

Internship at University of Vienna

Michal Olbrich
michal.olbrich@ufal.mff.cuni.cz



Charles University
Faculty of Mathematics and Physics
Institute of Formal and Applied Linguistics



unless otherwise stated

University of Vienna

- 87,000 Students
- Established 1365
- Second oldest in Central Europe
- associated with 17 Nobel Prize laureates



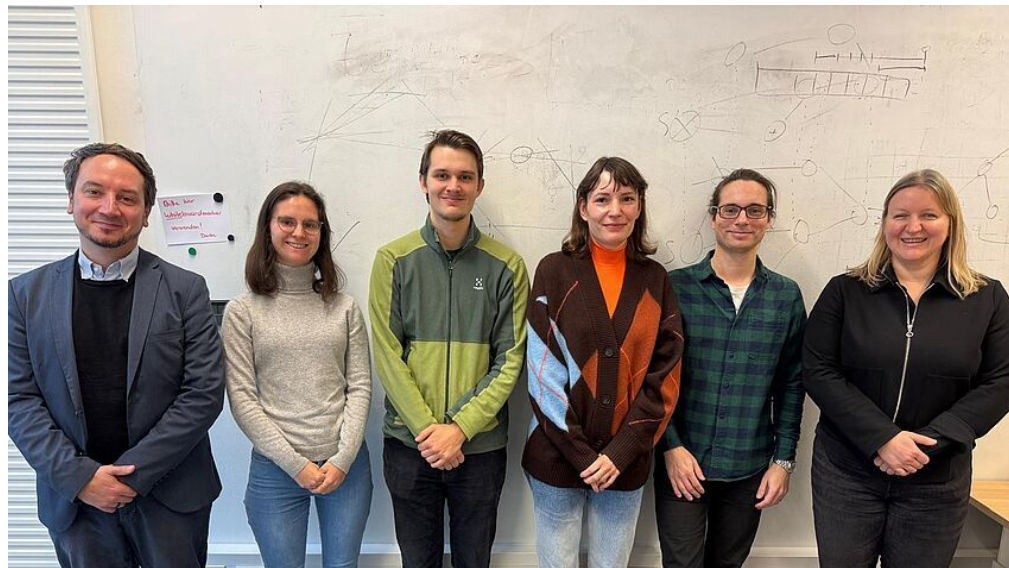
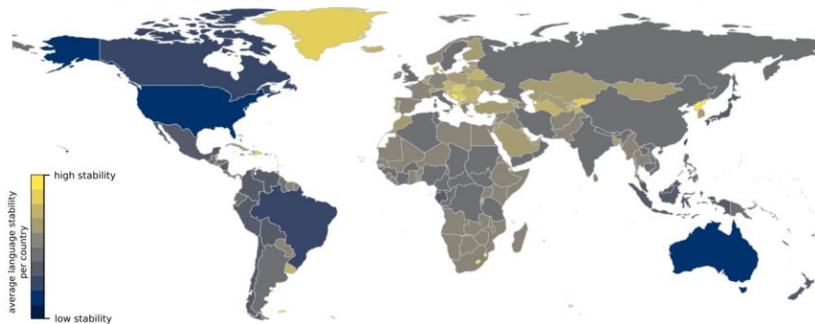
universität
wien



DIGILINGDIV project

- First visit: 15th November - 15th December 2025
- Second visit March-May 2026
- Ass. Prof. Mag. Mag. Dr. **Andreas Baumann**
- Disentangling effects of digitization on linguistic diversity (DIGILINGDIV)

Disentangling effects of digitization on linguistic diversity



Digital Humanities (Master's programme)

- Andreas Baumann: Assistant Professor, Digital Linguistics, Department of German Studies
 - Cognitive and evolutionary linguistics (phonology, phonotactics, lexicon, semantics, emotion)
 - Quantitative linguistics (mathematical, computational, and statistical modeling)
 - Epidemiological, ecological and evolutionary models in diachronic linguistics
 - Linguistic diversity and drivers thereof (digitalization, colonialism)
- Hannes Fellner: Univ.-Prof. Mag. Hannes Fellner, M.A. PhD
 - Deputy Head - Department of European and Comparative Literature and Language Studies
 - Historical-comparative linguistics and digital philology
 - Focus on nominal morphology, theoretical and digital approaches to language change



My work in Vienna

- Application of Evolutionary Game Theory (EGT) on diachronic data
- Application of Replicator dynamics to population of speakers in digital vs. non-digital domain

Population-dynamic modeling of inter-language competition in the digital vs. non-digital domain: a proof of concept

Michal Olbrich¹, Hannes Essfors², Andreas Baumann³

¹ Charles University, ² TU Wien, ³ Universität Wien

DIGLINGDIV Panel Discussion, 15 June 2025

Background

- Language ecologies are **biased**: only a few languages, such as English, are increasingly used by a large share of speakers, and this particularly holds true for the digital domain (Lewin, 2024; Essfors et al., 2025).
- We want to study such **inter-language competition** by means of population-dynamic modeling.
- Different **population-dynamic modeling** frameworks have been already employed to study competition among languages, such as generalized Lotka-Volterra models (Lewin et al., 2021; Zhang & Long, 2023; Lewin et al., 2024), the replicator equation (Lewin, 2024), and the replicator-mutator equation (Lewin, 2024; Essfors et al., 2025).
- Rather than just examining models theoretically, we fit them to rich **language-specific time-series data** in the digital and the non-digital domain.
- Our aim is to estimate inter-language competition coefficients to learn about **which languages suffer/benefit from competition** and **the difference between the digital and the non-digital domain** in terms of competitive interactions.

Data

Finland

- Linguistic situation
 - Finnish (Co-official)
 - Swedish (Co-official)
 - 8 minority languages
 - 10.8% speakers of foreign L1 (European Union 2024)
- English is widespread L2
- Part of school curriculum
- Used for inter-cultural communication and in functional domains

Non-digital

- Finnish Population Information System
- Native language of every resident, aggregated on the national level
- Speaker distributions across 108 languages (1990 – 2024)

Digital

- Wikilingua: Wiktionary of Wikipedia page views across language editions (Essfors & Baumann, 2025)
- National language distributions (2015 – 2024)
- Proxy for digital language use

Methods

- The **replicator equation** is a mathematical model that describes how different strategies compete between each other in certain environment (Maynard Smith, 1982).
- For our purpose, these strategies are represented by different languages that compete in the non-digital and the digital domain, respectively.
- The equation models the change in relative abundances of all competing languages.
- Rate of growth depends linearly on the difference between payoffs gained in interactions and mean fitness in the population:
- Here, A is the **payoff matrix** that holds all information about how languages compete with each other (competition coefficients; see heatmaps above; positive entry = l_j benefits from l_i , blue); negative entry = l_j suffers from l_i ; red); diagonal entries refer to intra-language interactions).
- We solved this equation numerically for three languages (fi, swe, eng) using scipy odeint package and least squares method as a loss function.
- For assessing statistical robustness, confidence intervals were derived (* indicates significant non-zero competition coefficients, $p < .05$).

Results & discussion

- Our preliminary findings are:**
 - Proof-of-concept: our modeling approach works in principle given the available data
 - **Non-digital**:
 - English benefits from Finnish and Swedish
 - Swedish benefits from Finnish; Swedish suffers from Swedish
 - **Digital**:
 - English benefits only from Finnish
 - There seems to be a **stronger competition in the non-digital domain** than in the digital domain
 - **Open issues**:
 - Account for different time spans across domains
 - Integrate more languages
 - Systematically compare competition coefficients across domains
 - Methodological robustness: test other population-dynamic modeling frameworks and optimization procedures

Competition coefficients, Population

	Language 1	Language 2
Language 1	11.18*	-4.87*
Language 2	4.9*	5.83*

Competition coefficients, Wikipedia

	Language 1	Language 2
Language 1	-25.7*	-10.86
Language 2	3.39*	3.18*

Population ratios (non-digital)

Wikipedia use (digital)

Abstract

The relative frequencies of a word's senses change over time as a reflex of interactions between them. Although lexical semantic change has been extensively studied from both theoretical and, more recently, computational perspectives, the nature of interactions between individual word senses and how they evolve has not yet been systematically investigated. Here, we model this evolutionary interplay within the framework of Evolutionary Game Theory using replicator dynamics, a system of ordinary differential equations modeling scenarios such as competition, predation, and cooperation, governed by a matrix of interaction coefficients. To estimate these coefficients from empirical data, we employ methodology based on non-linear least-squares optimization for fitting the evolutionary model to observed trajectories of word-sense frequencies. Our experiments show that cooperation is the most common interaction type for words with two senses, whereas predation dominates among words with three senses. Surprisingly, direct competition is not the prevailing interaction type, despite being one of the most frequently discussed mechanisms in theories of lexical semantic change.

Figure 1: Phase portrait depicting oscillating interactions among three senses as represented by the word 'affair'. (1) 'event'; (2) 'romantic relationship'; (3) 'object of a particular type' as per Oxford English Dictionary.

1 Introduction

The evolution of word meaning is perhaps one of the most salient aspects of language change. It is no wonder, then, that considerable efforts went into mapping out the semantic change of words both through traditional lexicographic and historical-linguistic approaches as well as by means of computational methods. In particular the latter field has skyrocketed since the establishment of modern NLP technology more than a decade ago (Kutuzov et al., 2018; Tahmasebi et al., 2021; Schlechtweg et al., 2019, 2025; Liétard and Loiseleur, 2026). Thus, lexical semantic change detection (LSCD) has emerged as a significant sub-field of computational linguistics and NLP, as witnessed by regu-

This Is Not a Competition: Modeling Interactions Between Word Senses in a Diachronic Polysemy Dataset Using Replicator Dynamics

Anonymous ACL submission

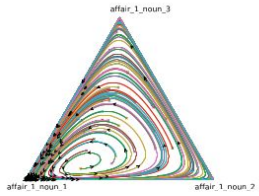


Figure 1: Phase portrait depicting oscillating interactions among three senses as represented by the word 'affair'. (1) 'event'; (2) 'romantic relationship'; (3) 'object of a particular type' as per Oxford English Dictionary.

lar competitions and workshops, transferring state-of-the-art NLP methods to the study of lexico-semantic change (Schlechtweg et al., 2020).

Much of this research is concerned with the question of which words change, and whether some word senses become more dominant than others. Questions like this are, of course, central for our understanding of how linguistic cognition works. Indeed, research on language cognition and evolution has shown that whether interactions between words referring to the same meaning, i.e., instances of onomasiological ambiguity, are competitive or cooperative is modulated by properties of a word's semantic neighborhood (Mirman and Magnuson, 2008; Chen and Mirman, 2012) and communicative need (Karjane et al., 2020a).

Less is known, however, about how word senses interact in lexico-semantic change, i.e., semasiological ambiguity in polysemous words such as *affair*, which can mean 'event' or 'romantic relationship' (Figure 1).¹ A study that is important in

¹In the present study, we adopt a broad definition of se-

My workplace (first stay)

- Kolingasse



My workplace (second stay)

- Rotunde - U-bahn Spittelau



Opportunities to meet other PhD students

- DigHum Fridays
 - Once a month a lecture relevant to the field of Digital Humanities like
 - Responsible Research in Digital Humanism: Ethics

Additional programme to co-working:



Colloquium (monthly)

- "Lernraum" 1st floor from 09:00h - 12:00h,
Talk starts at 09:30h
- Light Breakfast available



Workshops (monthly)

- "Lernraum" 1st floor from 09:00h - 12:00h
Workshop starts at 09:30h
- Coffee available in the reserved room



- Linguistikstammtisch
 - Once a month
 - Informal meeting of students of linguistics from various departments
 - 19:00 at Cafe Merkur



Food

- Lunch at menza
 - Costs 9-12 Euro
 - Quality 2/5
 - Quantity 3/5



Food



Food



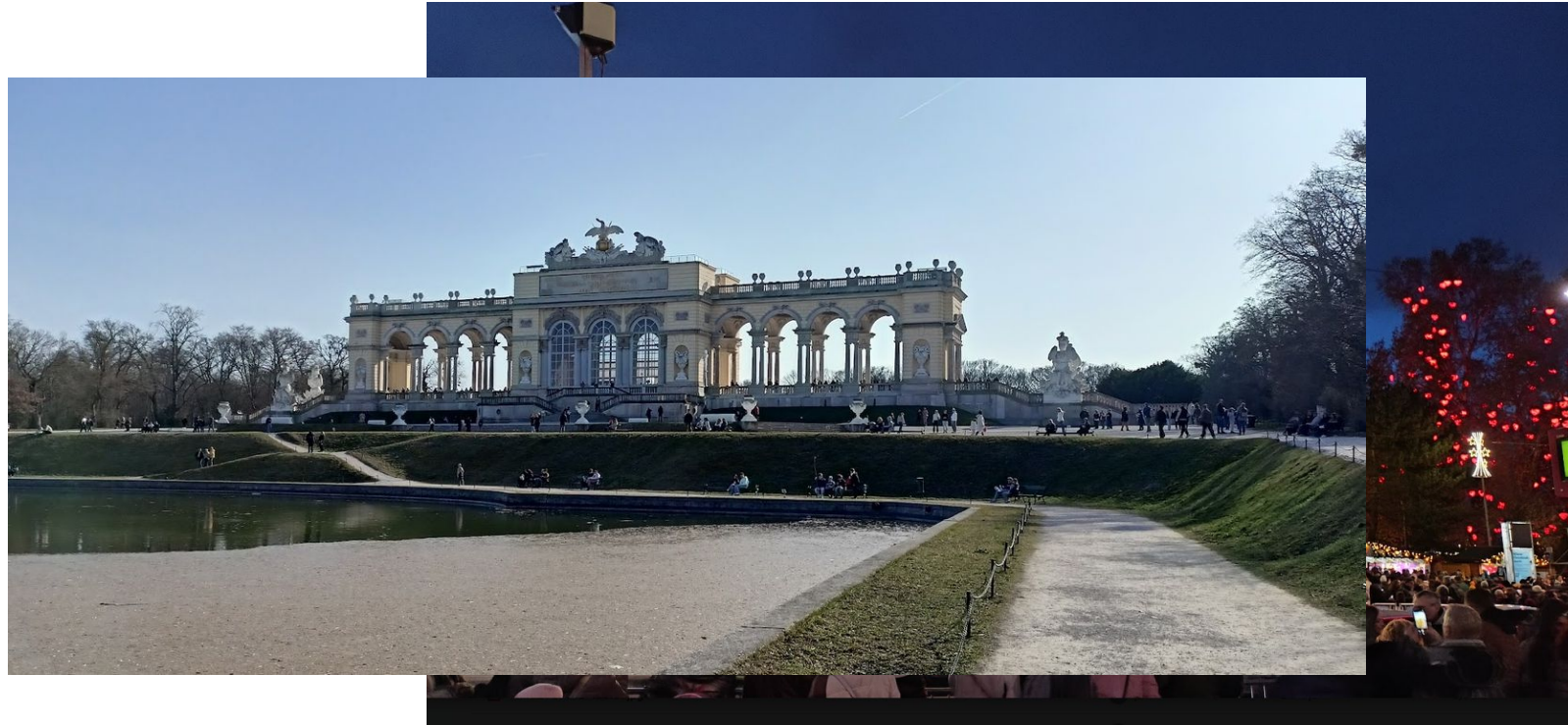
Food



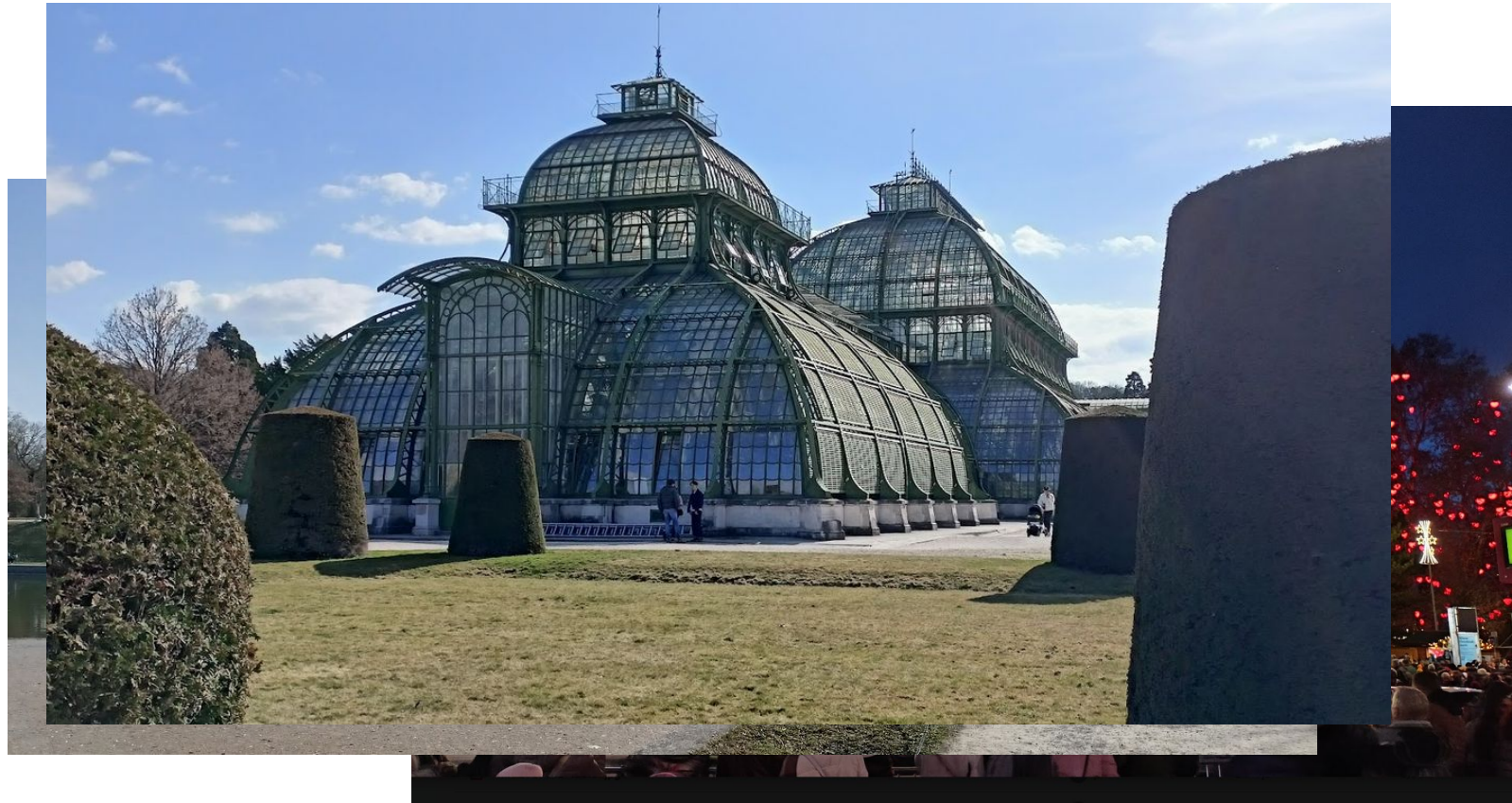
Free time in Vienna



Free time in Vienna



Free time in Vienna



Free time in Vienna

