Creoles in contrast

Exploring the typological similarity of creoles and their lexifiers with neighbornet

Danae Perez & Melanie Röthlisberger
Introduction – what are phylogenetic trees good for?

(Bakker et al. 2011)
Introduction – the relevant questions

1) Do creoles constitute a class on their own (i.e. *Creole Exceptionalism*, cf. e.g. McWhorter 2001; Bakker et al. 2011)?

2) To what extent does the superstrate determine the typological similarities between contact languages (cf. Mufwene 2000, 2008; DeGraff 2005)?

3) Does a shared social history and/or geographical closeness determine typological similarities between contact languages?

4) How to classify “intermediate” varieties (Winford 2000) or “semi-creoles” (Holm 2004)?
Previous studies on creoles in comparison with their lexifiers: English in the Atlantic

(Daval-Markussen & Bakker 2011)
Previous studies on creoles in comparison with their lexifiers: Spanish world-wide

(Perez et al. 2017)
Previous studies on creoles in comparison with their lexifiers: Iberoromance in the Atlantic
Aims

1) To compare contact languages of three different lexifiers on the basis of 48 typological features in phylogenetic networks

2) To cast light on the question what determines the grouping of lexifiers and their contact varieties

3) To better understand how to compare and classify contact languages

4) To question and revisit the quantitative comparative method in creolistics based on statistical tools
The varieties in comparison
Data sources

- THE ATLAS OF PIDGIN AND CREOLE LANGUAGE STRUCTURES ONLINE

- THE WORLD ATLAS OF LANGUAGE STRUCTURES ONLINE

- dialectological and descriptive literature for under-documented varieties (e.g. Lipski 2005, 2007; Lucchesi et al. 2009; Perez in prep.)
Feature catalogue – an example

<table>
<thead>
<tr>
<th>No.</th>
<th>Feature/Variable</th>
<th>Value/Variant form</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Order of subject, object, and verb</td>
<td>1. SOV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. SVO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. VSO</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. OSV</td>
</tr>
<tr>
<td>2</td>
<td>Order of possessor and possessum</td>
<td>1. Possessor-possessum</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Possessum-possessor</td>
</tr>
<tr>
<td>3</td>
<td>Order of adjective and noun</td>
<td>1. Adj-N</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. N-Adj</td>
</tr>
<tr>
<td>4</td>
<td>Order of adposition and noun phrase</td>
<td>1. postposition</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. preposition</td>
</tr>
<tr>
<td>5</td>
<td>Order of demonstrative and noun</td>
<td>1. dem-N</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. N-dem</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. dem-N-dem</td>
</tr>
</tbody>
</table>
# Methodology

- 2 ways of coding the data

## Dataset A:

<table>
<thead>
<tr>
<th>Varieties</th>
<th>Feature: Word order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variety A</td>
<td>2</td>
</tr>
<tr>
<td>Variety B</td>
<td>2</td>
</tr>
<tr>
<td>Variety C</td>
<td>1</td>
</tr>
<tr>
<td>Variety D</td>
<td>2</td>
</tr>
<tr>
<td>Variety E</td>
<td>3</td>
</tr>
<tr>
<td>Variety F</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

## Dataset B:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Variety A</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Variety B</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Variety C</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Calculating the distance w/o variation

Dataset A:

• do two varieties share the same variant form for feature X?
  → yes: distance = 0
  → no: distance = 1

• calculate the number of shared variant forms divided by the number of comparisons (i.e. 48 features), ignoring NAs

<table>
<thead>
<tr>
<th>Varieties</th>
<th>Feature: Word order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variety A</td>
<td>2</td>
</tr>
<tr>
<td>Variety B</td>
<td>2</td>
</tr>
<tr>
<td>Variety C</td>
<td>1</td>
</tr>
<tr>
<td>Variety D</td>
<td>2</td>
</tr>
<tr>
<td>Variety E</td>
<td>3</td>
</tr>
<tr>
<td>Variety F</td>
<td>1</td>
</tr>
<tr>
<td>…</td>
<td>…</td>
</tr>
</tbody>
</table>
Calculating the distance with variation

Dataset B:

- Do two varieties share the same variant form?
- How many variant forms are possible per feature?
- Ignore NAs

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Variety A</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Variety B</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Variety C</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

distance = \sum \frac{number \ of \ non-shared \ variant \ forms}{number \ of \ variant \ forms}
Some technicalities

- our own distance measures in R
- phangorn library (Schliep 2011)
- visualizations with neighbornet (Bryant & Moulton 2004)
Results

1. without variation
   • Iberoromance creoles
   • English-based creoles
   • Creoles in contrast A

2. with variation
   • Iberoromance creoles
   • English-based creoles
   • Creoles in contrast B
Iberoromance creoles w/o variation

Based on Perez (under revision)
English-based creoles w/o variation

- Jamaican.Creole
- Nigerian.Pidgin
- Cameroon.Pidgin.English
- Ghanaian.Pidgin.English
- Nengee
- Saramaccan
- Sranan
- Creolese
- Vincentian.Creole
- Pichi
- Krio
- Nicaraguan.English.Creole
- San.Andres.Creole
- Trinidad.English.Creole
- Bahamian.Creole
- Gullah
- AAVE
- Standard.English
- Belizean.Creoles
- Jamaican.Creole

Creoles in contrast
Creoles in contrast A: w/o variation
Iberoromance creoles with variation

Diagram showing the relationships between different creoles and languages.
English-based creoles with variation
Creoles in contrast B: with variation
Summary of results

• varieties cluster according to lexifier AND according to geographic proximity

• standard-near varieties cluster close together (Spanish/Portuguese, AAVE/Standard English/Bahamian Creole)

• methodology: taking variation into account augments the accuracy in the description and comparison of varieties
Issues to consider

- frequency of variant forms cannot matter
  - Register specificity
  - some variant forms occur more frequently in spoken languages
  - documentation of basilectal (infrequent / non-standard) forms
thank you

danae.perez@es.uzh.ch
melanie.roethlisberger@es.uzh.ch
Selected references


Perez, D. under revision. Tracing the long-lost “missing Spanish creole”: Afro-Yungueño Spanish compared.

Perez, D. in prep. From the Atlantic to the Andes: The History and Evolution of Afro-Yungueño Spanish. Amsterdam: Benjamins.