Linguistic diversity and the South American perspective

Antrittsvorlesung
Rik van Gijn
The (South-)American puzzle

How can it be that South America, the last continent to be populated, is so diverse linguistically?
The (South-)American puzzle
The (South-)American puzzle

How can it be that South America, the last continent to be populated, is so diverse linguistically?

Three claims:

1. the Americas were populated relatively recently
2. the initial settlers formed a (small, coherent) subset of the source population
3. the Americas are linguistically more diverse than expected
Populating the Americas: geology

Broad consensus: entry via Beringia

In genetic terms, native Americans are closest to northeast Asian populations.

Competing hypothesis (Atlantic migration) is much less well supported by archaeology and genetics.

Populating the Americas: geology

Bering Land Bridge
Open 60-36kya and 30-13kya

Ice-free land corridor
Impassible from ca. 24-14kya.

Coastal route
open from ca. 15kya

Museum of natural and cultural history
Populating the Americas: geology

**Bering Land Bridge**
Open 60-36kya and 30-13kya

**Ice-free land corridor**
Impassible from ca. 24-14kya.

**Coastal route**
open from ca. 15kya

Museum of natural and cultural history
Populating the Americas: archaeology

When were the Americas populated?

Monte Verde is a human settlement (there is even a human footprint) widely accepted to be at least 14,500 years old.

The Buttermilk-Creek complex in Texas is claimed to be between 15,500 and 13,200 years old.

Broken Mammoth dates back some 14,000 years.
Populating the Americas: archaeology

When were the Americas populated?

"Luzia" is the oldest human skeleton found in the Americas, dating back some 12,000 years.

Minas Gerais Museo Nacional
"The archaeological evidence of human entry into South America before 15,000 B.P. is weak and only presumed at this time. However, given the evidence confirming the presence of humans probably centuries before 12,500 B.P., the likely entry date should lie between 15,000 and 13,500 B.P."
Populating the Americas: genetics

Mitochondrial and Y-chromosomal DNA

Passed on directly through the maternal and paternal line, respectively, without recombination during meiosis.

→ All variation is caused by copying mistakes (random mutations), which are sufficiently rare to assume common ancestry for people who share the same mutation. Each these mutation defines a haplogroup named by a capital letter, and further numbers or smaller letters for subtypes.
Populating the Americas: genetics

Mitochondrial DNA

Estimated coalescence times for haplogroups present throughout the Americas lie between 15,000-17,000 BP. The spread over the Americas must have happened after that period.
Populating the Americas: genetics

Mitochondrial DNA

Estimated coalescence times for haplogroups present throughout the Americas lie between 15,000-17,000 BP. The spread over the Americas must have happened after that period.
Populating the Americas: summary

Geological, archaeological, and genetic evidence seem to converge on an entry date around 15,000 years ago.
The (South-)American puzzle

How can it be that South America, the last continent to be populated, is so diverse linguistically?

Three claims:

1. the Americas were populated relatively recently
2. the initial settlers formed a (small, coherent) subset of the source population
3. the Americas are linguistically more diverse than expected
Populating the Americas: initial diversity

Subset/population bottleneck

1. Haplogroups
2. Allele frequencies
Populating the Americas: initial diversity

Size of founder population

Estimates for the Americas differ quite substantially between 70 (Hey), 450 (Fagundes et al.) and 1800 (Mulligan et al) but are generally low, and all models assume a subsequent population explosion.
Populating the Americas: initial diversity

Multiple migrations

Genetic evidence suggests at least three migrations (the other two much later, and confined to North America.

>> All the South American native population stems from a single migration.
Populating the Americas: conclusions

Most evidence suggests that

- The Americas were first populated post-glacially ca. 15,000 years ago
- The source population came from Asia, and was probably small
- The entry point was Beringia, subsequent spread must have been very fast
- There is a single source population for all South American Native groups
Populating the Americas: missing data

However...

- We lack coastal data because of shifting coastlines
- Archaeological data in the tropical lowlands generally preserves badly, and excavations are challenging.
- There are some suggestions for (much) earlier archaeological dates, but they don't meet with broad consensus.
- The arrival of the Europeans in 1492 decimated the native peoples of the Americas, leaving a large gap in the DNA, cultural, and linguistic record
- There are relatively few well-preserved skeletons in South America, making the analysis of ancient DNA a challenge.
The (South-)American puzzle

How can it be that South America, the last continent to be populated, is so diverse linguistically?

Three claims:

1. the Americas were populated relatively recently
2. the initial settlers formed a (small, coherent) subset of the source population
3. the Americas are linguistically more diverse than expected
Linguistic diversity

1. Phylogenetic diversity:
   How many different groups of related languages are there in an area?

2. Structural diversity:
   What is the bandwidth of linguistic structures (grammatical characteristics) in an area?

Nichols 1992, Nettle 1999b
Phylogenetic diversity

South America represents about 13% of the inhabitable land mass

114 out of 430 top-level families (27%) are represented in South America

64 out of 188 isolates (43%) are in South America

Image: Taras Zakharko
Phylogenetic diversity

Nichols (1990)

The Americas together have 195 top-level language families

- If we apply the rates of language splits we know from other areas of the world to the Americas, it would require at least 50,000 years to develop the diversity found in the Americas.

- Only at - in her opinion - unrealistic rates of splitting events we come close to the proposed data of entry (20,000)

>> Proposed colonization dates of 15-18k BP are inconsistent with the linguistic facts. Allowing for initial linguistic diversity, the figure of 50,000 can be reduced to about 35,000, which would still mean a pre-glacial entry into the Americas.
Explaining the patterns

2-stage model: initial diversity increase, then decrease

Phylogenetic diversity as a function of time, size of area, population size, and an extinction rate.
Explaining the patterns

2-stage model: initial diversity increase, then decrease

Initial swift, rake-like dispersion, settlement in ecological niches, and rapid differential development of languages in relative isolation.
Explaining the patterns

Although it does not rest as much as Nichols’ account on observed patterns of change, it fits in with a more general pattern in South America.

Initial swift, rake-like dispersion, settlement in ecological niches, and rapid differential development of languages in relative isolation.

Nichols 1990

Nettle 1999b
Explaining the patterns

Although it does not rest as much as Nichols’ account on observed patterns of change, it fits in with a more general pattern in South America.

Fast spread

- Dates of archaeological sites compared to estimated entry date
- The continent-wide distribution of founder MtDNA and Y Chromosome haplogroups.
Explaining the patterns

Although it does not rest as much as Nichols’ account on observed patterns of change, it fits in with a more general pattern in South America.

Swift local developments

- Early ecological adaptations (maritime, highland, savannas, riverine, etc.) resulting in different technologies.
- Early domestication processes (from 10,000 BP)
- Genetic variation tends to be very local (corresponding to language rather than language family), suggesting isolated developments, especially in the Amazon.
Explaining the patterns

2-stage model: initial diversity increase, then decrease

Population pressure due to expanding families after adoption of agriculture.
Explaining the patterns

The development of agriculture in South America

- Early start domestication (from 10,000 BP)
- Slow development towards intensive agriculture (only from around 2,000 BP)
Explaining the patterns

Comparatively limited time depth of agriculture

1. Crops (maize, beans, manioc): low protein crops, difficult domestication process
2. Individual planting (rather than broadcast)
3. Lack of animals for agricultural tasks and for manuring
Explaining the patterns

Geographic and ecological factors hindering spread

1. North-south axis
2. Fragmented ecology
3. Physical barriers (highland v. lowland, Amazon river)

Explaining the patterns

Slow development of agriculture as the reason for surviving diversity in (South) America

Diversity in South America is a left-over from an earlier phase that has not been ‘eaten up’ by expanding agricultural populations.

Diamond (1997: 370)
"Had any food-producing Native American peoples succeeded in spreading far with their crops and livestock and rapidly replacing hunter-gatherers over a large area, they would have left legacies of easily recognized language families, as in Eurasia, and the relationships of Native American languages would not be so controversial."
Peopling of South America and phylogenetic diversity: conclusions

- A rapid dispersion of people not too long ago
- Followed by local differentiations
- Agricultural expansions ’eating up’ diversity, while leaving some islands of isolated languages
Linguistic diversity

1. Phylogenetic diversity:
   How many different groups of related languages are there in an area?

2. Structural diversity:
   What is the bandwidth of linguistic structures (grammatical characteristics) in an area?
**Structural diversity**

- A rapid dispersion of people not too long ago
  - Followed by local differentiations
  - Agricultural expansions 'eating up' diversity, while leaving some islands of isolated languages

South-American traits?
Structural diversity

Continent-wide features:
Nominalized subordinate clauses

Huallaga Quechua

\[
\begin{align*}
\text{wawqe-n-ta} & \quad \text{musya-:} \\
\text{brother-his-ACC} & \quad \text{know-I} \\
\text{I know his brother.}'
\end{align*}
\]

\[
\begin{align*}
\text{maqa-sha-n-ta} & \quad \text{musya-:} \\
\text{hit-ing-his-ACC} & \quad \text{know-I} \\
\text{I know that he hit (someone) - Lit. I know his hitting (someone).}'
\end{align*}
\]
Structural diversity

Continent-wide features:
Nominalized subordinate clauses

- 90% of sample languages have nominalized subordinate clauses
- Much more than expected on basis of Cristofaro (2003) - p=3.26e-06
- A lot of internal variation
Structural diversity

Continent-wide features:
Desideratives

Yurakaré

che-nta-y
eat-want-l
‘I want to eat.’
Structural diversity

Continent-wide features:
Desideratives

- 56% of the languages in the South American sample have a morphological strategy to mark a want construction.
- In the global sample of Haspelmath (2013), only 16% of the languages, a highly significant difference (p=5.176e-15)
Explaining the patterns

Continent-wide features: basic design features?

Some other widely shared South American features:

Clear noun-verb distinction, all other distinctions weak (Krasnoukhova 2012, Aikhenvald 2012)
Explaining the patterns

- South American families are very coherent wrt their stability profile, much more so than expected on the basis of chance frequency.
- Siberian languages cluster significantly with American languages, especially N American.

>> Consistent with the population history on the assumption that these profiles represent very slowly changing entities.
Structural diversity

- A rapid dispersion of people not too long ago
  Followed by local differentiations
  Agricultural expansions ‘eating up’ diversity, while leaving some islands of isolated languages

Exuberant diversity in the isolated parts (West Amazon)?
Structural diversity
Structural diversity

Features with wide, contiguous distribution in the western Amazon
Switch reference

Yurakaré

tibëjta-ja         tilmala         samu
it.sees.me-SS     it.runs.away.from.me  jaguar

‘When the jaguar saw me, it ran away from me.’

bëjtay-ti         tilmala         samu
l.see.it-DS       it.runs.away.from.me  jaguar

‘When I saw the jaguar, it ran away from me.’
Structural diversity

Features with wide, contiguous distribution in the western Amazon
Switch reference

Yurakaré

tibējta-ja      tilamala      samu
it.sees.me-SS  it.runs.away.from.me  jaguar

‘When the jaguar saw me, it ran away from me.’

bējtay-ti      tilamala      samu
I.see.it-DS    it.runs.away.from.me  jaguar

‘When I saw the jaguar, it ran away from me.’
Explaining the patterns

Features with wide, contiguous distribution
Switch reference

• Variation in SR systems tends to be small within families, big between families
• SR was well established in these languages before their families started to diverge (genealogical stability)
• Also areal sensitivity
• Possibly functional pressure (Van Gijn 2012)
Explaining the patterns

Features with wide, contiguous distribution
Switch reference

More ‘western’ features

Birchall 2014: double marking, suffixed person markers, indexing of R argument in ditransitives, accusative case alignment, marked neutral pattern in ditransitives, applicatives

Krasnoukhova 2012: pre-head position for all modifiers, Absence of gender and classifiers, property words are nouny, lack of inalienable nouns

Possibly much larger ancient contact area (Bickel & Nichols 2006)
Explaining the patterns

Pathways and barriers

In a series of papers, we have explored the Upper Amazon area, which can be regarded as a (highly diverse) zone between the Andes and the Amazon, which both have been regarded as (separate) contact regions, each associated with their own areal features.
Explaining the patterns

Pathways and barriers

Correlation elevation and closeness to Andean profile

Neighbornet of linguistic distances
Explaining the patterns

Pathways and barriers
Explaining the patterns

Pathways and barriers

A cline of «Andean» versus «Amazonian» features
Explaining the patterns

Pathways and barriers

A method that
1. generates random potential diffusion areas,
2. evaluates these on the basis of how well they explain the linguistic data.
3. selects the best diffusion area
Explaining the patterns

Identity preservation

South American hunter gatherer populations (as well as populations in contact zones more generally) tend to borrow few words.

>> Maintained diversity through identity preservation, emblematic function of language.
Reconstructing contact dynamics

Geographical context

Cultural-historical context

Communicative context

Preconditions for contact, likely routes, enabler and inhibitor

What kind of contact? Exchange, intermarriage, incorporation of groups into larger groups? Historical changes in the area

How did people communicate within the given context(s)? (multilingualism, lingua franca, language shifts?)

Geographical models of likely movement, potential contact

Genetics: admixture, archaeology, anthropology, history: cultural exchange, mobility of groups, extinction of groups, etc.

Linguistics: phylogenetics, feature distributions, theories of language change, lexicon

Geography → Sociohistory → Communication
Thank you!
References

Gijn, R. van 2016a. Want constructions in indigenous South American languages. Presentation given at the 49th SLE meeting September, Naples.
References


Krasnoukhova, O. 2012. *The noun phrase in the languages of South America*. Utrecht: LOT.


References

Roewer et al. 2013. Continent-Wide Decoupling of Y-Chromosomal Genetic Variation from Language and Geography in Native South Americans. *PLOS Genetics* 9 (4) e1003460
Tamm et al. 2007. Beringian Standstill and Spread of Native American Founders. *PLOS One* 2(9) e829