

CCN (“Corpus of Cyber-Nigerian”): Concordancing, annotation and visualisation in a very large web- derived CMC corpus

Christian Mair

Englisches Seminar, Universität Freiburg

DFG MA 1652/9-1: *'Cyber-Creole'*



Deutsche
Forschungsgemeinschaft

DFG

Roadmap

1. Introduction: scenes from contemporary life – mobile speakers, mobile languages
2. The research agenda: from "English as a World Language" (EWL) to the "Sociolinguistics of Globalisation" / The World System of (Standard and Non-Standard) Englishes
3. www.nairaland.com and CCN: from the community to the corpus (and back)
4. Conclusion and outlook

**1. Introduction: scenes from
contemporary life – mobile speakers,
mobile languages**

(i) the Italian from England and the Nigerians from the US meeting in cyberspace ...

(Topic: "How to convince an Ibo girl that I am truly in love with her," www.XXX.com, 2008)

XXX ["sono di CITY IN SOUTHERN ITALY ma vivo in UK"]: please please please i beg you to help me!!!
I am an Italian white man who has met this beautiful Ibo girl. [...] this girl i am in love with, she is married too to a **English man**, but he doesn't treat her the way she should be. we are very closed friends and [...]
any help from you Ibo people will be really appreciated.
i thank you all in advance for **all your advises**.

Houston TX and New York City joining the conversation ...

YYY: Oka what now?

Are you sure you are not on the rebound? Cause you sound sort of desperate. Also the chick is married and you dare disrespect her marriage by involving herself in it? C'mon now fella, i know you smarter than that. My diagnosis, is for you to **lay off the chick**, and really think to yourself that you ACTUALLY LOVE her, not infatuated by her.

For your future reference, **it is Igbo not ibo.**

ZZZ: if you married her you'D be headed for divorce soon ... maybe you'll be seeking help from Wollof folks this time.

(ii) The Nigerian from Treviso saving the day

...

XXX: nonostante tu dici che il tuo Italiano non sia buono, ti esprimi benissimo e sai scrivere bene. complimenti. [...]

AAA [= "Gospel from CITY IN SOUTHERN ITALY"?]:

Grazie per il tuo complemento.

una domanda solo **quanto anni sei** e quanto anni e lei?

XXX: io 33 e lei 26

2. The research agenda:

from "English as a World Language" (EWL) to the "Sociolinguistics of Globalisation" / The World System of Standard and Non-standard Englishes

From "varieties of English" to Englishes in a multilingual global language ecology ...

- varieties → styles → ideologies
- structural features → communicative resources
- (CMC) spelling → (digital) literacy practice
- (monolingual) native-speakers → multilingual and / or truncated repertoires
- local / vernacular community → deterritorialised community of practice
- face-to-face interaction → mediated communication

Working hypotheses ...

The more dominant English is globally, the more heterogeneous it becomes internally. The farther the language spreads, the more it is affected by the multilingual settings in which it is being used.

"Natural" links between vernaculars and their territories and primary communities are becoming weaker, as migrations and media encourage the transnational and global flow of linguistic resources.

Even in a world in which "all Englishes are everywhere", there is no egalitarian polyphony of voices. Unequal relationships and restricted access to linguistic resources persist.

The World Language System (de Swaan 2002, 2010)

- the World Language System as "a surprisingly efficient, strongly ordered, hierarchical network, which ties together – directly or indirectly – the 6.5 billion inhabitants of the earth at the global level" (de Swaan 2010: 56)

English – contact language for the world ...

- **hyper-central** language: English, the hub of the world language system
- **super-central** languages: French, Hindi, Mandarin, Spanish, Turkish, Arabic, ...
- **central** languages: Dutch, Finnish, Cambodian, ...
- **peripheral** languages: 6,000+

The World System of Standard and Non-standard Englishes (Mair 2013)

... the English Language Complex as ... a surprisingly efficient, strongly ordered, hierarchical constellation of varieties, styles and registers which ties together – directly or indirectly – the 1-billion-plus regular users of English at the global level ...

- **the hub / hyper-central variety:**
 - Standard American English
- **super-central varieties:**
 - **standard:** BrE, AusE, IndE, NigE, ...
 - **non-standard:** AAVE, Jam CreoleE, popular LondonE, ...
 - **further domain-specific ELF uses:**
 - science, business, international law
- **central varieties:**
 - **standard:** IrE, NZE, JamEng, ...
 - **non-standard:** US "Southern," ...
- **peripheral varieties:** all traditional rurally based non-standard dialects, plus a large number of ex-colonial varieties including pidgins and creoles

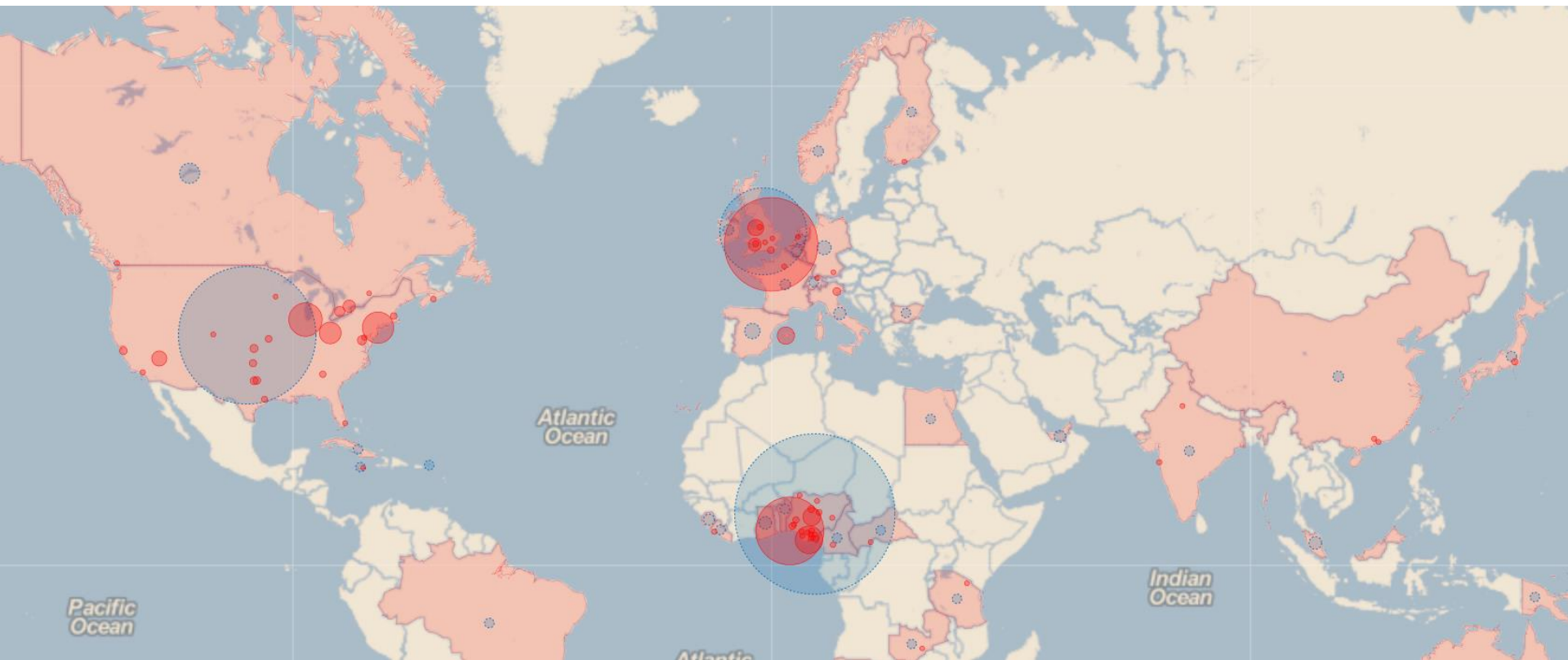
3. www.nairaland.com and CCN: from the community to the corpus (and back)

Data:

- Corpus of "Cyber-Jamaican" (CCJ)
www.jamaicans.com
2,128 members; 252,015 posts
16.9 million tokens
2000 – 2008
- Corpus of "Cyber-Cameroonian" (CCC)
www.cameroon-info.net
3,140 members; 179,563 posts
22.1 million tokens
2000 – 2008
- Corpus of "Cyber-Nigerian" (CCN)
www.nairaland.com
11,718 members; 244,048 posts
17.3 million tokens
2005 – 2008

N-CAT (Net Corpora Administration Tool) – concordancing and visualisation: CCN

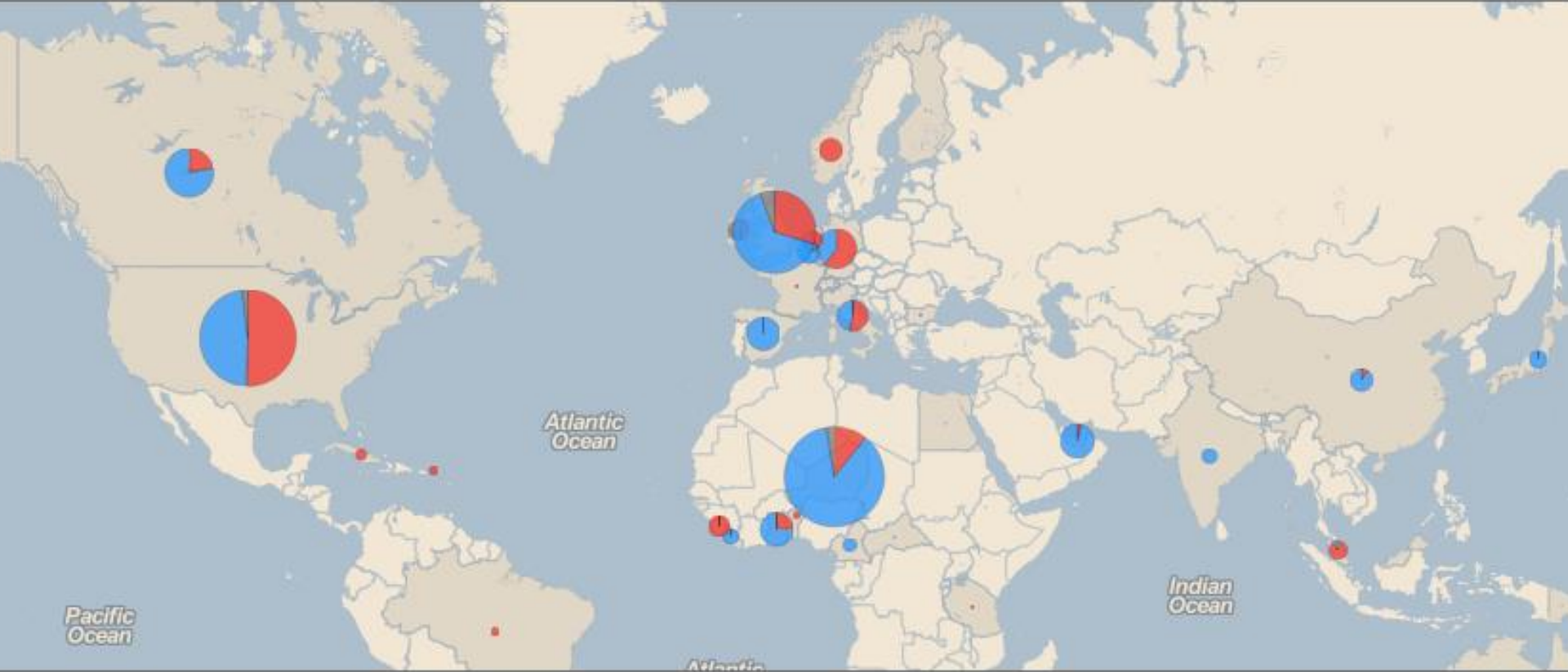
Search item a: Nairaland is a globally dispersed digital diaspora.



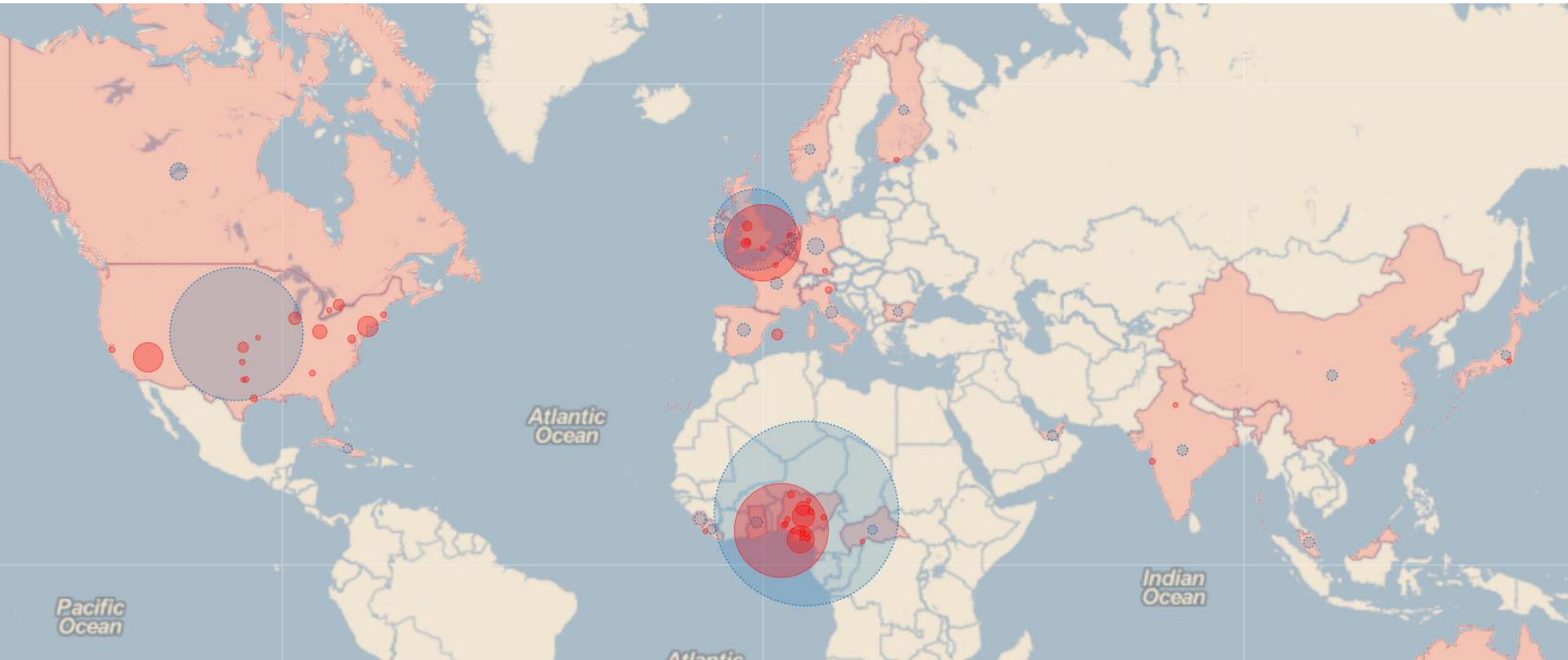
A look at the underlying raw data ...

docs / country		docs / city	docs / month
"none" - 70539	...	"none" - 158917	...
Nigeria - 63811	Jamaica - 47	London - 19646	04.2007 - 11797
US - 54570	Pap. N. G. - 47	Lagos - 17128	05.2007 - 13925
UK - 33411	Egypt - 44	Chicago - 5969	06.2007 - 11429
Canada - 4548	Switzerland - 37	New York - 5328	07.2007 - 10159
Spain - 2515	Tanzania - 36	Pt Harcourt - 4592	08.2007 - 7651
Germany - 1881	Portugal - 28	Abuja - 3590	09.2007 - 10077
Malaysia - 1397	Togo - 27	Columbus - 3370	10.2007 - 7999
Ghana - 1360	Botswana - 13	Palma - 2491	11.2007 - 7147
Italy - 1206	Cote d'I. - 9	Liverpool - 2238	12.2007 - 12632
Australia - 1112	Namibia - 8	Las Vegas - 2058	01.2008 - 10937
UAE - 960	Romania - 8	Cardiff - 1629	02.2008 - 8284
Fiji - 747	Kenya - 7	Toronto - 1574	03.2008 - 6905
Sierra L. - 741	Senegal - 6	Washington - 880	04.2008 - 7456
Netherl. - 618	Korea (S.) - 4	Houston - 773	05.2008 - 7402
Ireland - 476	Poland - 4	Suva - 747	06.2008 - 11052
China - 475	Zambia - 4	Brisbane - 744	07.2008 - 10833
Belgium - 455	Grenada - 3	Ibadan - 700	08.2008 - 7857
Norway - 440	Congo - 2	Boston - 635	10.2008 - 4
...		...	12.2008 - 365

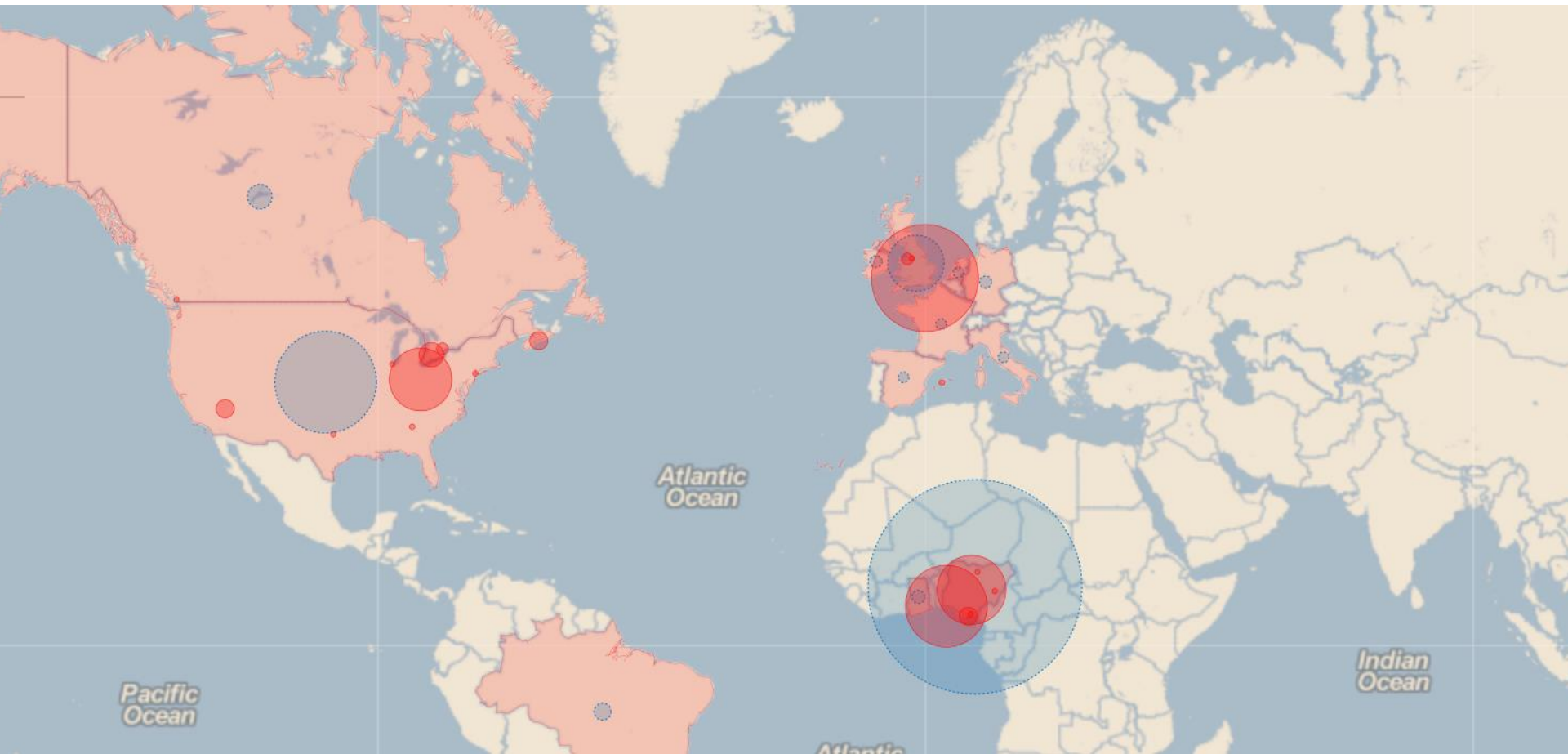
Search item *a* (# posts by country): Nairaland is a globally dispersed *and gendered* digital diaspora.



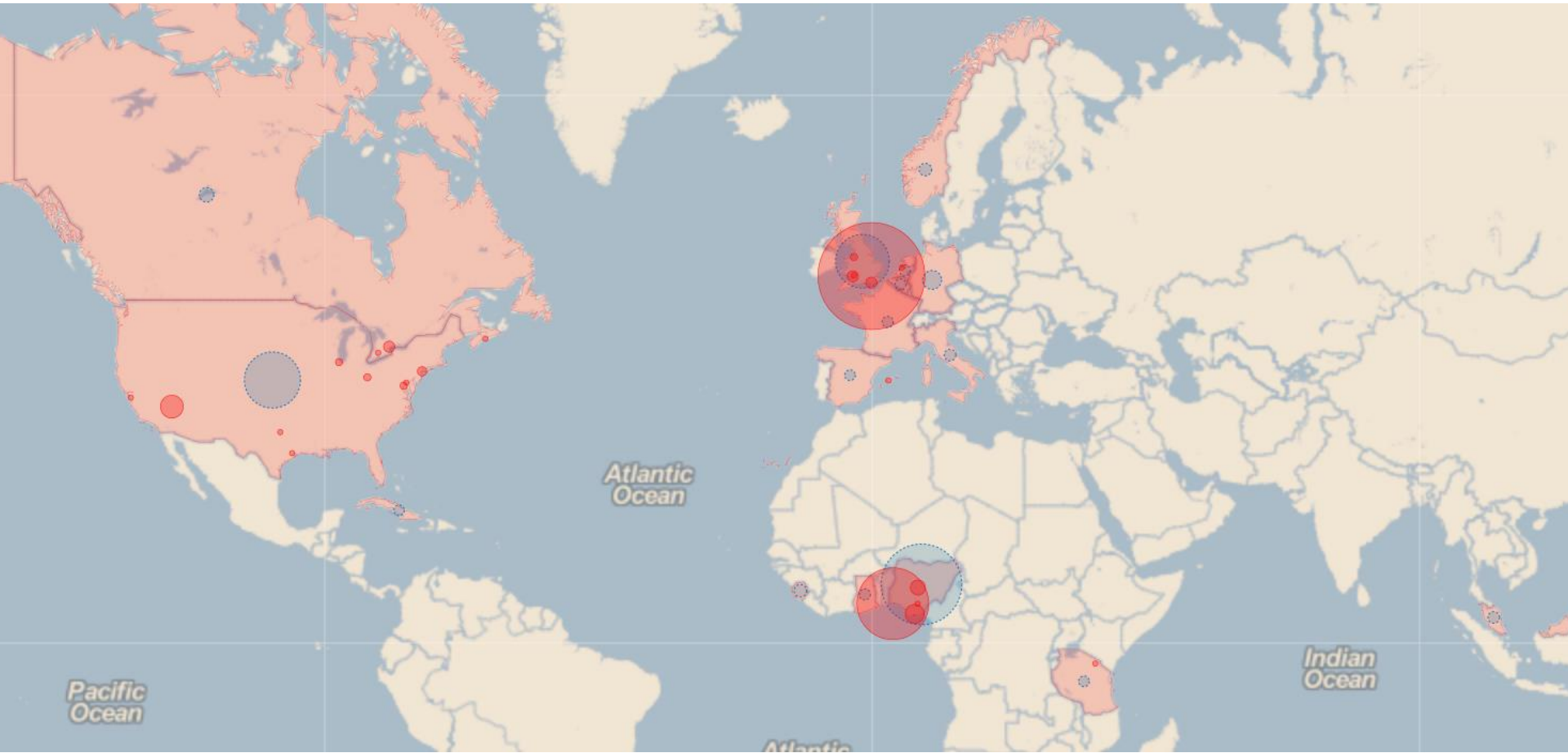
Search item *abi*: Nigerian Pidgin is globally dispersed ...



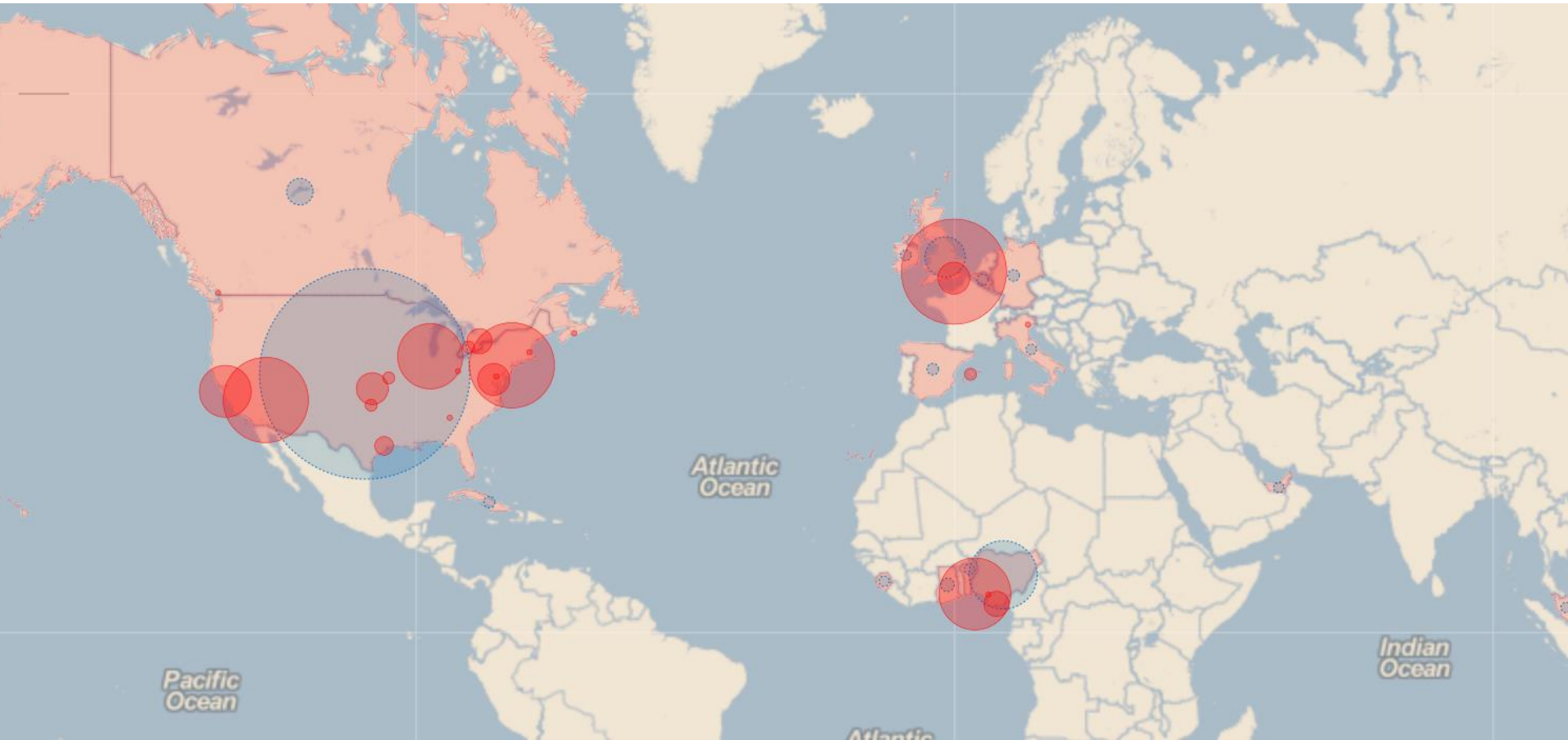
Search item *stuffs*: ... more dispersed in fact than Nigerian English.



Search item *Mum and Dad*: Empire still casts its shadow.



Search item *Mom and Dad* shows selective assimilation of Nigerian immigrants to US norms.



- (1) if you are over 25 and your dad talks to you in a rude manner... Dude, it's high time you "win your own bread" just like seun said. **Get your ass out of the house.**
(CNN, hot-angel [5754])
- (2) GOOD FOR EMEKA, IF I WERE HIM NA POLICE I GO TAKE **ARREST YOUR DUMB ASS ! NEXT TIME YOUR STUPID ASS WILL NOT BET ON UNCERTAINTY, ESPECIALLY ON A TEAM THAT BARELY MADE IT TO THE SECOND ROUNDS.** (CCN, Sweet T [5409])
- (3) My views on divorce are: I am willing to divorce **any stupid-ass man** that messes with my heart (CNN, hot-angel [5754])

4. Conclusion & outlook: discourse analysis of CMC

"[...] networked writing questions the adequacy of the feature-based approach and spoken language bias that have dominated conceptions of language change in sociolinguistics. [...] a change of scale in the volume and publicness of vernacular writing; a diversification of old and new vernacular patterns; an extension of written language repertoires, and a concomitant pluralisation of written language norms." (Androutsopoulos 2011: 153)

"In domains of unregimented writing, stylistic appropriateness is opened up to localised negotiation, for example with regard to spelling and punctuation or the representation of regional dialects." (Androutsopoulos 2011: 155)

→ *What is it that we "hear" when we "read" the digitally mediated vernacular voices?*

Conclusion & outlook: World Englishes research

weakness of the postcolonial nation state $\leftarrow \rightarrow$ weakness of the emerging national standard?

inversion of prestige of Standard New Englishes and the corresponding pidgins and creoles in the diaspora and the global mediascape?

Crisis of the Outer Circle?

References

- Alim, H. Samy, Awad Ibrahim, and Alastair Pennycook, eds. 2009. *Global Linguistic Flows: Hip Hop Cultures, Youth Identities, and the Politics of Language*. New York NY: Routledge.
- Androutsopoulos, Jannis. 2011. "Language change and digital media: a review of conceptions and evidence." In Tore Kristiansen and Nikolas Coupland, eds. *Standard Languages and Language Standards in a Changing Europe*. Oslo: Novus Press. 145-159.
- Androutsopoulos Jannis . 2010. "The study of language and space in media discourse." In Peter Auer and Jürgen E. Schmidt, eds. *Language and Space: An International Handbook of Linguistic Variation*. Volume 1: *Theory and Methods*. Berlin, New York: de Gruyter. 740-758.
- Androutsopoulos, Jannis. ed. 2006. Special issue on *The Sociolinguistics of Computer-Mediated Communication*. *Journal of Sociolinguistics* 10(5).
- Appadurai, Arjun. 1996. *Modernity at Large: Cultural Dimensions of Globalisation*. Minneapolis MN: University of Minnesota Press.
- Beck, Rose Marie. 2010. "Urban languages in Africa." *Africa Spectrum* 45: 11-41.
- Benor, Sarah Bunin. 2010. "Ethnolinguistic repertoire: Shifting the analytic focus in language and ethnicity." *Journal of Sociolinguistics* 14: 159-183.
- Blommaert, Jan. 2010. *The Sociolinguistics of Globalization*. Cambridge: Cambridge University Press.
- Cameron, Deborah. 2012. "The commodification of language: English as a global commodity". In Terttu Nevalainen and Elizabeth Traugott, eds. *The Oxford Handbook of the History of English*. Oxford: Oxford University Press. 352-364.
- Castells, Manuel. 2010. *The information age*. 3 vols. I: *The rise of the network society*. II: *The power of identity*. III: *End of millennium*. Oxford: Blackwell.
- Coupland, Nikolas. 2007. *Style: language variation and identity*. Cambridge: CUP.
- Coupland, Nikolas. 2010. "Language, ideology, media, and social change." In Karen Junod and Didier Maillat, eds. *Performing the self*. Tübingen: Narr. 55-79.
- Coupland, Nikolas, ed. 2010. *The Handbook of Language and Globalization*. Malden, MA: Blackwell.
- de Swaan, Abram. 2002. *The World Language System: A Political Sociology and Political Economy of Language*. Cambridge: Polity.
- de Swaan, Abram. 2010. "Language systems". In Nikolas Coupland. *The Handbook of Language and Globalization*. Malden, MA: Blackwell. 56-76.
- Goglia, Francesco. 2009. "Communicative strategies in the Italian of Igbo-Nigerian immigrants in Italy: a contact-linguistic approach." *Sprachtypologie und Universalienforschung* 62: 224-240.

- Heller, Monica. 2003. "Globalization, the new economy, and the commodification of language and identity." *Journal of Sociolinguistics* 7: 473-492.
- Heller, Monica. 2011. *Paths to post-nationalism: a critical ethnography of language and identity*. Oxford: OUP.
- Heyd, Theresa, and Christian Mair. 2014. "From Vernacular to Digital Ethnolinguistic Repertoire: The case of Nigerian Pidgin." In Véronique Lacoste, Jakob Leimgruber, and Thiemo Breyer, eds. *Indexing Authenticity: Sociolinguistic Perspectives*. FRIAS *Linguae & Litterae* Series. Berlin: de Gruyter. 244-268.
- Jaffe, Alexandra, Jannis Androutsopoulos, Mark Sebba, and Sally Johnson, eds. 2012. *Orthography as social action: scripts, spelling, identity and power*. Berlin and New York: Mouton de Gruyter.
- Mair, Christian. 2011. "Corpora and the New Englishes: Using the 'Corpus of Cyber-Jamaican' (CCJ) to explore research perspectives for the future". In Fanny Meunier, Sylvie De Cock, Gaëtanelle Gilquin and Magalie Paquot, eds. *A Taste for Corpora: In Honour of Sylviane Granger*. Amsterdam: Benjamins. 209-236.
- Mair, Christian. 2013a. "The World System of Englishes: Accounting for the transnational importance of mobile and mediated vernaculars." *English World-Wide* 34: 253-278.
- Mair, Christian. 2013b. "Corpus-Approaches to the Vernacular Web: Post-Colonial Diasporic Forums in West Africa and the Caribbean." In Katrin Röder & Ilse Wischer, eds. *Anglistentag 2012: Proceedings*. Trier: WVT, 2013. 397-406. [erweiterter Nachdruck in *Covenant Journal of Language Studies* [Ota, Nigeria] 1: 17-31. Digitale Version: <http://journals.covenantuniversity.edu.ng/jls/published/Mair2013.pdf>
- Mair, Christian, and Stefan Pfänder. 2013. "Vernacular and multilingual writing in mediated spaces: web forums for post-colonial communities of practice." In Peter Auer, Martin Hilpert, Anja Stukenbrock & Benedikt Szmrecsanyi, eds. *Space in Language and Linguistics: Geographical, Interactional, and Cognitive Perspectives*. Berlin/New York: de Gruyter, 2013. 529-556.
- Mesthrie, Rajend, and Rakesh M. Bhatt. 2008. *World Englishes: The Study of New Varieties*. Cambridge: Cambridge University Press.
- Moll, Andrea. 2012. *Jamaican Creole Goes Web: Sociolinguistic Styling and Authenticity in a Digital Yaad*. PhD University of Freiburg.
- Rampton, B. 1995. *Crossing: language and ethnicity among adolescents*. London: Longman.
- Sebba, Mark. 2007. *Spelling and Society: The Culture and Politics of Orthography around the World*. Cambridge: CUP.
- Sebba, Mark. 2012. *Language Mixing and Code-Switching in Writing: Approaches to Mixed-Language Written Discourse*. London: Routledge.
- Sharma, Devyani. 2011. "Style repertoire and social change in British Asian English." *Journal of Sociolinguistics* 15: 464-492.
- Vertovec, Steven. 2007. "Super-diversity and its implications". *Ethnic and Racial Studies* 30: 1024-1054.