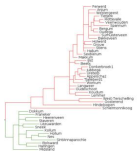
Agent-based modelling of language change

Guest lecture "Current trends in AI" 2024-2025

Peter Dekker

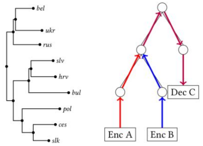
BSc Artificial Intelligence | Utrecht University

Thesis: Reconstruction phylogenetic tree Dutch dialects



MSc Artificial Intelligence | Universiteit of Amsterdam

Thesis: Reconstruction ancestry tree by predicting words using machine learning



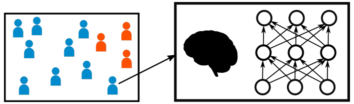
/instituut voor de Nederlandse taal/

Systems developer | Instituut voor de Nederlandse Taal

among others: crowdsourcing

PhD researcher | AI Lab, Vrije Universiteit Brussel

Agent-based modelling of language change (supervisor: Bart de Boer)



Postdoctoral reseacher | AI Lab, Vrije Universiteit Brussel

Industrial AI application including domain knowledge (with Johan Loeckx)

Federale Overheidsdienst FINANCIEN

Business analyst | Federal Public Service Finance (from March 2025)

Identifying drivers of language change using agent-based models

PETER DEKKER

Promotor: prof. dr. Bart de Boer
Vrije Universiteit Brussel
2024

ARTIFICIAL INTELLIGENCE RESEARCH GROUP
AI FLANDERS
BUILDING OUR DIGITAL FUTURE

Feel free to ask questions (or discuss)!

Motivation

What are **driving forces** behind **language change** when **languages are in contact?**

Study using **agent-based computer simulations**

Two perspectives:

- What can computer models teach us about **language change**?
- What can language change teach us about **computational modelling?**

Language important research line in AI Lab since 2005

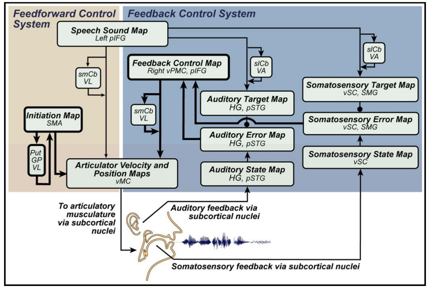
Two perspectives on AI

Engineering

- Goal: perform intelligent task well
- Build to improve accuracy
- Use representation and process that work well for task

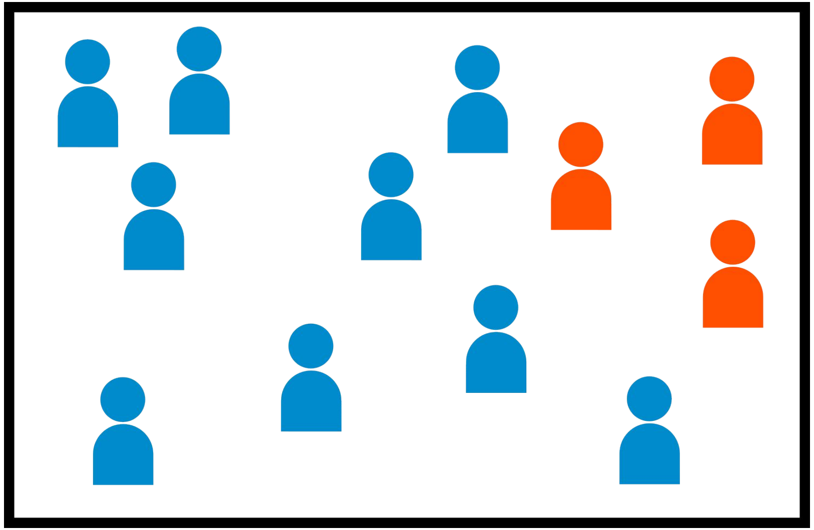
Cognitive

- **Goal: understand human cognition**
- **Understanding by building**
- **Learn how humans represent and process information**



Computational modelling

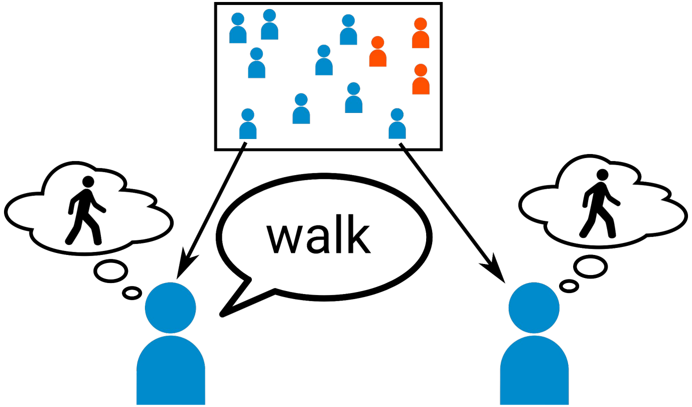
- **Cognitive** perspective on AI
- Build computer model to **better understand system (cognition & social system)**
- Speaker as agent in **multi-agent system**



Through which AI models did you learn more about **human cognition**?

Is there a fundamental difference between **single-agent** and **multi-agent** models?

Agent-based modelling



- Computer simulations of interactions between individuals
- Behaviour of **individual is defined**
- How does individual behaviour lead to global behaviour?
- Allows to make **hypotheses explicit**
  - Explicit modelling of linguistic representations: **interpretable**
  - Possibility to include **real-world data** from case studies

Mesa tutorial: Model code

https://mesa.readthedocs.io/en/stable/tutorials/intro_tutorial.html

```
class MoneyAgent(Agent):
    """ An agent with fixed initial wealth."""
    def __init__(self, unique_id, model):
        super().__init__(unique_id, model)
        self.wealth = 1

    def move(self):
        possible_steps = self.model.grid.get_neighborhood(
            self.pos,
            moore=True,
            include_center=False)
        new_position = self.random.choice(possible_steps)
        self.model.grid.move_agent(self, new_position)

    def give_money(self):
        cellmates = self.model.grid.get_cell_list_contents([self.pos])
        if len(cellmates) > 1:
            other = self.random.choice(cellmates)
            other.wealth += 1
            self.wealth -= 1

    def step(self):
        self.move()
        if self.wealth > 0:
            self.give_money()
```

```
class MoneyModel(Model):
    """A model with some number of agents."""
    def __init__(self, N, width, height):
        self.num_agents = N
        self.grid = MultiGrid(width, height, True)
        self.schedule = RandomActivation(self)
        # Create agents
        for i in range(self.num_agents):
            a = MoneyAgent(i, self)
            self.schedule.add(a)
            # Add the agent to a random grid cell
            x = self.random.randrange(self.grid.width)
            y = self.random.randrange(self.grid.height)
            self.grid.place_agent(a, (x, y))

    def step(self):
        self.schedule.step()
```

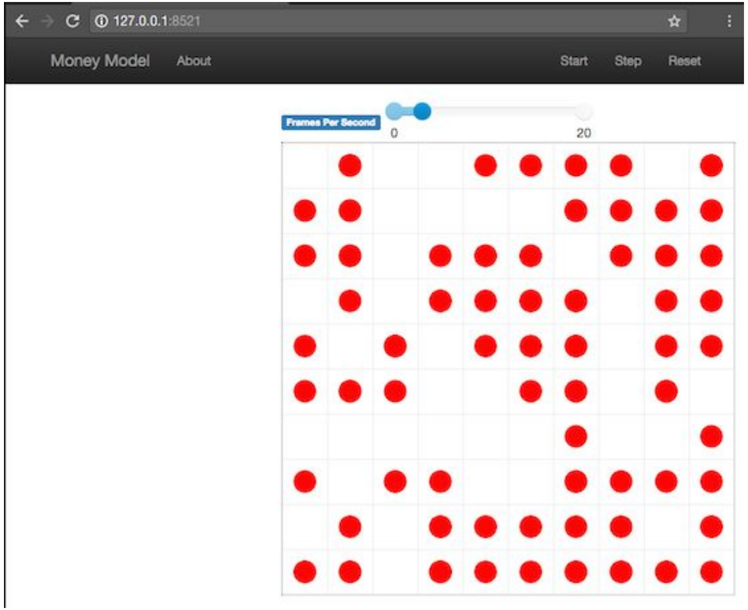
Mesa tutorial: Visualization

https://mesa.readthedocs.io/en/stable/tutorials/adv_tutorial.html

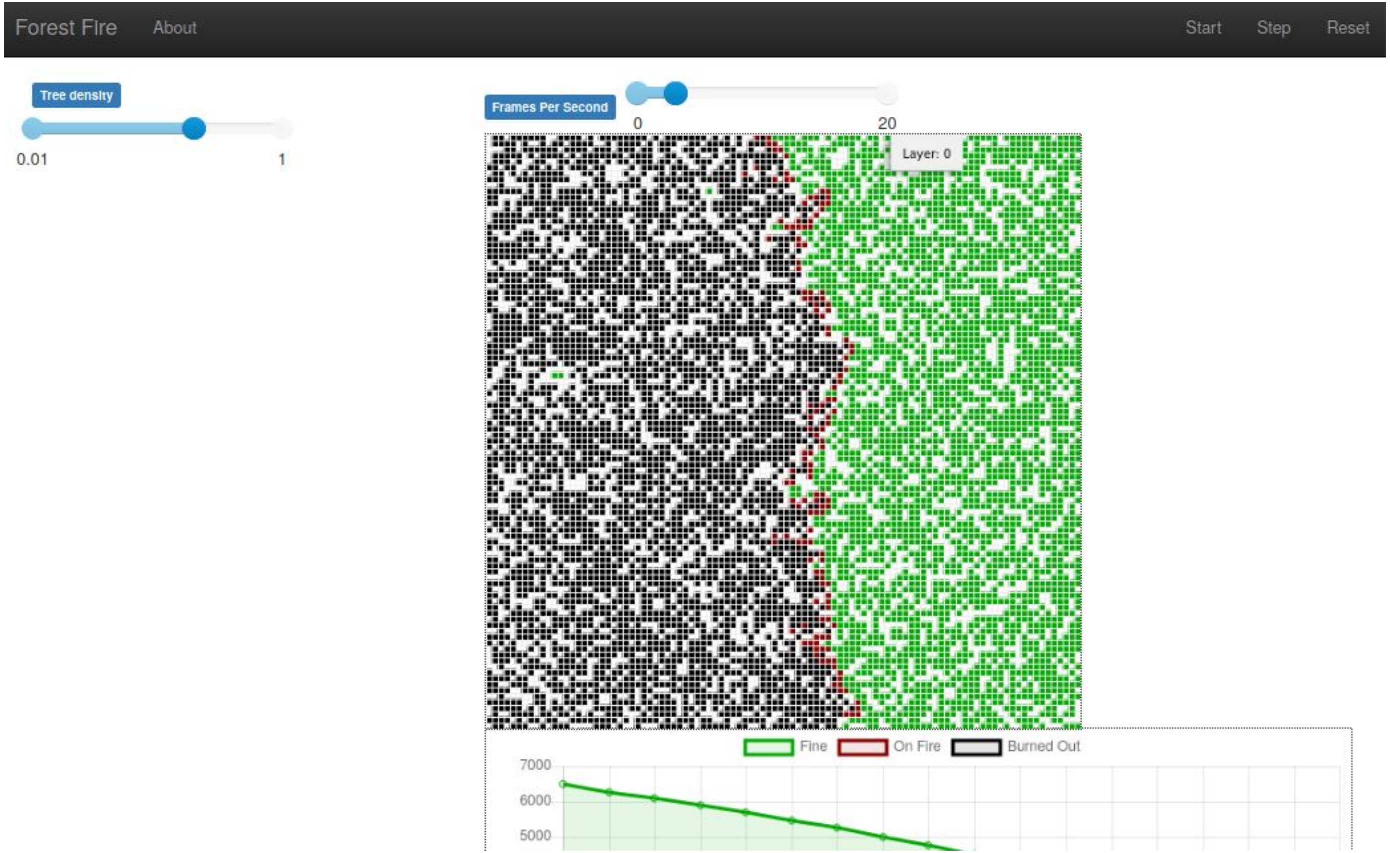
```
from MoneyModel import *
from mesa.visualization.modules import CanvasGrid
from mesa.visualization.ModularVisualization import ModularServer


def agent_portrayal(agent):
    portrayal = {"Shape": "circle",
                 "Filled": "true",
                 "Layer": 0,
                 "Color": "red",
                 "r": 0.5}
    return portrayal

grid = CanvasGrid(agent_portrayal, 10, 10, 500, 500)
server = ModularServer(MoneyModel,
                       [grid],
                       "Money Model",
                       {"N":100, "width":10, "height":10})
server.port = 8521 # The default
server.launch()
```

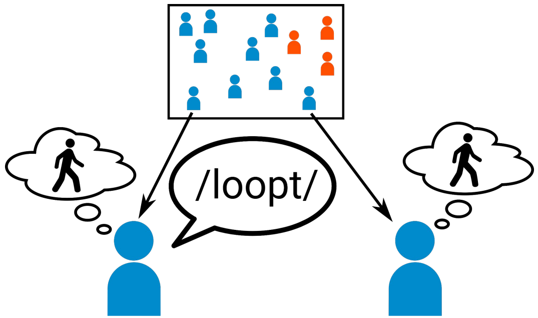


Forest fire example



Agent-based models and AI

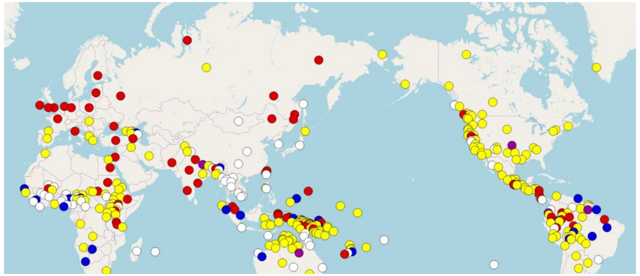
- AI as a **tool** to analyze a real-world system
  - Like AI used to analyze e.g. biology, economy
- Speaker as **agent** in multi-agent system
  - How do **individual** choices (e.g. sloppiness) lead to **collective behaviour** (e.g. language losing a word ending)?
  - Useful for human-human, human-robot and robot-robot interaction
  - Insights from my "descriptive" agents used for "acting" agents (e.g. chatbot)



Morphology

(he) walk-s

- Morphology: how words are built up from meaningful parts
- Specifically: marking of **subject** on **verb**
    - Variation across languages
    - Sensitive to change by social factors



World Atlas of Linguistic Structures
Verbal Person Marking

Social dynamics

- Groups of speakers with different characteristics
- Language contact, dialect contact, different groups within society
- Situation and outcome different in every language: **use case studies**

Research questions

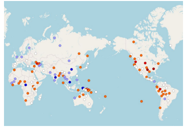
Linguistic: What are the factors that lead to **change** in **morphology** in situations of **social dynamics**?

Computational: How can language change be modelled using computer models?

- Integrate **real-world data**
- Plausible **cognitive model** ('brain') of agent
- Insights for **multi-agent AI**

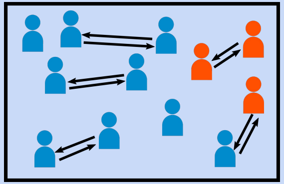
Implications: How can agent-based models be used to model **language preservation** efforts?

Chapter 3
study of subject markers in world's languages



provides information

Agent-based models

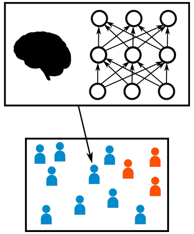


Chapter 2
Morphological simplification

Chapter 4
Conversational priming

'brain' for agent
(future work)

Chapter 5
Neural network



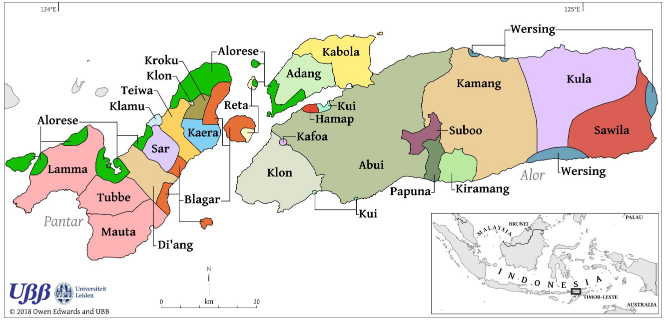
Chapter 2: Morphological simplification

What are interactions between **phonology** (system of sounds in language and their order) and **morphology** (how words are built from smaller parts) during **language contact**?

Case study: Alorese

(photo: Yunus Sulistyono)

- **Alorese** (Austronesian). Alor & Pantar islands, Eastern Indonesia
- Alorese **lost verb morphology** (Klamer, 2020)
- Contact with unrelated Alor-Pantar languages
- Language contact with adult speakers can lead to morphological simplification (Lupyan & Dale, 2010)
- **But could the phonology of Alorese also play a role?**
  - Two types of verbs: Suffixing verbs lost suffix, prefixing verbs kept prefix
  - Prefixing verbs start with vowel, suffixing verbs sometimes end in consonant
  - **Avoidance of sequences of consonants factor in morphological simplification?**



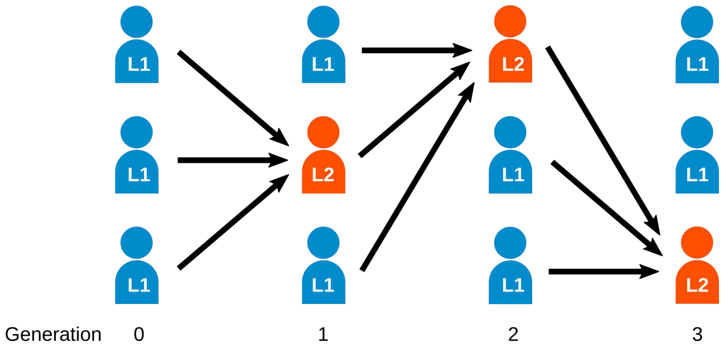
(map: Owen Edwards and UBB)

	before contact		**after contact**
prefixing	*n-enung* he-drink 'he drinks'	no simplification →	*n-enung* he-drink 'he drinks'
suffixing	*balik-na* ? return-he 'he returns'	simplification →	*balik* return-he 'he returns'

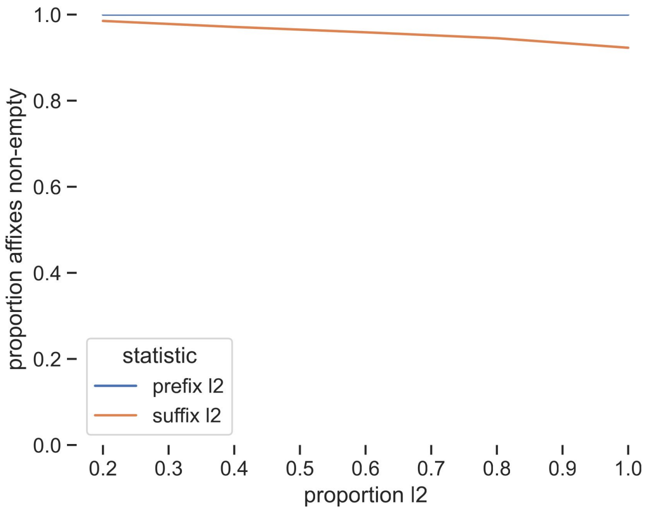
Agent-based model

- Model of **intergenerational transmission**
  - **Mother tongue speakers (L1) initialised with full language**
  - **Second language learners (L2) learns** from previous generation (L1 & L2)
  - **Meanings:** verb + person (e.g. to go-he)
  - **Signals:** verb affixes (e.g. k-, t-, -ko)
  - Update when successful interaction
- Test mechanisms:
  - **Phonological reduction** mechanism: L2 speakers drop full affix when consonant cluster arises
    `balik-na`
  - **Generalisation** mechanism: use affixes from all concepts during production

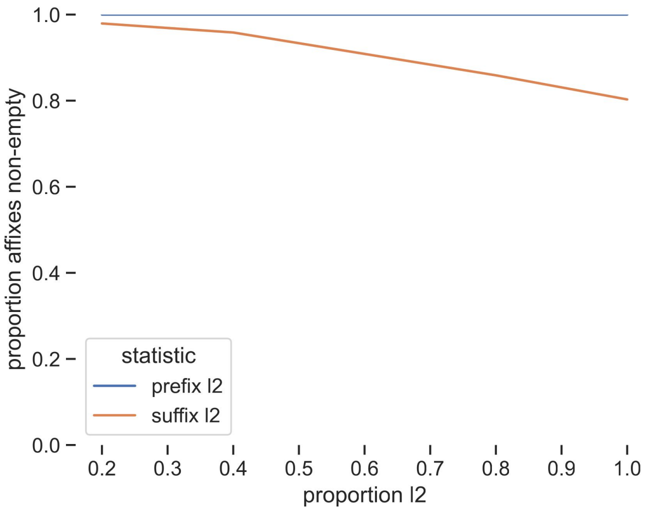
Model initialised with data from language before simplification.
Which mechanisms lead to simplified situation in Alorese?



Results: Phonological reduction and generalisation



No generalisation



Generalisation 10%

- **Phonological reduction leads to modest morphological simplification**
- **Generalisation reinforces effect**

Conclusion

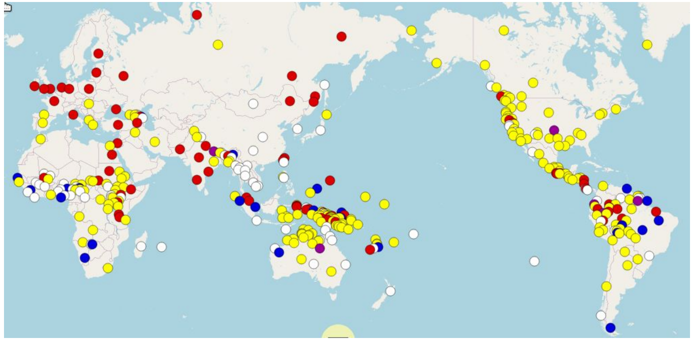
- **Some evidence for effect of phonological reduction** on morphological simplification in Alorese
- Model surprisingly **resilient to strong reduction mechanism** (through **meaning**)
- **Generalisation needed to spread** empty affix to verbs without consonant clusters
- Small effect suggests role of other interacting factors:
  - Mismatch with affixes in Alor-Pantar languages
  - Frequency of use
  - Sociodemographic situation

How do you use **prior knowledge** to **inform an AI model**?

Chapter 3: Rate of change of subject markers

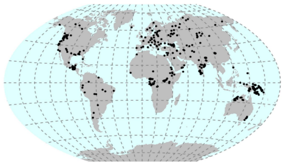
Do some subject markers on verb (e.g. I walk-∅, he walk-s) change faster than others?

→ Find mechanisms behind subject marking, for agent-based model in next chapter



WALS, Siewierska (2013)
Verbal subject marking in the world

Exploratory data analysis



Sample of languages
(Seržant & Moroz, 2022)

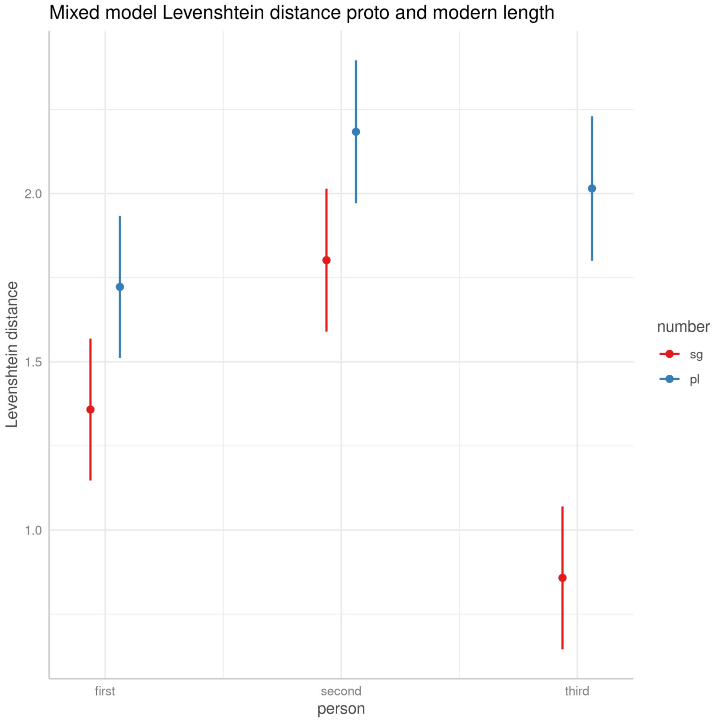
Subject markers from 310 languages from database Seržant & Moroz (2022)

6 persons: I, you, he/she/it, we, you (plural), they

language	proto_language	person_number	modern_form	proto_form
Lithuanian	Proto-Indo-European	1sg	u	ō, oh2
Lithuanian	Proto-Indo-European	2sg	i	e-s-i
Lithuanian	Proto-Indo-European	3sg	a	e-t-i

- **Proxy for rate of change: difference between form in ancestor language (proto-form) and modern form**
- Metric: **Levenshtein distance** (number of edits between forms)
- Mixed linear model
  - **levenshtein_distance ~ person*number + (1|family)**
  - Interdependence of data points from same language family

Results



Most robust finding: **3SG (he, she, it) is most conservative**

Conclusion

- Exploratory study suggests: **3SG (he/she/it) most conservative subject marker**
- Mechanisms behind change of subject markers:
    - **Frequency of use**
        - Evidence from spoken language: **3SG most frequent subject marker** (Bybee, 1985; Scheibman, 2001; Seržant & Moroz, 2022)
        - **Frequency has conserving effect** on morphology
    - Interaction with other factors, such as community size (Nettle, 1999)
- Re-use of datasets for new questions
- Study **frequency** as concurrent factor to **conversational priming**

Chapter 4: Conversational priming

💬 Not much research on the role of **conversation** in language change

🚊 Languages change, and as conversation is the "habitat" of language, it must provide **infrastructures for linguistic innovations to spread** from one speaker to another.

HYPOTHESIS:

💡 ***Conversational priming in repetitional responses*** leads to faster spread of innovative forms

In **multi-agent AI** models, when does **communication** play a role?

And when does the **structure of the communication/protocol** matter?

Conversational priming in repetitional responses

***Are you** there?*

- Yes, **I am** there.

1st/2nd person: different forms in question and answer

No priming effect

***Es he** there?*

- *Yes, **he es** there.*

3rd person: same form in question and answer

Priming effect:
Answerer may **copy innovative form** from interlocutor and **use it in the future**

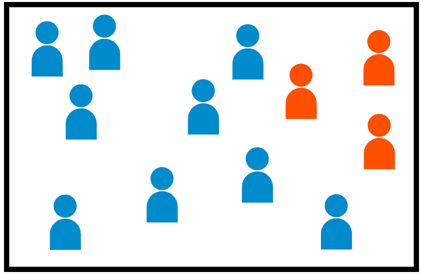
Language change

Agent-based model

Zietela, Belarus (Wikipedia)
Inspired by data from Lithuanian dialects Zietela and Lazūnai, where 3SG=PL changes fastest.

Computer model of repetitional responses in conversations

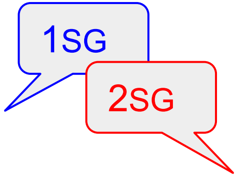
- Interaction in population of agents: **conservator** (0% innovative form) vs. **innovator** (90% innovative form) agents
- Meanings: **1SG** (I)**, 2SG** (you)**, 3SG** (he)
- Forms: **conservative** vs. **innovative**
- Agent updates when speaking and listening

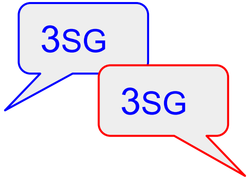


Priming:

If person different (1/2SG) ⟶ Sample form from own distribution

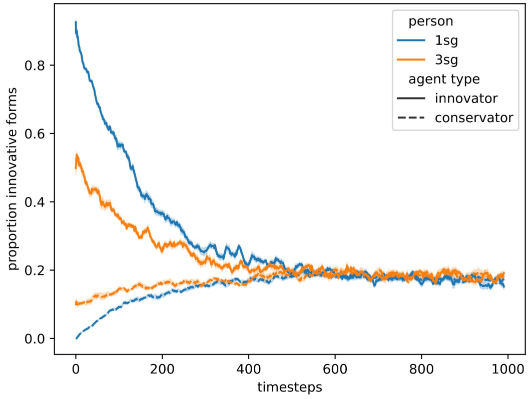
If person same (3SG) ⟶ **Use same form as questioner**



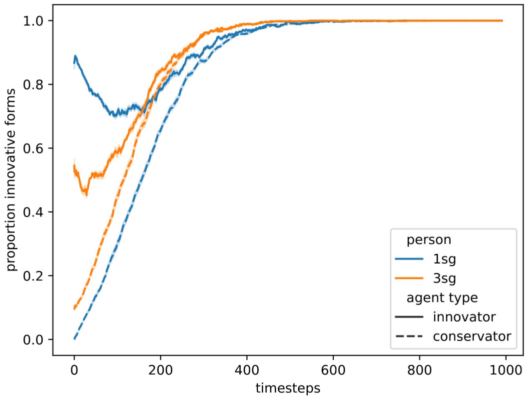


Results

- Basic model:
  - **Languages of innovators and conservators converge**: innovative form does not take over
  - **Priming (3SG) converges faster** than non-priming (1SG)
- **Innovative form favoured** (Blythe & Croft, 2012) & **stronger increase in production** during conversational priming (Cleland & Pickering, 2003)
  - **Priming (3SG) makes innovative form take over faster**
- **→ Specific circumstances in the language needed**



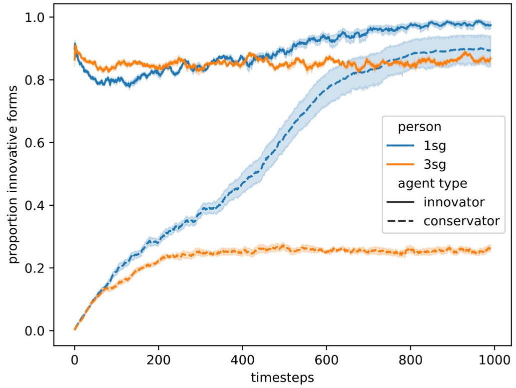
Basic



Favour innovative + stronger production increase

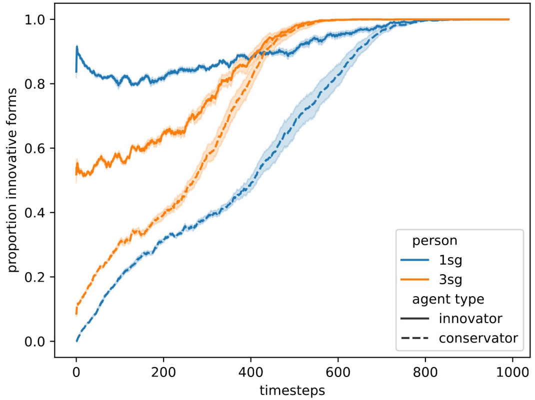
Conversational priming versus frequency

- Conversational priming spreads innovation for 3SG, frequency has conservative effect on 3SG (data study)
- **Can conversational priming (3SG) overcome conservative effect of frequency**?
- Frequency implemented by decay mechanism: variant (innovative/conservative) with lowest probability decreased



No conv. priming
(for none of the persons)

- Conserving effect of frequency neutralises S-curve



Conv. priming
(only for 3SG, as in other experiments)

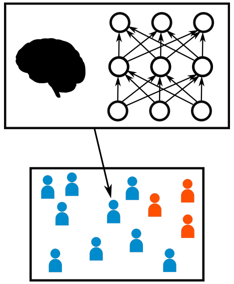
- **Conversational priming overcomes frequency: innovative form takes over again for 3SG**

Conclusion

- Conversational priming leads to **faster convergence**, and, when innovative form has inherent benefit, can lead to **faster take-over**
- Conversational priming can, for certain parameter settings, **overcome conserving effect of frequency**
- Highlights importance of timescales (Enfield, 2014)

Chapter 5: Neural network model of learning inflection classes

- Domain of study: **Inflection classes**
  - Groups of words inflected in same way
  - Inflection classes help language users deduce unseen word forms (Milin et al., 2009; Veríssimo & Clahsen, 2014)
- **How do inflection classes change over time?**
- How does **individual language user** acquire **system of inflection classes**?
  - Because learnability influences change (Elsner et al. 2019; Kusters, 2003)
- Which role does **generalisation** play?



Recent computer models of morphological processing (Elsner et al., 2019; Kodner et al., 2022):

- Generation of inflected forms, E.g. 1SG I work → 3SG he …..

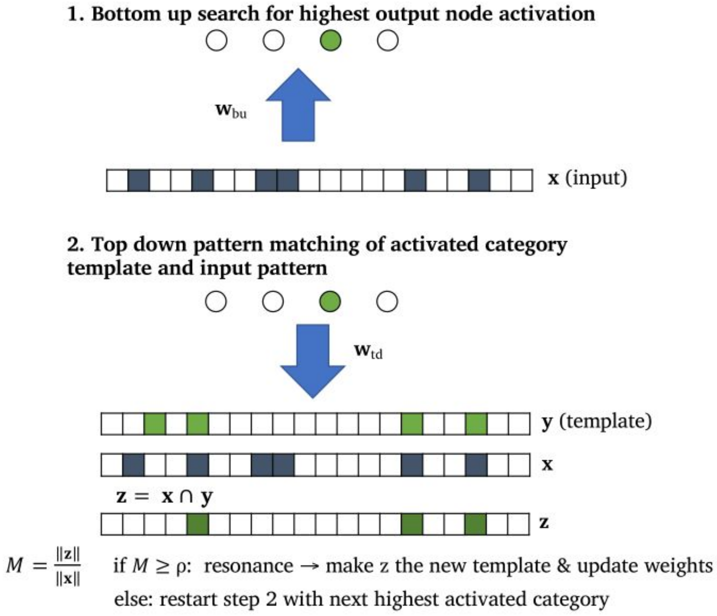
Our task: **Unsupervised inflection class clustering**

- **Explicitly predict inflection classes**
- **Learning abstraction** of inflection class carries information
- Task: Cluster verbs (1 datapoint = all forms for 1 verb) into inflection classes

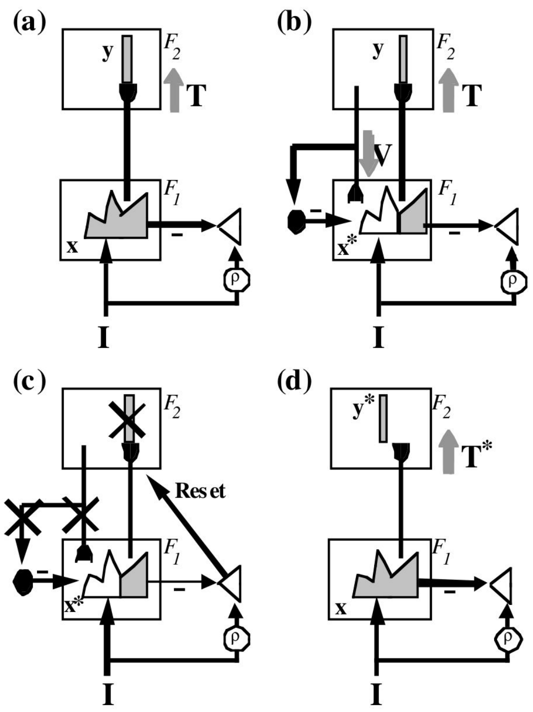
Method

Adaptive Resonance Theory 1 (Carpenter & Grossberg, 1987):

- Cognitively inspired neural network of category learning
- Input layer (new stimuli) & perception layer (learned categories)
- Vigilance parameter: degree of **generalisation**
- Interpretability: weights represent cluster features (Grossberg, 2020)



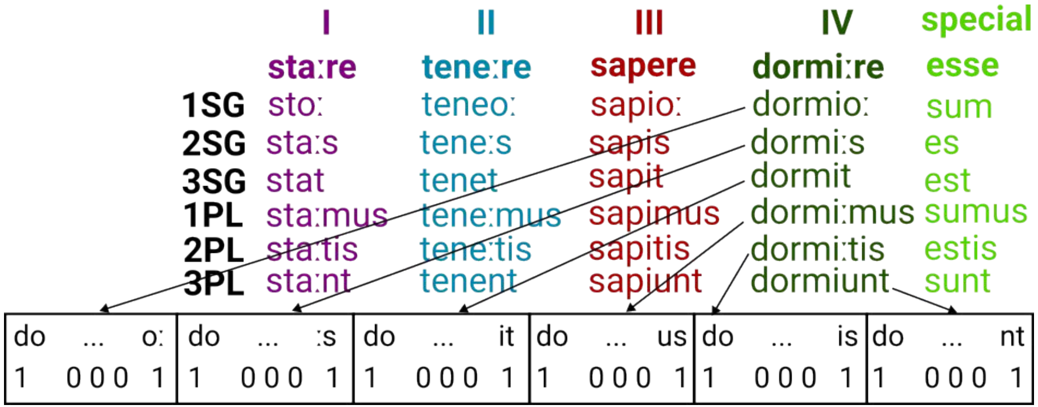
Adaptive Resonance Theory



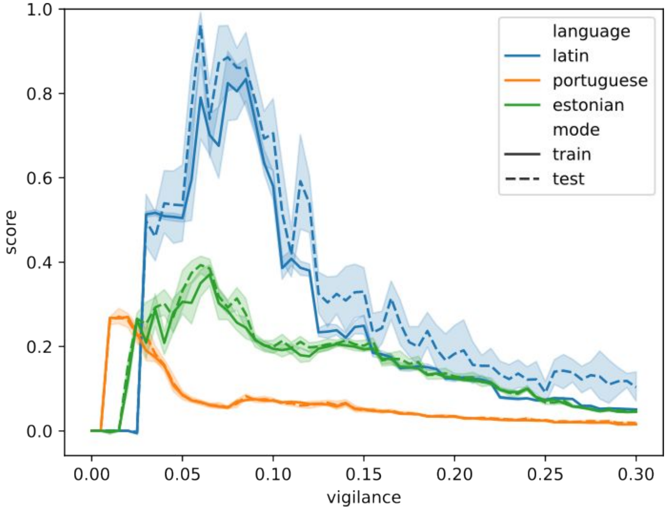
Carpenter & Grossberg
(2013)

Data

- Phonetic verb forms with inflection classes
- Languages, in increasing complexity:
  - Latin (Beniamine et al., 2020): predictable
  - Portuguese (Beniamine et al., 2021): less predictable
  - Estonian (Beniamine et al., 2024): complex morphophonological alternations
- Datapoint: concat of n-grams of all forms for verb



Results



Clustering similarity to real inflection classes (Adjusted Rand Index)
Highest performance in narrow vigilance range
Good generalisation to test set

Cluster analysis for Latin (bars=clusters, colours=real ICs)
Good correspondence between clusters and inflection classes.
Fundamental difficulties for Portuguese and Estonian

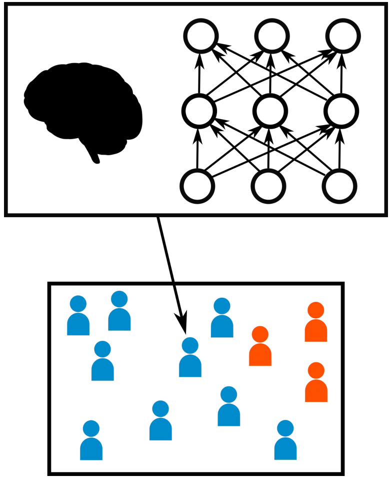
Conclusion

- Similarity between ART and real inflection classes varies based on complexity of inflectional system
- Best similarity for **narrow range of generalisation parameter**
- Extracted features correspond to real inflection classes

Future work: ART in agent-based model

Study **change of inflection classes** using agent-based model with ART as comprehension & production model ('brain' of speaker)

Task (clustering) has to be adapted to communication setting



Conclusion

Linguistic: Factors behind language change

- **Mechanisms interact**
  - **Phonological factors** play role in **morphological simplification**
  - **Conversational priming** may speed up language change and counter **frequency**
- **Short and long time scales connected** in model of conversational priming

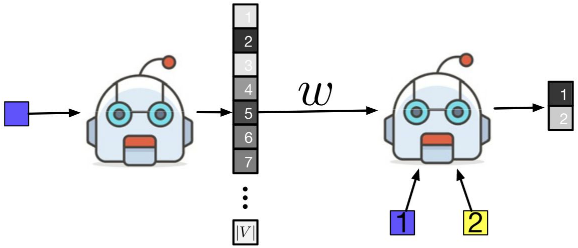
Computational: Small amount of data

- Different ways of using data: **close to data** or **data as inspiration**
- **Choice of models** that work with small amounts of data (ABM & shallow NN)
- Use of **data-driven methods** (such as statistical study subject markers) to **inform computer models**

Computational: Cognitive model

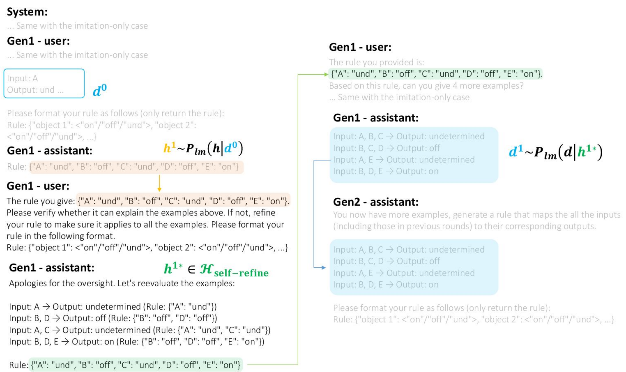
- ABMs with **explicit linguistic representations** to model linguistic data
- ART1 as possible **neural cognitive model** in ABM
- Simpler models lead to more **interpretability**?
  - Large language models (e.g. ChatGPT) also used to study language change, but
    - Biased by language use in training data
    - Computationally heavy
    - Not always interpretable

Computational: Insights for other multi-agent AI systems



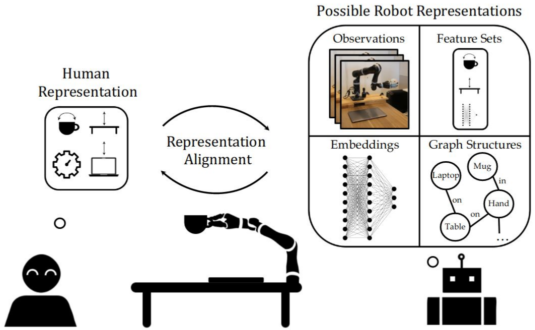
Chaabouni et al. (2021)

Emergent communication



Ren et al. (2024)

Communication between LLMs



Bobu et al. (2023)

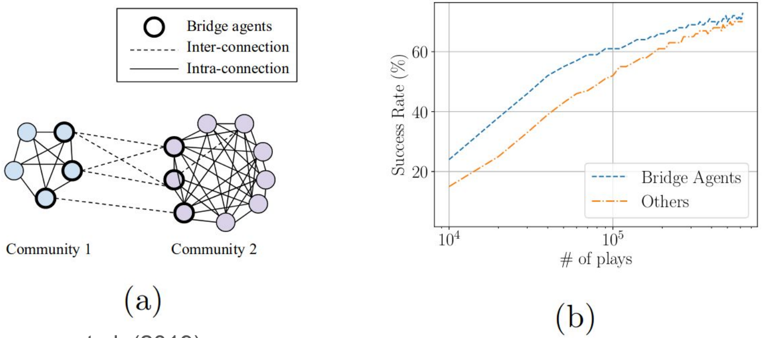
Human-machine interaction

Computational: Insights for other multi-agent AI systems

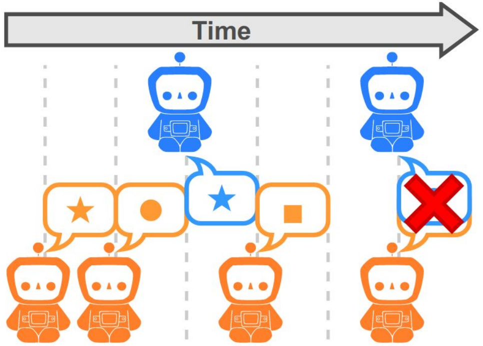
Languages in contact: two types of agents

→ insights for multi-agent AI with different types of agents (such as human-machine interaction)

- **Different objectives and characteristics** of agents (e.g. L1 vs L2)
- Different **layers of communication system may interact** (e.g. phonology & morphology)
- **Temporal structure** of communication system (e.g. conversational priming)



Graesser et al. (2019)
Contact between different groups in neural emergent communication



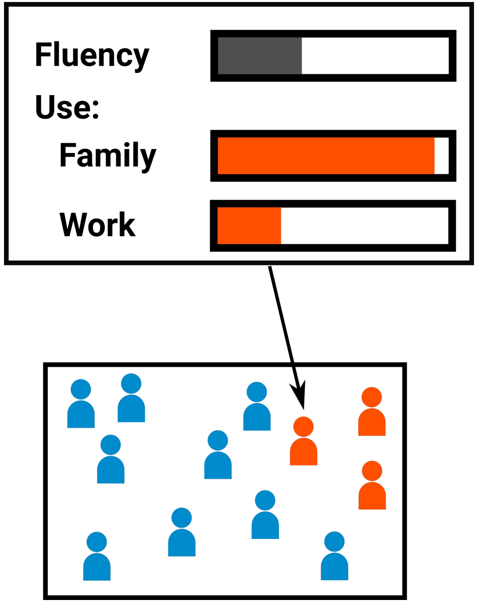
Taillandier et al. (2023)
Turn-taking in neural emergent communication

Implication: Language preservation

- Model **vitality of entire language**, rather than phenomena within language
- Evaluate **policy alternatives**
- Factors
  - Bilingualism
  - Education
  - Economic status of language

Explore with organisations/researchers working on language preservation?

How do you connect your research with society?



Concluding remarks

- **Agent-based models** can shed light on **interactions between mechanisms** in morphological change
- **Simple, interpretable computer models** suitable to study real-world processes, such as language change
- **Insights for multi-agent AI systems**: heterogeneous agents
- Agent-based modelling as tool for **language preservation**

- Dekker, P. (2024). *Identifying drivers of language change using agent-based models* [Doctoral dissertation, Vrije Universiteit Brussel]. [pdf]
- Dekker, P., Gipper, S. & de Boer, B. (2024). Conversational priming in repetitional responses as a mechanism in language change: evidence from agent-based modelling. *Linguistics Vanguard*, *10*(1), 549-564. https://doi.org/10.1515/lingvan-2023-0187
- Dekker, P., Gipper, S. & de Boer, B. (2024). 3SG is the most conservative subject marker across languages: An exploratory study of rate of change. In *The Evolution of Language: Proceedings of the 15th International Conference (Evolang XV).* Madison, WI, USA. [pdf] [SI]

Feel free to contact me: research@peterdekker.eu