Doing Subregular Linguistics

LIN 655.02, Fall 2023

Time:	Fri 10:00–12:50	
Place:	SBS N250 (Compling Lab)	
Website:	lin655.kennethhanson.net	
Instructors:	Thomas Graf	Kenneth Hanson
	lin655@thomasgraf.net	kenneth.hanson@stonybrook.edu
	SBS N-249, office hours TBA	SBS N-210, office hours TBA

Course Description

This course will introduce subregular formal languages and their application to linguistics. Though the use of formal language theory in linguistics was out of fashion for some time, recent research has discovered that the overwhelming majority of linguistic patterns fall into a small number of subregular classes. These findings bring new insights into linguistic structure as well as opportunities for connections to language typology and cognition. This course aims to enable you to understand the mathematics behind subregular linguistic analyses as well as to produce your own. We will focus our attention on syntax to a large extent, but will also spend a fair bit of time on phonotactics and other topics along the way.

Course Objectives

- Understand the hierarchy of subregular languages, the kinds of patterns they are able to express, and their relevance to linguistic theory
- Become familiar with recent findings regarding the formal complexity of linguistic patterns, including those in phonotactics, phonological maps, and syntax
- Develop the ability to both decode and formulate formal analyses of linguistic patterns using subregular languages

Prerequisites

It is recommended that students have taken the following courses, or their equivalents.

- LIN 521 (Syntax I)
- LIN 523 (Phonology I)
- LIN 539 (Mathematical Methods in Linguistics)

Experience with formal language theory and automata theory as taught in LIN 637 (Computational Linguistics II) is helpful but not required.

Requirements, Grading

The course will consist mainly of readings and in-class discussions and exercises. Students are expected to actively participate. Those enrolled for credit will lead the discussion for at least one paper. The second major component of the course is a project in which you will produce a subregular analysis of data from one or more languages. We will ask you to submit a project proposal (< 1 page) mid-way through the semester. Later, you will present your work

to the class to gather feedback and write up your analysis as an abstract or paper. The exact requirements by number of credits are follows:

- 0 credits: Do the readings and participate in class activities.
- 1 credit: All of the above, plus present a paper.
- 2 credits: All of the above, plus develop and present an analysis.
- 3 credits: All of the above, plus write an abstract or paper.

Grading The course will be letter graded: A, A-, B+, etc.

About the Project Your project proposal must be approved by us by the date shown below. We are flexible regarding the format of the final write-up. Your options include: a 2 page conference abstract, a 4 or 8 conference paper, or a full paper. Beware that writing a good conference abstract is not necessarily easier than writing a short paper.

Important Dates

Fri. 11/3 (Week 10)	Deadline for Project Proposal Approval
Fri. 11/17 (Week 12)	Project Workshop
Fri. 12/15 (Finals Week)	Deadline for Abstract/Paper (due by 11:59pm)

- You need to submit your project proposal well in advance of the approval deadline in case the initial version is not approved. Ideally, you should submit it as soon as possible so you can start working on your analysis.
- We are flexible regarding presentation times. If you want to present earlier, let us know and we will try to fit you in. Presenting later is not ideal as the next class is on 12/1, leaving you with little time to write your paper/abstract.

Policies

- The best way to contact us is by email. Please allow up to 24 hours for a reply.
- Kenneth should be your primary go-to person for most matters. Administrative issues such enrollment, extensions, and grades should go to Thomas.
- If you want to come to office hours and anticipate a longer meeting, please send an email so that we can set aside enough time and avoid conflicts.

Semester Plan

We will start the course with subregular string languages and their use in modeling phonotactics. The reasons for this are twofold: first, because subregular linguistics started in this area, and second, because subregular string languages are easier to understand than tree languages or maps. We will also begin studying Minimalist Grammars (MGs), which are used to define the tree structures used in subregular syntax.

In the second part of the course, we will take a deep dive into syntax. Topics to be discussed include selection, movement, case, agreement, and licensing of NPIs and anaphora. To date there have been two approaches to extending subregular languages to trees, the first using trees directly, and the second using constraints over path strings; we will study both of them.

In the remainder of the semester, we will cover a selection of other other topics such as maps (transductions), learning, and others according to the interests of the class. We can also revisit past topics in more detail, or conversely, we can also tackle one or more of these earlier if anyone is interested in doing a project on that topic.

Preliminary Schedule

Note: We will add, remove, and shuffle readings according to the interests of the class, and the schedules of presenters.

Week 1–2: Subregular Languages, MGs (9/1, 9/8)

9/4 – Labor Day

Graf, T. (2022). "Subregular linguistics: bridging theoretical linguistics and formal grammar". In: *Theoretical Linguistics* 48.3-4, pp. 145–184.

Heinz, J. (2018). "The computational nature of phonological generalizations". In: *Phonological Typology, Phonetics and Phonology*, pp. 126–195.

Week 3–4: Phonotactics, Logic, Learning (9/15, 9/22)

Graf, T. (2017). "The Power of Locality Domains in Phonology". In: *Phonology* 34, pp. 385–405.

Heinz, J. (2010). "Learning Long-Distance Phonotactics". In: *Linguistic Inquiry* 41.4, pp. 623–661.

Heinz, J. (2014). "Culminativity Times Harmony Equals Unbounded Stress". In: *Word Stress: Theoretical and Typological Issues*. Chap. 8.

McMullin, K. and G. Ó. Hansson (2016). "Long-Distance Phonotactics as Tier-Based Strictly 2-Local Languages". In: *Proceedings of the Annual Meetings on Phonology*. Vol. 2.

Week 5–6: Selection, Movement, Case (9/29, 10/6)

Graf, T. (2018). "Why movement comes for free once you have adjunction". In: *Proceedings of CLS* 53, pp. 117–136.

Graf, T. (2022). "Typological implications of tier-based strictly local movement". In: *Proceedings of the Society for Computation in Linguistics 2022*, pp. 184–193.

Hanson, K. (2023). "A TSL Analysis of Japanese Case". In: *Proceedings of the Society for Computation in Linguistics 2023*.

Vu, M. H. et al. (2019). "Case assignment in TSL syntax: A case study". In: *Proceedings of the Society for Computation in Linguistics 2019*, pp. 267–276.

Week 7: Structure-Sensitive Tier Projection, Command Strings (10/13)

10/9–10 – Fall Break

Graf, T. and C. Mayer (2018). "Sanskrit n-Retroflexion is Input-Output Tier-Based Strictly Local". In: *Proceedings of SIGMORPHON 2018*, pp. 151–160.

Graf, T. and N. Shafiei (2019). "C-command dependencies as TSL string constraints". In: *Proceedings of the Society for Computation in Linguistics 2019*, pp. 205–215.

Week 8: Agreement, Another Take on Command Strings (10/20)

Graf, T. and A. De Santo (2019). "Sensing Tree Automata as a Model of Syntactic Dependencies". In: *Proceedings of the 16th Meeting on the Mathematics of Language*, pp. 12–26.

Hanson, K. (2023). *A Computational Perspective on the Typology of Agreement*. Unpublished manuscript.

Week 9: Phonological Maps (10/27)

Burness, P. A. et al. (2021). "Long-distance phonological processes as tier-based strictly local functions". In: *Glossa: a journal of general linguistics* 6.1.

Chandlee, J. and J. Heinz (2018). "Strict Locality and Phonological Maps". In: *Linguistic Inquiry* 49.1, pp. 23–60.

Week 10: Syntactic Maps (11/3)

Project proposal approval deadline

Graf, T. (2020). "Curbing Feature Coding: Strictly Local Feature Assignment". In: *Proceedings of the Society for Computation in Linguistics 2020*, pp. 224–233.

Graf, T. (2023). "Subregular Tree Transductions, Movement, Copies, Traces, and the Ban on Improper Movement". In: *Proceedings of the Society for Computation in Linguistics 2023*.

Week 11: Logic and Learning (11/10)

Heinz, J. (2010). "String Extension Learning". In: *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, pp. 897–906.

Lambert, D. et al. (2021). "Typology emerges from simplicity in representations and learning". In: *Journal of Language Modelling* 9.1, pp. 151–194.

Week 12: Project Workshop (11/17)

All students enrolled for 2-3 credits will present their work so far.

Week 13

11/22–24 – Thanksgiving – no class

Week 14–15: Choose Your Own Adventure (12/1, 12/8)

Topics may include: morphotactics, morphological maps, multiple tiers, probabilistic grammars, more on phonology/syntax/logic/learning. Some possible readings are listed below.

End of Semester

Abstract or paper due 12/15 by 11:59pm

Additional Readings

Aksënova, A. and S. Deshmukh (2018). "Formal Restrictions On Multiple Tiers". In: *Proceedings of the Society for Computation in Linguistics 2018*, pp. 64–73.

Aksënova, A., T. Graf, et al. (2016). "Morphotactics as Tier-Based Strictly Local Dependencies". In: *Proceedings of the 14th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology*, pp. 121–130.

Chandlee, J. (2017). "Computational locality in morphological maps". In: *Morphology* 27.4, pp. 599–641.

Chandlee, J. et al. (2018). "Input Strictly Local Opaque Maps". In: *Phonology* 35.2, pp. 171–205.

Gleim, D. and J. Schneider (2023). "Phonological processes with intersecting tier alphabets". In: *Proceedings of the Society for Computation in Linguistics 2023*.

Hanson, K. (2023). "Strict Locality in Syntax". In: *Proceedings of CLS 59*.

Heinz, J. et al. (2011). "Tier-based strictly local constraints for phonology". In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human language technologies*, pp. 58–64.
Jardine, A., J. Chandlee, et al. (2014). "Very efficient learning of structured classes of subsequential functions from positive data". In: *Proceedings of the Twelfth International Conference on Grammatical Inference (ICGI 2014)*. Vol. 34, pp. 94–108.
Jardine, A. and K. McMullin (2017). "Efficient Learning of Tier-Based Strictly k-Local Languages". In: *Language and Automata Theory and Applications*, pp. 64–76.
Mayer, C. (2021). "Capturing gradience in long-distance phonology using probabilistic tier-based strictly local grammars". In: *Proceedings of the Society for Computation in Linguistics 2021*, pp. 39–50.
Shafiei, N. and T. Graf (2020). "The Subregular Complexity of Syntactic Islands". In: *Proceedings of the Society for Computation in Linguistics 2020*, pp. 421–430.
Torres, C. et al. (2023). "Modeling island effects with probabilistic tier-based strictly local grammars over trees". In: *Proceedings of the Society for Computation in Linguistics 2023*.

More Policies

Student Accessibility Support Center Statement

If you have a physical, psychological, medical, or learning disability that may impact your course work, please contact the Student Accessibility Support Center, Stony Brook Union Suite 107, (631) 632–6748, or at sasc@stonybrook.edu. They will determine with you what accommodations are necessary and appropriate. All information and documentation is confidential.

Students who require assistance during emergency evacuation are encouraged to discuss their needs with their professors and the Student Accessibility Support Center. For procedures and information go to the following website: https://ehs.stonybrook.edu/programs/fire-safety/emergency-evacuation/evacuation-guide-disabilities and search Fire Safety and Evacuation and Disabilities.

Academic Integrity Statement

Each student must pursue his or her academic goals honestly and be personally accountable for all submitted work. Representing another person's work as your own is always wrong. Faculty is required to report any suspected instances of academic dishonesty to the Academic Judiciary. Faculty in the Health Sciences Center (School of Health Professions, Nursing, Social Welfare, Dental Medicine) and School of Medicine are required to follow their school-specific procedures. For more comprehensive information on academic integrity, including categories of academic dishonesty please refer to the academic judiciary website at http://www.stonybrook.edu/commcms/academic_integrity/index.html.

Critical Incident Management

Stony Brook University expects students to respect the rights, privileges, and property of other people. Faculty are required to report to the Office of Student Conduct and Community Standards any disruptive behavior that interrupts their ability to teach, compromises the safety of the learning environment, or inhibits students' ability to learn. Faculty in the HSC Schools and the School of Medicine are required to follow their school-specific procedures. Further information about most academic matters can be found in the Undergraduate Bulletin, the Undergraduate Class Schedule, and the Faculty-Employee Handbook.