ORIGINAL ARTICLE

Epicentral influence via agent-based modelling

Marianne Hundt[1] | **Laetitia Van Driessche[1]** | **Dirk Pijpops[2]**

[1]English Department, University of Zurich, Switzerland

[2]RU Lilith, University of Liège, Belgium

Correspondence
Marianne Hundt, English Department, University of Zurich, Switzerland.
Email: m.hundt@es.uzh.ch

Abstract

One limitation of corpus-based research into the epicentre hypothesis is that it only provides information on structural similarity/difference of varieties but not on the role that attitudes may play in the choice of variants in a pluricentric language. In our case study on verb complementation patterns, we use simulation as a complementary methodology. We build an agent-based model for two speaker communities with English as a second language (ESL) which allows for attitudes towards American and British English to affect the choice of verb complementation. This approach enables us to gauge the likely effect that attitudes towards potential epicentres might have on the choice of variants and to predict patterns of variation we should see in language use. We provide preliminary data from the *News On the Web* corpus to test whether the predictions are, in fact, borne out.

1 | INTRODUCTION

Previous research into the epicentre hypothesis has mostly relied on synchronic corpus data (Hoffmann, Hundt, & Mukherjee, 2011; Gries & Bernaisch, 2016), which only provides information on the relative similarity or difference between regional varieties of English. Even diachronic corpus data would merely supply evidence on convergence, divergence or other patterns of differential change (Hundt, 2009), that is, corpus data alone do not allow us to distinguish between 'ordinary' processes of change and epicentral influence. Thus, one of the methodological challenges that purely corpus-based research is unable to address is the role that attitudes may play in the choice of variants in a pluricentric language like English (Hundt, 2013). Agent-based modelling (ABM) offers a complementary methodology (Livet, Phan, & Sanders, 2014) that can be used to provide corpus linguists with an experimental approach to

predicting variation in a speaker community (Pijpops, forthcoming; Pijpops, Beuls, & Van de Velde, 2015). We build an ABM for two speaker communities where English is an institutionalised second language (ESL), one with American English (AmE), the other with British English (BrE) as their matrilect. Our model allows for attitudes towards the two matrilects to affect the choice of verb complementation (between an NP and a PP complement) for the two verbs *protest* and *appeal*. There is regional variation among national standard varieties of English in the complementation patterns of these two verbs: AmE shows a clear preference for NP complements as in examples (1) and (2), whereas BrE prefers a PP complement as in examples (3) and (4):

(1) Initially, I wore a camouflaged vest and helmet given to me by George Kennedy, the first of the US officials to resign from the State Department *to protest* __ the government's Bosnia policy. (Power, 2019, p. 85.)[1]

(2) He *appealed* __ the ruling, but the conviction was upheld (COCA (*Corpus of Contemporary American English)*, NEWS)

(3) … Britain could protect its own interests within the European Community without having *to protest about* threats to national sovereignty (ICE-GB, S2B-007)

(4) Uh the charterers sought leave *to appeal against* that decision … (ICE-GB, S2A-065)

Preliminary corpus evidence from the *International Corpus of English* (ICE) shows that both patterns are attested in Indian English (IndE) and Philippine English (PhilE), with a slight tendency of IndE to follow in the footsteps of BrE and PhilE to align with AmE, but the figures are too small to allow for any predictions. In fact, *appeal* is not attested with either the NP complement or with *against* in ICE-IND:[2]

(5) Manufacturers in developed countries *have protested against* such imitations … . (ICE-IND, W2E-008)

(6) … they are not able *to protest* __ it (ICE-IND, S1A-057)

(7) She *protested against* the inhumane demolition of squatters' homes in the city. (ICE-PHI, S2B-023)

(8) … their action moves all the fishing communities in the land *to protest* __ their plight. (ICE-PHI, W2A-002)

(9) BIR commissioner Beethoven Rualo last week asked the High Tribunal for a 30-day extension *to appeal* __ an earlier dismissal by the Court of Appeals. (ICE-PHI, W2C-009)

English is a pluricentric language with several national varieties of English, which are codified to different degrees. Among these national varieties, there is a 'pecking order', with AmE and BrE perceived at the top of the hierarchy and other varieties ranking much lower (for example McArthur in Rubdy and Saraceni, 2006). According to Schneider's (2007) dynamic model on the evolution of Postcolonial Englishes (PCEs), emerging national varieties of English continue orienting towards an external or 'exonormative' model, typically the matrilect that served as the main input during colonisation. Exonormative orientation of PCEs continues well into stage three of Schneider's (2007) model, that is, the stage when nativised, local usages emerge. With AmE having superseded BrE in the world hierarchy of Englishes in the twenty-first century (Mair, 2013), it is possible that a variety with BrE as its matrilect may be shifting towards AmE as its new epicentre. Individual speakers are likely to behave differently, though, also depending on differences in exposure and attitudes towards the different epicentre varieties. An important factor in different attitudes towards BrE and AmE is their historical connection with the former colonial powers. For India, these were more negative following independence than for the Philippines. However, multilingualism in both countries and tensions between different ethnicities and social groups have led to a strengthening of the role of English (more generally) in India and a backlash against English in the Philippines (compare Lange, 2020 and Wee, 2020). Moreover, shifting epicentral influence assumes that speakers are aware of regionalisms. We make the assumption that this is the case for the speakers of our case study.[3]

In the following, we review the scant evidence on different complementation patterns of *protest* and *appeal*. We also review previous research into the epicentre hypothesis for English with the aim of providing a backdrop for the choice of methodology adopted in this paper (section 2). The relation between corpus-based research and ABM is addressed
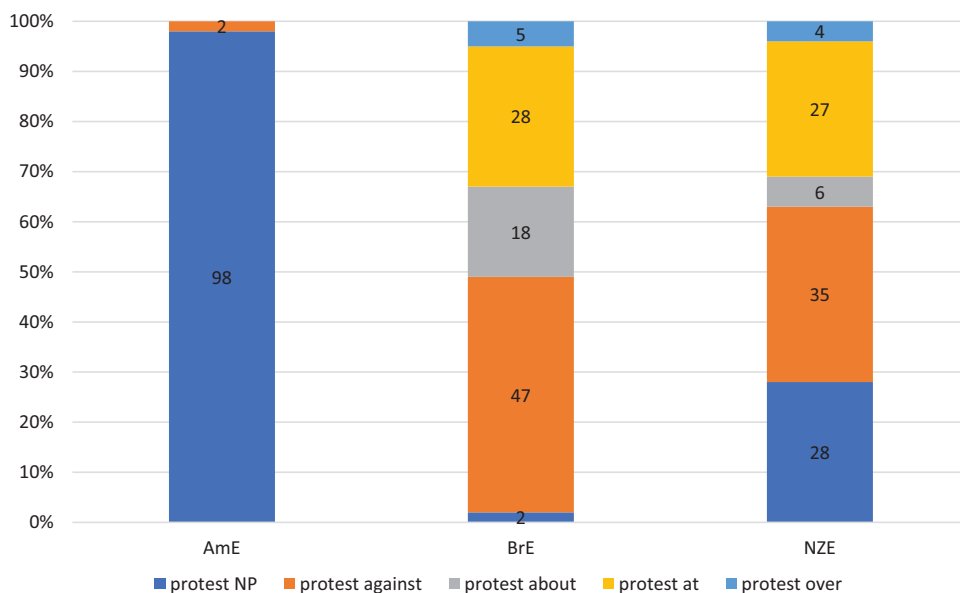
FIGURE 1 Complementation of *protest* in Newspaper databases (based on Hundt, 1998; N = 100 per variety from the *Miami Herald* (US), the *Guardian* (GB) and the *Dominion/Evening* Post (NZ), respectively) [Colour figure can be viewed at wileyonlinelibrary.com]

in section 3. The research questions we aim to address with our simulation and subsequent corpus analysis are presented in section 4, followed by the model design and implementation in section 5, with the results of the simulation discussed in section 6. We provide preliminary corpus evidence in section 7 and discuss how simulation and corpus data relate to each other and which directions further research should take on the basis of this discussion (section 8).

2 | BACKGROUND

2.1 | Previous research into complementation of *protest* and *appeal*

Variation in complementation with an NP vs. PP of verbs like *protest* and *appeal* has received little attention so far. Callies (2018, pp. 151–153) gives a useful summary of studies that have commented on differences in this respect between BrE and AmE and recent change. With the exception of Hundt (1998) and Rohdenburg (2009), these rely on anecdotal evidence and do not provide corpus data on regional variation or change. Hundt (1998, p. 110) uses evidence from newspaper databases for AmE, BrE and New Zealand English (NZE), showing that AmE has a clear preference for the NP variant whereas BrE prefers the PP-complement; NZE falls between the two varieties. The figures for the more frequently attested verb *protest* are given in Figure 1.

According to Gordon and Deverson (1989), attitudes towards American complementation patterns tend to be negative, regardless of whether a bare or a PP-complement is chosen:

> 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, p. 75f.)

In his investigation of a set of antagonistic verbs including *protest* and *appeal*, Rohdenburg (2009, pp. 198–200) in the *Times* and *Washington Times* newspaper archives finds that there is a general tendency for BrE to use prepositional complements, which is particularly marked (that is at or close to 100%) for *protest* and *appeal* (Rohdenburg, 2009, p. 199), whereas AmE tends to opt for the bare NP complement (with similarly extreme tendencies). The clear regional difference is one that Rohdenburg (2009, p. 200) says to have developed through AmE divergence in the twentieth century.

Seeing that complementation patterns in NZE fall between those of the two northern-hemisphere metropolitan standard varieties, the two verbs provide a good candidate to simulate the role that attitudes towards AmE and BrE would play in a case of shifting epicentre influence from BrE towards AmE. Moreover, since it is also possible that BrE might have been shifting towards a greater use of PP complements with the two verbs since the turn of the century, the case study prompts us to simulate possible patterns of epicentral influence, one with a stable situation in BrE and one with ongoing change in the direction of AmE (whether this would be as a result of drift or globalising influence of AmE is not something we will address here). There is no study, to date, on complementation patterns of *appeal* and *protest* in varieties of English as an institutionalised second language (ESL) variety.

2.2 | The epicentre hypothesis

One of the reasons why corpus-based studies that set out to test the epicentre hypothesis have invariably run into difficulties is that corpus data allow researchers to collect evidence on the relative closeness of/distance between varieties, be it among varieties of English as a first language (ENL), between traditional ENL epicentres and ESL varieties (Biewer, 2015) or among varieties of ESL (Hoffmann et al., 2011 or Gries & Bernaisch, 2016 for IndE as a regional epicentre in South Asia). The cautious conclusion of most of this research is that corpus evidence does not allow us to reject the possibility of epicentre influence, but it is not enough to prove it. A further obstacle of most previous studies has been that it is typically based on synchronic corpus evidence, whereas the verification of epicentral influence is likely to require diachronic data. The only diachronic study so far is Peters (2009), who uses evidence from historical dictionaries – notably first attestations of regional vocabulary – to test whether Australian English served as a regional epicentre for New Zealand English.[4] For epicentral influence on ESL varieties, historical evidence is rarely available, and the few diachronic corpora that are being compiled tend to be too small to study lexico-grammatical variation of the kind that we are interested in (see Hundt, 2020, p. 522 for references).

Hundt's (2013) critical assessment of the epicentre hypothesis also concludes that corpus evidence is not enough because, by itself, it does not allow us to verify that speakers in the respective (ESL) countries are shifting their usage preferences under the influence from a different epicentre than the one that previously provided the exonormative model. Hundt (2013, p. 202) points out that 'we lack studies that combine evidence on usage and attitudes towards potential exonormative model varieties'. Just like diachronic corpus studies into ESL varieties, investigation of attitudes towards different Englishes in these countries is still relatively sparse. Moreover, attitudinal research into PCEs faces various difficulties such as the applicability and local meaning of variety terms in direct approaches (Hundt, Zipp, & Huber, 2015) or the need to contextualise the stimuli in indirect approaches (Wilson & Westphal, 2021). In this paper, we therefore do not use empirical data from surveys into attitudes of speakers. Instead, we use simulation to model the potential effect that attitudes to exonormative models (that is, ENL epicentres) would have on lexico-grammatical variation in two PCEs.
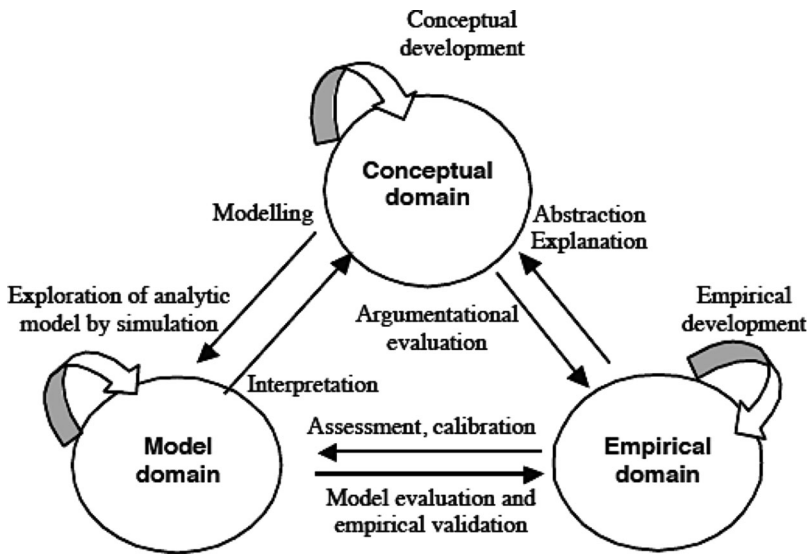
FIGURE 2   Relating simulation to theory and data (from Livet et al., 2014)

3 | AGENT-BASED MODELLING AND CORPUS DATA

ABM does not aim at capturing the actual complexities of language use and speakers' choices in individual communicative situations, as these are also determined by various cultural, historical and individual factors that are beyond the aims of this paper. Doing so would require extremely complex simulation models that rapidly become intractable and uninterpretable – and hence useless (Abrams, 2013, p. 2). Instead, ABMs ought to adhere to Occam's razor in striving to be as simple as possible while still implementing the mechanisms under scrutiny (Landsbergen, 2009, p. 18–19; van Trijp & Steels, 2012, p. 9). Livet et al. (2014, p. 466) visualise the relationship between simulation (the model domain), corpus evidence (the empirical domain) and theoretical model (conceptual domain) as in Figure 2. The theoretical model – in our case the epicentre theory – provides the analytic model that feeds into the building of the simulation model. The three aspects (theory, simulation and data) interact in a knowledge-advancing triangulation process. Essentially, simulation provides a more rigorous (theory-driven) way of predicting behaviour rather than just using verbal argumentation.

4 | RESEARCH QUESTIONS

Our first and main research question concerns the influence of attitudes on epicentre influence (on the assumption that speakers are aware of the regional origin of the variants). We assume that speakers of English in India and the Philippines have acquired their English in the educational context but also to some extent (and increasingly so) through more informal contacts with English via various media, including language use on the internet. The traditional exonormative models or matrilects of English in India and the Philippines are BrE and AmE, respectively. Through online newspapers and forums, speakers in both ESL contexts are exposed to both metropolitan standard varieties. Attitudes towards a growing influence of AmE in India are not uniform, with some speakers resisting it and preferring to adhere to the 'correct' use found in BrE, some being neutral (or unaware of the regional differences) and yet others openly endorsing a shift away from the model of the former colonial power to a variety that is used by speakers of a global economic power. The main research question is whether and how these differing speaker attitudes impact the shift

in epicentral influence from BrE towards greater (or joint) influence of AmE, the assumed 'centre of gravity' in this ongoing change.

We assume that speakers of PhilE will start off from usage patterns that closely align with AmE but are not identical with them. One reason why PhilE speakers might be more conservative would be that PhilE has an older AmE variety from which it developed than the variety currently used in the US, and with change having recently taken place in AmE, PhilE speakers are likely to still use more PP complements (this would then be an instance of post-colonial conservatism). Another reason why PP-complementation of *protest* and *appeal* are likely to be used slightly more often in PhilE than in AmE is that the PP complements are structurally more explicit and transparent and thus cater to a general preference of ESL varieties (Biewer, 2015). This general tendency may, in fact, be an overall conservative influence in both ESL varieties. This could be a potential cause of difference between our ABM, where we abstract away from this factor, and the corpus evidence we aim to compare to our simulation in section 7. We predict that the negative attitudes and resulting conservative behaviour by speakers who are against a growing Americanisation of IndE will significantly decrease epicentral influence in the population. We further predict that speakers of PhilE, whose matrilect is (an older form of) AmE, anyway, will further converge on the AmE model despite (moderate) exposure to the BrE model via media.

Our secondary research question is linked to the fact that the case studies look into lexico-grammatical variation, that is, variation at the interface between grammar and the lexicon. This aspect is potentially relevant for the practical aspects of modelling epicentral influence in our case study because different lexical items may be prone to epicentre influence to different degrees, either because the verb-specific complementation options differ or because the lexical items are attested with different type frequencies, which in turn might influence openness or resistance to epicentral influence. The fact, for instance, that *protest* has more than one PP complement option might slow down epicentral attraction from AmE towards the NP complement. The lower overall text frequency of *appeal*, on the other hand, might prompt speakers to opt for complementation patterns that are frequently available for semantically related verbs.

5 | MODEL DESIGN

5.1 | Components of the model

Our simulation of the impact that attitude has on variant choice (and thus results in epicentral influence) makes the following assumptions:

- There are regional differences in the complementation patterns of *protest* and *appeal* such that AmE has a preference for the bare NP complement whereas BrE prefers the PP complement (Hundt, 1998; Rohdenburg, 2009).
- IndE and PhilE have different matrilects, which means that speakers in the communities start off from different usage preferences that are related to (but not identical with) their matrilects.
- We assume that all speakers are aware of the regional differences between BrE and AmE in the complementation of the two verbs. Speakers in the ESL communities differ with respect to their attitudes towards AmE: some speakers are open to influence from AmE (pro-Am), some are opposed to it (con-Am) and some are neutral.
- Speakers are exposed to both BrE and AmE variants through media (news on the web). We take the differences in size of the two regional sub-corpora in *News On the Web* (NOW) corpus as suggestive, that is, ESL readers are twice as likely to encounter an AmE complementation pattern than a BrE complementation pattern in their reading.

We build three different models to simulate potential influence of attitude on variant choice that assume the following conditions:

Condition 1: There is no ongoing language change and the input from the matrilects therefore 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.

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.

We predict that under condition 1, there will be relatively slow change in speakers of IndE towards the AmE complementation patterns, as the impact of pro-Am speakers is counter-balanced by con-Am and neutral speakers. Under condition 2, the epicentral influence of AmE will be counterbalanced by the con-Am speakers actively avoiding American input. In the third scenario (condition 3), the 'cushioning' effect of the con-Am speakers present in condition 2 will be neutralised by an increasingly Americanised BrE input, so convergence on the AmE complementation pattern should be most marked under condition 3. Under all conditions, PhilE speakers will show an overall higher (and increasing) preference of AmE complementation patterns than IndE speakers, due to the different matrilectal origins.

5.2 | The implementational level of model design

The simulation is implemented in Python code. It consists of two communities of agents, the Indians and the Filipinos, with a respective population size of i and f. Each agent retains a memory of language forms (this implements component 2 of the model), with for each language form an associated entrenchment score e and an attitude factor a. There are two verbs in the simulation, namely *appeal* and *protest*, that may take two types of complement, a bare NP or a prepositional variant. The prepositional variant of *appeal* always contains the preposition *against*, whereas the prepositional variant of *protest* may contain any of four prepositions, namely *about, against, at* or *over*. This yields seven language forms in total, namely *appeal* NP, *appeal against* NP, *protest* NP, *protest about* NP, *protest against* NP, *protest at* NP and *protest over* NP. One assumption the simulation makes is that the agents are aware of the bare NP variant being typical of AmE; another assumption is that agents have an attitude towards AmE. For agents with a negative attitude towards AmE, the attitude factor for both *appeal* NP and *protest* NP, is set to a_{neg}, with $a_{neg} < 1$. For agents with a positive attitude towards AmE, the attitude factor for the bare language forms is set to a_{pos}, with $a_{pos} > 1$ and for agents with a neutral attitude it is set to 1. All attitude factors for the other language forms are always set to 1. The attitude factors do not change during a simulational run. There are i_{neg} Indian agents that have a negative attitude towards AmE, i_{pos} Indian agents that have a positive attitude towards AmE, and $(i - i_{neg} - i_{pos})$ Indian agents that have a neutral attitude towards AmE. For the Filipino agents, the respective values are f_{neg}, f_{pos} and $(f - f_{neg} - f_{pos})$. (This implements components 1 and 3, section 5.1). Finally, the simulation contains two media sources, one American and one British. Each media source has a frequency distribution over the seven language forms. This distribution contains frequency counts c for each language form. The distribution determines how often the media source produces each language form (see below, Equation 2). (This implements component 4, see section 5.1)

During each simulational run, t consecutive points in time are executed. Each point in time, i interactions are executed between Indian agents, and f interactions between Filipino agents. For each interaction, a speaker agent and a hearer agent is randomly selected from the community at hand. The speaker agent is then set to produce a form of either *protest* or *appeal*, with a probability v that it will produce a form of *appeal*, and a probability $(1 - v)$ it will produce a form of *protest*. This probability v indicates how relatively often language users use the verb *protest* and *appeal*, which is assumed to be consistent across the Indian and Filipino communities. The speaker agent then produces a language form l of the selected verb with probability p_l, which is calculated according to Equation (1), based on the entrenchment scores e and the attitude factors a of all n language forms of the selected verb that are found in the memory of the speaker agent. Once the speaker agent has produced a language form, the hearer agent reacts by adding 1 to the

entrenchment score of that form in its memory:

$$p_l = \frac{a_l e_l}{\sum_{i=1}^{n} (a_i e_i)} \, ,$$ (1)

Equation 1: where probability p_l with which a speaker agent produces a language form l of the set verb, based on the attitude factors a and the entrenchment scores e in its memory.

After each interaction, there is a probability m that one of the agents of the community at issue receives linguistic input from a media source. If that occurs, an agent is randomly selected as a hearer agent from the community at issue. Next, the probability s that the American source is selected is dependent upon the community of the hearer agent in Condition 1. Meanwhile, in Condition 2, it is both dependent upon the community and the attitude of the hearer agent. This introduces six parameters, namely $s_{i,pos}$, $s_{i,neu}$ and $s_{i,neg}$ for the Indian agents, and $s_{f,pos}$, $s_{f,neu}$ and $s_{f,neg}$ for the Filipino agents. In Condition 1, the following always hold: $s_{i,pos} = s_{i,neu} = s_{i,neg}$ and $s_{f,pos} = s_{f,neu} = s_{f,neg}$. In Condition 2, the following always hold: $s_{i,pos} \geq s_{i,neu} \geq s_{i,neg}$ and $s_{f,pos} \geq s_{f,neu} \geq s_{f,neg}$. When the American media source is not selected, the British media source is used.

The selected media source will then produce a language form s with probability p_s, which is calculated according to Equation 2, based on the frequency counts c of the seven forms in their distribution in the media source. Once the media source has produced a language form, the hearer agent reacts by adding 1 to the entrenchment score of that form in its memory.

$$p_s = \frac{c_s}{\sum_{i=1}^{7} c_i} \, ,$$ (A2)

Equation 2: where probability p_s with which a media source produces a language form s based on the frequency counts c in the frequency distribution of that media source.

This is how the simulation functions in the first two conditions. In these conditions, the frequency distributions of the media sources do not change during a single simulational run. In the third condition, however, the distribution of the British media source slowly changes towards the distribution of the American media source. This is implemented as follows. Every c point in time, the American media source will produce a language form according to Equation (2). The British media source will then add 1 to the frequency count of that language form in its distribution. In this way, the distribution of the British media source will slowly evolve towards the distribution of the American source, but never become exactly identical to it.

When executing a simulational run, we keep track of how often each language form is produced in each community during certain time intervals. Because the simulation makes intense use of probabilities, the results of all simulational runs will be slightly different, even when executed with the exact same parameter settings. We therefore always execute a batch of 10 simulational runs with the exact same parameter settings, and plot the means, minima and maxima over those 10 simulational runs.

6 | RESULTS OF THE SIMULATION

The following sections present the results of the ABM for a stable situation (6.1), a situation where the con-Am agents actively avoid the American complementation pattern (6.2) and one that combines this avoidance with a situation where the BrE media sources become more similar to the American sources, that is, have a higher prevalence of NP complementation (6.3).
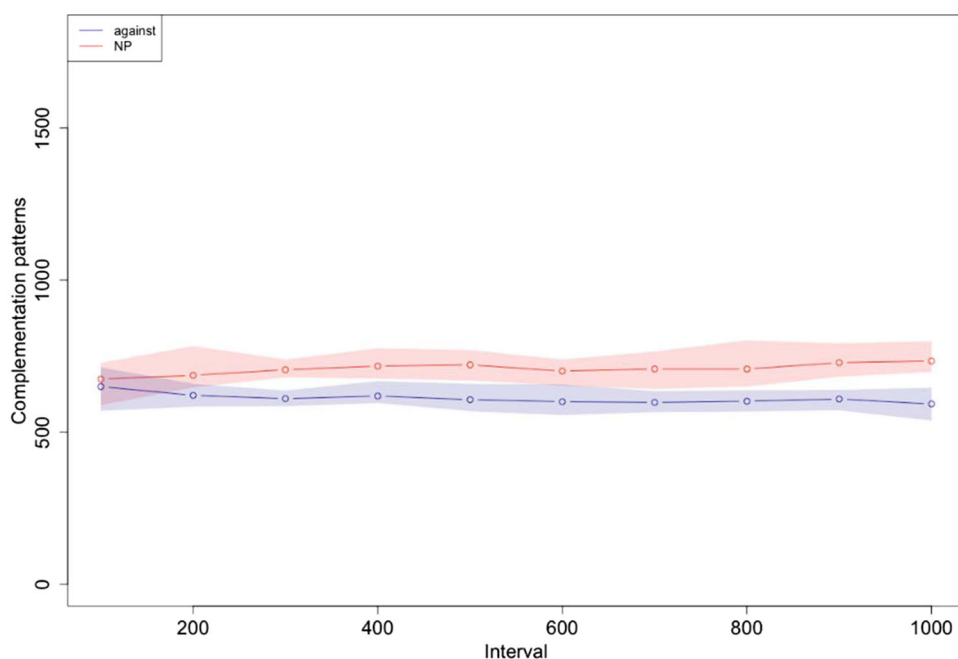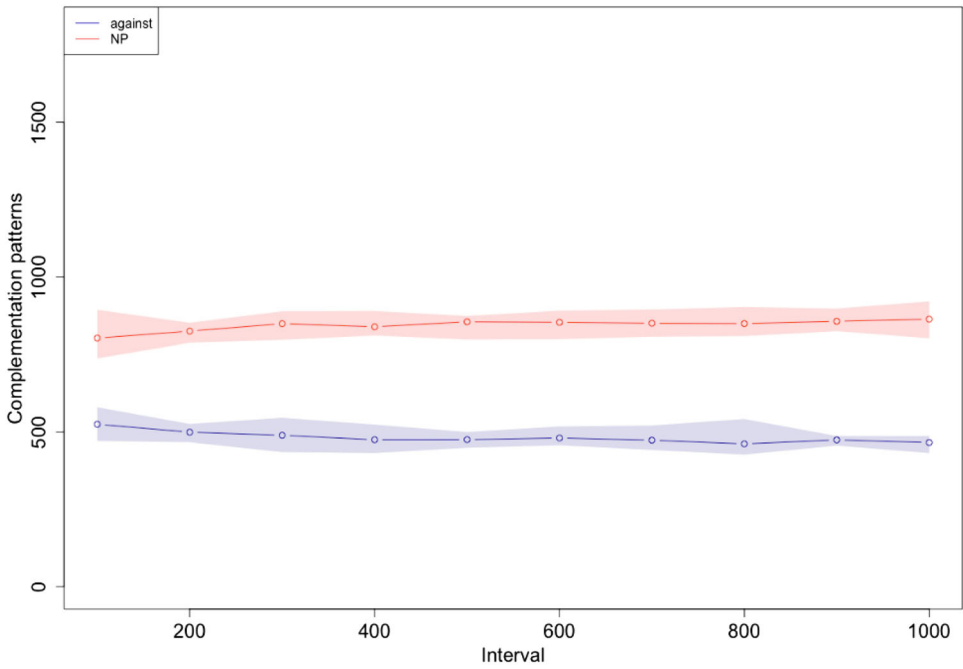
FIGURE 3 Complementation patterns for *appeal* (con-Am Indian agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]

6.1 | Condition 1: Stable matrilectal input[5]

This section presents the results for the model in the stable condition. The *x*-axis marks the development over 1000 points in time, while the *y*-axis shows the prevalence of the different complementation patterns during this time. The lines represent the mean over 10 simulational runs, with the corresponding colour surrounding the line presenting the minima and the maxima for these figures. Figures 3 and 4 show the complementation patterns for *appeal* for the con-Am and pro-Am Indian agents, respectively. For both populations, the (American) NP variant is the preferred option and increases over time. The prepositional variant (*against*), on the other hand, decreases over time for both populations. The difference between the two populations is that the divergence is far clearer for the pro-Am agents: they start out and end with a higher frequency for the NP variant and a lower frequency for the prepositional variant. The neutral agents show the same tendency as the pro-Am agents albeit to a lesser degree (Figure 5 and Figure A in the Appendix).[6]

A similar tendency can be seen for the verb *protest* (Figures 6 and 7). The prepositional variants *over*, *about* and *at* are infrequent and decrease even further over time for both populations. The competition is mostly between *against* and the NP variant, which are used similarly frequently in the con-Am population, compared to the clear preference for the NP complementation in the pro-Am (and the neutral[7]) Indian populations. The difference with *appeal* is that complementation with *against* is also on the rise for *protest*, albeit less prominently so than for the NP variant. The prepositional variant *against* thus encroaches on the domain of the other prepositional variants, which decrease over time.

For the Filipino agents, the NP variant becomes even more preferred over time than in the Indian population for both *appeal* (Figure 8) and *protest* (Figure 9). As was the case for the Indian population, *appeal against* decreases over time whereas *protest against* increases slightly.[8]
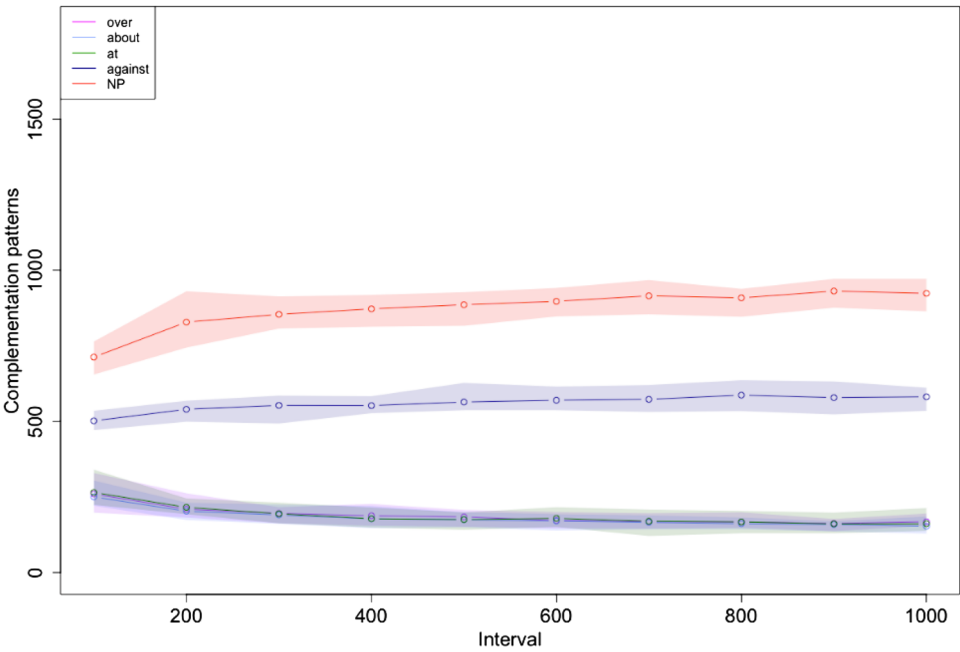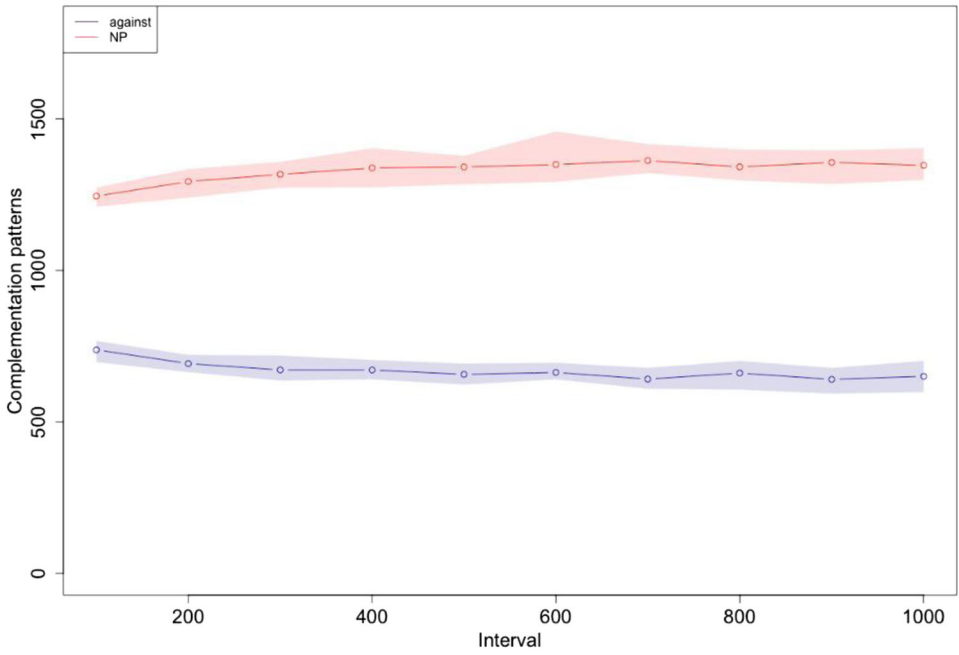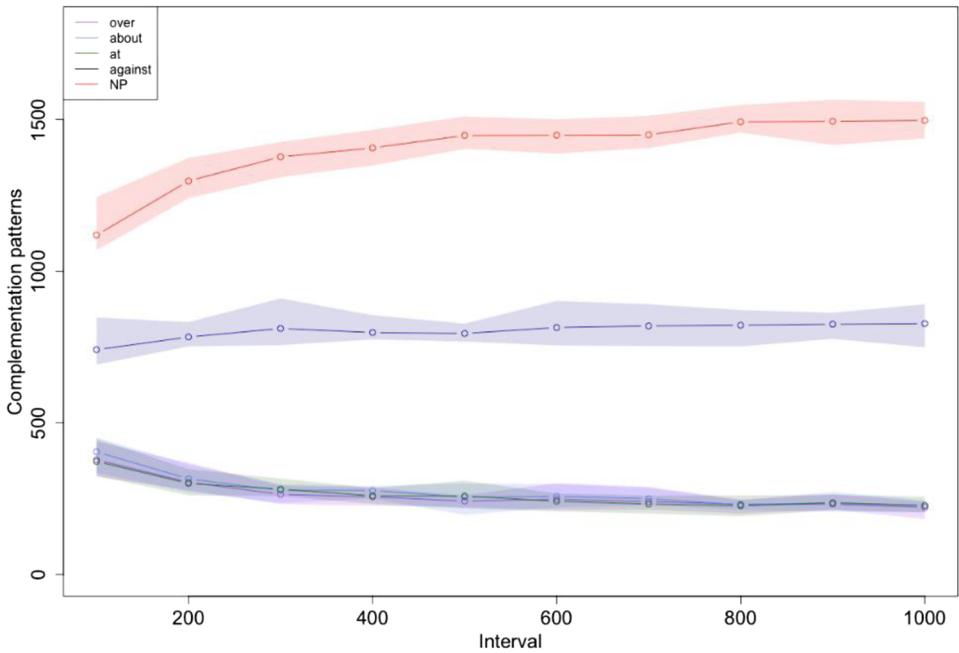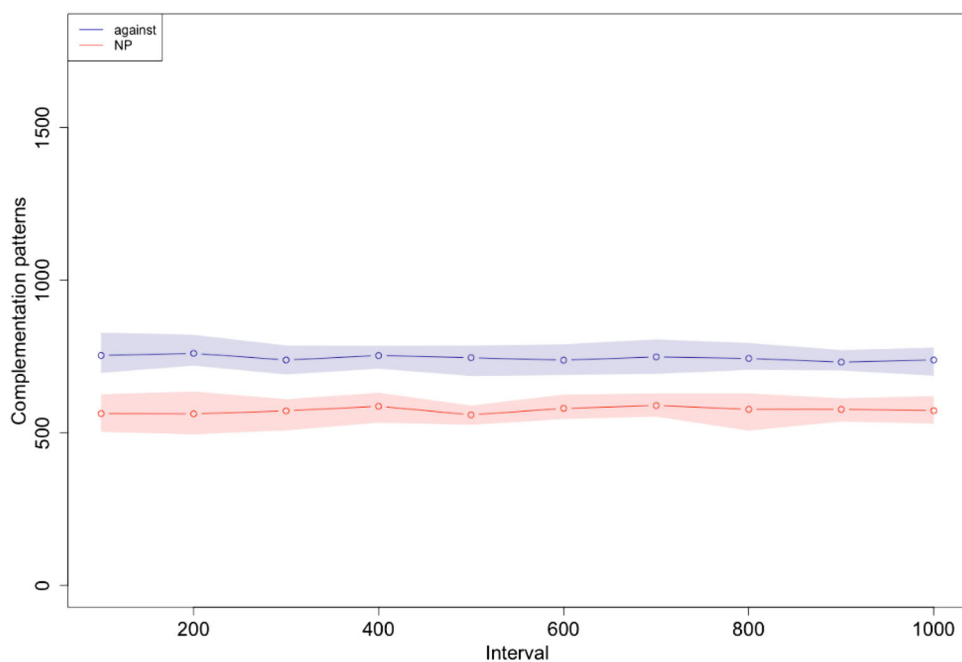
FIGURE 4 Complementation patterns for *appeal* (pro-Am Indian agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]
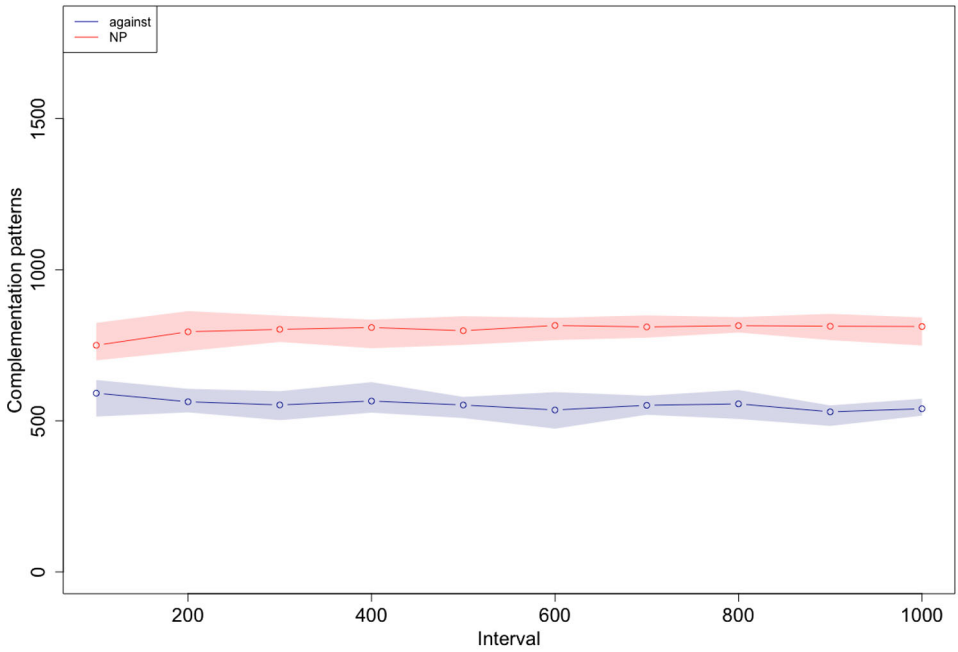
FIGURE 5 Complementation patterns of appeal (neutral Indian agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]
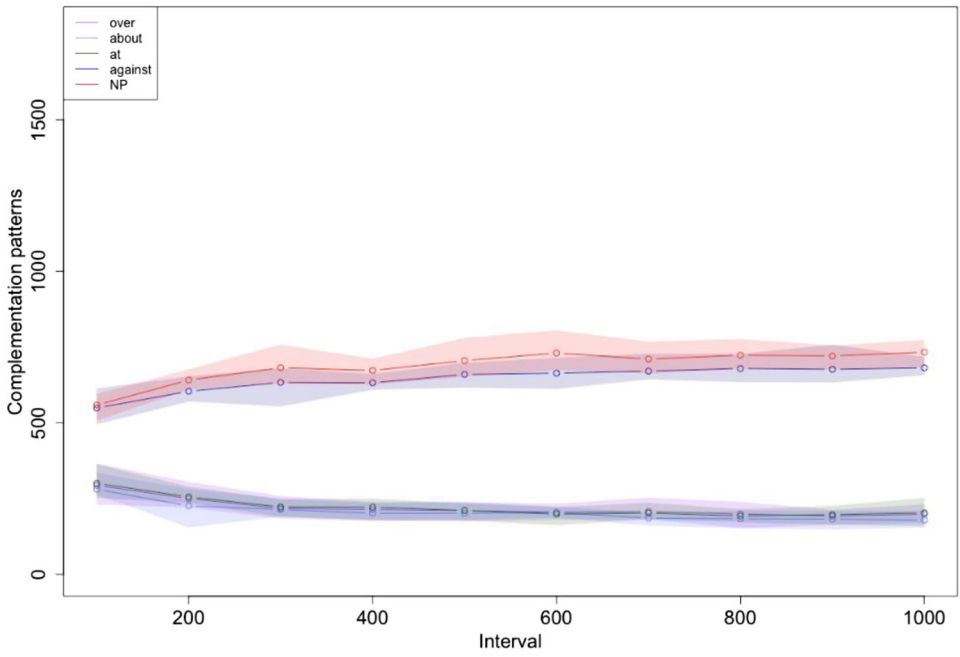
FIGURE 6 Complementation patterns for *protest* (<u>con-</u>Am Indian agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 7 Complementation patterns for *protest* (<u>pro-</u>Am Indian agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 8 Complementation patterns for *appeal* (pro-Am Filipino agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 9 Complementation patterns for *protest* (pro-Am Filipino agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 10   Complementation patterns for *appeal* (con-Am Indian agents, strong avoidance) [Colour figure can be viewed at wileyonlinelibrary.com]

6.2 | Condition 2: Active avoidance by the con-American agents

This section looks at active avoidance on the part of the speakers who hold negative attitudes towards AmE. The alteration to the model only changes the input from the British and American media sources and not the interaction between the agents. In the previous conditions, the agents were more likely to read or listen to an American media source than a British one, which mirrors the distribution in the NOW corpus, where most of the online content is American. In this version of the model, however, the agents with a negative attitude towards AmE get less input from the American source, simulating a situation in real life where such language users actively avoid American newspapers or websites. We first modelled a situation where the American content was still more frequent but this did not show a substantial difference from the stable situation in Condition 1. We therefore present the version where the agents are so opposed to American English that they actively seek out – and thus have more input from – British than American content despite the latter's higher online presence.[9] In the following, we focus on the Indian agents as the con-Am Filipino agents are very low in number in the model.[10]

Figures 10 and 11 give the results of Condition 2 under 'strong avoidance' for the verbs *appeal* and *protest*. In this condition, the PP variant with *against* is more frequent than the NP variant for the entire duration of the model. As complementation of *appeal*, both variants fluctuate between slight increases and decreases during the model. However, the general trend is for a very slight increase of the NP variant and decrease of *against*. Under condition of 'strong avoidance', the PP variant with *against* is the most frequent complement for *protest*, too, but both the NP variant and *against* are on the rise for this verb whereas other prepositions decrease. However, the NP variant rises slightly more than the PP complement with *against*.

6.3 | Condition 3: Ongoing Americanisation[11]

This section presents the results for the model where the con-Am speakers actively avoid American English input and the British media source becomes more like the American source (starts using more NP complements) over time. As
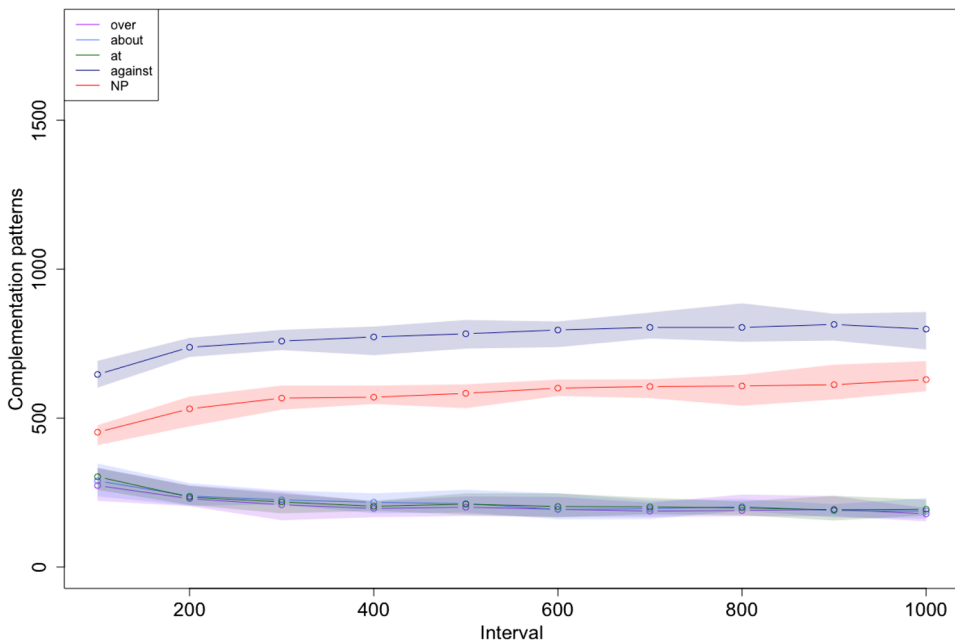
FIGURE 11 Complementation patterns for *protest* (<u>con-</u>Am Indian agents, strong avoidance) [Colour figure can be viewed at wileyonlinelibrary.com]

American input in the model with the Filipino agents is relatively strong to begin with, Americanisation predictively does not show a strong effect[12] and is therefore not discussed in detail.

For the use of *appeal* by con-Am Indian agents, there was a decrease for the prepositional variant (*against*) and an increase for the NPs in the stable condition; as the NP was the preferred variant, the graph showed divergence between the two variants (Figure 3). However, if the speakers actively avoid the American content to such an extent that BrE becomes more prevalent for these agents, the two complementation patterns show convergence (Figure 12).

For *protest*, both the prepositional variant and the NP variant increase when BrE becomes the more frequent content for the con-Am speakers (active avoidance). However, the difference in frequency between the two alternatives is smaller than without the influence of Americanisation on the British source (Figure 11), as Figure 13 shows.

6.4 | Interpretation of the model

Our initial prediction was that negative attitudes and resulting conservative behaviour by speakers who are against a growing Americanisation of IndE would significantly decrease epicentral influence in the population. However, our ABMs have shown that speakers with a negative attitude towards epicentral influence from AmE would have a hard time to counteract ongoing trends in the community. This was evident not only under Conditions 1 and 3 (stable input and ongoing Americanisation) but also under Condition 2, where con-Am Indians actively avoided AmE input and sought out BrE reading materials. Thus, even in the case of strong avoidance, the American variant might become the preferred complementation pattern over time as the NP rises slightly more strongly than the PP variant for *protest*, and the latter seems to decrease slightly for *appeal* (see Figures 10 and 11). With respect to speakers of PhilE our prediction was that they would further converge on the AmE model despite (moderate) exposure to the BrE model via media. As it turns out, the influence from BrE via the media or through con-Am agents is so minimal under all three conditions that it does not have a noticeable effect in our model at all. In other words, PhilE cannot be affected by
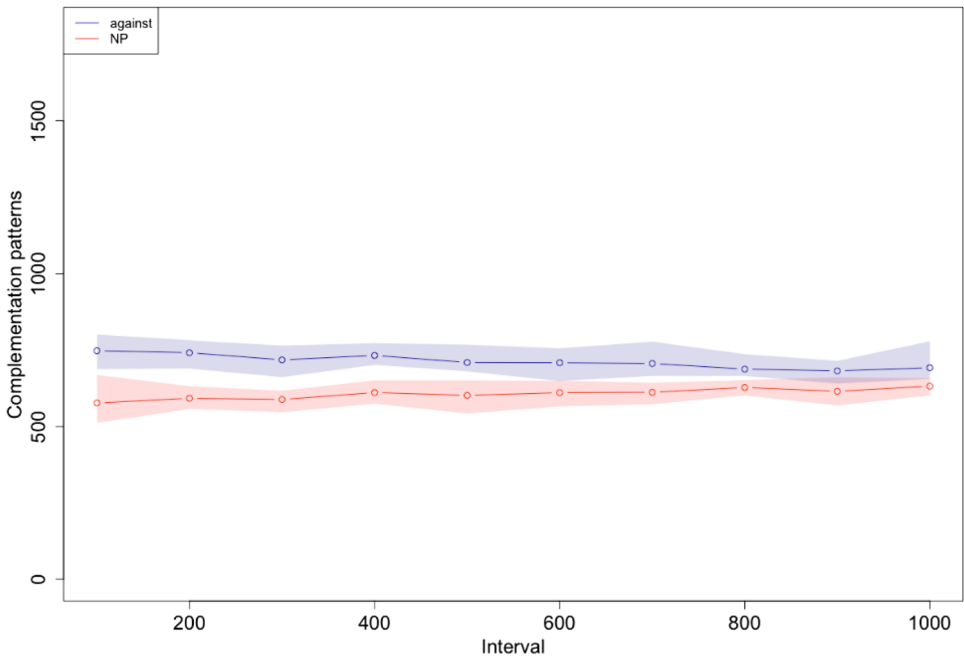
FIGURE 12 Complementation patterns for *appeal* (con-Am Indian agents; Americanisation and strong active avoidance) [Colour figure can be viewed at wileyonlinelibrary.com]
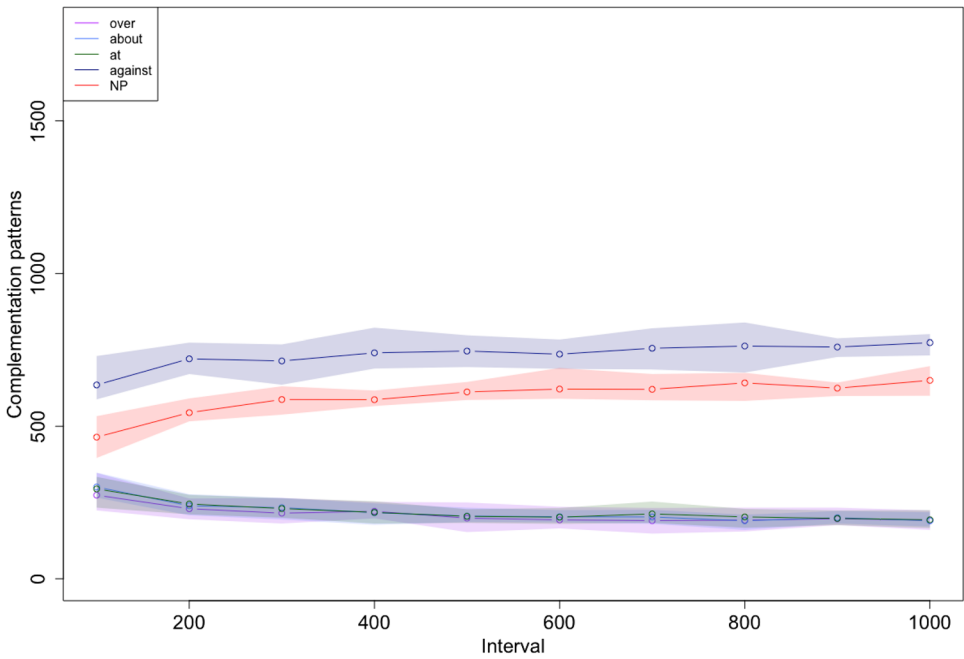
FIGURE 13 Complementation patterns for *protest* (con-Am Indian agents, Americanisation and strong active avoidance) [Colour figure can be viewed at wileyonlinelibrary.com]

epicentral influence because it already is modelled on AmE. As far as differences between the two verbs is concerned, we expected the greater range of PP complements for *protest* to slow down epicentral attraction from AmE towards the NP complement. However, since *against* is the dominant preposition, other PPs did not play a discernible role in slowing down epicentral pull from AmE towards the NP in any of the conditions for IndE.

7 | CORPUS EVIDENCE

In the following, we will compare the outcomes of our ABM with recent micro-changes in a large corpus of online newswriting, that is, a subset of the kind of input data that we were assuming to play a role in the (changing) exonormative orientation for the two ESL varieties we include in our case study. Before presenting the results of the corpus study, we briefly comment on the choice of corpus and mode of data extraction.

7.1 | Methodology

As pointed out in the introduction, studies on lexico-grammatical phenomena typically require large enough datasets. For long-term studies, these are only available for AmE and BrE so far. Our study, therefore, is limited to the investigation of micro-changes, that is, more short-term developments within a generation of users. A suitably large corpus for world Englishes research is the online web-based NOW corpus. The earliest data available in NOW date from 2010, the most recent are from 2021 at the time of writing. This means that only very recent change can be studied on the basis of this corpus, but it provides sufficiently large amounts of data for the four varieties under investigation.[13] Part-of-speech (POS) annotation of the NOW corpus allows us to target the variable patterns we are interested in (thus potentially increasing precision albeit at the expense of recall). Our retrieval algorithm includes all forms of the respective verbs followed immediately either by the definite or indefinite article (NP variant) or by the respective preposition(s).

7.2 | Micro-changes in the NOW corpus

With respect to the possibility of ongoing change the data from NOW in Figure 14 indicate that epicentral influence from AmE on the two ESL varieties in the complementation of *protest* is an unlikely scenario. In fact, PhilE complementation patterns are diverging from the matrilect and moving towards a relatively equal distribution of bare NP complement and the PP complement with *against*. Similarly, divergence is also what we can observe in IndE newswriting, with a recent tendency showing complementation with *against* on the rise and the NP complement in decline. If anything, it is BrE that is (slowly) converging on the AmE model: the NP complement is gaining at the expense of PP variants, including the dominant one with *against*. The situation is quite different for *appeal* (Figure 15), where we see the expected (ongoing) convergence of both IndE and BrE on the complementation patterns that are dominant in AmE and PhilE. Moreover, IndE is much more advanced in this ongoing change than BrE. To sum up, the data from NOW on micro changes in complementation patterns are very different for the two verbs: IndE and PhilE (but not BrE) are resisting Americanisation in the complementation patterns of the frequently used verb *protest*; for the overall more infrequent verb *appeal*, micro changes in the NOW data show the expected pattern of epicentral convergence on AmE complementation patterns of IndE and its matrilect.
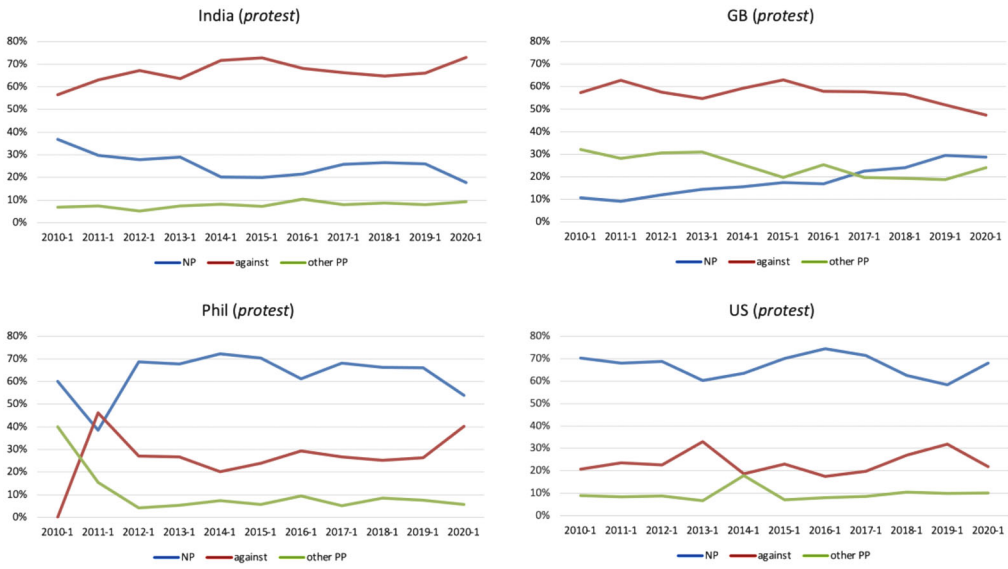
FIGURE 14 Complementation patterns for *protest* across subsections in NOW [Colour figure can be viewed at wileyonlinelibrary.com]
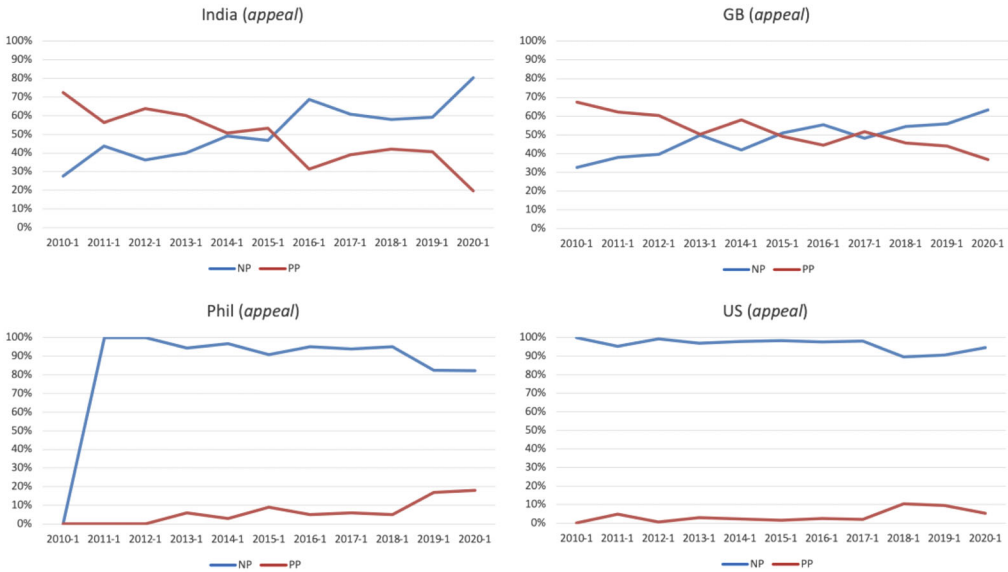
FIGURE 15 Complementation patterns for *appeal* across subsections in NOW [Colour figure can be viewed at wileyonlinelibrary.com]

8 | CONCLUSION

Our ABMs showed that 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. The results of the simulation are therefore similar in outcome to what Anderwald (2016) found for the negligible overall long-term effect of prescriptive grammars on speakers' usage. Simulation of the impact that 'attitude' would have on epicentral influence is a useful backdrop for

the interpretation of corpus data. Having established that con-Am attitudes get drowned out by pressure from input and general drift towards AmE complementation patterns in our model, we stand on safer grounds to postulate that the diverging trends we see in the complementation of *protest* in IndE and PhilE are unlikely to be the result of negative attitudes towards AmE as an epicentre. We pointed out in section 4 that our ABM abstracts away from other factors that might be at play in the ESL varieties, notably the preference for structurally more explicit patterns (Biewer, 2015). It might be the case that this factor could play a more important role for frequent verbs (such as *protest*). This aspect could be tested in further corpus-based studies or controlled experimentation. Future research could supplement our findings with actual attitude data in the respective communities and evidence on how aware speakers really are of the regional differences in complementation.

NOTES

[1] Samantha Power is an American national with an Irish background, who emigrated to the United States as a young teenager (Power, 2019). Emphasis in this and all other examples has been added.

[2] The verb is mostly used in the sense of appealing to an institution or a group of people for something, as in 'The Rajya Sabha *has appealed to* the oil sector not to go on strike today' (ICE-IND, S2b-010), that is, not as the 'antagonistic' verb that is the focus of the present study (Rohdenburg, 2009, p. 198).

[3] Possible sources of speaker awareness of regionalisms are online meta-comments in grammar forums or in-house style guides of newspapers. Whether these are available for the complementation patterns of *protest* and *appeal* is difficult to verify as the former are hard to find/target and the latter are rarely available.

[4] The specific problem for this study is that there are close historical connections between New Zealand and Australian English and that speakers moved relatively freely between the two countries, so that it remains a moot point whether NZE actually oriented towards its close relative across the Tasman or whether the regional vocabulary first attested in Australia might not simply have been brought to New Zealand by speakers who took that migration route.

[5] The population sizes (i, f) were 100 speakers for both populations, with $i_{neg} = 33$, $i_{pos} = 33$, $(i - i_{neg} - i_{pos}) = 34$, $f_{neg} = 5$, $f_{pos} = 50$ and $(f - f_{neg} - f_{pos}) = 45$. The attitudes are regulated with an attitude factor a, which is 0.8 for s_{neg}, 1 for S_{neu} and 1.2 for S_{pos}. The frequency distribution for the sources (c) is *protest* NP = 15, *protest about* = 5, *protest against* = 70, *protest at* = 5, *protest over* = 5, *appeal* NP = 10, *appeal against* = 90 for the British source and *protest* NP = 70, *protest about* = 5, *protest against* = 20, *protest at* = 5, *protest over* = 5, *appeal* NP = 90 and *appeal against* = 10 for the American source. The parameter that regulates the input that the speakers get from the sources is set to 1 for British and 2 for American across all attitudes. The parameter that regulates the likelihood that a speaker will interact with another speaker, as opposed to reading or hearing a source (probability m), is set at 7 for the peers and 3 for the sources. The starting memory for the speakers is 1 for all language forms (l) for both populations. The verb frequency (v) is set at 60 for *protest* and 40 for *appeal*. The simulation lasts for 1000 points in time (t); there are 100 interactions for each point in time. We ran one batch with 10 series for each condition.

[6] The graphs for the populations and verbs that are not mentioned in this section can be found in the Appendix.

[7] See Figure B in the Appendix.

[8] See Figures C and D in the Appendix for the con-American Filipino agents, and Figures E and F for the neutral Filipino agents.

[9] The parameter settings for this condition are the same as for the previous one, apart from the parameter that regulates the input that the speakers get from the sources, which is changed for the con-Am speakers. For the 'weaker' type of avoidance, where American English stays the more prevalent source, the values are changed to 1.2 for the British and 1.8 for the American source (compared to 1 and 2 for the original condition, respectively). For the stronger type of avoidance, the values are changed to 1.8 for the British and 1.2 for the American source.

[10] See Figures G and H in the Appendix.

[11] The parameter settings are the same as in the previous condition, apart from the parameter that regulates when the British source gets input from the American source, which was set to 10 (points in time). In the original model, the value for this factor is set to a very high number, so that the British source does not get input from the American source in that condition.

[12] See Figures I and J in the Appendix.

[13] Note that the searches do not include the second half of the 2020s as complex searches by subsection (US, GB and so on) is no longer enabled in NOW.

REFERENCES

Abrams, M. (2013). A moderate role for cognitive models in agent-based modeling of cultural change. *Complex Adaptive Systems Modeling, 1*, 16.

Anderwald, L. (2016). *Language between description and prescription*. Oxford: Oxford University Press.

Beuls, K., & Steels, L. (2013). Agent-based models of strategies for the emergence and evolution of grammatical agreement. *PLoS ONE, 8*, e58960.

Biewer, C. (2015). *South Pacific Englishes. A sociolinguistic and morphosyntactic profile of Fiji English, Samoan English and Cook Islands English*. Amsterdam: John Benjamins.

Callies, M. (2018). Patterns of direct transitivization and differences between British and American English. In M. Kaunisto, M. Höglund, & P. Rickmann (Eds.), *Changing structures: Studies in constructions and complementation* (pp. 151–167). Amsterdam: John Benjamins.

Gordon, E., & Deverson, T. (1989). *Finding a New Zealand voice*. Auckland: New House.

Gries, S. Th., & Bernaisch, T. (2016). Exploring epicentres empirically: Focus on South-east Asia. *English World-Wide, 37*, 1–25.

Hoffmann, S., Hundt, M., & Mukherjee, J. (2011). Indian English – an emerging epicentre? A pilot study on light verbs in web-derived corpora of South Asian Englishes. *Anglia, 129*, 258–280.

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

Hundt, M. (2009). Colonial lag, colonial innovation, or simply language change? In G. Rohdenburg & J. Schlüter (Eds.), *One language, two grammars: Morphosyntactic differences between British and American English* (pp. 13–37). Cambridge: University Press.

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: Cambridge University Press.

Hundt, M. (2020). Corpus-based approaches to world Englishes. In D. Schreier, M. Hundt, & E. W. Schneider (Eds.), *The Cambridge handbook of world Englishes* (pp. 506–533). Cambridge: Cambridge University Press.

Hundt, M., Zipp, L., & Huber, A. (2015). Attitudes in Fiji towards varieties of English. *World Englishes, 34*, 688–707.

Landsbergen, F. (2009). *Cultural evolutionary modeling of patterns in language change: Exercises in evolutionary linguistics*. Utrecht: LOT.

Lange, C. (2020). English in South Asia. In D. Schreier, M. Hundt & E. W. Schneider (Eds.), *The Cambridge handbook of world Englishes* (pp. 236–262). Cambridge: Cambridge University Press.

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.

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

Peters, P. (2009). Australian English as a regional epicenter. In T. Hoffmann & L. Siebers (Eds.), *World Englishes: Problems, properties and prospects: Selected papers from the 13th IAWE Conference* (pp. 107–124). Amsterdam: John Benjamins.

Pijpops, D., Beuls, K., & Van de Velde, F. (2015). The rise of the verbal weak inflection in Germanic: An agent-based model. *Computational Linguistics in the Netherlands Journal, 5*, 81–102.

Pijpops, D. (Forthcoming). Lectal contamination. Evidence from corpora and from agent-based simulation. *International Journal of Corpus Linguistics*.

Power, S. (2019). *The education of an idealist*. London: William Collins.

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.

Rubdy, R., & Saraceni, M. (2006). An interview with Tom McArthur. In R. Rubdy & M. Saraceni (Eds.), *English in the world: Global rules, global roles* (pp. 21–39). London: Continuum.

Schneider, E. W. (2007). *Postcolonial Englishes: Varieties around the world*. Cambridge: Cambridge University Press.

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

Wee, L. (2020). English in Southeast Asia. In D. Schreier, M. Hundt, & E. W. Schneider (Eds.), *The Cambridge handbook of world Englishes* (pp. 263–281). Cambridge: Cambridge University Press.

Wilson, G., & Westphal, M. (2021). New Englishes, new methods: Attitudinal research in New Englishes. *English World-Wide, 42*, 175–199.
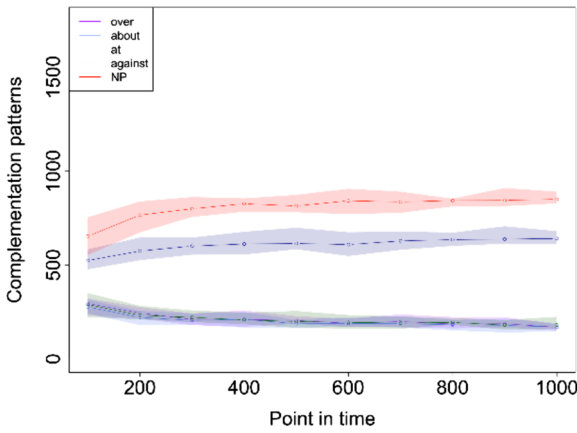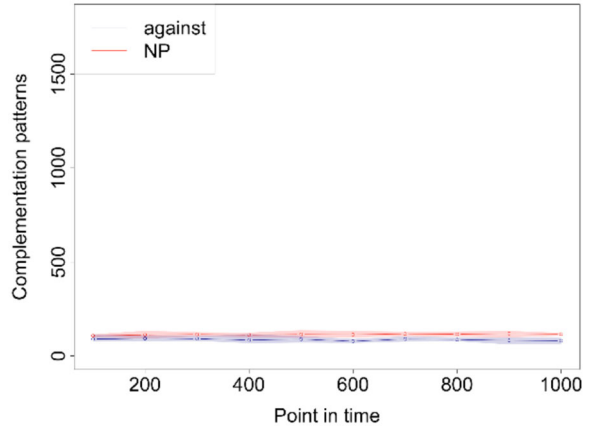
APPENDIX A

STABLE CONDITION

Indian agents
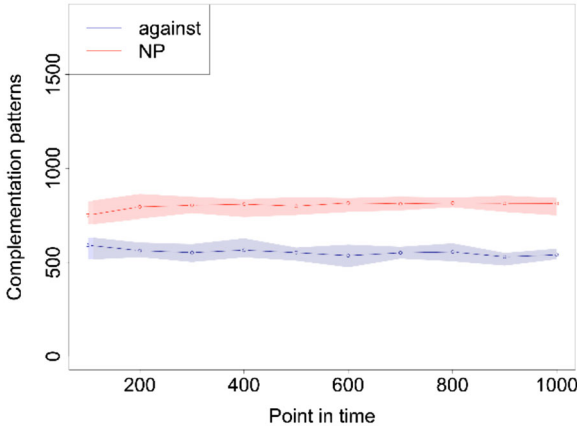
FIGURE A Complementation patterns of appeal (neutral Indian agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]
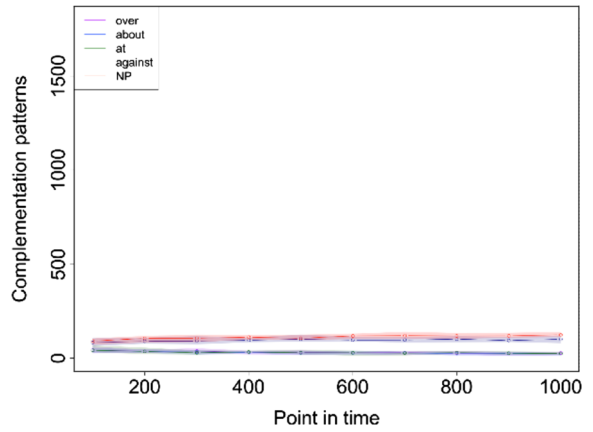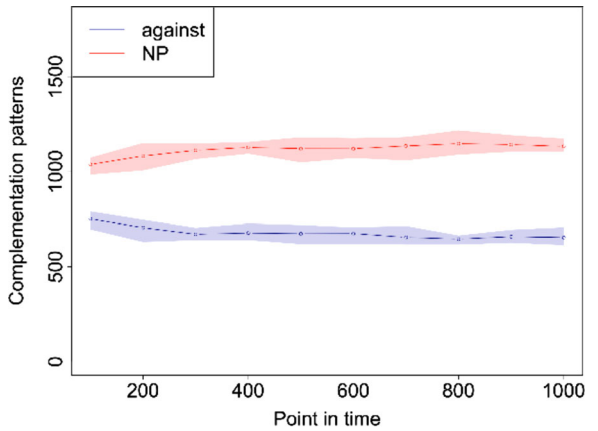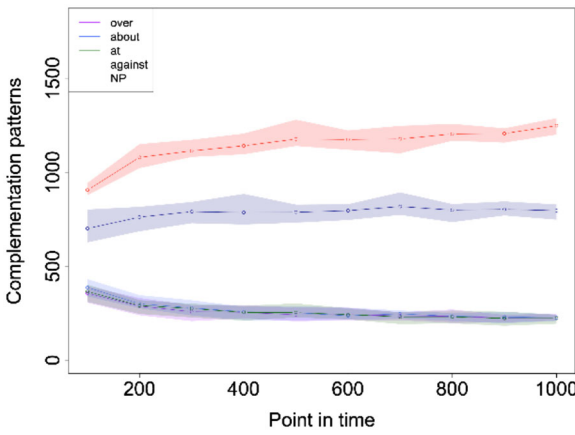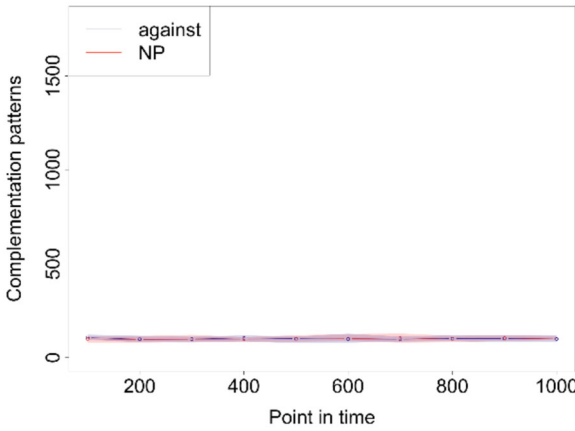

FIGURE B Complementation patterns of protest (neutral Indian agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]

Filipino agents

FIGURE C Complementation patterns of appeal (con-Am Filipino agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]
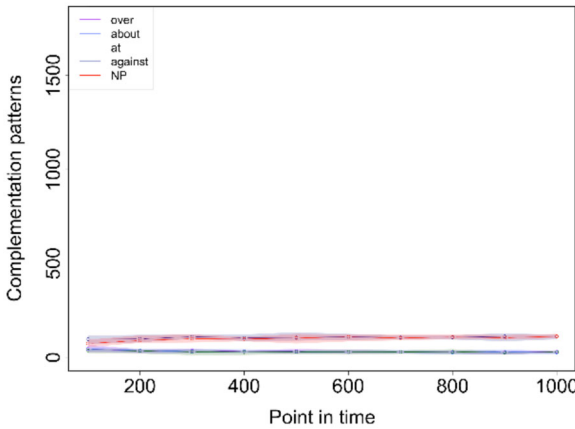

FIGURE D Complementation patterns of protest (con-Am Filipino agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE E Complementation patterns of appeal (neutral Filipino agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE F Complementation patterns of protest (neutral Filipino agents, stable condition) [Colour figure can be viewed at wileyonlinelibrary.com]
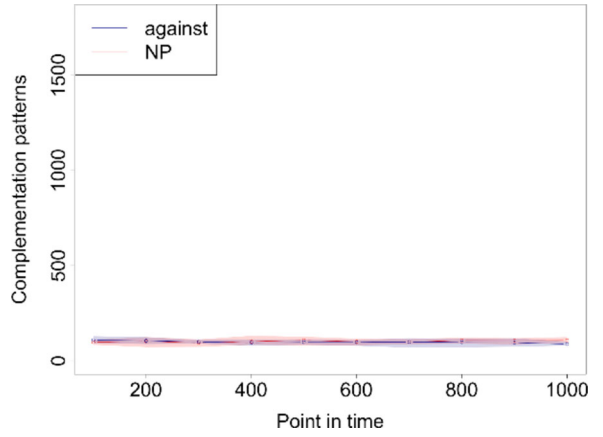
STRONG AVOIDANCE

Filipino agents

FIGURE G Complementation patterns of appeal (con-Am Filipino agents, strong active avoidance) [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE H Complementation patterns of protest (con-Am Filipino agents, strong active avoidance) [Colour figure can be viewed at wileyonlinelibrary.com]

STRONG AVOIDANCE AND AMERICANISATION

FIGURE I Complementation patterns of appeal (con-Am Filipino agents, Americanisation and strong active avoidance) [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE J Complementation patterns of protest (con-Am Filipino agents, Americanisation and strong active avoidance) [Colour figure can be viewed at wileyonlinelibrary.com]