University of Zurich[UZH]

English Department

KU LEUVEN

Approaching Epicentral Influence with Agent-Based Modelling

Marianne Hundt, Laetitia van Driessche and Dirk Pijpops

Workshop on

'Modelling Constructional Variation and Change: Agents, Networks and Vectors'

What is meant by an ‘epicentre’ and why use ABM to study it?

Norm orientation in pluricentric languages

“Language spreads from the centre to the periphery; the periphery develops independent ‘standards’ that first compete and then coexist with those of the homeland, and these new standards may in their turn become new centres of radiating influence.”

Bailey (1990: 85)

Motivation

- Usage-based evidence on epicentre influence requires diachronic data
- Corpus data only do not give us the whole picture (lack of information on attitudes, see Hundt 2013)
- Use complementary approach – Agent-Based Modelling (ABM): simulation of variation to predict variation *should* attitudes play an important role for epicentral influence

Objectives

- Case study on verb complementation in two metropolitan varieties and two ESL
- Build agent-based model to gauge potential effect of attitudes vs. dialect contact
- Compare results of ABM with corpus data
- Feed results into theory on modelling epicentre in World Englishes

Caveat

ABM does not aim at modelling the actual complexities of language users/speakers' choices (e.g. by including historical, cultural, individual factors)

ABM adhere's to Occam's Razor: aim is to build a model that is as simple as possible while still simulating the mechanisms under scrutiny (Landsbergen, 2009: 18-19; van Trijp & Steels, 2012: 9)

ABM – corpus data – theory

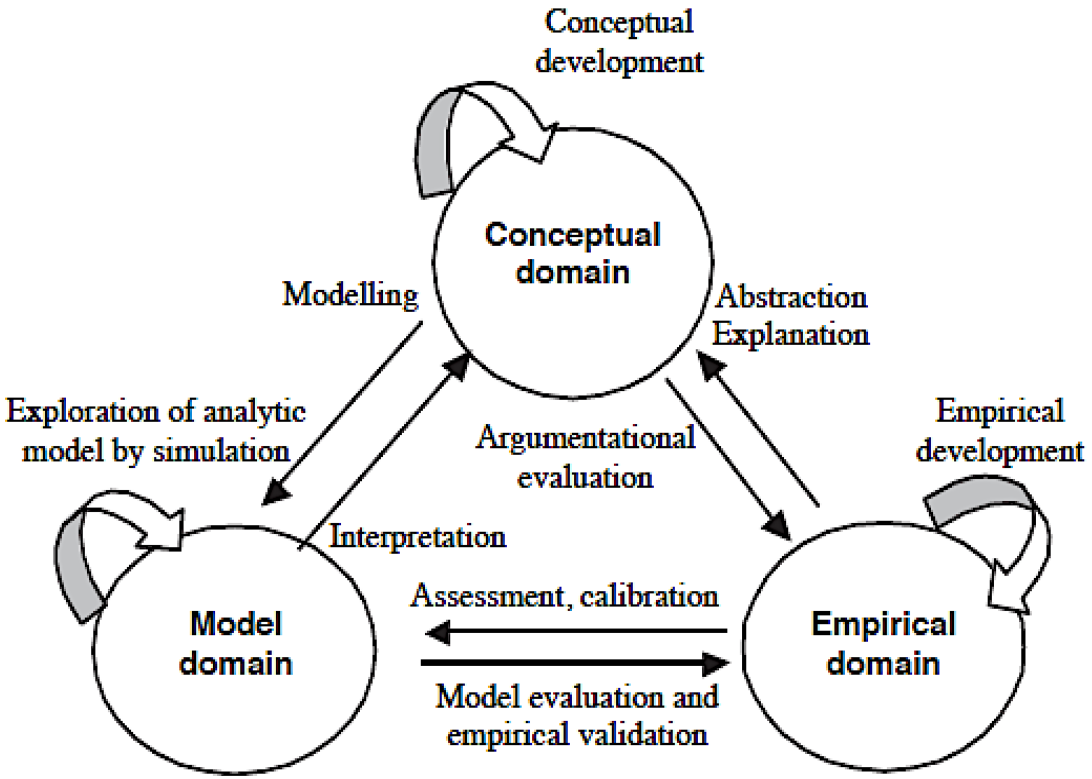


Fig 1. Relating simulation to theory and data
(from Livet et al., 2014)

Case study: ASCs with *protest* and *appeal*

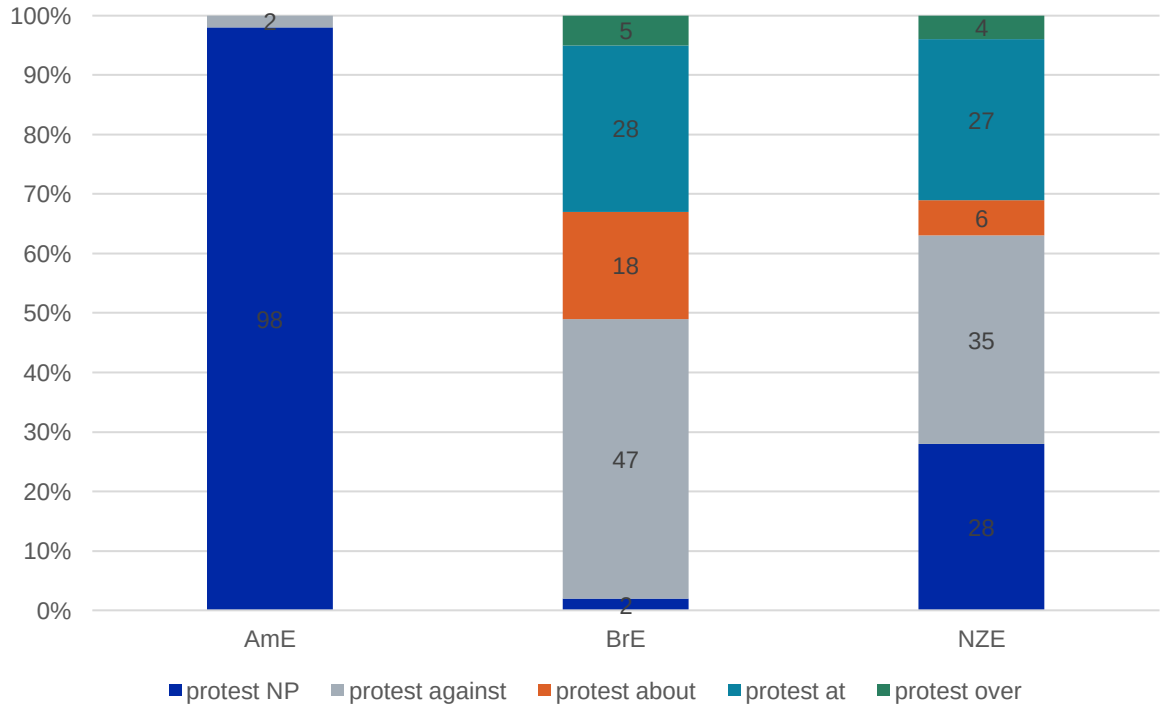


Fig. 2. Complementation of *protest* in Newspaper databases

(based on Hundt 1998; N = 100 per variety from the *Miami Herald* (US), the *Guardian* (GB) and the *Dominion/Evening* Post (NZ), respectively)

ESL varieties and the dynamics of WEs

Indian English (IndE) – with BrE as matrilect/traditional exonormative model

Philippine English (PhilE) – with AmE as matrilect/traditional exonormative model

- Acquisition of English as second language (education, online media)
- AmE at the top of the hierarchy (Mair 2013)
- Individual speakers differ in their attitude towards the historical matrilect / AmE

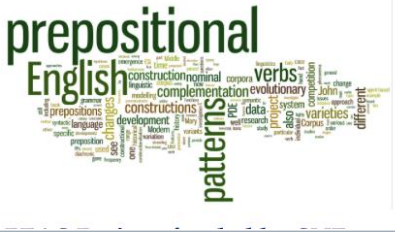
Attitudes and ASCs

"It is often claimed that American English uses more words than are strictly necessary … . However, it is interesting that complaints are just as common when an American innovation is in fact more concise, more economical than the traditional British form, as in *to protest* or *to appeal*, as opposed to *to protest against* or *to appeal against.* … This inconsistency suggests that hostility to American usages may have less to do with the linguistic forms themselves than with a rather deeper dislike of the society and culture from which these forms derive."

Gordon & Deverson (1989: 75f.)

Research questions

- Do attitudes towards the matrilects play a role in epicentre influence (shift from BrE to AmE ASC)?
  - Assumption that PhilE speakers are more conservative than AmE speakers in their ASCs (shift occurred in the second half of the twentieth century)
  - PP complements are structurally more explicit and therefore generally favoured by ESL speakers.
- Does lexical variation play a role in epicentral influence?
  - Does frequency have an influence on epicentral influence?
  - Does paradigmatic variation in the ASCs have an influence?



Components of the model

- 2 populations of agents (Indian, Filipino), 2 sources (British, American)
- The populations have different distributions according to their attitudes towards AmE
  - IndE and PhilE have different matrilects, i.e. speakers **start off from different usage preferences** related to (but not identical with) their matrilects
  - speakers are **aware of the regional differences** between BrE and AmE in the complementation of the two verbs and speakers in the ESL communities differ with respect to their attitudes towards AmE (some speakers are pro-Am, some are con-Am, some are neutral)

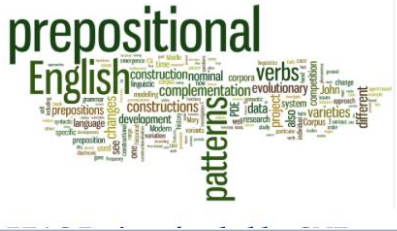
Population	con-Am (<1)	Neut (=1)	pro-Am (>1)
Indian	33	34	33
Philippino	5	45	50

Components of the model

– There are **regional differences** in the ASCs of *protest* and *appeal* (AmE prefers bare NP complement, BrE prefers PP complement, see Hundt 1998, Rohdenburg 2009)

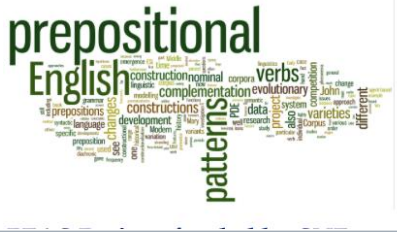
Language form	BrE	AmE
Protest + NP	15	70
Protest + against	70	20
Protest + other preps	5	5
Appeal + NP	10	90
Appeal + against	90	10

– There is a bias towards protest (60-40) because of the higher frequency of this verb



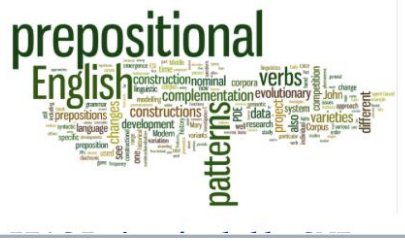
Components of the model

- Our simulation = 1000 points in time
- Each point in time = several interactions between agents
- For every interaction: one random speaker + one random hearer is selected
- A verb is selected (probabilities)
  - The speaker selects all forms for that verb in their memory
  - The speaker consults the entrenchment scores (= counts of forms in their memory)
  - If the form is NP (typical for AmE): entrenchment score * attitude factor
  - A form is produced according to probabilities based on the entrenchment scores
  - +1 for the entrenchment score of the verb form in the memory of the hearer agent



Components of the model

- Speakers are exposed to BrE and AmE variants through the media (news on the web)
  - After several interactions: opportunity for agents to read a source
  - How often they read an American vs. a British source depends on settings that we can change
    - Default settings: agents are twice more likely to read an American than a British source (based on distributions NOW corpus)
  - A random agent is chosen to read the source
  - The source generates a form (probabilities)
  - The agent adds + 1 to the entrenchment score in their memory



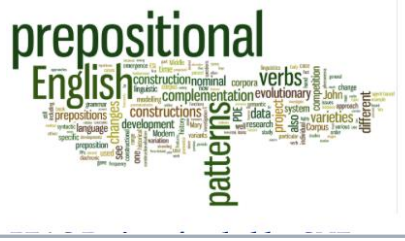
Three different models

Condition 1: Stable matrilectal input

= There is no ongoing language change and the input from the matrilects is stable.

Condition 2: Speakers in the two communities differ in their reading preferences in that con-Am speakers actively avoid American sources and thus get lower Am input.

- We change the factor that regulates the influence from the sources
- It will then become more likely for con-Am agents to read a British source



Three different models

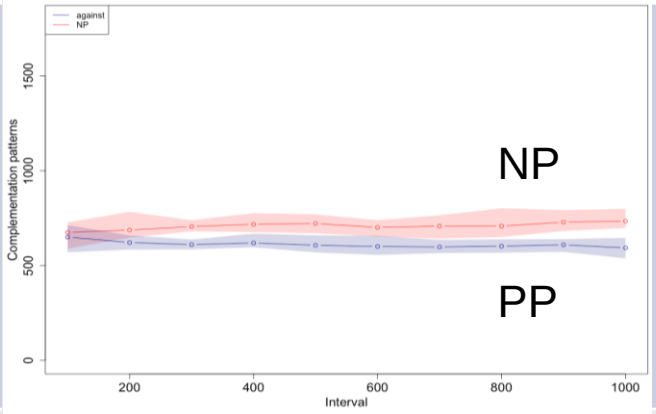
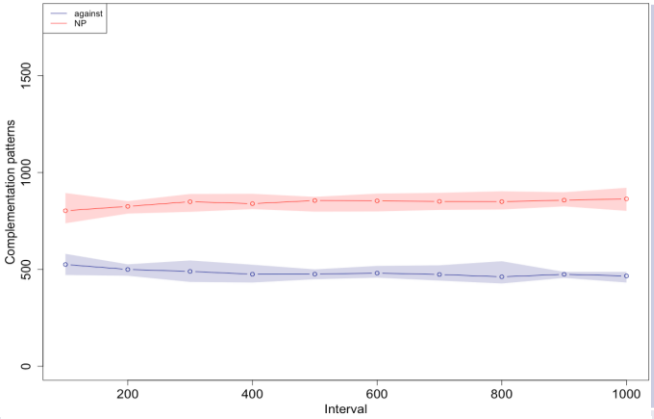
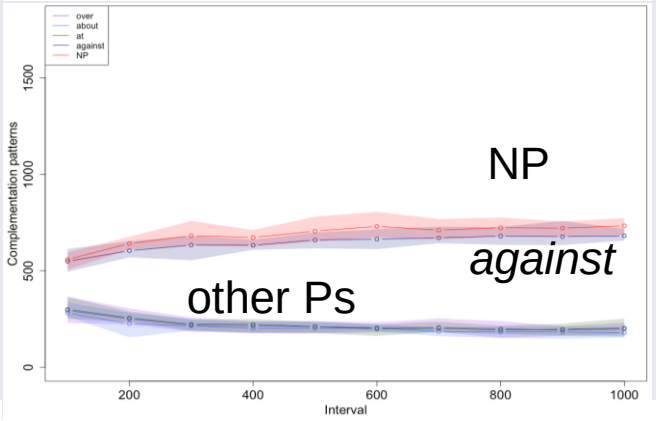
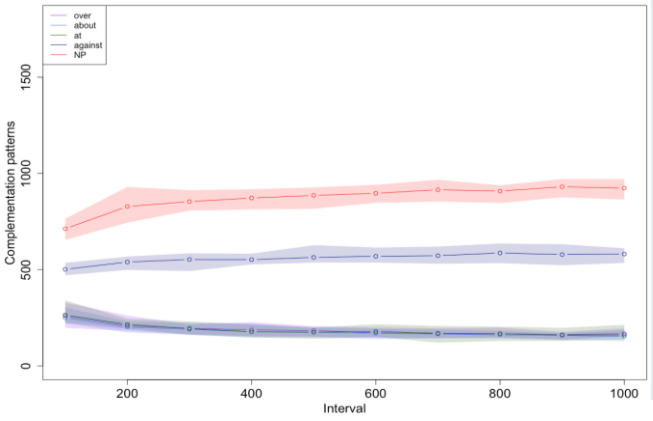
Condition 3: In addition to the parameters under condition 2, there is ongoing language change such that BrE is slowly shifting towards AmE complementation patterns for *protest* and *appeal;* this is the 'global Americanisation' condition.

= There will be a gradual rise in NPs for the British source

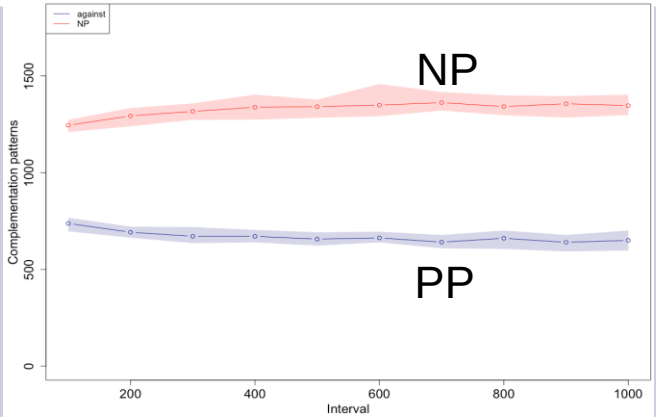
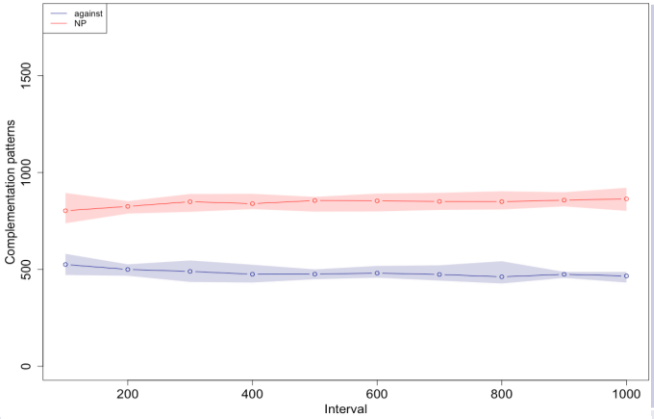
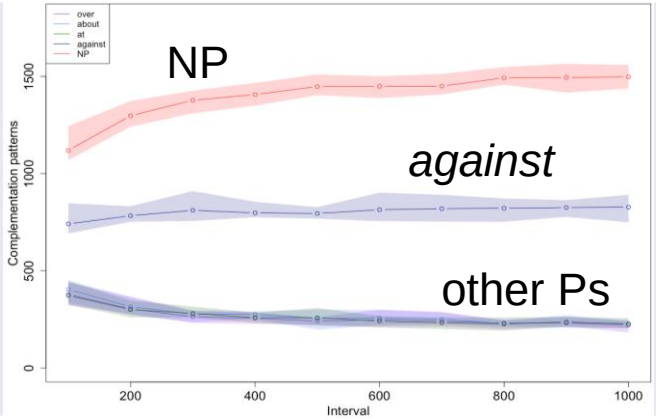
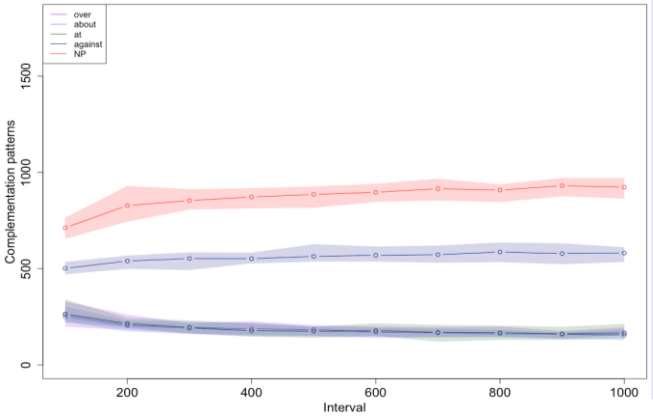
- After several points in time, the American source produces a language form
- The British source adds + 1 to that form in their distribution
- The British source will start looking more like the American source over time

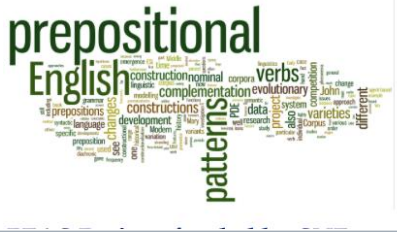
Implementation in Python (Laetitia van Driessche & Dirk Pijpops)

Results of the simulation: stable matrilectal input

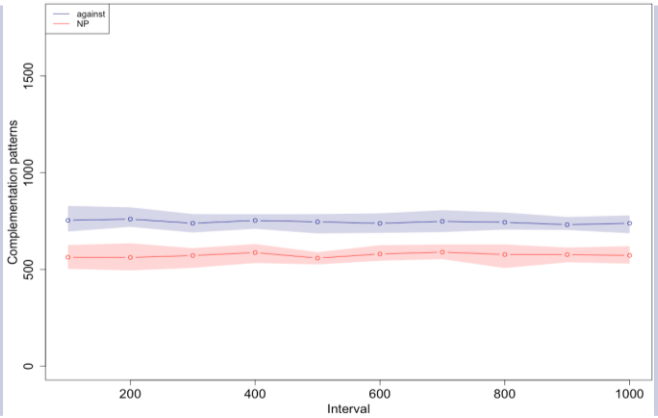
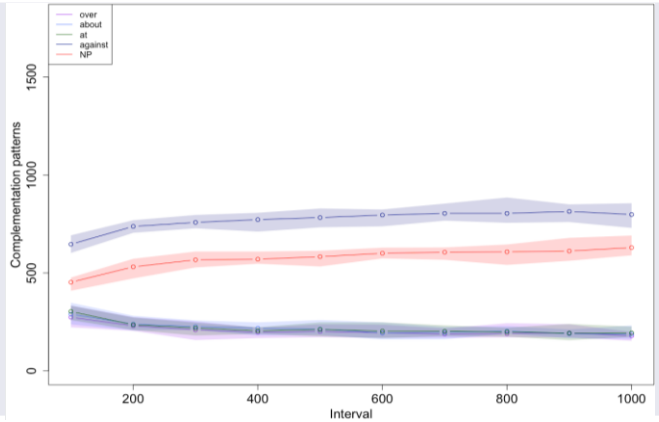
	IndE Con-American	IndE Pro-American
appeal		
protest		

Results of the simulation: stable matrilectal input

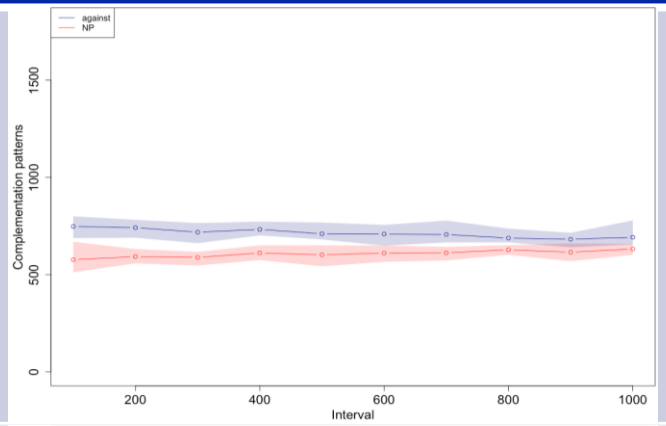
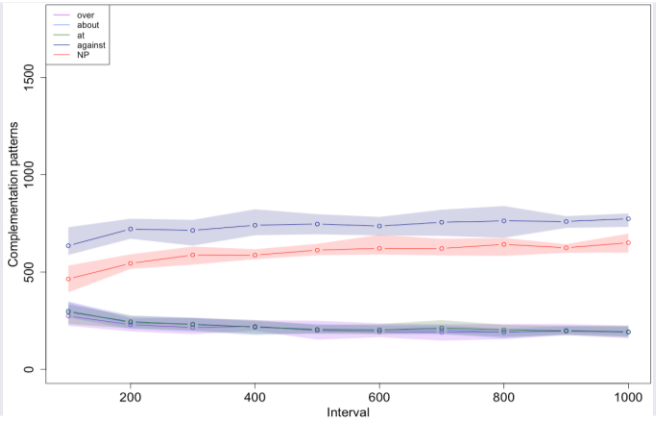
	PhilE Pro-American	IndE Pro-American
appeal		
protest		



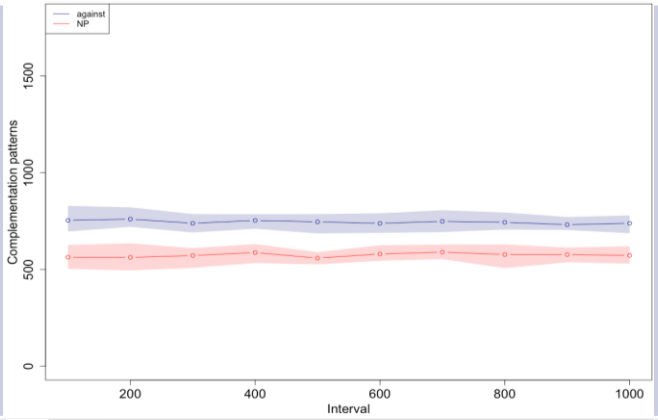
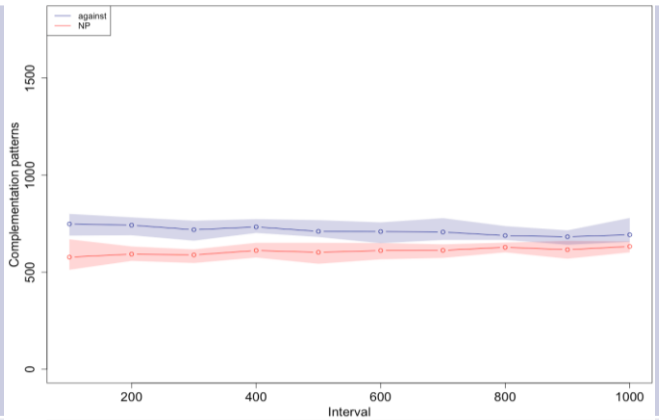
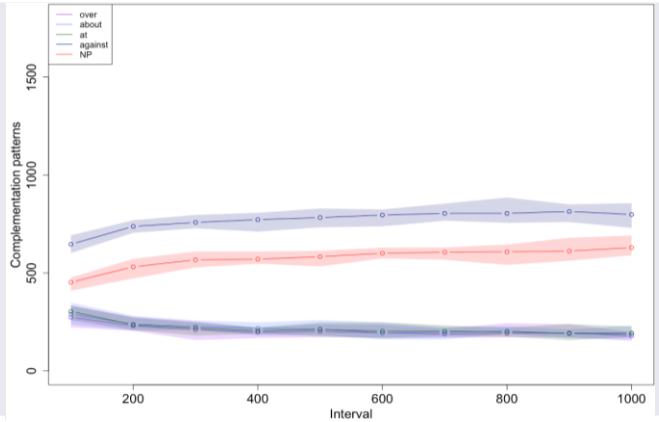
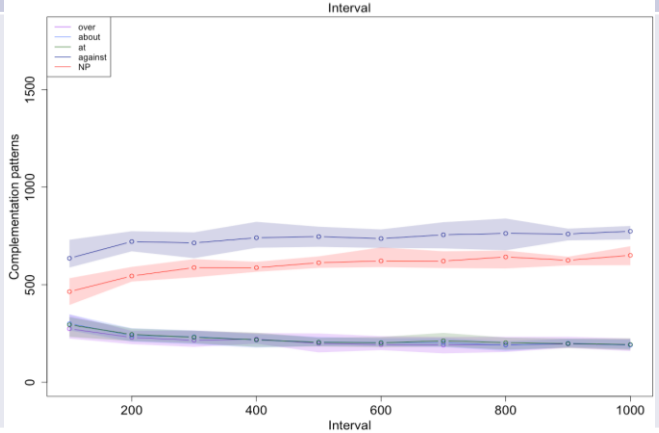
Results of the simulation: avoidance

	IndE Con-American
appeal	
protest	

Results of the simulation: avoidance + ongoing Americanisation

	IndE Con-American
appeal	
protest	

Results of the simulation: condition 2 vs. 3 (con-American IndE)

	Avoidance	Avoidance + Americanisation
appeal		
protest		

Results of the simulation

Condition 1: No ongoing change and stable matrilectal input

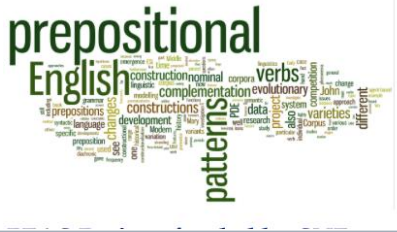
- NPs are preferred and ⬆ over time (PhilE > IndE)
- PPs *(against)* ⬇ for *appeal* and ⬆ for *protest*

Avoidance (con-American Indian agents)

- PPs (*against*) are preferred, but there is a slightly higher ⬆ for NPs

Avoidance + Americanisation (con-American Indian agents)

- *appeal*: PPs (*against*) are preferred, but ⬇ ⇔ NPs ⬆ , leading to convergence
- *protest*: NPs and PPs ⬆, but with a smaller difference in frequency than in condition 2



Interpretation of the ABM

Prediction: Negative attitudes towards AmE decreases epicentral influence for Indian agents

- In all conditions (even strong avoidance of AmE sources), the American variant will become the preferred complementation pattern over time.
- A conservative attitude will not stop ongoing trends of Americanisation

Prediction: Philippine agents further converge on the AmE complementation pattern

- but PhilE closely resembles AmE from the outset – epicentral influence???

Prediction: *protest* has a greater range of PPs => this will slow down epicentral attraction from AmE

- In all PP complements, *against* is the dominant preposition
- The other PPs do not play a discernible role in slowing down the epicentral pull from AmE

Conclusion

- ABM: negative attitudes towards a variety that is assumed to have epicentral ‘pull’ on speakers usage is unlikely to play out as a significant factor in language change, not only under conditions of stable variation but also with agents actively avoiding the pattern they dislike.
- Simulation of the impact that ‘attitude’ would have on epicentral influence is a useful backdrop for the interpretation of corpus data.

References

Bailey, R.W. (1990). 'English at its twilight' In C. Ricks and L. Michaels, eds. The State of the Language. London: Faber & Faber, 83-94.

Hundt, M. (1998). *New Zealand English grammar – Fact or fiction? A corpus-based study in morphosyntactic variation*. Amsterdam: John Benjamins.

Hundt, M. (2013). The diversification of English: Old, new and emerging epicentres. In D. Schreier & M. Hundt (Eds.), *English as a contact language* (pp. 182-203). Cambridge: University Press.

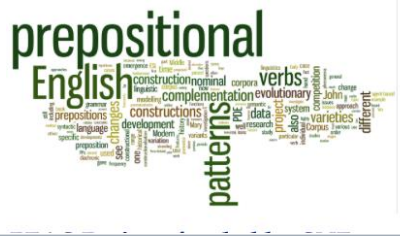
Landsbergen, F. (2009). *Cultural evolutionary modeling of patterns in language change: exercises in evolutionary linguistics*. Utrecht: LOT. Livet, P., Phan, D. & Sanders, L. (2014). Diversity and complementarity of agent-based models in the social sciences. *Revue francaise de sociologie, 55*, 463-500.

Hundt, M. (2013). The diversification of English: Old, new and emerging epicentres. In D. Schreier & M. Hundt (Eds.), *English as a contact language* (pp. 182-203). Cambridge: University Press.

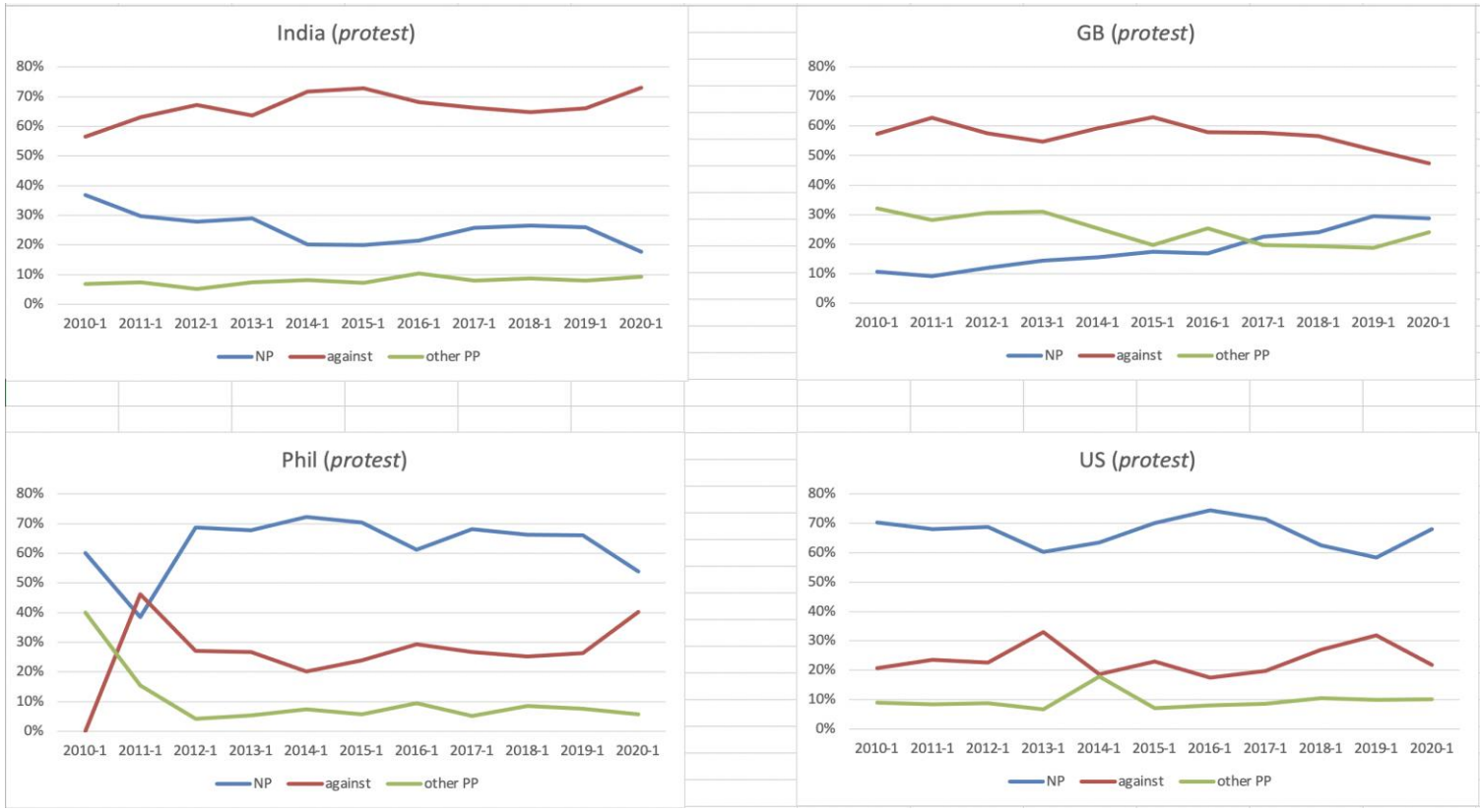
Mair, C. (2013). The world system of Englishes. Accounting for the transnational importance of mobile and mediated vernaculars. *English World-Wide*, 34, 253-278.

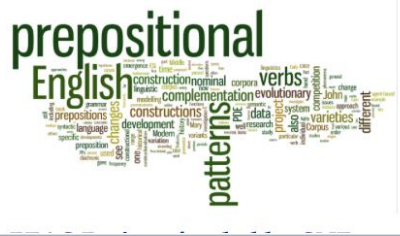
Rohdenburg, G. (2009). Nominal complements. In G. Rohdenburg & J. Schlüter (Eds.), *One language, two grammars? Differences between British and American English* (pp. 194-211). Cambridge: Cambridge University Press.

van Trijp, R. & Steels, L. (2012). Multilevel alignment maintains language systematicity. *Advances in Complex Systems, 15* (3–4).

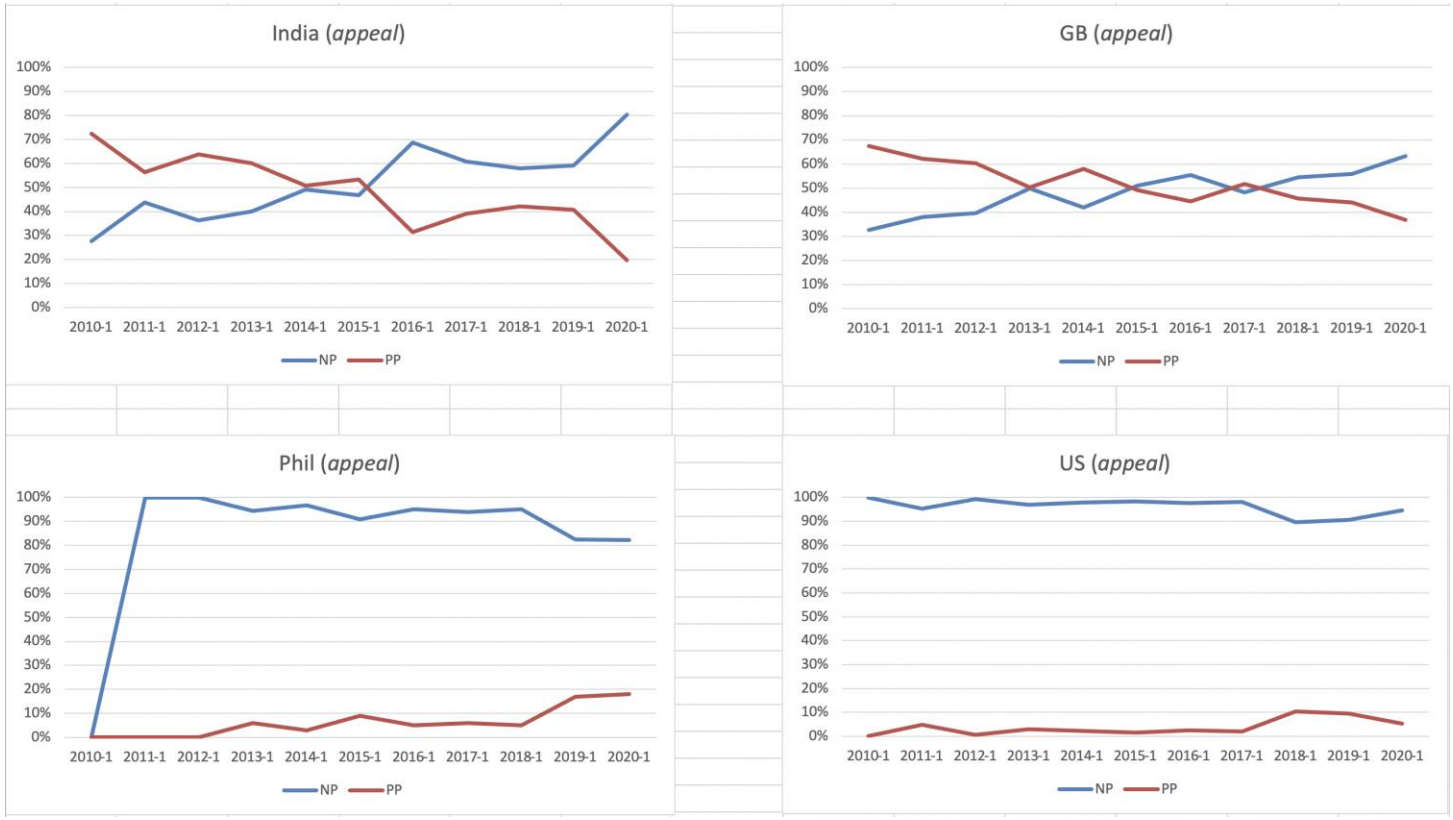


Complementation for *protest* in NOW





Complementation for *appeal* in NOW



Diachronic trajectory for *protest* in AmE

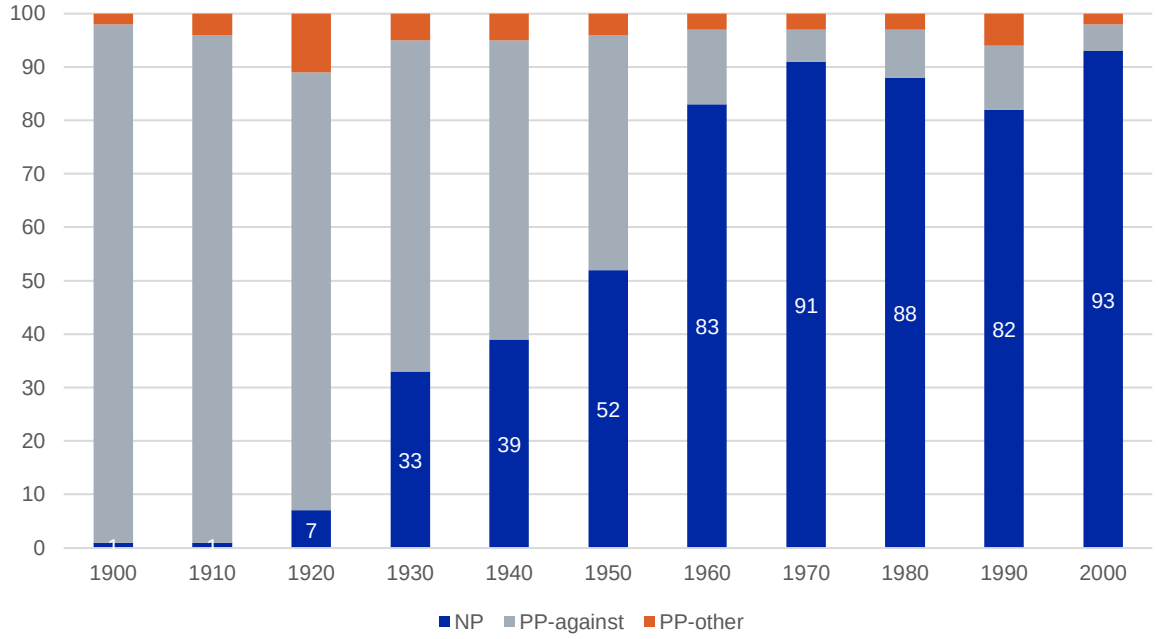


Fig. A-1: Complementation of *protest* in COHA (set of 100 random relevant hits per decade).

Constructional Americanization?

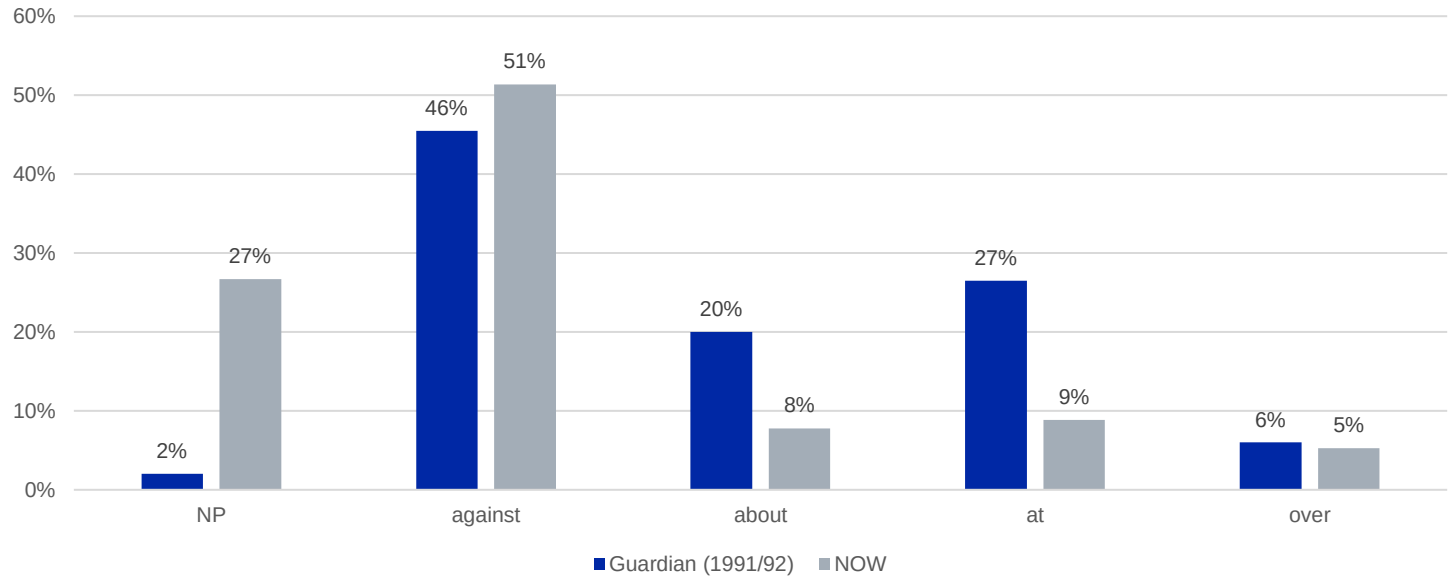


Fig. A-2: Complementation of *protest* in BrE: 1990s *Guardian* (N=400) vs. 2010+ NOW (N=757)**.**

Constructional Americanization?

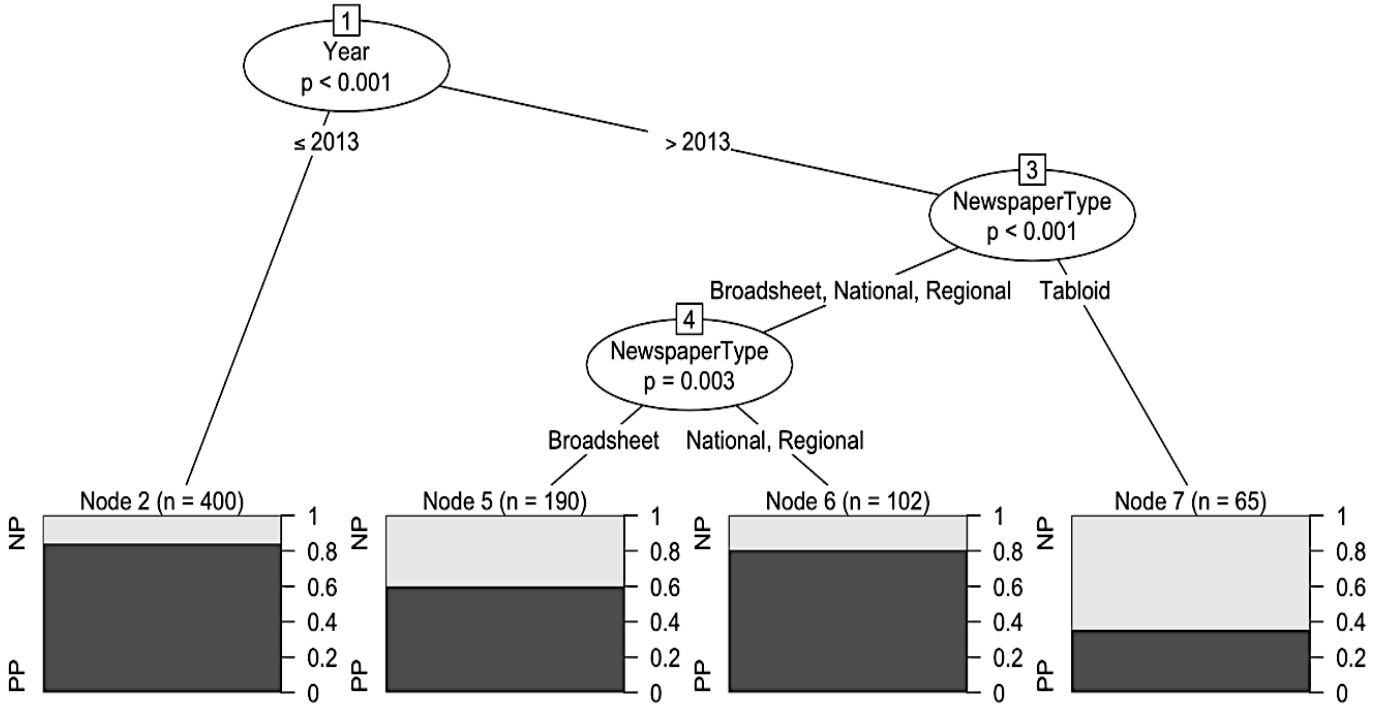


Fig. A-3: Complementation of *protest* in BrE: 1990s *Guardian* vs. 2010+ NOW (Somers2 C = 0.583)**.**

Python code

```
parameter_dict = { 'population_sizes':
                        {'indian': {'con-Am':33, 'neut':34, 'pro-Am':33}, # distributions of attitudes in the population: relatively equal in Indian English, showing effect of recent globalization
                     'philippino': {'con-Am':5, 'neut':45, 'pro-Am':50}}, # same thing, but for historical reasons, the con-American population is really small, the neutral population is higher, because American English is the language of the former colonial power

                                'attitude_factors_dict': {'con-Am':0.8, 'neut':1, 'pro-Am':1.2},

                                'starting_memories': {'indian': {'protest_NP': 1, 'protest_about': 1,'protest_ag':1, 'protest_at':1, 'protest_over':1, 'appeal_NP':1, 'appeal_ag':1},
                                                                 'philippino': {'protest_NP': 1, 'protest_about': 1, 'protest_ag':1, 'protest_at':1, 'protest_over':1, 'appeal_NP':1, 'appeal_ag':1}},

                  'sources_distributions':

                              {'british': OrderedDict([('protest_NP', 15), ('protest_about', 5), ('protest_ag', 70), ('protest_at', 5), ('protest_over', 5), ('appeal_NP', 10), ('appeal_ag', 90)]), # very rough approximation of the distributions based on Hundt 1998
                                          'american': OrderedDict([('protest_NP', 70), ('protest_about', 5), ('protest_ag', 20), ('protest_at', 5), ('protest_over', 5), ('appeal_NP', 90), ('appeal_ag', 10)])},

                  'verb_freq': {'protest': 60, 'appeal': 40}, # Approximation of how often verbs occur (in the simulation: how often an agent needs to express a form of protest or appeal)

                  'influence_from_sources': {'indian': {'con-Am': {'british': 1.8, 'american': 1.2}, ## parameters op 1 zetten om het gewoon te houden
                                                        'neut':{'british': 1, 'american': 2},
                                                        'pro-Am': {'british': 1, 'american': 2}}, # distributions of attitudes in the population: relatively equal in Indian English, showing effect of recent globalization
                                             'philippino': {'con-Am': {'british': 1.8, 'american': 1.2},
                                                            'neut':{'british': 1, 'american': 2},
                                                            'pro-Am': {'british': 1, 'american': 2}}}, # Taken from NOW-corpus as proportion of varieties in online media

                  'peers_vs_sources': {'peers': 7, 'sources': 3},

                                      'british_source_gets_american_input_every': 10,

                  'points_in_time': 1000,
                  'record_every': 100
                  }

batch_id = 'batch-1'
number_of_series = 10

run_batch(batch_id=batch_id, number_of_series=number_of_series, parameter_dict=parameter_dict)
```