això Amèrica del sud és la zona del món on apareixen més llengües amb ordre Objecte-Verb aïllades i que no poden agrupar-se en cap de les 7 zones descrites.

Malgrat que els *genus* implicats no pertanyen a famílies pròximes, cal destacar que 5 *genus* de la quinzena aproximada que formarien part de la regió pertanyen a la mateixa família, l'arawak. I només cinc dels vuit *genus* d'aquesta família en formen part, mentre que la resta no ha patit cap dels canvis descrits. A partir d'això, es pot conjecturar que els canvis han hagut de ser posteriors a l'aparició de la família i datar-la en només uns pocs mil·lennis.

Els canvis han estat més nombrosos a les variables Genitiu i Objecte, mentre que entre les adposicions hi ha més *genus* sense un ordre dominant. Pel que fa a l'ordre del subjecte els *genus* que han participat del canvi són només una part del total.

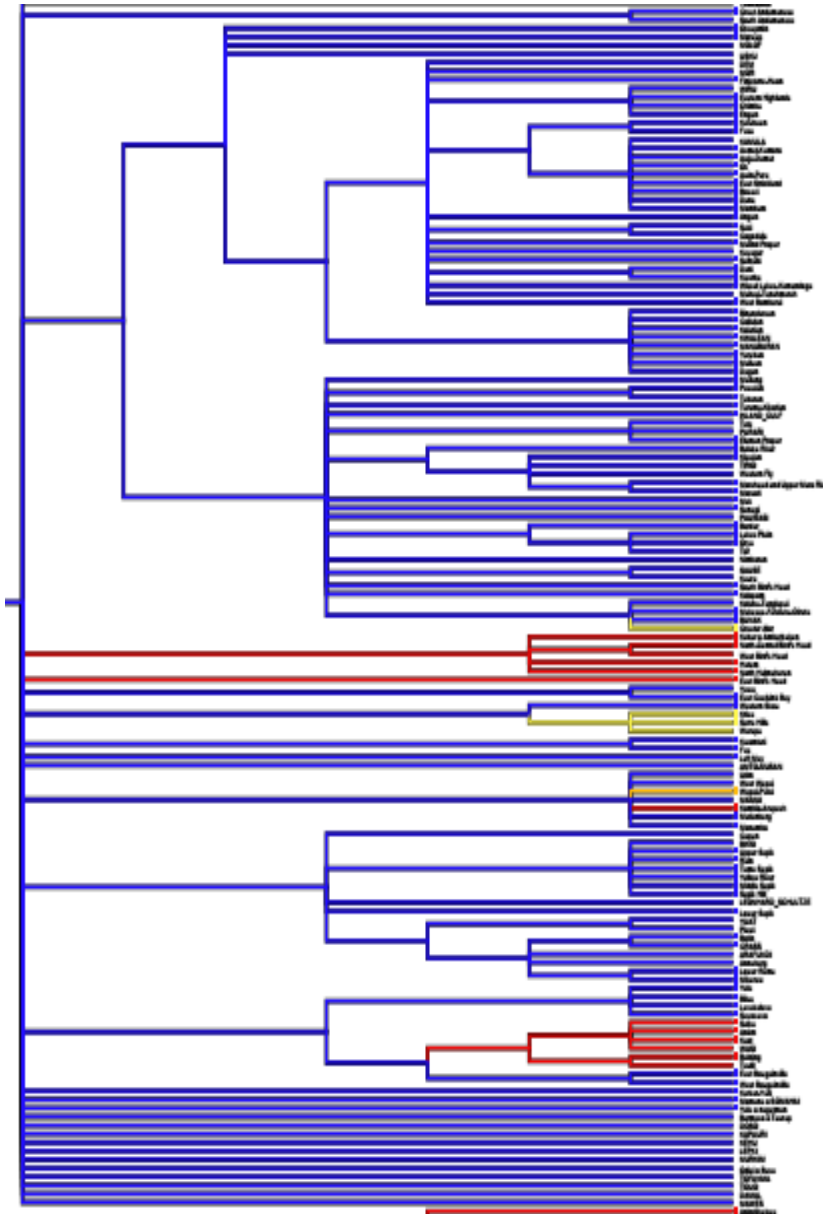
És rellevant que la zona coincideixi amb una àrea de domesticació d'espècies vegetals, concretament de mandioca, xili i cacauet, definida prèviament a [112], en concret la zona D3 del mapa 32.

Alguns dels *genus* involucrats semblen haver migrat posteriorment fora de la zona, escampant les noves disposicions, especialment els de la família arawak que s'ha escampat arreu de l'Amazònia i més enllà fins Hondures i el Paraguai.

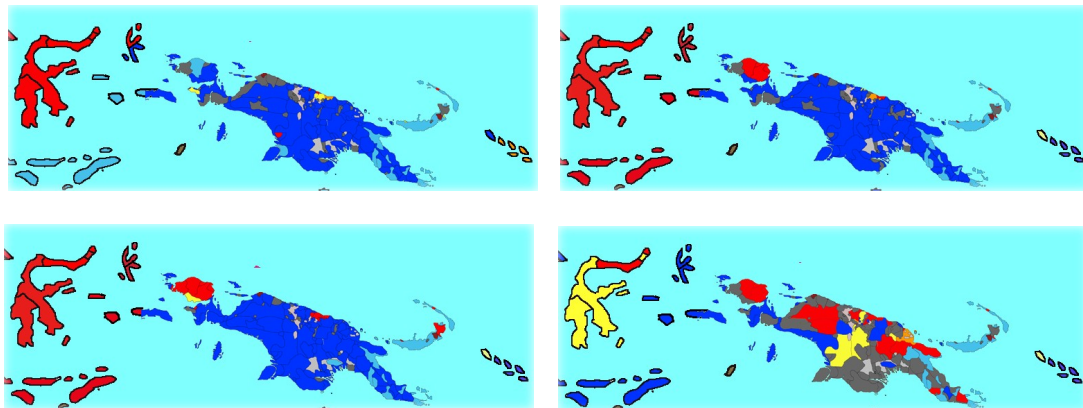


Mapa 32. Zones inicials d'Agricultura sudamericana. Font [112].

## 7.7 NOVA GUINEA

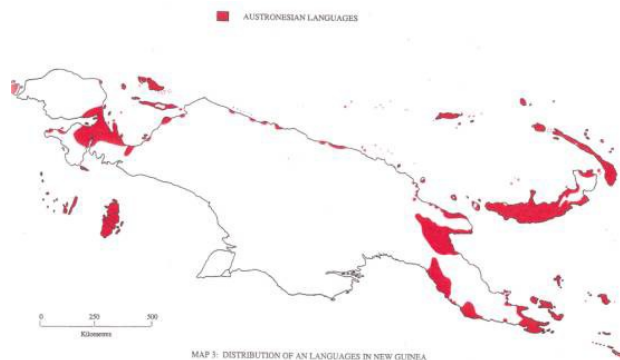


Arbre 16 dels valors dels de quatre variables a Nova Guinea



Mapa 33 Multimapa dels valors dels de quatre variables a Nova Guinea





Mapa 34 Distribució de llengües oceàniques a Nova Guinea

Aquesta vuitena àrea s'afegeix perquè malgrat no ser una àrea on s'ha produït l'aparició de llengües de nucli inicial, el cert és que l'expansió de les llengües austronèsiques les ha fet força presents. La base de la població de Nova Guinea està formada per la població anomenada papú que va arribar fa més de 40.000 anys i que probablement va donar origen a les diferents famílies de llengües papús, encara que no hi ha cap prova lingüística que siguin monolítiques. Hi ha autors que postulen una remota immigració d'elements australoides que encara que dissolts genèticament entre els papús es veurien representats per les llengües Sepik-Ramu ([113]).

L'arribada molt posterior de parlants austronèsics a l'est va donar lloc un *genus* autòcton (South halmahera-New Guinea). Ja en temps històrics, poblacions de parlants de llengües oceàniques ha arribat a les costes occidentals.

Nova Guinea és un escenari interessant per estudiar la interacció entre tipologies perquè tenim una base antiga de llengües papús posposicionals, OV i GN que en molts llocs han conviscut com a veïnes de llengües austronèsiques originalment preposicionals, VO i NG, donant lloc a interaccions tan interessants com diverses.

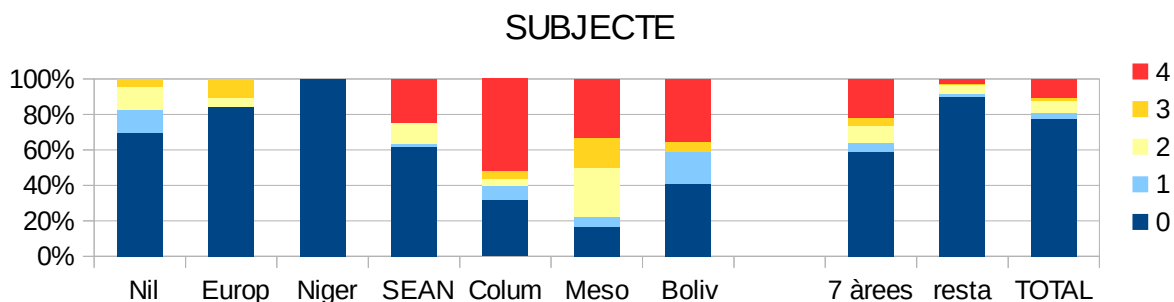
La interacció s'ha donat preferentment a 4 zones diferents. Una zona occidental (la més pròxima a Àsia), una central (la costa nord prop de la actual frontera internacional entre Indonèsia i Papua Nova Guinea), una oriental (la costa nord en el punt més proper a les illes Salomons) i la quarta que serien les illes de Melanèsia.

A la zona occidental, l'arribada és tan antiga [113] que les llengües austronèsiques ja s'han diferenciat prou per a constituir un *genus* endèmic a Nova Guinea i les illes de Halmahera. Les llengües d'aquest *genus* han adoptat en el genitiu l'ordre GN propi de les llengües papús veïnes, mentre que alguns *genus* de llengua papúq han abandonat el seu ordre ancestral (nucli final) pel adoptar el de les llengües oceàniques.

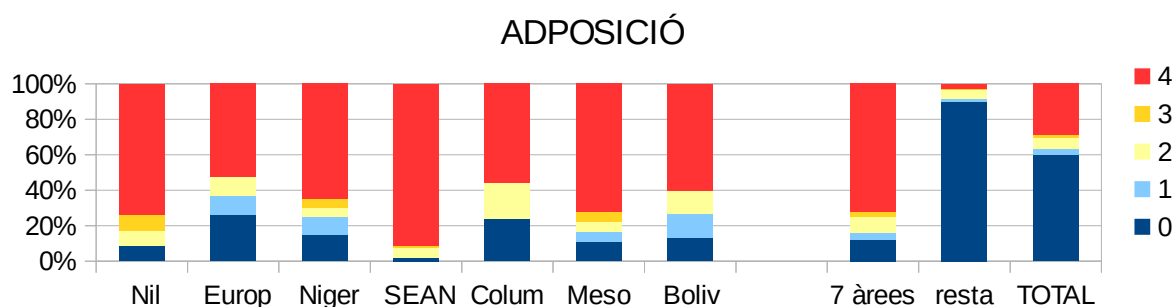
A la zona central de la costa no hi ha hagut canvi del genitiu entre les llengües oceàniques establertes i han estat majoritàriament les llengües papús les han canviat l'ordre de totes tres variables. Aquesta zona és pròxima al nucli de formació d'un crioll malai format al segle XVII [114].

Finalment, a la zona oriental i a les illes de la Melanèsia són les llengües oceàniques les que han canviat sovint el seu ordre original pel de les veïnes llengües papús

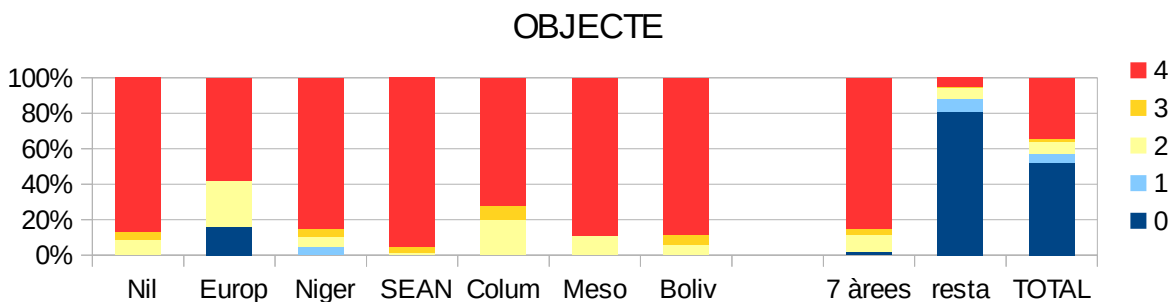
## 7.8 LES SET ÀREES: DADES COMPARATIVES



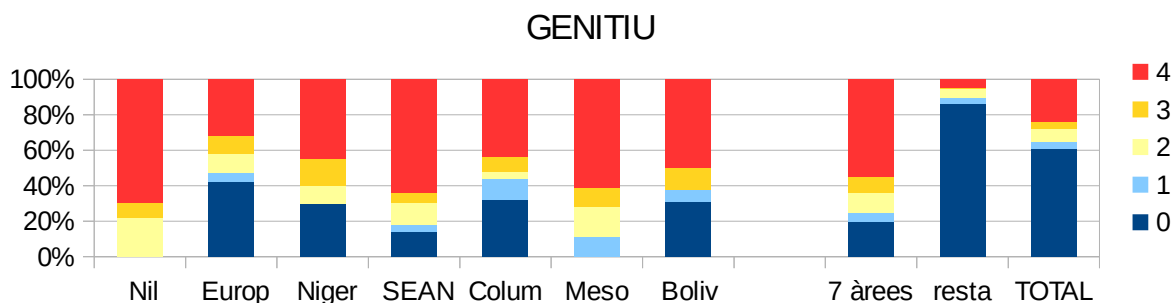
**Gràfic 22** La variable Subjecte segons les set àrees especials i total.



**Gràfic 23** La variable Objecte segons les set àrees especials i total.



**Gràfic 24** La variable Adposició segons les set àrees especials i total.



**Gràfic 25** La variable Genitiu segons les set àrees especials i total.

Els arbres i els mapes de distribució, analitzats en apartats anteriors, apunten a que el comportament i distribució de les diferents variables és diferent entre sí i al mateix temps entre els *genus* en general i els de set zones concretes que ja s'han tractat.

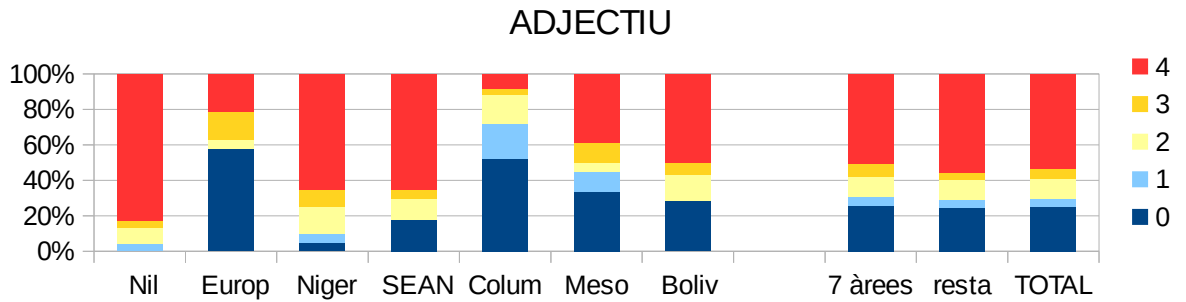
A la pàgina anterior i a la següent es mostren les gràfiques de distribució per cada variable dels quatre variables d'ordre que s'han fet servir fins ara. Hi ha una columna per cada àrea (a la del sud-est asiàtic s'inclouen els *genus* de nucli inicial de Nova Guinea). A les columnes de la dreta hi ha una columna pel total de les set àrees, una per la resta de *genus* i la del total.

La variable Subjecte té una presència molt superior a les set zones comentades que a la resta, però això es deu bàsicament al sud-est asiàtic i a les zones americanes. A la regió del Nil i Euromediterrània les freqüències són força similars al total del món. A la zona del Níger ni tant sols hi ha cap *genus* de nucli inicial.

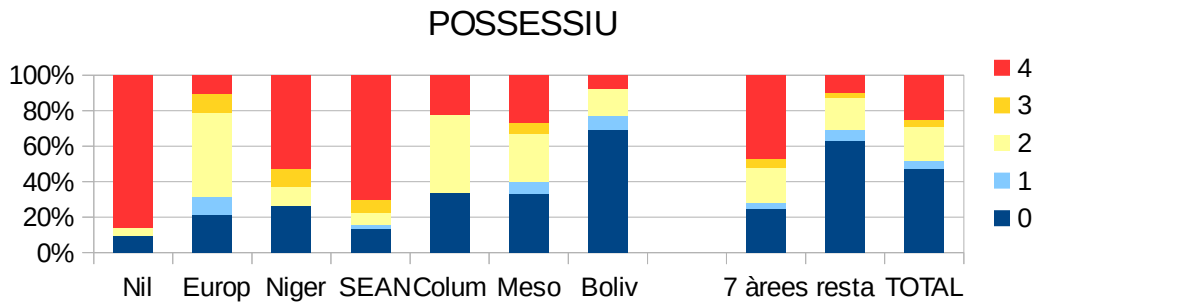
Les variables Objecte, Adposició i Genitiu mostren una distribució tant similar que no es veu cap diferència a primer cop d'ull. A les tres variables es pot veure com les llengües de nucli inicial són clarament majoritàries a les set àrees seleccionades, mentre que a la resta del món pràcticament són inexistentes. Els *genus* de llengües de nucli final són totalment dominants a la columna TOTAL on es compuen els *genus* que no pertanyen a cap de les set regions diferenciades. En canvi, es confirma que fora de les set regions esmentades, les llengües de nucli inicial són molt minoritàries.

Malgrat aquestes semblances, també hi ha alguna diferència. Per exemple, la variable Objecte és molt més homogènia, mentre que a l'Adposició i al Genitiu hi ha molta més presència de *genus* de nucli final i d'intermedis.

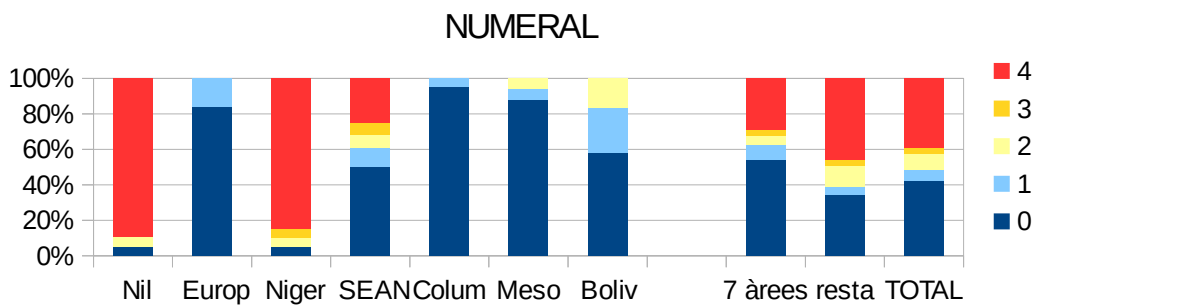
Per àrees, el sud-est asiàtic és segurament la que presenta menys excepcions a l'ordre dominant a la zona i l'àrea Euromediterrània la que té més excepcions. En aquest cas, això es degut a que s'han inclòs alguns *genus* que possiblement han patit canvis degut a la influència de llengües europees com el rus sobre les llengües de l'antiga Unió Soviètica.



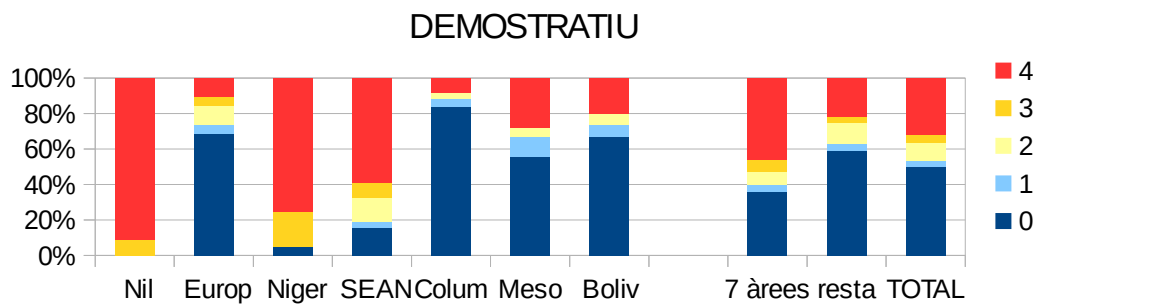
**Gràfic 26** La variable Adjectiu segons les set àrees especials i total.



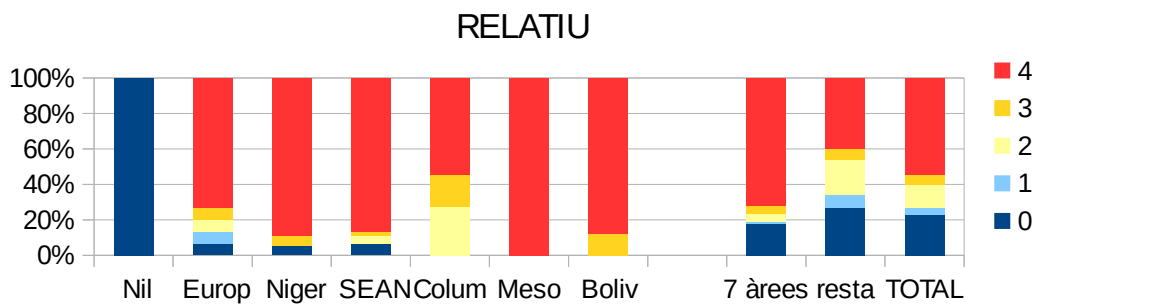
**Gràfic 27** La variable Possessiu segons les set àrees especials i total.



**Gràfic 28** La variable Numeral segons les set àrees especials i total.



**Gràfic 29** La variable Demostratiu segons les set àrees especials i total.



**Gràfic 30** La variable Relatiu segons les set àrees especials i total.

A la pàgina anterior es veu com les variables representades mostren uns patrons força diferents a les vistes prèviament. La primera cosa que destaca és que la freqüència de les llengües amb ordre complement-nucli no varia de forma significativa entre les set zones definides i la resta del món. La segona constatació és que cadascuna de les set àrees es comporta de manera diferent a la resta.

L'adjectiu confirma el que ja s'ha comentat, les columnes de la dreta són pràcticament idèntiques. Malgrat que cada zona té percentatges molt diferents, la mitjana d'elles coincideix amb la de la resta del món.

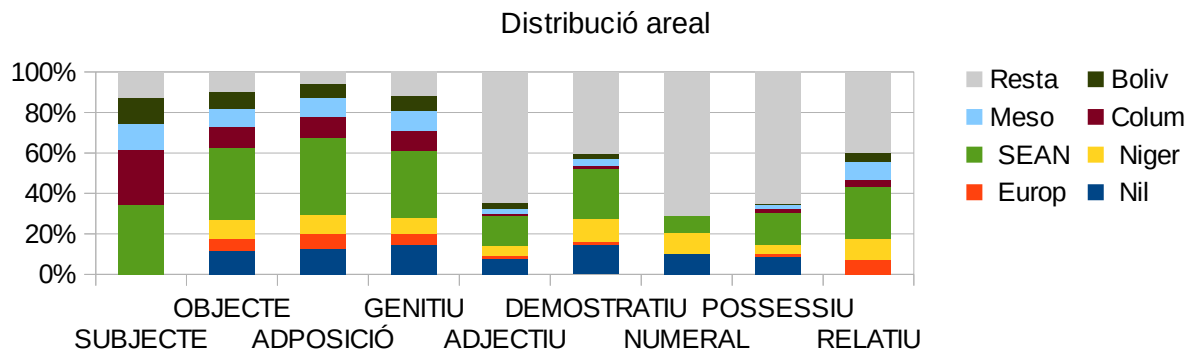
El cas del possessiu és el més dubtós. Cada zona mostra freqüències diferents, però en general força superiors a la de la resta del món. La columna corresponent a l'Amazonia boliviana té un patró pràcticament oposat al de la regió del Nil. La resta mostra distribucions intermèdies.

En el cas de la variable Demonstratiu, les zones d'Amèrica i Euromediterrània mostren columnes similars, contràriament a les africanes i del sud-est asiàtic que també són coincidents entre sí. Aquest alineament per zones també s'insinua a l'adjectiu i als numerals. La mitjana de les set àrees, però, no és clarament diferent al de la resta.

El gràfic del relatiu mostra que a totes les àrees, tret de la zona del Nil, totalment de nucli final, el predomini de l'ordre de nucli inicial és general. Però al contrari que a les variables com l'Objecte, l'Adposició o el Genitiu, aquest ordre a la resta del món és prou comú.

El darrer gràfic és el corresponent als numerals i es manté el patró de zones africanes per un cantó i americanes amb euromediterrànies per l'altre. En aquest cas, però, la situació de les llengües del sud-est asiàtic és intermèdia i s'allunya de les africanes. Però també en aquest cas el resultat de les tres columnes de la dreta és força similar.

## LLENGÜES DE NUCLI INICIAL



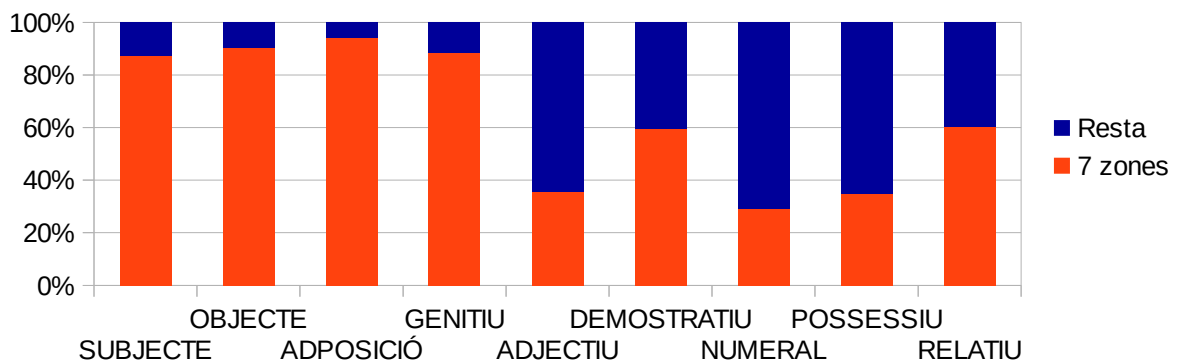
Gràfic 31 La variable Relatiu segons les set àrees especials i total

El gràfic superior permet veure el pes de cadascuna de les set àrees definides, diferenciades amb un color propi, al conjunt de cada variable pel que fa als total de *genus* de tipus 4 (nucli inicial clarament dominant). I també el de les llengües de la resta del món en color gris.

Altres cop queda clar la diferència de comportament entre les quatre variables de les columnes de l'esquerra (Subjecte, Objecte, Adposició i Genitiu) i les cinc de la dreta (Adjectiu, Demonstratiu, Numeral, Possessiu i Relatiu). A les quatre primeres, els *genus* de les set àrees estudiades representen la pràctica totalitat, mentre que a les altres cinc són una part significativa, però els *genus* de la resta del món representen una proporció pràcticament igual o fins i tot clarament majoritària.

En altres paraules, els *genus* de tipus 4 que s'han pintat de vermell als mapes pràcticament només es donen a les set àrees definides i a les seves zones d'expansió (com és el cas de les llengües austroasiàtiques o bantus)

## LLENGÜES DE NUCLI INICIAL



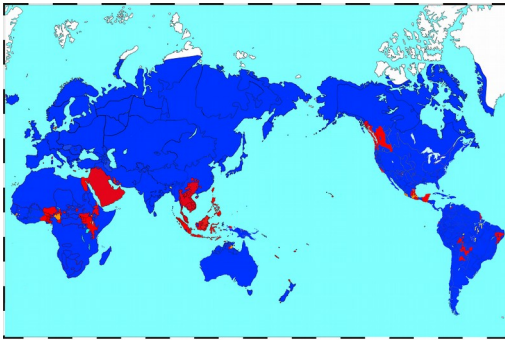
Gràfic 32

El gràfic superior és igual al gràfic 34, però s'han ajuntat els *genus* de les set zones especials per veure més clarament la seva rellevància a les quatre variables de les columnes de l'esquerra. Per aquestes variables, pot dir-se que el pes dels *genus* de la resta del món és insignificant, al contrari que a les altres cinc variables.

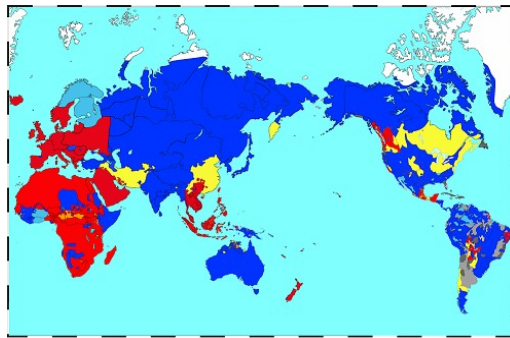
Aquest gràfic constitueix un resum de la tesi sencera al deixar clar que no totes les variables segueixen un mateix comportament i que les llengües amb nucli inicial per aquestes variables només es donen a poques zones del món.

## 7.9 VISIÓ DIACRÒNICA

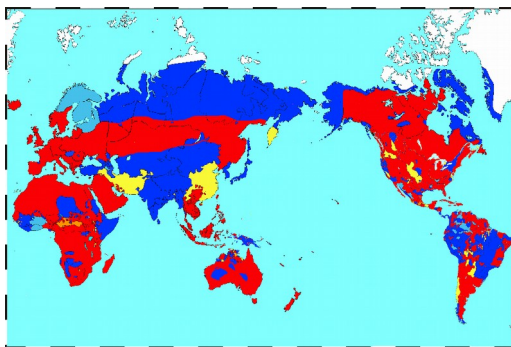
Les dades sincròniques, amb l'ajuda dels mapes i arbres, permeten inferir quin ha estat el canvi global que s'ha donat pel que fa als ordres de constituents.



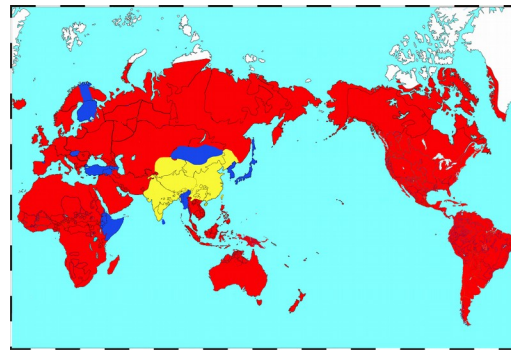
A: Fase inicial



S.XV



S.XIX



Llengües oficials al segle XXI.

**Mapa 35** Multimapa que permet mostrar el canvi temporal de la distribució de llengües de nucli inicial.

La seqüència de mapes permet visualitzar el que es conclou a partir de les dades de la tesi. Les llengües de nucli inicial van néixer lligats a unes set zones concretes del planeta i des d'allà es van anar estenen fins representar quasi bé la meitat del món i des d'allà es van estendre gràcies a l'ajuda de les llengües colonitzadores al segle XIX.

El recent procés de desaparició de llengües minoritzades i el fet que la pràctica totalitat de llengües oficials siguin de nucli inicial, permet especular que el seu percentatge s'incrementarà de forma dràstica en el futur.

Givon, al 1979 [49] va defensar que la majoria de les famílies lingüístiques eren predominantment SOV o es podien derivar d'un estat ancestral SOV tirant enrere fins a 7000 o 8000 que es pot considerar el límit temporal del mètode comparatiu. D'aquí va deduir que fins fa uns 8.000 anys l'ordre predominant era justament el SOV .

Aquesta tesi intenta explicar la correlació en temps històrics de les llengües de nucli final i nucli inicial a partir d'uns procés concret, però no per això s'exclou la possibilitat que altres fenòmens puguin portar a l'aparició de llengües de nucli inicial. De fet, aquestes llengües es troben també, si bé en un nombre molt reduït, entre les parlades per pobles caçadors-recolectors d' Austràlia o el Kalahari, entre altres.





## 8 DISCUSSIÓ I MODEL EXPLICATIU

Aquesta tesi ha adoptat una aproximació multidisciplinària sobre la qüestió de la variabilitat en l'ordre de constituents a arreu del món.

El capítol 1 és sobre l'interès i objectius de la tesi. Al capítol 2, la introducció, s'ha fet una descripció somera de l'estat de la qüestió, la distribució a grans trets dibuixada pels experts i els principals patrons que s'han proposat per explicar-la. També s'han examinat els darrers treballs publicats últimament per proposar un model explicatiu i la crítica dels mateixos. S'ha posat especial atenció a les propostes que buscaven lligams entre la tipologia lingüística i factors socials com la població o el caràcter agrícola i sedentari de les comunitats de parlants.

La metodologia (capítol 3) emprada s'ha basat en reunir la informació pertinent de les bases de dades més complertes de lliure accés per tal de d'aplicar-hi estadística bàsica i representar-les en mapes areals i arbres filogenètics. La major part de les dades s'ha aplicat als *genus*, en lloc de les llengües, per tal d'optimitzar-ne l'ús.

El capítol 4 mostra com l'aplicació de la tècnica estadística coneguda com Anàlisi de Components Principals, en aquest cas a les llengües i no als *genus*, ha demostrat que els diferents constituents sintàctics no varien de forma solidària sinó que mostren patrons particulars. Malgrat l'anterior afirmació també s'ha demostrat que aproximadament la meitat de la variabilitat total és atribuïble a un únic factor principal (F1) que impacta de forma distinta a cada variable. La identitat d'aquest factor és discutible però coincideix en prop d'un 90% (la xifra varia lleugerament segons la base de dades empleada) amb el valor de l'adposició o del Verb-Objecte.

La presentació de les dades en diagrames de barres (capítol 5) confirma que les diferents variables no segueixen un patró únic, però que l'objecte, l'adposició i el genitiu si que es comporten de forma similar, tal com ja apuntava el fet que totes tres siguin proporcionals a un mateix factor principal F1.

L'estudi particular de la variable adposició mostra una asimetria en la seva distribució segons els factors. Així, el percentatge de llengües estrictament preposicionals suposa un 50% de la mostra i un 75% dels parlants, però només un 30% dels *genus* i un 20% de les famílies. En concret, pel que fa al nombre de parlants de cada *genus*, els estrictament preposicionals van augmentant el seu percentatge de forma proporcional d'un 20% entre els de menys de 10.000 parlants fins ser majoritaris entre els de més de 10 milions de parlants. També hi ha percentatges molt més alts per les grans famílies que entre les llengües aïllades de manera que quan més ramificada és una família més presents hi són les preposicions.

El percentatge de *genus*, nombre de parlants i llengües preposicionals també varia enormement segons el continent sense que hi hagi cap patró universal.

Els capítols 6 i 7 representen els valors de la variable Verb-Objecte en mapes areals i arbres filogenètics, demostrant que un 90% dels *genus* VO s'han originat en set zones concretes ( Delta del Níger, tram alt del Nil, regió Euromediterrània, sud-est asiàtic, Amazònia boliviana, Mesoamèrica i Colúmbia britànica) que no representen ni un 10% de la superfície dels continents. I que des d'allà s'han expandit fins cada cop més. Cal destacar que els *genus* implicats de cada regió no estan tots emparentats filogenèticament sinó que el fenomen és areal i afecta a grups d'origen distint.

Aquestes set zones coincideixen amb aquelles en què s'han donat importants increments de població a causa de l'agricultura intensiva i una barreja de comunitats lingüístiques pertanyents a un elevat nombre de *genus*. La única excepció serien poblacions pescadores de la Colúmbia britànica que usaven la farina de peix en lloc de la vegetal. També s'han trobat correlacions entre el percentatge de l'alimentació d'origen agrícola, el nombre de parlants i el percentatge de preposicions. Les comunitats no agrícoles tendeixen a usar llengües de nucli final i aquelles que depenen de l'agricultura en més d'un 85% per les de nucli inicial de manera clara (gràfic 18).

Al final de la tesi, s'han trobat també correlacions entre l'ordre Subjecte-Objecte i factors fonològics com el nombre de fonemes, la longitud de paraula. I entre el nombre de fonemes i la població dels *genus*. La significació d'aquestes correlacions és dubtosa.

Dels nombre de fonemes i la longitud de paraula, s'han confeccionat mapes per grans àrees que mostren com a més de les set àrees implicades en l'ordre de constituents, on augmenta el nombre de fonemes i disminueix la longitud de paraula, sembla clar que hi juga un paper la distància a les llengües khoisanides africanes, de manera que es confirmaria parcialment la hipòtesis que la dispersió de la humanitat ha anat acompanyada per un decreixement en la complexitat fonètica.

Compatible amb aquesta tendència, el principal resultat de la tesi és que aquesta primera expansió va ser protagonitzada per parlants de llengua de nucli final pel que fa a les adposicions, l'objecte i el genitiu. No seria fins molts mil·lennis després, que a unes àrees concretes es va donar les condicions perquè es donés un canvi radical d'ordre que va portar al domini de llengües de nucli inicial que posteriorment s'han escampat de forma progressiva.

Les set àrees originals tenen en comú un increment galopant de la població, bàsicament degut a processos d'adopció de l'agricultura i juntament amb una alta diversitat de *genus* diferents pertanyents a famílies no emparentades.

Mentre que a partir de les dades sembla clar que l'àmplia distribució de les llengües preposicionals Objecte-Verb és el resultat de l'expansió recent a partir de set petites àrees que han sofert processos similars lligats a l'increment de població i l'agricultura. No és fàcil explicar perquè es donen canvis similars reiterats a punts tant llunyans.

En contrast amb el fet que les llengües VO semblen predominar en el cas de pidgins i criolls, entre les llengües de signes, la tendència es cap a l'ordre OV fins i tot quan la llengua de la comunitat on viuen els sords és VO.

L'única explicació versemblant és que la confluència d'una població elevada amb important barreja de comunitats lingüístiques generi un elevat nombre de parlants L2 que aprenguin la llengua com adults, sense els avantatges d'un aprenentatge precoç, i això afecti a la gramàtica com s'ha demostrat en casos com els de les llengües franques, criolls o llengües en vies d'extinció. A totes elles hi ha una tendència a la simplificació morfològica que podria afectar la sintaxi.

## MODEL EXPLICATIU

	SE Àsia	Índia	Japó	Europa	Nil	Níger	Perú	Mèxic	Oregon
OV → VO	+	-	-	+	+	+	-	+	+
δ població	+	+	+	+	+	+	+	+	+
cultius	+	+	+	+	+	+	+	+	+?
diversitat	+	-	-	+	+	+	-	+	+

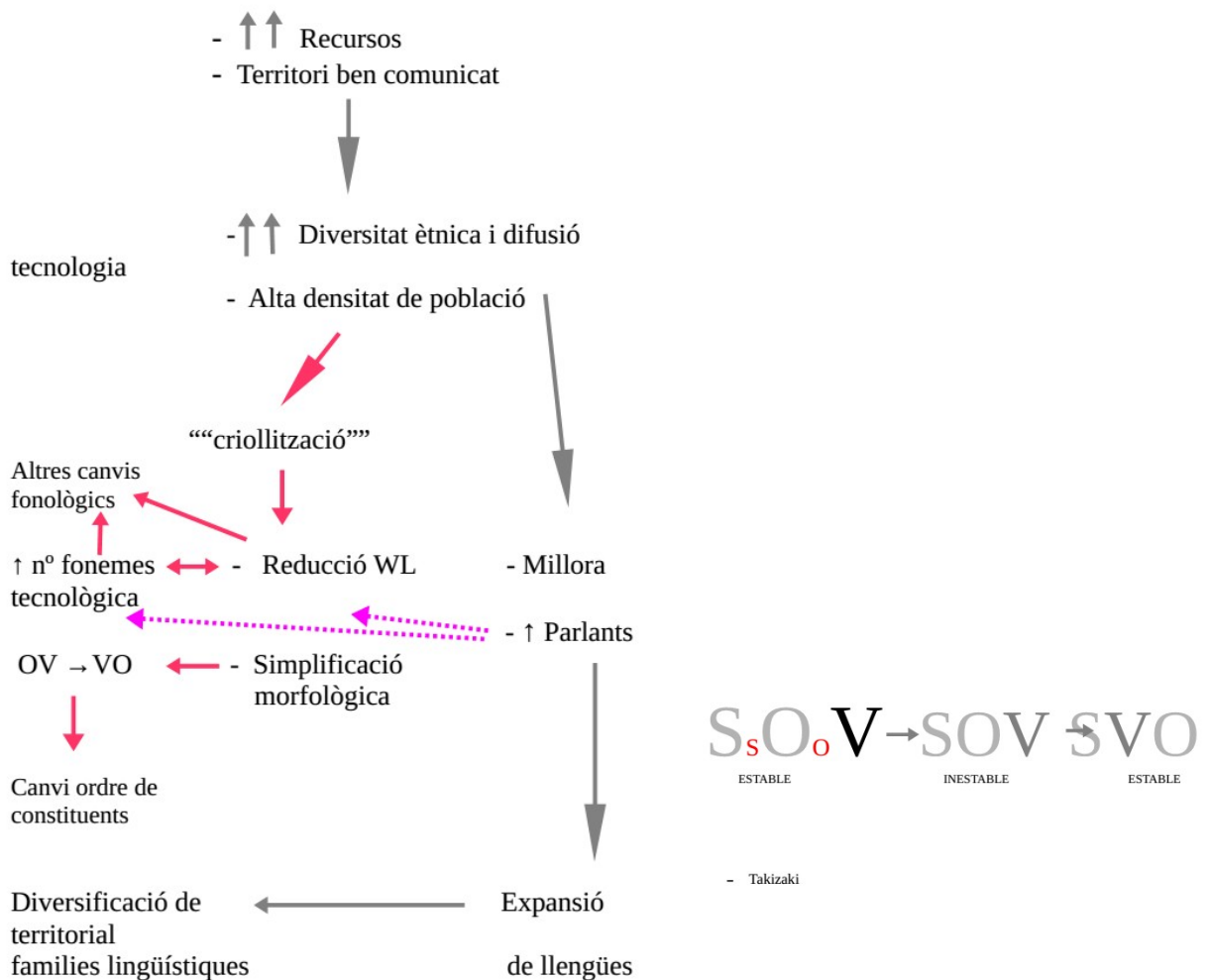
	L.França	Obsolescència	Criolls	OV → VO
Canvi d'ordre constituents	-	+	+?	+
No hi ha tons	-?	?	+	+
→ Sil.labes simples	-	+?	+	+
Reducció WL	-	?	+	+
→ Morfologia simple	+	+	+	+
Sintètic → Analític	+	+	+	+

**Taula 23** Adalt i **Taula 24** abaix.

No era objectiu d'aquesta tesi concebre el mecanisme concret que expliqui els patrons de variació tipològica dels constituents, sinó només trobar els patrons de distribució i alguns possibles factors que poguessin servir com a candidats per explicar-los. Però un cop acabada la tesi ha estat difícil no intentar proposar un model explicatiu.

La taula 23 mostra que el canvi OV a VO no s'ha donat sempre allà on hi ha grans civilitzacions agràries. Com es pot veure a la taula, no s'ha donat ni a la Índia ni al Perú ni al Japó, i si a la Columbia britànica on teòricament vivien pescadors paleolítics, encara que amb una densitat de població altíssima. El canvi només s'ha donat allà on també s'ha donat una trobada i/o barreja entre un munt de *genus* diferents.

La taula 24 mostra que el canvi OV a VO coincideix amb altres canvis que també es produeix a les llengües franques, els criolls i les llengües en procés d'obsolescència. I entre aquests tres casos, l'element comú és el fet que hi ha una interrupció de la transmissió lingüística ancestral entre unes generacions i les següents, de manera que molts adults parlen una llengua que no és la que van aprendre de nens.



Gràfic 33. Model explicatiu

El gràfic 33 resumeix un possible mecanisme d'aparició de zones amb predomini de llengües de nucli inicial.

Si en un territori ben comunicat hi ha un augment de recursos important és raonable esperar un increment de la densitat de població i de la diversitat ètnica. Això és el que hauria passat als territoris on van néixer societats agrícoles. L'acumulació d'excedents en forma de farina permet fer front a les èpoques d'escassetat, sedentaritzar la població i atreu a la població circumdant, possiblement de llengua i cultura diferent.

A nivell social, aquestes noves societats tenen força possibilitats d'expandir-se a costa de les veïnes per la seva major població i la suma de tecnologies que acumulen. El resultat pot ser una diversificació territorial de les llengües i famílies lingüístiques involucrades i també un augment de parlants L2 que haurien après aquestes llengües ja d'adults..

A nivell lingüístic, les dificultats d'un aprenentatge d'aquests L2, majoritàriament adults, i el seu elevat nombre, propícia l'aparició dels canvis similars als d'una criollització llistats a la taula 24 com l'increment de fonemes, la reducció de WL una simplificació de la morfologia que esdevé més analítica. En el cas que la reducció afecti a les marques de cas, a les llengües OV es poden produir casos d'ambigüitat entre subjecte i objecte que només desapareixen si canvien a VO. Una llengua OV sense marca sempre genera dubtes per distingir el subjecte de l'objecte. Posteriorment aquest canvi de l'ordre de l'objecte acaba arrossegant a les preposicions i genitius i acaba arribant altres constituents.

## 9 CONCLUSIONS

Les principals aportacions de la tesi són:

- La utilització del *genus* com unitat d'anàlisi és adequada per fer-se una idea de conjunt allà on la desaparició de llengües o la manca d'informació deixaria forats. També permet fer mapes areals i estudiar les variables a través dels arbres filogenètics de forma que seria impossible de fer si es fes servir la llengua com unitat.
- Creació de mapes areals de la distribució de nou variables tal com serien abans de l'expansió de les llengües colonials.
- Proposta d'història dels canvis soferts al conjunt de grups lingüístics per nou variables des dels orígens fins l'actualitat d'acord al principi de màxima parsimònia .
- Un model explicatiu de les principals variables tipològiques en base a elements extralingüístics com la població i l'agricultura que concorda amb les dades acumulades fins avui dia.

Les conclusions són:

- 1) Un 50% de la variabilitat de 9 ordres de constituents s'expliquen per un factor F1.
- 2) Els valors de les variables d'ordre de constituents es distribueixen geogràficament en àrees generalment consistents.
- 3) No tots els constituents es comporten de la mateixa manera, però si que ho fan les adposicions i/o el Subjecte-Objecte i el Genitiu-Nom que correlacionen fortament amb el factor F1 obtingut a la prova d'ACP.
- 4) Els constituents que correlacionen amb F1 presenten una asimetria de manera que aquells *genus* on l'ordre dominant és el de nucli final correlacionen amb un nombre de parlants baix, famílies poc ramificades, no agrícoles.
- 5) Els *genus* de nucli inicial han aparegut pràcticament només en set petites regions on es dona la coincidència de l'agricultura i una alta diversitat lingüística i des elles s'han expandit fins esdevenir majoritàries.
- 6) L'equilibri aproximat entre el nombre de llengües de nucli inicial i el de nucli final és fruit de la casualitat, al donar-se un decreixement històric.

La dispersió dels humans fora d'Àfrica sembla haver estat protagonitzada per grups petits de caçadors-recol·lectors que parlaven llengües OV, amb posposicions i predomini de l'ordre Complement-Nucli. L'aparició de l'agricultura no sembla haver canviat les coses allà on poblacions agrícoles es van superposar a caçadors recol·lectors (l'Índia, Andes, Japó o l'Europa preromana) i van mantenir l'ordre ancestral OV.

A set zones del món es va produir un canvi de l'ordre OV a VO i aquest sembla haver arrossegat a molts altres constituents i potser s'ha produït un reequilibri que hauria afectat al nombre de fonemes, la longitud de paraules o la complexitat morfològica.



## BIBLIOGRAFIA

- [1] GIRARD, Gabriel. 1747. *Les Vrais Principes de la Langue Française*. Lebreton, Paris, 1982.
- [2] DRYER, Matthew S., et al. *The world atlas of language structures*. 2005.
- [3] GREENBERG, Joseph H. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of language*, 1963, vol. 2, p. 73-113.
- [4] DUNN, Michael, et al. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 2011, vol. 473, no 7345, p. 79-82.
- [5] LEHMAN, Jill Fain, et al. *The role of language in cognition: A computational inquiry*. 1993.
- [6] VON HUMBOLDT, Wilhelm Freiherr. *Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*. (Lettre à M. Jacquet sur les alphabets de la Polynésie Asiatique.). 1836.
- [7] VON DER GABELENTZ, Georg. *Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherigen Ergebnisse*. Weigel, Leipzig, 1891.
- [8] WEIL, Henri. 1978. *The Order of Words in the Ancient Languages Compared with that of the Modern Languages*. 1844.
- [9] TESNIÈRE, Lucien. *Eléments de syntaxe structurale*. Librairie C. Klincksieck, 1959.
- [10] HAWKINS, John A. *Word order universals: Quantitative analyses of linguistic structure*. 1983.
- [11] LEHMANN, Winfred P. A structural principle of language and its implications. *Language*, 1973, p. 47-66.
- [12] VENNEMANN, Theo. Analogy in generative grammar: the origin of word order. *En Proceedings of the eleventh international congress of linguists (1972)*. 1974. p. 79-83.
- [13] TOMLIN, R. S. *Basic Word Order: Functional Principles* (Croom Helm, London). NClS–Noun Classifier NNum–Noun Numeral Nom–Nominative Rel–Relative NRel–Noun Relative RelN–Relative Noun, 1986.
- [14] DRYER, Matthew S. The Greenbergian word order correlations. *Language*, 1992, p. 81-138.
- [15] EHARA, Terumasa. Relation among Word Order Parameters Analyzed by Multi-Dimensional Scaling. *En Proceedings of The first Annual Meeting of The Association for Natural Language Processing*. 1995. p. 173-176.
- [16] EHARA, Terumasa. *Word Order Characteristics Analyzed by Multi Dimensional*

Scaling. En Proceedings of The 13th Annual Meeting of The Association for Natural Language Processing, A1-3. 2007.

[17] DUNN, Michael, et al. Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 2011, vol. 473, no 7345, p. 79-82.

[18] CROFT, William, et al. Greenbergian universals, diachrony and statistical analyses. *Linguistic Typology*, 2011, vol. 15, no 2, p. 433-453.

[19] EVANS, Nicholas; LEVINSON, Stephen C. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and brain sciences*, 2009, vol. 32, no 05, p. 429-448.

[20] GIVÓN, Talmy. Historical syntax and synchronic morphology: an archaeologist's field trip. En *Chicago Linguistic Society*. 1971. p. 394-415.

[21] GIVÓN, Talmy. *Syntax: A functional-typological approach*. Amsterdam: J. Benjamins, 1984.

[22] DIXON, Robert MW. *The rise and fall of languages*. Cambridge University Press, 1997.

[23] GELL-MANN, Murray; RUHLEN, Merritt. The origin and evolution of word order. *Proceedings of the National Academy of Sciences*, 2011, vol. 108, no 42, p. 17290-17295.

[24] GOULD, Niles Eldredge-Stephen Jay. *Punctuated equilibria: an alternative to phyletic gradualism*. 1972.

[25] NETTLE, Daniel. Language diversity in West Africa: An ecological approach. *Journal of Anthropological Archaeology*, 1996, vol. 15, no 4, p. 403-438.

[26] HOLMAN, Eric W. Why are language families larger in some regions than in others?. *Diachronica*, 2004, vol. 21, no 1, p. 57-84.

[27] BELLWOOD, Peter. The Austronesian dispersal and the origin of languages. *Scientific American*, 1991, vol. 265, no 1, p. 88-93.

[28] BELLWOOD, Peter S.; RENFREW, Colin. Examining the farming/language dispersal hypothesis. McDonald Institute for Archaeological Research, 2002.

[29] RENFREW, Colin. Problems in the general correlation of archaeological and linguistic strata in prehistoric Greece: the model of autochthonous origin. *Bronze Age migrations in the Aegean*, 1973, p. 263-76.

[30] RENFREW, C. *Archaeology and language: the puzzle of Indo-European origins*. 1987.

[31] RENFREW, Colin. At the edge of knowability: towards a prehistory of languages. *Cambridge Archaeological Journal*, 2000, vol. 10, no 01, p. 7-34.

[32] RENFREW, Colin. World linguistic diversity. *Scientific American*, 1994, vol.



270, no 1, p. 104-111.

[33] BELLWOOD, Peter. Early agriculturalist population diasporas? Farming, languages, and genes. *Annual Review of Anthropology*, 2001, p. 181-207.

[34] RENFREW, Colin. World linguistic diversity and farming dispersals. *Archaeology and language I: theoretical and methodological orientations*, 1997, p. 82-90.

[35] CAMPBELL, Lyle; POSER, William J. *Language classification. History and method*. Cambridge, 2008.

[36] FENK, August; FENK-OCZLON, Gertraud. Crosslinguistic computation and a rhythm-based classification of languages. En *From Data and Information Analysis to Knowledge Engineering*. Springer Berlin Heidelberg, 2006. p. 350-357.

[37] FENK-OCZLON, Gertraud; FENK, August. Crosslinguistic correlations between size of syllables, number of cases, and adposition order. *Sprache und Natürlichkeit, Gedenkband für Willi Mazerthaler*, 2005, p. 75-86.

[38] FENK-OCZLON, Gertraud; FENK, August. Cognition, quantitative linguistics, and systemic typology. *Linguistic Typology*, 1999, vol. 3, no 2, p. 151-178.

[39] HAY, Jennifer; BAUER, Laurie. Phoneme inventory size and population size. *Language*, 2007, 83.2: 388-400.

[40] LUPYAN, Gary; DALE, Rick. Language structure is partly determined by social structure. *PloS one*, 2010, 5.1: e8559.

[41] ATKINSON, Quentin D. Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, 2011, 332.6027: 346-349.

[42] HOLMAN, Eric WICHMANN., et al. Automated dating of the world's language families based on lexical similarity. *Current Anthropology*, 2011, vol. 52, no 6, p. 841-875.

[43] WICHMANN, Søren; MÜLLER, André; VELUPILLAI, Viveka. Homelands of the world's language families: a quantitative approach. *Diachronica*, 2010, vol. 27, no 2, p. 247-276.

[44] WICHMANN, Søren; RAMA, Taraka; HOLMAN, Eric W. Phonological diversity, word length, and population sizes across languages: The ASJP evidence. *Linguistic Typology*, 2011, 15.2: 177-197.

[45] WICHMANN, Søren. On the power-law distribution of language family sizes. *Journal of Linguistics*, 2005, vol. 41, no 01, p. 117-131.

[46] HOLMAN, Eric WICHMANN. Why are language families larger in some regions than in others?. *Diachronica*, 2004, vol. 21, no 1, p. 57-84.

[47] WICHMANN, Søren. *Neolithic linguistics. Language and prehistory of the Indo-European peoples*, Budapest: *Archae-olingu*. Forthcoming, 2010.

- [48] YAMAMOTO, Hideki. A survey of areal distribution of word order around the world. *STUF-Language Typology and Universals*, 1999, vol. 52, no 1, p. 64-77.
- [49] YAMAMOTO, Hideki. A historical view on areal distribution of word order around the world. *STUF-Sprachtypologie und Universalienforschung*, 2005, vol. 58, no 4/2005, p. 374-391.
- [50] GIVÓN, Talmy. From Discourse to Syntax: Grammar as a Processing Strategy in Discourse and Syntax. *Syntax and Semantics Ann Arbor, Mich.*, 1979, vol. 12, p. 81-112.
- [51] Mapa de las migraciones humanas fuera de África, versión de Naruya Saitou y Masatoshi Nei (2002) del Instituto Nacional de la Genética del Japón. Es una traducción del japonés original. <http://www.museum.kyushu-u.ac.jp/WAJIN/113.html>
- [52] GORDON JR, Raymond G. *Ethnologue: Languages of the World*. Dallas, Tex: SIL International, 2005. Online version: <http://www.ethnologue.com>, 2007.
- [53] CALLAWAY, Ewen; ADAM, Genetic. Eve did not live too far apart in time. *Nature International Weekly Journal of Science*, 2013.
- [58] Cavalli-Sforza, L. L., Piazza, A., Menozzi, P., & Mountain, J. (1988). Reconstruction of human evolution: bringing together genetic, archaeological, and linguistic data. *Proceedings of the National Academy of Sciences*, 85(16), 6002-6006.
- [59] GEPTS, Paul. *Origins of plant agriculture and major crop plants. Our fragile world: Challenges and opportunities for sustainable development*. Oxford: EOLSS Publishers, 2001, p. 629-637.
- [60] HALSTEAD, Paul. The development of agriculture and pastoralism in Greece: when, how, who and what. The origins and spread of agriculture and pastoralism in Eurasia, 1996, p. 296-309.
- [61] ZONG, Yongqiong, et al. Fire and flood management of coastal swamp enabled first rice paddy cultivation in east China. *Nature*, 2007, vol. 449, no 7161, p. 459-462.
- [62] RAO, Ch Krishna. *Adoption of Tissue Culture in Horticulture: A Study of Banana-Growing Farmers from a South-Indian State*. Cambridge Scholars Publishing, 2014.
- [63] RUHLEN, Merritt. *A guide to the world's languages: Classification*. Stanford University Press, 1991.
- [64] GIMBUTAS, Marija Alseikaitė; HENCKEN, Hugh O'Neill. *The Prehistory of Eastern Europe: Mesolithic, Neolithic and Copper Age Cultures in Russia and the Baltic Area*. Peabody Museum, 1956.
- [65] FULLER, Dorian Q. Pathways to Asian civilizations: Tracing the origins and spread of rice and rice cultures. *Rice*, 2011, vol. 4, no 3-4, p. 78-92.
- [66] DIXON, Robert MW. *Australian languages: their nature and development*. Cambridge University Press, 2002.

[67] MONA, Stefano, et al. Patterns of Y-chromosome diversity intersect with the Trans-New Guinea hypothesis. *Molecular biology and evolution*, 2007, vol. 24, no 11, p. 2546-2555.

[68] WURM, S. A. Linguistic prehistory in the New Guinea area. *Journal of Human Evolution*, 1983, vol. 12, no 1, p. 25-35.

[69] GREENBERG, Joseph Harold. *Language in the Americas*. Stanford University Press, 1987.

[70] ARNAIZ-VILLENA, A. The Origin of Amerindians and the Peopling of the Americas According to HLA Genes: Admixture with Asian and Pacific People (vol 11, pg 103, 2010). *Current Genomics*, 2010, vol. 11, no 6, p. 481-481.

[71] MCWHORTER, John H. *The language hoax: Why the world looks the same in any language*. Oxford University Press, 2014.

[72] MCWHORTER, John H. Identifying the creole prototype: Vindicating a typological class. *Language*, 1998, p. 788-818.

[73] KOUWENBERG, Silvia; SINGLER, John Victor (ed.). *The handbook of pidgin and creole studies*. John Wiley & Sons, 2009.

[74] MUFWENE, Salikoko. *Pidgin and Creole Language*, Humanities.uchicago.edu, 2010

[75] ANDERSEN, Henning. Center and periphery: adoption, diffusion, and spread. *Historical dialectology: Regional and social*, 1988, p. 39-84.

[76] CAMPBELL, Lyle. Language death. *The Encyclopedia of language and linguistics*, 1994, vol. 4, p. 1960-1968.

[77] CAMPBELL, Lyle. Language death. *The Encyclopedia of language and linguistics*, 1994, vol. 4, p. 1960-1968.

[78] KUSTERS, Wouter. *Linguistic complexity*. Netherlands Graduate School of Linguistics, 2003.

[79] LEES, Robert B.; CHOMSKY, N. Syntactic Structures. *Language*, 1957, vol. 33, no 3 Part 1, p. 375-408.

[80] WHORF, Benjamin Lee. *Linguistics as an exact science*. Massachusetts Institute of Technology, 1940.

[81] KUSTERS, Wouter. *Linguistic complexity*. Netherlands Graduate School of Linguistics, 2003.

[82] TRUDGILL, Peter. Dialect typology and social structure. *Language contact: theoretical and empirical studies*, 1992, vol. 60, p. 195.

- [83] TRUDGILL, Peter. Dialect typology: Isolation, social network and phonological structure. *Towards a social science of language*, 1996, vol. 1, p. 3-21.
- [84] TRUDGILL, Peter. Typology and sociolinguistics: Linguistic structure, social structure and explanatory comparative dialectology. *Folia linguistica*, 1997, vol. 31, no 3-4, p. 349-360.
- [85] ANDERSEN, Henning. Center and periphery: adoption, diffusion, and spread. *Historical dialectology: Regional and social*, 1988, p. 39-84.
- [86] THURSTON, William R. Processes of change in the languages of north-western New Britain. Department of Linguistics, Research School of Pacific Studies, The Australian National University, 1987.
- [87] THURSTON, William R. Sociolinguistic typology and other factors effecting change in northwestern New Britain, Papua New Guinea. *Culture Change, Language Change: Case Studies from Melanesia*, 1992, p. 123-39.
- [88] NICHOLS, Johanna. *Language diversity in space and time*. Chicago: University of Chi, 1992.
- [89] BARTH, Fredrik (ed.). *Ethnic groups and boundaries: The social organization of culture difference*. (Results of a symposium held at the University of Bergen, 23rd to 26th February 1967.). Universitetsforlaget, 1969.
- [90] GILES, Howard. Ethnicity markers in speech. *Social markers in speech*, 1979, p. 251-289.
- [91] LEVINE, Robert A.; CAMPBELL, Donald T. *Ethnocentrism: Theories of conflict, ethnic attitudes, and group behavior*. 1972.
- [92] WALLMAN, Sandra. Ethnicity and the boundary process in context. *Theories of race and ethnic relations*, 1986, p. 226-45.
- [93] URCIUOLI, Bonnie. Language and borders. *Annual Review of Anthropology*, 1995, p. 525-546.
- [94] SIEWIERSKA, Anna (ed.). *Constituent order in the languages of Europe*. Walter de Gruyter, 1998.
- [95] HASPELMATH, Martin; DRYER, Matthew S. *The world atlas of language structures online*. 2008.
- [96] MATSUMOTO, Katsumi *Word Types and Their Geographical Distribution*, 1987.
- [98] MURDOCK, George Peter. *Ethnographic atlas*. 1967.
- [99] ASHER, Ron. E. and Christopher Moseley. eds. 1993. *Atlas of the World's Languages*.
- [100] GREENBERG, Joseph Harold. *The languages of Africa*. Indiana Univ., 1963

- [101] MITHUN, Marianne. The languages of native North America. Cambridge University Press, 2001.
- [102] PAGE, Roderic DM; HOLMES, Edward C. Molecular Evolution: A Phylogenetic Approach. 1998.
- [103] MADDISON, D. R.; MADDISON, W. P. Sunderland, MA. 2000.
- [104]
- [105] MOHAN REDDY, B., et al. Austro-Asiatic tribes of Northeast India provide hitherto missing genetic link between South and Southeast Asia. PLoS One, 2007, vol. 2, no 11, p. e1141\_1-e1141\_12.
- [106] FULLER, Dorian Q. Pathways to Asian civilizations: Tracing the origins and spread of rice and rice cultures. Rice, 2011, vol. 4, no 3-4, p. 78-92.
- [107] BERMAN, Howard. Proto-Algonquian-Ritwan Verbal Roots. International Journal of American Linguistics, 1984, p. 335-342.
- [108] KROEBER, A. L., et al. Cultural and natural areas of native North America. 1939. School for Advanced Research Press, 2012.
- [109] HARRIS, Marvin. Cultural materialism: The struggle for a science of culture. AltaMira Press, 2001.
- [110] KIRCHHOFF, Paul. 1943 Mesoamerica: its geographic limits, ethnic composition and cultural characteristics. Sol Tax, Heritage of conquest, Cooper Square Publishers, New York, 198
- [111] STALLER, John E.; TYKOT, Robert H.; BENZ, Bruce F. Histories of maize. Left Coast Press, 2010.
- [112] PIPERNO, Dolores R. The origins of plant cultivation and domestication in the New World tropics. Current Anthropology, 2011, vol. 52, no S4, p. S453-S470.
- [113] WURM, S. A. Linguistic prehistory in the New Guinea area. Journal of Human Evolution, 1983, vol. 12, no 1, p. 25-35.
- [114] DONOHUE, Mark. Papuan Malay of New Guinea: Melanesian influence on verb and clause structure. Creoles, their substrates, and language typology, 2011, p. 413-435.



## **ANNEXOS**

### **ANNEX 1**

Aquesta tesi resulta incompleta sense els annexos que l'acompanyen, però degut a la mida dels fitxers, resulta impossible la seva impressió en folis DIN4. Per això s'adjunten en un CD i només s'imprimirà l'annex 1 amb els valors de les variables dels 590 *genus* utilitzats.

ara	genus	Subj.	Obj.	Adp.	Adje.	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Àrea
1	Hadza	2	4	4	4	4	-	-	-	-	AFR	F1 Nil
2	Sandawe	0	0	0	4	0	0	4	0	4	AFR	resta
3	Northern Khoisan	0	4	0	4	2	4	4	2	4	AFR	F2 Niger
4	Central Khoisan	0	0	0	0	0	0	0	0	1	AFR	resta
5	Southern Khoisan	0	4	4	3	0	4	4	0	4	AFR	F2 Niger
6	#Hoan	0	4	2	4	0	4	4	-	4	AFR	F2 Niger
7	Kadugli	1	4	4	4	4	4	4	4	4	AFR	F1 Nil
8	Katla-Tima	0	4	4	4	4	4	-	4	4	AFR	F1 Nil
9	Heiban	0	4	4	4	4	4	4	4	4	AFR	F1 Nil
10	Talodi Proper	1	2	4	4	3	4	4	4	4	AFR	F1 Nil
11	Tegem	0	0	0	4	0	4	-	0	-	AFR	resta
12	Rashad	0	0	0	4	4	4	4	4	-	AFR	resta
13	Eastern Mande	0	0	0	4	0	4	4	0	4	AFR	resta
14	Western Mande	0	1	0	4	0	2	4	0	2	AFR	resta
15	BIJAGO	-	-	-	-	-	-	-	-	-	AFR	resta
16	Northern Atlantic	0	4	4	4	4	4	3	4	4	AFR	resta
17	Southern Atlantic	0	4	3	4	3	3	4	3	4	AFR	resta
18	Kru	0	2	0	3	0	3	2	0	4	AFR	F2 Niger
19	Dogon	0	0	0	4	0	4	4	2	2	AFR	resta
20	Gur	0	3	1	4	0	3	4	0	3	AFR	F2 Niger
21	Adamawa	0	4	3	4	3	4	4	2	4	AFR	F2 Niger
22	Gbaya-Manza-Ngbaka	0	4	4	1	4	4	4	4	4	AFR	F2 Niger
23	Ubangi	0	4	4	2	3	4	4	4	4	AFR	F2 Niger
24	Ijoid	0	1	0	0	0	0	0	0	0	AFR	F2 Niger
25	Kwa	0	4	1	4	0	4	4	0	4	AFR	F2 Niger
26	Nupoid	0	4	4	4	2	4	4	3	4	AFR	F2 Niger
27	Idomoid	0	4	4	4	4	4	4	4	-	AFR	F2 Niger
28	Defoid	0	4	4	4	4	4	4	4	4	AFR	F2 Niger
29	Edoid	0	4	4	4	4	4	4	4	4	AFR	F2 Niger
30	Igboïd	0	4	4	2	4	4	4	4	4	AFR	F2 Niger
31	JUKUNOID	0	4	4	4	4	4	4	4	-	AFR	F2 Niger
32	Cross River	0	4	4	2	4	4	3	3	4	AFR	F2 Niger
33	Kainji	0	4	4	4	4	3	4	4	4	AFR	F2 Niger
34	Platoid	0	4	4	4	3	4	4	4	4	AFR	F2 Niger
35	Bantoid	0	4	4	4	4	3	4	4	4	AFR	F2 Niger



ara	genus	Subj.	Obj.	Adp.	Adje,	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Àrea
36	Songhay	0	2	0	4	0	4	4	0	4	AFR	resta
37	Western Saharan	0	0	0	4	3	4	4	4	4	AFR	resta
38	Eastern Saharan	0	0	0	4	1	4	4	4	4	AFR	resta
39	Maban	0	0	0	4	2	4	0	4	4	AFR	resta
40	Fur	0	0	0	4	0	0	4	0	4	AFR	resta
41	Nubian	0	1	0	4	0	0	4	1	4	AFR	resta
42	Surmic	2	3	2	4	4	4	4	4	4	AFR	F1 Nil
43	Shabo	0	0	0	0	-	-	0	-	4	AFR	resta
44	Nara	0	0	0	4	1	4	0	2	4	AFR	resta
45	Eastern Jebel	0	4	2	4	2	4	0	4	4	AFR	F1 Nil
46	Nyimang	0	0	0	4	1	4	-	0	-	AFR	resta
47	Temein	0	4	4	4	4	4	4	4	-	AFR	F1 Nil
48	Taman	0	0	0	4	0	-	4	2	-	AFR	resta
49	Daju	0	2	4	4	4	4	4	4	4	AFR	F1 Nil
50	Nilotic	2	4	4	4	4	4	4	4	4	AFR	F1 Nil
51	Kuliak	3	4	4	4	3	4	4	4	4	AFR	F1 Nil
52	Kresh	0	4	4	1	4	4	2	4	4	AFR	F1 Nil
53	Bongo-Bagirmi	0	4	3	4	4	4	4	4	4	AFR	F1 Nil
54	Moru-Ma'di	0	2	0	4	2	4	4	2	4	AFR	resta
55	Mangbetu	0	4	0	4	0,5	4	4	2	-	AFR	F1 Nil
56	Mangbutu-Efe	0	4	0	4	2	4	4	0	-	AFR	F1 Nil
57	Lendu	0	2	0	0	0	0	1	2	2	AFR	resta
58	Berta	0	4	3	4	2	4	4	4	4	AFR	F1 Nil
59	Kunama	0	0	0	4	0	3	4	2	4	AFR	resta
60	Gumuz	0	4	4	2	4	4	-	-	4	AFR	F1 Nil
61	Koman	0	4	4	3	4	3	-	4	-	AFR	F1 Nil
62	Egyptian-Coptic	0	4	4	4	4	1	1	3	4	AFR	F3 Eurom
63	Berber	3	4	4	4	4	3	0	4	4	AFR	F3 Eurom
64	Masa	0	4	4	4	4	4	4	4	4	AFR	F1 Nil
65	East Chadic	0	4	4	4	2	4	4	4	4	AFR	F1 Nil
66	KUJARGE	-	-	-	-	-	-	-	-	-	AFR	resta
67	Biu-Mandara	1	4	4	4	4	4	4	4	4	AFR	F1 Nil
68	West Chadic	0	4	4	2	4	3	4	4	4	AFR	F1 Nil
69	North Omotic	0	0	0	0	0	1	1	1	1	AFR	resta
70	South Omotic	0	0	0	3	0	3	2	2	4	AFR	resta
71	Beja	0	0	0	1	0	2	2	2	4	AFR	resta
72	Central Cushitic	0	0	0	1	0	0	0	0	0	AFR	resta
73	Eastern Cushitic	0	0	0	2	2	2	2	2	2	AFR	resta

ara	genus	Subj.	Obj.	Adp.	Adje,	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Àrea
74	Southern Cushitic	0	2	2	4	4	4	2	4	4	AFR	resta
75	Semitic	2	2	4	3	3	2	1	4	3	AFR	F3 Eurom
76	Bangime	0	0	0	4	5	0	4	0	4	AFR	resta
77	Laal	0	4	5	4	5	5	4	0	4	AFR	F1 Nil
78	AKPES	-	-	-	-	-	-	-	-	-	AFR	resta
79	BIRAYLE	0	0	0	4	0	4	4	5	4	AFR	resta
80	MBRE	-	-	-	-	-	-	-	-	-	AFR	resta
81	OKO	-	-	-	-	-	-	-	-	-	AFR	resta
82	UKAAN	-	-	-	-	-	-	-	-	-	AFR	resta
83	Armenian	0	2	1	0	0	0	0	1	3	EU	resta
84	Indic	0	0	0	0	0	0	0	0	2	EU	resta
85	Iranian	0	0	2	2	2	0	0	2	4	EU	F3 Eurom
86	Albanian	0	4	4	4	4	0	0	2	4	EU	F3 Eurom
87	Greek	0	4	4	0	4	0	0	1	4	EU	F3 Eurom
88	Romance	0	4	4	3	4	0	0	1	4	EU	F3 Eurom
89	Celtic	3	4	4	3	4	4	0	2	4	EU	F3 Eurom
90	Germanic	0	4	4	0	2	0	1	2	4	EU	F3 Eurom
91	Baltic	0	4	4	0	0	0	0	0	4	EU	F3 Eurom
92	Slavic	0	4	4	0	3	0	0	0	4	EU	F3 Eurom
93	Yukaghir	0	0	0	0	0	0	0	2	0	EU	resta
94	Samoyedic	0	0	0	0	0	0	0	3	-	EU	resta
95	Ugric	0	2	0	0	0	0	0	2	1	EU	F3 Eurom
96	Permic	0	2	0	0	0	0	0	2	-	EU	F3 Eurom
97	Mari	0	0	0	0	0	0	0	2	-	EU	resta
98	Mordvin	0	2	0	0	0	2	0	2	-	EU	F3 Eurom
99	Saami	0	4	1	0	0	0	0	3	-	EU	F3 Eurom
100	Finnic	0	4	1	0	0	0	0	2	4	EU	F3 Eurom
101	Turkic	0	0	0	0	0	0	0	2	1	EU	resta
102	Mongolic	0	0	0	0	0	0	0	2	0	EU	resta
103	Tungusic	0	0	0	0	0	0	0	2	1	EU	resta
104	Korean	0	0	0	0	0	0	1	0	0	EU	resta
105	Ainu	0	0	0	0	0	0	0	0	0	EU	resta
106	Japanese	0	0	0	0	0	0	0	0	0	EU	resta
107	Northern Chukotko-Kamchatkan	0	2	0	0	0	0	0	0	2	EU	F3 Eurom
108	Southern Chukotko-Kamchatkan	0	0	2	0	1	0	0	2	-	EU	F3 Eurom
109	Northern Dravidian	0	0	0	0	0	0	0	0	4	EU	resta
110	Central Dravidian	0	0	0	0	0	0	0	0	0	EU	resta

ara	genus	Subj.	Obj.	Adp.	Adje.	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Àrea
111	South-Central Dravidian	0	0	0	0	0	0	0	0	0	EU	resta
112	Southern Dravidian	0	0	0	0	0	0	0	0	1	EU	resta
113	Kartvelian	0	0	0	1	0	0	0	1	4	EU	resta
114	Northwest Caucasian	0	0	0	2	0	0	2	0	0	EU	resta
115	Nakh	0	0	0	0	0	0	0	0	0	EU	resta
116	Avar-Andic-Tsezic	0	1	0	0	0	0	0	0	1	EU	resta
117	Lak-Dargwa	0	0	0	0	0	0	0	0	0	EU	resta
118	Lezgcic	0	0	0	0	0	0	0	0	0	EU	resta
119	Kusunda	0	0	0	0	0	0	0	0	-	EU	resta
120	Nahali	0	0	0	0	0	0	0	0	-	EU	resta
121	Burushaski	0	0	0	0	0	0	0	0	0	EU	resta
122	Nivkh	0	0	0	0	0	0	2	0	0	EU	resta
123	Basque	0	0	0	4	0	4	0	0	0	EU	F3 Eurom
124	Yeniseian	0	0	0	0	0	0	0	0	-	EU	resta
125	Bai	0	4	2	0	0	4	4	-	0	SEAN	F4 SEAN
126	Chinese	0	4	2	0	0	0	0	0	0	SEAN	F4 SEAN
127	Karen	0	4	4	4	0	4	3	0	4	SEAN	F4 SEAN
128	Hruso/DHAM MAI_UNCLASSIFIED	-	-	-	-	-	-	-	-	-	SEAN	resta
129	Tani	0	0	0	2	0	2	4	-	0	SEAN	resta
130	Digaroan	0	0	0	2	0	0	4	-	0	SEAN	resta
131	Dhimalic	0	0	0	0	0	0	0	-	0	SEAN	resta
132	Lepcha	0	0	0	4	0	0	4	0	-	SEAN	resta
133	rGyalrong	0	0	0	4	0	0	4	-	1	SEAN	resta
134	Bodic	0	0	0	3	0	2	3	0	1	SEAN	resta
135	Mahakiranti	0	0	0	1	0	0	1	0	0	SEAN	resta
136	Qiangic	0	0	0	4	0	2	4	-	0	SEAN	resta
137	Bodo-Garo	0	0	0	2	0	0	4	0	0	SEAN	resta
138	Northern Naga	0	0	0	2	0	0	2	0	1	SEAN	resta
139	Kuki-Chin	0	0	0	4	0	2	4	0	2	SEAN	resta
140	MRU	0	0	-	-	-	-	-	-	-	SEAN	resta
141	Jingpho	0	0	0	3	0	0	4	-	0	SEAN	resta
142	Nungish	0	0	0	4	0	2	4	0	0	SEAN	resta
143	Naxi	0	0	-	4	-	0	4	-	-	SEAN	resta
144	Burmese-Lolo	0	0	0	2	0	2	2	0	0	SEAN	resta
145	Hmong-Mien	0	4	4	3	2	3	0	0	4	SEAN	F4 SEAN
146	Munda	0	0	0	0	0	0	0	2	1	SEAN	resta
147	Khasian	0	2	2	2	4	0	0	4	4	SEAN	F4 SEAN
148	Palaung-Khmuic	0	4	4	4	3	4	4	4	4	SEAN	F4 SEAN
149	Viet-Muong	0	4	4	4	4	4	0	4	4	SEAN	F4 SEAN
150	Katuic	0	4	4	3	4	4	0	4	-	SEAN	F4 SEAN

ara	genus	Subj.	Obj.	Adp.	Adje.	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Àrea
151	Bahnaric	0	4	4	4	4	4	1	4	4	SEAN	F4 SEAN
152	Khmer	0	4	4	4	4	4	2	4	4	SEAN	F4 SEAN
153	Pearic	0	4	-	4	-	4	4	-	-	SEAN	F4 SEAN
154	Monic	0	4	4	4	4	4	4	4	-	SEAN	F4 SEAN
155	Aslian	0	4	4	4	4	4	0	4	4	SEAN	F4 SEAN
156	Shompen	4	4	-	0	-	-	0	-	-	SEAN	F4 SEAN
157	Nicobarese	0	4	4	3	4	1	0	4	-	SEAN	F4 SEAN
158	Hlai	0	4	4	4	4	4	0	-	-	SEAN	F4 SEAN
159	Kadai	0	4	4	4	4	4	0	-	-	SEAN	F4 SEAN
160	Kam-Tai	0	4	4	4	4	4	3	4	4	SEAN	F4 SEAN
161	Atayalic	4	4	4	4	4	4	0	4	4	SEAN	F4 SEAN
162	Tsouic	4	4	4	0	4	2	0	4	-	SEAN	F4 SEAN
163	Rukai	4	4	4	0	4	0	0	4	2	SEAN	F4 SEAN
164	Paiwanic	4	4	4	2	2	0	1	0	4	SEAN	F4 SEAN
165	East Formosan	4	4	4	-	4	0	0	-	0	SEAN	F4 SEAN
166	Western Plains Austronesian	4	4	4	0	2	2	-	1	-	SEAN	F4 SEAN
167	Chamorro	2	4	4	0	4	0	0	4	4	SEAN	F4 SEAN
168	Palauan	0	4	4	0	4	0	0	4	4	SEAN	F4 SEAN
169	Yapese	4	4	4	4	4	4	0	4	4	SEAN	F4 SEAN
170	Northern Luzon	4	4	4	2	4	1	0	4	4	SEAN	F4 SEAN
171	Batanic	4	4	4	2	4	2	0	4	-	SEAN	F4 SEAN
172	Central Luzon	4	4	4	0	4	2	0	2	4	SEAN	F4 SEAN
173	Greater Central Philippine	4	4	4	2	4	2	0	3	4	SEAN	F4 SEAN
174	Bilic	4	4	4	2	4	4	0	3	4	SEAN	F4 SEAN
175	Celebic	2	4	4	4	4	3	1	4	3	SEAN	F4 SEAN
176	Sangiric	4	4	4	4	4	4	-	4	-	SEAN	F4 SEAN
177	Minahasan	2	4	4	4	4	4	1	4	4	SEAN	F4 SEAN
178	South Sulawesi	2	4	4	2	4	3	0	4	4	SEAN	F4 SEAN
179	Land Dayak	-	-	-	-	-	-	-	-	-	SEAN	F4 SEAN
180	North Borneo	2	4	4	4	4	4	0	4	4	SEAN	F4 SEAN
181	Barito	2	4	4	4	4	2	4	4	4	SEAN	F4 SEAN
182	Sama-Bajaw	2	4	4	4	4	4	1	4	4	SEAN	F4 SEAN
183	Javanese	0	4	4	4	4	4	2	4	-	SEAN	F4 SEAN
184	Lampungic	0	4	4	4	4	4	0	4	4	SEAN	F4 SEAN
185	Enggano	0	4	4	4	4	4	2	-	4	SEAN	F4 SEAN

ara	genus	Subj.	Obj.	Adp.	Adje.	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Àrea
186	NW Sumatra-Barrier Islands	4	4	4	4	4	4	1	4	4	SEAN	F4 SEAN
187	Malayo-Sumbawan	0	4	4	4	4	4	0	4	4	SEAN	F4 SEAN
188	MOKLEN	0	4	-	-	-	-	-	-	-	SEAN	F4 SEAN
189	Rejang	0	4	-	-	-	-	-	-	-	SEAN	F4 SEAN
190	Central Malayo-Polynesian	0	4	4	4	1	4	3	3	4	SEAN	F4 SEAN
191	South Halmahera - West New Guinea	0	4	4	4	0	4	4	4	4	SEAN	F4 SEAN
192	Oceanic	1	3	4	4	3	3	2	2	4	SEAN	F4 SEAN
193	Western Arunachal	-	0	-	2	-	-	-	-	-	SEAN	resta
194	Sulung	-	0	-	-	-	-	-	-	-	SEAN	resta
195	Anindilyakwa	0	0	4	0	2	0	2	2	-	AUSTRAL	resta
196	Ndjébbana	0	0	0	4	-	0	4	-	-	AUSTRAL	resta
197	Gaagudju	-	-	0	4	-	0	2	-	-	AUSTRAL	resta
198	Mangarrayi	4	0	4	4	2	0	2	-	4	AUSTRAL	resta
199	Nakkara	-	-	-	4	-	-	-	-	-	AUSTRAL	resta
200	Nunggubuyu	0	0	0	0	0	0	-	-	4	AUSTRAL	resta
201	Tiwian	0	4	4	0	0	0	0	0	4	AUSTRAL	resta
202	Waray	0	4	0	-	-	-	-	-	-	AUSTRAL	resta
203	Limilngan	2	2	-	4	-	2	4	-	-	AUSTRAL	resta
204	UMBUGARL A-NGUMBUR	-	-	-	-	-	-	-	-	-	AUSTRAL	resta
205	Murrinh-Patha	0	0	2	4	4	4	-	-	2	AUSTRAL	resta
206	Iwaidjan	0	4	2	1	1	0	0	0	4	AUSTRAL	resta
207	GIIMBIYU	-	-	0	-	-	-	-	-	-	AUSTRAL	resta
208	DJAUAN	-	-	-	-	-	-	-	-	-	AUSTRAL	resta
209	Rembarnga	0	0	-	-	-	0	-	-	-	AUSTRAL	resta
210	Ngandi	0	0	-	4	2	0	4	-	-	AUSTRAL	resta
211	Ngalakan	2	2	2	4	2	0	-	-	4	AUSTRAL	resta
212	Gunwinygic	0	2	2	3	3	2	2	2	4	AUSTRAL	resta
213	Wagiman	0	1	0	4	4	-	4	4	-	AUSTRAL	resta
214	Yangmanic	2	2	0	4	2	0	0	-	-	AUSTRAL	resta
215	Burarran	0	1	0	4	0	0	-	-	-	AUSTRAL	resta
216	Maran	1	3	2	3	2	1	-	2	-	AUSTRAL	resta
217	West Barkly	1	2	0	4	2	1	4	2	-	AUSTRAL	resta
218	Garrwan	3	4	0	0	0	0	2	0	-	AUSTRAL	resta
219	Laragiyan	0	0	0	4	4	-	4	-	-	AUSTRAL	resta
220	Southern Daly	1	1	0	4	-	4	4	4	2	AUSTRAL	resta

ara	genus	Subj.	Obj.	Adp.	Adje,	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Area
221	Western Daly	0	0	0	4	1	4	4	4	4	AUSTRALIA	resta
222	Anson Bay	0	1	0	4	0	-	4	4	-	AUSTRALIA	resta
223	Eastern Daly	1	1	0	4	0	4	4	4	-	AUSTRALIA	resta
224	Northern Daly	0	0	0	4	1	4	4	3	-	AUSTRALIA	resta
225	Jaminjungan	2	2	-	2	-	2	2	-	-	AUSTRALIA	resta
226	Djeragan	0	0	0	3	0	-	-	-	-	AUSTRALIA	resta
227	Bunuban	0	1	2	2	1	0	0	-	-	AUSTRALIA	resta
228	Nyulnyulan	2	2	0	3	2	1	0	-	-	AUSTRALIA	resta
229	Wororan	1	1	0	2	0	3	1	2	4	AUSTRALIA	resta
230	Tangkic	0	3	2	0	0	0	0	-	4	AUSTRALIA	resta
231	Pama-Nyungan	0	1	0	2	1	2	2	2	3	AUSTRALIA	resta
232	Tasmanian	-	-	0	4	0	4	-	4	-	SAHUL	resta
233	Great Andamanese	0	0	0	0	0	0	-	0	-	SAHUL	resta
234	South Andamanese	0	0	-	4	0	0	-	-	-	SAHUL	resta
235	Oksapmin	0	0	0	2	0	4	2	0	-	SAHUL	resta
236	Morwap	0	0	0	4	0	-	-	-	-	SAHUL	resta
237	MOLOF	-	-	-	-	-	-	-	-	-	SAHUL	resta
238	USKU	-	-	-	-	-	-	-	-	-	SAHUL	resta
239	DEM	-	-	-	-	-	-	-	-	-	SAHUL	resta
240	MOR	0	4	-	-	-	-	-	-	-	SAHUL	F4 SEAN
241	Finisterre-Huon	0	0	0	4	0	4	4	3	4	SAHUL	resta
242	WIRU	-	-	-	-	-	-	-	-	-	SAHUL	resta
243	Eastern Highlands	0	0	0	1	0	0	2	0	1	SAHUL	resta
244	Chimbu	0	0	0	4	0	3	4	2	4	SAHUL	resta
245	Engan	0	0	0	2	0	2	4	0	2	SAHUL	resta
246	Kutubuan	0	0	0	4	0	4	4	0	-	SAHUL	resta
247	Fasu	0	0	0	4	0	4	4	0	-	SAHUL	resta
248	KAMULA	0	0	-	-	-	-	4	-	-	SAHUL	resta
249	Asmat-Kamoro	0	0	0	4	0	0	4	0	0	SAHUL	resta
250	Awju-Dumut	0	0	0	2	0	2	3	0	2	SAHUL	resta
251	Ok	0	0	0	4	0	2	4	0	2	SAHUL	resta
252	Awin-Pare	0	0	0	4	0	4	4	0	-	SAHUL	resta
253	East Strickland	0	0	0	4	0	-	-	0	-	SAHUL	resta
254	Bosavi	0	0	0	4	0	4	4	-	-	SAHUL	resta
255	Duna	0	0	0	4	0	4	4	0	-	SAHUL	resta

ara	genus	Subj.	Obj.	Adp.	Adje,	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Area
256	Mombum	0	0	0	-	0	-	4	0	-	SAHUL	resta
257	Angan	0	0	0	3	0	3	4	1	-	SAHUL	resta
258	Suki	0	0	0	0	0	0	2	0	-	SAHUL	resta
259	Gogodala	0	0	0	-	0	4	4	0	-	SAHUL	resta
260	Marind Proper	0	0	0	0	0	0	3	0	-	SAHUL	resta
261	Kayagar	0	0	-	4	4	0	-	4	-	SAHUL	resta
262	Sentani	0	0	0	4	0	0	4	0	-	SAHUL	resta
263	Dani	0	0	0	-	0	4	-	0	4	SAHUL	resta
264	Kwerba	0	0	0	4	-	-	4	-	-	SAHUL	resta
265	Wissel Lakes-Kemandoga	0	0	0	0	0	2	4	-	-	SAHUL	resta
266	Mairasi-Tanahmerah	0	0	-	-	-	-	-	-	-	SAHUL	resta
267	West Bomberai	0	0	0	4	0	4	4	2	-	SAHUL	resta
268	Binanderean	0	0	0	4	0	4	4	2	-	SAHUL	resta
269	Goilalan	0	0	0	3	0	4	4	0	-	SAHUL	resta
270	Koarian	0	0	0	4	0	2	4	1	4	SAHUL	resta
271	KWALEAN	0	0	0	2	-	0	-	0	-	SAHUL	resta
272	MANUBARAN	0	0	0	3	0	0	4	0	-	SAHUL	resta
273	Yareban	0	0	0	1	0	0	4	0	0	SAHUL	resta
274	Mailuan	0	0	0	2	0	0	4	0	-	SAHUL	resta
275	Dagan	0	0	0	4	0	4	4	4	4	SAHUL	resta
276	Madang	0	0	0	4	0	4	4	2	3	SAHUL	resta
277	Pawaiian	0	0	0	1	0	0	4	0	-	SAHUL	resta
278	Teberan	0	0	0	0	0	0	4	0	-	SAHUL	resta
279	Turama-Kikorian	0	1	0	4	0	4	4	2	0	SAHUL	resta
280	INLAND_GULF	0	0	0	-	0	-	4	-	-	SAHUL	resta
281	Tate	0	0	-	4	0	0	4	-	-	SAHUL	resta
282	PURARI	-	-	-	-	-	-	-	-	-	SAHUL	resta
283	Eleman Proper	0	0	0	2	0	0	4	0	-	SAHUL	resta
284	Bulaka River	0	0	0	-	1	-	4	0	-	SAHUL	resta
285	Kiwaiian	0	0	0	0	0	0	0	0	-	SAHUL	resta
286	TIRIO	-	-	-	-	-	-	-	-	-	SAHUL	resta
287	Western Fly	0	0	-	0	0	-	0	-	-	SAHUL	resta
288	Morehead and Upper Maro Rivers	0	0	0	0	0	-	0	0	-	SAHUL	resta
289	Moraori	0	0	0	-	0	-	4	-	-	SAHUL	resta
290	Mek	0	0	0	2	0	2	4	-	0	SAHUL	resta

ara	genus	Subj.	Obj.	Adp.	Adje.	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Àrea
291	Senagi	0	0	0	2	0	2	2	0	3	SAHUL	resta
292	PAUWASI	-	-	-	-	-	-	-	-	-	SAHUL	resta
293	Border	0	0	0	4	0	1	4	2	2	SAHUL	resta
294	Lakes Plain	0	0	0	-	0	4	4	-	4	SAHUL	resta
295	Orya	0	0	0	4	0	4	-	-	-	SAHUL	resta
296	Tor	-	-	-	-	-	-	-	-	-	SAHUL	resta
297	Nimboran	-	-	-	-	-	-	-	0	-	SAHUL	resta
298	SAUSE	-	-	-	-	-	-	-	-	-	SAHUL	resta
299	Kaure	0	0	-	-	-	-	-	-	-	SAHUL	resta
300	South Bird's Head	0	2	0	4	0	-	4	-	-	SAHUL	resta
301	Kolopom	0	0	0	-	0	-	4	-	-	SAHUL	resta
302	Kolana-Tanglapui	-	-	-	-	-	-	-	-	-	SAHUL	resta
303	Makasae-Fataluku-Oirata	0	0	0	4	0	4	4	0	4	SAHUL	resta
304	BUNAK	0	0	0	4	0	-	-	-	-	SAHUL	resta
305	Greater Alor	0	0	2	4	0	2	4	0	4	SAHUL	resta
306	Kebar o Amberbaken	0	4	4	4	-	4	4	-	4	SAHUL	F4 SEAN
307	North-Central Bird's Head	0	4	4	4	1	4	4	-	4	SAHUL	F4 SEAN
308	West Bird's Head	0	4	-	-	-	-	-	-	-	SAHUL	F4 SEAN
309	Hatam	0	4	4	4	0	4	4	-	4	SAHUL	F4 SEAN
310	North Halmaheran	0	3	4	4	0	4	4	0	4	SAHUL	F4 SEAN
311	East Bird's Head	0	4	4	4	0	2	4	2	4	SAHUL	F4 SEAN
312	Yawa	-	-	-	-	-	-	-	-	-	SAHUL	resta
313	East Geelvink Bay	0	0	0	-	-	-	-	-	-	SAHUL	resta
314	Western Skou	0	0	0	4	0	4	4	4	-	SAHUL	resta
315	Krisa	0	0	2	4	4	4	4	-	4	SAHUL	resta
316	Serra Hills	0	1	2	4	0	4	4	-	-	SAHUL	resta
317	Warapu	0	0	-	4	4	4	4	4	4	SAHUL	resta
318	Kwomtari	0	0	0	2	0	0	3	-	2	SAHUL	resta
319	Fas	0	0	0	-	0	-	-	-	-	SAHUL	resta
320	Left May	0	0	0	2	0	-	4	-	-	SAHUL	resta
321	AMTO-MUSAN	0	0	-	-	-	-	-	-	-	SAHUL	resta
322	Urim	0	4	-	-	-	-	-	-	-	SAHUL	F4 SEAN
323	West Wapei	0	4	-	-	2	2	-	-	-	SAHUL	F4 SEAN
324	Wapei-Palei	0	4	3	4	2	4	4	4	4	SAHUL	F4 SEAN
325	MAIMAI	0	4	-	-	32	-	-	-	-	SAHUL	F4 SEAN



ara	genus	Subj.	Obj.	Adp.	Adje,	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Àrea
326	Kombio-Arapesh	0	4	4	4	3	3	3	-	4	SAHUL	F4 SEAN
327	Marienberg	0	0	0	4	0	4	4	-	2	SAHUL	resta
328	Monumbo	0	0	0	4	0	-	-	-	-	SAHUL	resta
329	Gapun	-	-	-	-	-	-	-	-	-	SAHUL	resta
330	BIKSI	-	-	-	0	-	-	-	-	-	SAHUL	resta
331	Upper Sepik	0	0	0	4	0	2	4	0	-	SAHUL	resta
332	Ram	0	0	0	4	0	0	4	-	0	SAHUL	resta
333	Tama Sepik	0	0	0	4	0	-	-	0	-	SAHUL	resta
334	Yellow River	0	0	0	4	0	4	4	-	2	SAHUL	resta
335	Middle Sepik	0	0	0	1	0	0	2	1	0	SAHUL	resta
336	Sepik Hill	0	0	0	3	0	3	3	2	0	SAHUL	resta
337	LEONHARD_SCHULTZE	-	-	-	4	-	-	-	-	-	SAHUL	resta
338	Lower Sepik	0	0	0	3	0	4	-	-	-	SAHUL	resta
339	YUAT	-	-	-	-	-	-	-	-	-	SAHUL	resta
340	Piawi	0	0	-	4	0	-	-	-	2	SAHUL	resta
341	Botin	0	0	0	4	-	4	-	0	-	SAHUL	resta
342	GRASS	0	0	0	-	0	-	-	-	-	SAHUL	resta
343	ARAFUNDI	-	-	-	-	-	-	-	-	-	SAHUL	resta
344	Annaberg	0	0	-	4	0	-	-	-	-	SAHUL	resta
345	Lower Ramu	0	0	0	-	-	-	-	-	-	SAHUL	resta
346	Mikarew	0	0	0	4	-	0	4	0	-	SAHUL	resta
347	Yele	0	0	0	4	0	4	4	0	-	SAHUL	resta
348	Bilua	2	4	0	0	0	0	0	0	2	SAHUL	F4 SEAN
349	Lavukaleve	0	0	0	4	0	4	4	0	2	SAHUL	resta
350	Savosavo	0	0	0	0	0	0	0	0	0	SAHUL	resta
351	Sulka	0	4	4	4	2	4	4	-	4	SAHUL	F4 SEAN
352	Anêm	0	4	4	4	2	4	4	-	-	SAHUL	F4 SEAN
353	Kuot	5	4	4	4	4	0	0	-	4	SAHUL	F4 SEAN
354	WASI	-	-	-	-	-	-	-	-	-	SAHUL	resta
355	Baining	0	4	4	-	-	-	-	-	-	SAHUL	resta
356	Taulil	0	4	-	-	-	-	-	-	-	SAHUL	resta
357	East Bougainville	0	0	0	4	0	0	1	2	4	SAHUL	resta
358	West Bougainville	0	0	0	-	0	-	-	-	-	SAHUL	resta
359	Karkar-Yuri	0	0	0	2	0	0	-	-	-	SAHUL	resta
360	Momuna o SOMAHA	0	0	0	-	-	-	-	-	-	SAHUL	resta

ara	genus	Subj.	Obj.	Adp.	Adje,	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Àrea
361	Yale o nagatman	0	0	0	2	-	-	-	-	-	SAHUL	resta
362	BURMESO o TAURAP	-	-	-	-	-	-	-	-	-	SAHUL	resta
363	DOSO	-	-	-	-	-	-	-	-	-	SAHUL	resta
364	KAPAURI	-	-	-	-	-	-	-	-	-	SAHUL	resta
365	KEHU	-	-	-	-	-	-	-	-	-	SAHUL	resta
366	LEPKI	-	-	-	-	-	-	-	-	-	SAHUL	resta
367	MURKIM	-	-	-	-	-	-	-	-	-	SAHUL	resta
368	ODIAI o Busa	-	-	-	-	-	-	-	-	-	SAHUL	resta
369	TOFANMA	-	-	-	-	-	-	-	-	-	SAHUL	resta
370	TOUO	-	-	-	-	-	-	-	-	-	SAHUL	resta
371	DAMAL uhunduni?	0	0	-	-	-	-	-	-	-	SAHUL	resta
372	MAWES	-	-	-	-	-	-	-	-	-	SAHUL	resta
373	Aleut	0	0	0	4	0	0	0	2	-	NAM	resta
374	Eskimo	0	0	0	2	0	0	0	3	4	NAM	resta
375	Haida	0	0	0	4	0	0	4	2	2	NAM	resta
376	Tlingit	0	0	0	0	0	0	0	0	0	NAM	resta
377	Eyak	0	0	-	-	-	-	-	-	-	NAM	resta
378	Athapaskan	0	0	0	4	0	0	1	0	3	NAM	resta
379	Kutenai	0	4	2	0	4	0	0	2	2	NAM	F5 Colum
380	Wiyot	0	1	0	4	0	0	2	0	-	NAM	resta
381	Yurok	0	3	4	0	0	0	0	0	2	NAM	F5 Colum
382	Beothuk	-	-	-	-	-	-	-	-	-	NAM	resta
383	Algonquian	1	3	2	1	1	0	0	2	3	NAM	F5 Colum
384	Chimakuan	4	4	4	0	4	0	0	-	-	NAM	F5 Colum
385	Northern Wakashan	3	4	4	0	4	0	-	-	4	NAM	F5 Colum
386	Southern Wakashan	4	4	4	1	3	0	1	4	3	NAM	F5 Colum
387	Bella Coola	4	4	4	1	4	4	-	4	4	NAM	F5 Colum
388	Central Salish	4	4	4	0	4	0	0	2	2	NAM	F5 Colum
389	Tillamook	4	4	4	2	0	0	0	2	-	NAM	F5 Colum
390	Tsamosan	4	4	4	0	4	0	0	2	-	NAM	F5 Colum
391	Interior Salish	4	4	4	1	3	1	0	2	-	NAM	F5 Colum
392	Keresan	0	2	0	0	0	0	2	0	-	NAM	resta
393	Yuchi	0	0	1	4	0	0	4	-	4	NAM	resta
394	Siouan	0	0	0	4	0	2	4	2	2	NAM	resta
395	Caddoan	1	0	0	4	0	0	0	1	4	NAM	resta

ara	genus	Subj.	Obj.	Adp.	Adje.	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Àrea
396	Northern Iroquoian	0	2	0	4	0	0	0	0	2	NAM	resta
397	Southern Iroquoian	1	0	0	0	1	0	0	0	0	NAM	resta
398	Tsimshianic	4	4	4	4	4	2	0	4	4	NAM	F5 Colum
399	Chinookan	4	4	4	3	4	0	0	-	-	NAM	F5 Colum
400	Takelma	0	0	0	4	0	0	2	2	-	NAM	resta
401	Kalapuyan	-	-	-	-	-	-	-	-	-	NAM	resta
402	Coosan	4	2	0	0	0	0	0	-	4	NAM	F5 Colum
403	Alsea	4	4	4	0	0	0	0	-	-	NAM	F5 Colum
404	Siuslawan	4	2	2	0	2	0	0	-	-	NAM	F5 Colum
405	Klamath-Modoc	1	1	0	0	0	0	0	-	4	NAM	resta
406	Sahaptian	4	4	2	0	0	0	0	0	4	NAM	F5 Colum
407	Molala	-	-	-	-	-	-	-	4	-	NAM	resta
408	Wintuan	1	1	0	-	0	-	-	-	-	NAM	resta
409	Maiduan		0	0	0	0	0	1	0	0	NAM	resta
410	Yokuts	1	2	2	1	0	0	0	2	-	NAM	F5 Colum
411	Costanoan	0	4	0	0	1	0	0	0	4	NAM	F5 Colum
412	Miwok	0	2	0	2	1	0	0	2	-	NAM	F5 Colum
413	Zuni	0	0	0	4	0	0	0	0	-	NAM	resta
414	Chitimacha	0	0	0	4	0	-	4	-	4	NAM	resta
415	Tunica	0	0	1	4	0	2	4	0	4	NAM	resta
416	Atakapa	0	0	0	2	0	0	2	0	-	NAM	resta
417	Yukian	-	-	-	-	-	-	-	-	-	NAM	resta
418	Wappo	0	0	0	4	0	0	4	0	2	NAM	resta
419	Natchez	0	0	0	4	-	-	-	-	4	NAM	resta
420	Muskogean	0	0	0	4	0	2	4	0	2	NAM	resta
421	Huavean	0	4	4	4	4	1	0	0	4	NAM	F6 Meso
422	Totonacan	2	4	4	0	4	0	0	0	4	NAM	F6 Meso
423	Mixe-Zoque	2	4	2	0	1	0	0	0	4	NAM	F6 Meso
424	Mayan	3	4	4	1	4	2	0	0	4	NAM	F6 Meso
425	Karok	0	2	0	4	0	0	0	0	-	NAM	F5 Colum
426	Shasta	0	0	0	4	0	4	-	0	-	NAM	resta
427	Chimariko	0	0	0	4	0	2	4	-	2	NAM	resta
428	Palaihnihan	0	4	0	0	0	0	0	0	-	NAM	F5 Colum
429	Yana	-	-	-	-	-	-	0	4	-	NAM	resta
430	Pomoan	0	0	0	4	0	0	1	0	-	NAM	resta
431	Washo	0	0	0	0	0	0	0	0	-	NAM	resta
432	Salinan	0	4	4	2	4	0	0	-	-	NAM	F5 Colum
433	Chumash	0,504	4	4	2	4	4	-	0	-	NAM	F5 Colum
434	Esselen	-	-	-	0	-	-	-	-	-	NAM	resta
435	Seri	0	0	0	4	0	4	4	-	2	NAM	resta

ara	genus	Subj.	Obj.	Adp.	Adje.	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Área
436	Yuman	0	0	0	4	0	4	4	1	2	NAM	resta
437	Tonkawa	0	0	0	4	0	4	-	-	-	NAM	resta
438	Karankawa	-	2	-	-	-	0	-	-	-	NAM	resta
439	Coahuiltecan	0	0	0	4	0	-	-	-	-	NAM	resta
440	Comecrudan	-	-	-	-	-	-	-	-	-	NAM	resta
441	Tequistlatecan	4	4	4	0	3	0	2	1	-	NAM	F6 Mesoa
442	Tol	0	0	0	4	0	-	4	0	-	NAM	resta
443	Yurimangí	-	-	-	4	0	-	-	-	-	NAM	resta
444	Subtiaba-Tlapanec	4	4	4	4	4	4	0	4	4	NAM	F6 Mesoa
445	Kiowa-Tanoan	0	0	2	4	0	0	0	0	2	NAM	resta
446	Numic	0	0	0	1	0	2	0	1	4	NAM	resta
447	Tubatulabal	0	4	0	0	4	0	0	4	-	NAM	F5 Colum
448	Takic	0	0	0	2	0	0	1	0	4	NAM	resta
449	Hopi	0	0	0	0	0	0	0	2	-	NAM	resta
450	Tepiman	2	2	1	1	1	0	0	2	4	NAM	F6 Mesoa
451	Tubar	-	-	-	-	-	-	-	-	-	NAM	resta
452	Tarahumaran	2	0	0	0	0	0	1	0	3	NAM	resta
453	Cahita	0	0	0	0	0	0	0	0	4	NAM	resta
454	Corachol	2	2	0	3	2	1	1	2	4	NAM	F6 Mesoa
455	Aztecan	1	4	3	2	2	0	0	0	4	NAM	F6 Mesoa
456	Amuzgoan	4	4	4	4	4	4	0	4	-	NAM	F6 Mesoa
457	Chichimec	0	0	-	4	0	2	0	2	-	NAM	resta
458	Pamean	0	4	4	4	4	0	0	-	-	NAM	F6 Mesoa
459	Otomian	2	4	4	0	4	0	0	2	4	NAM	F6 Mesoa
460	Mixtecan	4	4	4	4	4	4	0	4	4	NAM	F6 Mesoa
461	Popolocan	3	4	4	3	4	0	0	4	-	NAM	F6 Mesoa
462	Chinantecan	4	4	4	4	4	4	0	-	4	NAM	F6 Mesoa
463	Zapotecan	3	4	4	4	4	4	0	3	4	NAM	F6 Mesoa
464	CHIAPANEC-MANGUE	0	0	0	4	0	-	-	0	-	NAM	resta
465	Tarascan	0	4	0	0	2	0	0	2	4	NAM	F6 Mesoa
466	Cuitlatec	-	-	-	-	-	-	-	-	-	NAM	F6 Mesoa
467	XINCAN	4	4	4	0	3	0	-	-	-	NAM	F6 Mesoa
468	LENCAN	0	0	0	4	0	-	-	-	-	SAM	resta
469	Paya	0	0	-	-	-	-	-	-	-	SAM	resta
470	Yanomam	0	0	0	4	0	0	0	0	2	SAM	resta
471	Motilon	0	0	-	-	-	-	-	-	-	SAM	resta
472	Kuna	0	0	0	4	0	0	4	0	-	SAM	resta
473	Misumalpa	0	0	0	4	0	0	4	2	-	SAM	resta
474	Rama	0	0	0	4	0	0	4	-	-	SAM	resta

ara	genus	Subj.	Obj.	Adp.	Adje.	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Àrea
475	Talamanca	1	0	1	4	0	3	2	0	4	SAM	resta
476	Guaymi	0	0	0	4	0	4	4	0	-	SAM	resta
477	Aruak	0	1	0	4	0	-	1	-	2	SAM	resta
478	Chimila	0	4	-	-	-	-	-	-	-	SAM	resta
479	Chibchan Proper	0	0	0	4	0	1	4	0	0	SAM	resta
480	Timucua	0	0	-	-	-	-	-	-	-	SAM	resta
481	Warao	0	0	0	4	0	1	4	0	-	SAM	resta
482	Kunza	0	0	-	4	-	-	-	-	-	SAM	resta
483	Betoi	0	0	0	-	-	-	-	-	-	SAM	resta
484	Chimúan	0	1	0	0	0	0	0	0	-	SAM	resta
485	Itonama	0	4	4	4	4	0	0	0	-	SAM	F7 Bolivia
486	Huarpe	-	-	-	-	-	-	-	-	-	SAM	resta
487	Jirajaran	0	4	-	4	0	-	0	-	-	SAM	resta
488	Mura	0	0	0	4	0	4	4	-	-	SAM	resta
489	ANDAQUI	-	-	-	-	-	-	-	-	-	SAM	resta
490	Choco	0	0	0	4	0	0	4	0	3	SAM	resta
491	Páezan	0	0	0	1	0	0	0	0	-	SAM	resta
492	Barbacoan	0	0	0	0	0	0	0	0	0	SAM	resta
493	Leko	0	2	-	-	-	0	0	-	-	SAM	resta
494	Cholon	0	0	0	0	0	0	0	-	-	SAM	resta
495	Urarina	4	0	0	4	0	0	0	-	0	SAM	resta
496	Woorani	0	0	0	2	0	0	0	0	2	SAM	resta
497	Cahuapanan	0	0	-	0	4	0	-	-	4	SAM	resta
498	Zaparoan	0	2	0	0	0	0	0	0	-	SAM	resta
499	Quechuan	0	0	0	0	0	0	0	3	0	SAM	resta
500	Aymaran	0	0	0	0	0	0	0	4	-	SAM	resta
501	Puelche	2	-	2	2	2	-	0	2	-	SAM	resta
502	Araucanian	0	4	2	0	0	0	0	0	-	SAM	resta
503	Alacalufan	0	0	0	-	0	-	2	-	-	SAM	resta
504	Yámana	0	1	1	0	0	2	0	0	-	SAM	resta
505	Chon Proper		0	0	0	-	-	-	-	-	SAM	resta
506	Kwaza	0	0	2	4	0	-	-	-	4	SAM	resta
507	Kapixana	0	0	-	4	0	0	0	-	-	SAM	resta
508	Movima	4	4	4	0	4	0	0	2	-	SAM	F7 Bolivia
509	MUNICHI	-	-	-	0	-	-	-	-	-	SAM	resta
510	Iranxe	0	0	-	4	0	-	4	0	-	SAM	resta
511	Katukinan	0	0	0	4	0	0	2	-	4	SAM	resta
512	Aikaná	0	-	-	-	-	-	-	-	-	SAM	resta
513	ARUTANI o AWAKE	0	-	-	-	-	-	-	-	-	SAM	resta
514	SAPE o KALIANA	0	-	-	-	-	-	-	-	-	SAM	resta
515	Nambikuara n	0	0	-	-	-	-	-	-	-	SAM	resta

ara	genus	Subj.	Obj.	Adp.	Adje.	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Área
516	Puinave	-	-	-	-	-	-	-	-	-	SAM	resta
517	Nadahup	0	0	0	4	0	0	0	-	0	SAM	resta
518	Cacua-Nukak	0	0	-	-	-	-	-	-	-	SAM	resta
519	Ticuna	0	4	0	4	0	0	0	0	-	SAM	resta
520	Tucanoan	2	1	0	2	0	0	0	0	3	SAM	resta
521	Cayuvava	4	4	4	0	4	0	0	2	-	SAM	F7 Bolivia
522	Camsá	0	2	0	4	0	0	0	0	-	SAM	resta
523	Trumai	0	0	0	2	0	0	0	-	-	SAM	resta
524	Yuracare	0	0	-	-	-	-	-	-	-	SAM	resta
525	Piaroa	0	0	-	4	0	0	4	-	-	SAM	resta
526	Sáliba	0	0	0	-	-	-	-	-	-	SAM	resta
527	Timote-Cuica	0	4	4	2	-	-	0	-	-	SAM	resta
528	Cofán	0	0	0	0	0	-	2	0	-	SAM	resta
529	Tacame	4	4	4	2	4	-	-	-	-	SAM	resta
530	Yaruro	0	1	0	-	0	2	4	0	-	SAM	resta
531	Candoshi	0	0	0	4	0	-	4	1	-	SAM	resta
532	Taushiro	4	4	0	-	-	-	-	-	-	SAM	resta
533	Jivaroan	0	0	0	4	0	0	0	2	-	SAM	resta
534	Zamucoan	0	4	4	4	0	4	-	0	-	SAM	F7 Bolivia
535	Kariri	4	4	4	4	4	0	0	-	4	SAM	resta
536	PURUBORA	-	-	-	-	-	-	-	-	-	SAM	resta
537	Arikem	0	4	-	-	-	-	-	-	-	SAM	F7 Bolivia
538	Monde	0	0	0	4	0	0	-	-	-	SAM	resta
539	Ramarama	0	0	0	4	0	0	2	-	-	SAM	resta
540	Tupari	0	0	0	4	0	0	-	-	2	SAM	resta
541	Tupi-Guaraní	1	2	0	4	0	0	1	0	4	SAM	F7 Bolivia
542	AWETI	-	-	-	-	-	-	-	-	-	SAM	resta
543	MAWE-SATERE	-	-	-	-	-	-	-	-	-	SAM	resta
544	Munduruku	0	0	0	-	-	-	-	-	-	SAM	resta
545	Guahiban	0	0	0	0	0	0	-	0	4	SAM	resta
546	GUAMO	-	-	-	-	-	-	-	-	-	SAM	resta
547	Arauan	1	1	0	4	0	0	4	-	4	SAM	resta
548	Uru-Chipaya	0	0	-	1	0	1	0	-	0	SAM	resta
549	Chapacura-Wanham	4	4	4	-	4	4	2	4	4	SAM	F7 Bolivia
550	Central Arawakan	1	2	0	0	0	0	0	0	-	SAM	resta
551	Western Arawakan	4	4	-	0	4	0	0	0	-	SAM	F7 Bolivia
552	Purus	0	1	0	0	0	0	0	0	4	SAM	resta
553	Pre-Andine Arawakan	4	4	1	3	3	0	1	0	4	SAM	F7 Bolivia
554	Harakmbet	0	0	0	4	0	0	0	-	-	SAM	resta
555	Bolivia-Parana	3	4	2	4	4	1	1	1	4	SAM	F7 Bolivia

ara	genus	Subj.	Obj.	Adp.	Adje.	Gen.	Dem.	Num.	Poss.	Rel.	CONT	Àrea
556	Eastern Arawakan	0	4	4	4	0	0	0	-	-	SAM	F7 Bolivia
557	Northern Arawakan	1	3	1	2	1	2	0	0	3	SAM	F7 Bolivia
558	Andoke	0	0	0	0	0	0	-	0	-	SAM	resta
559	Peba-Yaguan	4	4	0	4	0	0	0	-	4	SAM	resta
560	Boran	0	0	0	0	0	0	4	0	-	SAM	resta
561	Huitoto	0	0	1	1	2	0	0	0	0	SAM	resta
562	Cariban	2	0	0	2	0	0	0	1	3	SAM	resta
563	Lule	0	0	-	4	0	-	0	0	-	SAM	resta
564	Mascoian	-	-	-	-	0	-	-	-	-	SAM	resta
565	Mosetenan	0	4	0	0	0	0	-	0	4	SAM	F7 Bolivia
566	Guaicuruan	1	4	2	2	3	0	0	0	4	SAM	F7 Bolivia
567	GUACHI	-	-	-	-	-	-	-	-	-	SAM	resta
568	Matacoan	0	4	4	4	0	4	2	0	4	SAM	F7 Bolivia
569	Panoan	0	0	0	4	0	1	3	0	-	SAM	resta
570	Tacanan	0	0	0	4	0	1	2	0	4	SAM	resta
571	Rikbaktsa	0	0	0	4	0	-	-	-	-	SAM	resta
572	Yatê	0	0	0	4	0	-	-	-	-	SAM	resta
573	Guató	4	4	4	4	4	0	-	-	-	SAM	F7 Bolivia
574	OPAYE	-	-	-	-	-	-	-	-	-	SAM	resta
575	Karajá	0	0	0	4	0	0	2	-	-	SAM	resta
576	Bororo	0	1	0	4	-	-	-	0	-	SAM	resta
577	Botocudo	0	0	-	4	0	4	4	-	-	SAM	resta
578	KAMAKAN	-	-	-	-	-	-	-	-	-	SAM	resta
579	PURI	-	-	-	-	-	-	-	-	-	SAM	resta
580	Jabutí	0	0	0	4	-	0	-	-	-	SAM	resta
581	Maxakalí	0	0	0	4	-	-	-	-	-	SAM	resta
582	Ge-Kaingang	0	0	0	4	0	3	4	0	3	SAM	resta
583	Chiquito	0	4	4	-	4	-	-	-	-	SAM	F7 Bolivia
584	Máku de Roraima	-	-	-	-	-	-	-	-	-	SAM	resta
585	YUWANA	-	-	-	-	-	-	-	-	-	SAM	resta
586	Puquina	-	-	0	-	-	-	-	-	-	SAM	resta
587	DORASKE	-	-	-	-	-	-	-	-	-	SAM	resta
588	OTOMACO	-	-	-	-	-	-	-	-	-	SAM	resta
589	PATA XO	-	-	-	-	-	-	-	-	-	SAM	resta
590	PAYAGUA	-	-	-	-	-	-	-	-	-	SAM	resta





