

Seminars for Guided Research in LEL in semester 2

Acoustic correlates of phonological contrasts (Rebekka Puderbaugh)

Phonemic contrasts and phonological categories form the basis of phonological theory, yet their relationship to the phonetic reality of speech sound is not yet fully understood. Due to phonetic variability across speaker characteristics, speech styles, dialects and languages, establishing phonetic descriptions of phonological categories is not straightforward. Despite this, phonemes are identified according to phonetic parameters such as voice, place, and manner and transcribed using symbols that represent these phonetic dimensions to some extent. Such definition suggests that sounds that are transcribed using the same symbols in different languages will sound similar to each other, but there are often substantial differences both in the acoustic detail of the sounds, as well as in the auditory impressions of them. This seminar will focus on the acoustic presentations of segmental phonemic contrasts, with particular attention on the variation and overlap of various acoustic correlates across categories that are nominally the same in different languages. Readings will include studies of cues to voice, place and manner in consonants, as well as dispersion and uniformity in vowel spaces. For the first assessment, students will review studies of a particular contrast (or series of contrasts) of their choice, then for the second assessment compare previously reported measures to those found in at least two further languages that also exhibit such contrasts.

Analysing and Transcribing Dialect Data (Warren Maguire)

In this module, students will learn how to orthographically transcribe audio recordings of non-standard dialect speech using the software ELAN (<https://archive.mpi.nl/tla/elan>), which creates time-aligned, multi-tiered transcriptions. The variety that will be analysed and transcribed is the Mid-Ulster English dialect of south-west Tyrone. Students will learn a transcription protocol, how to handle data ethically, and how to transcribe natural, non-standard conversational speech (both general principles and ELAN specifics). In addition, they will learn about the dialect exemplified in the recordings and about some of its key features. Each student will be given a different short sample of the dialect to transcribe and analyse. The first assignment will consist of a literature review of general studies of the dialect and of a particular feature or features that will be analysed in the second assignment. The second assignment will consist of the orthographic transcription of the audio data, a report on the methods and procedures followed, and an analysis of one or more selected features (reviewed in Assignment 1) in the recording that has been transcribed. This guided research module will provide training in a range of linguistic and more general skills that can be transferred to further study and to the workplace, including ethical data handling, accurate orthographic transcription using ELAN, and analysis non-standard variables in speech.

Bilingualism over the lifespan (Antonella Sorace)

How do people become bilingual at different stages in life, and what are the linguistic and cognitive effects of knowing more than one language? This course will examine and compare the main features of child bilingualism and adult bilingualism, from the perspective of language development and its interfaces with other cognitive domains. It will also focus on experimental methods for addressing specific research questions related to children and adults learning languages at different ages.

Community Studies in Sociolinguistics (Lauren Hall-Lew)

This course introduces students to the traditional methodology of variationist sociolinguistics: the community study. Students will be introduced to every aspect of classic research design, with particular attention paid to the role of social theory in both methodological and analytic

decisions. This is a highly reading-intensive and writing-intensive option, suited to those students whose primary interest is sociolinguistics. Each week requires over 100 pages of reading and the production of at least 200-300 words of writing for formative feedback. The first half of the semester entails a thorough reading of all of William Labov's landmark study, *The Social Stratification of English in New York City*. The second half of the semester will explore at least three other studies (a combination of books and journal articles, to be chosen by vote among the students). Every week, each student will submit a response to that week's reading. These responses will be circulated among other class members, and students will be required to comment on at least two other students' responses. The first assignment will be a partial summary of Labov's New York City study for a particular non-academic audience (to be decided by the student, in consultation with the instructor). The second assignment will be a project proposal for a new community study.

Descriptive analysis of intransitive verbs in Nuer (Bert Remijsen/Tatiana Reid)

In this data-driven research seminar, students engage with the descriptive analysis of intransitive verbs in Nuer, a West Nilotic language spoken in South Sudan and Ethiopia. This topic, on which there are no earlier publications, presents challenges and opportunities to develop expertise in a) syntax and semantics and b) morphophonology, phonology and phonetics. As for a), Nuer intransitive verbs divide into different form-based classes depending on whether they are intransitive roots, derived antipassive verbs or verbs derived from adjectives, and they also present derivations affecting valence. As for b), they present a rich system of morphological marking, which mostly involves stem-alternations: changes in vowel quality, voice quality, vowel length, tone and stem-final consonant of the predominantly monosyllabic stems. As such, this topic will allow you to develop expertise in phonology and phonetics, morphophonology, and syntax and semantics. In particular, this seminar builds on and extends your expertise in syntactic / semantic phenomena (cf. LEL2D), and will allow you to develop expertise in ear-based transcription and use of Praat to deal with tone, voice quality and vowel length (cf. LEL2B).

During the first three weeks you will gain understanding of key grammatical concepts, Nuer morphosyntax, and Nuer phonology. We will work through the usual fieldwork preparation procedures, including getting ethics clearance, collecting information about our linguistic consultant, obtaining informed consent, preparing elicitation sessions, recording the data and preparing to conduct elicitation sessions online with a speaker in South Sudan (or Kenya). From early on we will have audio elicitation sessions with a Nuer speaker to demonstrate the phenomena that you will read about in your assigned readings. From week four onwards you will participate, on a weekly basis, in a) a whole-group elicitation sessions; b) small-group elicitation sessions; and c) a whole-group sessions to discuss findings and problems. The course requires active involvement, and synchronous learning activities predominate. Willingness to spend considerable time working on your session notes is essential.

Dialect Syntax (Craig Sailor)

In this seminar we'll look at recent work that aims to bring together (a) detailed empirical study of dialectal variation within the syntax of a single language and (b) theoretical syntax in the generative tradition. This has recently become a very dynamic area of research that has broken down some traditional academic boundaries, and also has the potential to increase the understanding of syntactic concepts by non-academics. Topics will include appropriate methodologies; examinations of how empirical dialect study and syntactic theory building can be mutually beneficial; specific case-studies of syntactic phenomena. As well as readings from the literature, part of the work for the course will involve interacting with existing online resources. A central focus for this will be the recently developed Scots Syntax Atlas (SCOSYA), but participants will be encouraged to also look at other available corpora/databases, including the Yale Grammatical Diversity Project, and Multilingual

Manchester. The assessed work for the seminar will consist of an annotated bibliography and a project. The nature of the project is fairly open and can depend in part on the interests of each participant. It could for example be an analysis of a particular phenomenon drawing on data from available resources, such as those mentioned above, or a worked-out proposal (possibly including a pilot) for a study of a syntactic phenomenon that has not been fully investigated and that is expected to show dialectal variation (whether in English or in any other language).

Early Germanic Dialects (Bettelou Los)

The first part of this seminar introduces students to a number of Early Germanic Dialects (Gothic, Old Norse, Old Saxon, Old Frisian, Old English and Middle English) and significant developments in their phonology, morphology, syntax and lexicon; we will also discuss aspects of the surviving texts like their reception, i.e. what audience they were meant for, and what purpose they served; and what text type they represent (genres, translations) – these are all aspects that affect how useful these texts can be for linguistic research. The first assessment is a short assignment on the linguistic aspects. The second part of the seminar involves an independent piece of work from a range of topics on offer, on linguistic as well as more literary ("reception") aspects.

Experimental Approaches to Language Evolution (Simon Kirby)

The last ten years has seen an explosion of interest in empirical approaches to the origins and evolution of language. Whereas once it was seen as being primarily amenable only to speculative theories, it is now considered a topic that can be investigated in the experiment lab. Arguments about the origins of language based only on appeals to plausibility are now being replaced by ones that stand or fall on the results of experiments with human participants. We will look at why this has happened, and how it is possible to make hypotheses about language origins that make predictions which can be tested experimentally. We will cover two of the main paradigms in this new area - the silent gesture paradigm, and iterated learning. These recreate in a miniaturised lab setting the process of language emergence and language transmission respectively. Students will be organised into small teams and develop, run, and analyse their own silent gesture experiments. These experiments require participants to communicate a range of meanings (usually displayed as cartoon illustrations, videos, or written phrases) using only gesture and no speech. In doing so, they spontaneously improvise strategies (for example, word order) to convey information. These are remarkably consistent across participants but do not necessarily reflect the native language the participants speak. Because of this, silent gesture is seen as a window onto the cognitive biases that may be at play in the very early stages of language emergence.

Language and music (Graeme Trousdale)

In this seminar, we'll look at three topics in the study of the relationship between language and music as cognitive systems. These are:

- a. The relationship between phonological rhythm and rhythm in music
- b. Syntactic structures and harmonic structures
- c. Meaning in music and language

We'll look at a range of methodologies used to explore these topics, including corpus linguistics and corpus musicology, and experimental approaches, and draw on a wide variety of musical styles and languages. This will lead us to consider in detail the theoretical issues that both prompt and emerge from empirical work. You will be expected to submit an annotated bibliography for the first assessment, and an essay for the second. In order to get the most out of this seminar, you should have, or be willing to acquire before the seminar starts, a basic understanding of the fundamentals of music theory, and standard notation on musical staves. There are a number of print and online sources that you could make use of

in this regard, and if you have any questions or concerns about your level of musical knowledge, please contact the seminar leader who will be able to advise you.

Learning sounds and words (Mits Ota)

In this seminar, we will survey and discuss current research findings on how we learn the sounds and words of a language in two particular contexts: 1) when infants and young children are exposed to their first language and 2) when adult speakers learn a second language. In each of these contexts, we will first look at basic accounts of how phonological contrasts and phonological forms of words are acquired. We will then examine the relationship between the two, focusing on questions such as: Can sound categories be learned without knowledge of words? How do we store phonological information about individual words in our mental lexicon? How does our lexicon affect our phonological knowledge? Are there phonological biases in word learning? This seminar presupposes basic knowledge of phonetics and phonology (LEL2A + 2B or MSc Introduction to Phonology, or Phonetics and Laboratory Phonology, or equivalent). Familiarity with language acquisition is desirable though not a requirement. However, students should be warned that readings for the seminar will consist of articles that tend to be heavy on quantitative analysis related to acoustic phonetics, experimental data, and statistics. Activities will consist of independent reading of the assigned articles and discussion of key issues with the instructor and fellow students. Assessment will consist of an annotated bibliography and an essay. Students are encouraged to develop their own specific topic of investigation within the scope of the seminar.

Lexical Processing (Itamar Kastner)

This seminar will provide an overview of the central questions in lexical processing - the study of how humans comprehend individual words. We will examine the various processes taking place from the moment we first read or hear a word until we have a complete mental representation of it. This involves understanding why the past tense of a nonce word like "gleed" can be "gled" (alongside "gleeded"), whereas that of "bledge" would be "broge"; how exactly we figure out that a "writer" is someone who writes but a "corner" is not someone who corns; and under what circumstances do speakers of Arabic read words differently than speakers of English. Starting from the original wug study and working our way towards cutting-edge neurophysiological methods, we will see how technological advances have allowed researchers to give increasingly detailed answers to the central questions of morphological processing and word structure. Throughout the seminar, our focus will be on how this work can inform research in formal linguistics, while gaining a general overview of the psycholinguistic findings and models. Students will submit a project proposal as their main form of assessment. This proposal will outline an experiment that could be run in the future, letting students apply the methods surveyed in the seminar to their own interests.

Linguistic Complexity (Matt Spike)

All languages are equally expressive: any thought which can be expressed in one language can be expressed in any other. However, languages show huge variation in terms of what they *must* express, and in the ways in which they do it. *Linguistic complexity* is an emerging topic in contemporary linguistics, and attracts the attention of researchers from linguistic typology, formal linguistics, sociolinguistics, psycholinguistics, evolutionary linguistics, cognitive science, and beyond. This course will address some key questions involving linguistic complexity, including:

- i. Are languages spoken by smaller, more remote communities (e.g. Inuktitut, Archi) more morphologically complex than globally distributed languages (e.g. English, Mandarin)?
- ii. What does it mean for languages to be more or less complex, and how can we measure this?

- iii. Is greater simplicity in one aspect of a language's grammar (e.g. morphology) balanced by greater complexity in other aspects (e.g. syntax)?
 - iv. Can we use sociological, cultural, or cognitive factors to explain variation in linguistic complexity?
 - v. Are creoles - as some have argued - the world's simplest languages, or are they - as others have argued - of comparable complexity to more established languages?
 - vi. How can we integrate insights into language complexity from language description and comparison with work using theoretical, experimental, and computational approaches?
- This course will primarily consist of a curated selection of readings which provide an introduction to these topics (and others) and serve as a framework for students to conduct self-led research on a specific question (of their own choosing) in more depth.

Neuroscience of Speech Production (Alice Turk)

In this guided research seminar, we will explore neural anatomy & physiology as it relates to current theories of speech production, with a focus on sound production. The goal of the course is critical evaluation of existing theories of speech production in light of current understanding of the anatomy and physiology of the brain. Readings will include basic introductory neuroscience texts, as well as current theoretical and empirical work in the fields of phonetics and speech production.

Parsed historical corpora (Rob Truswell)

Parsed historical corpora are collections of historical texts annotated with a representation of syntactic structure. They have revolutionized research into syntactic and semantic change over the last 25 years, because they allow researchers to track the diachronic development of syntactic structures automatically, and independently of the lexical items that instantiate those structures. This makes them ideal for quantitative investigation of any research questions involving basic word order and/or unbounded dependencies, topics which are particularly resistant to efficient large-scale study in traditional corpora. In this seminar, we will read classic and recent studies which use parsed historical corpora. We will also cover the design of corpus studies (choice of corpus, design of queries) and analysis of corpus data. Students will also conduct their own corpus studies, technology permitting. Research with parsed historical corpora lies at the intersection of syntactic analysis, historical linguistics, and quantitative linguistics. Students will get the most out of this course if they are interested in these areas and are comfortable with simple programming tasks using the command line. However, nothing beyond Intro-level syntactic theory will be presupposed.

Predictability and Informativity (Hannah Rohde)

The premise of this course is that communication depends in part on the belief that people will tell us things that we're interested in hearing. In this sense, language can be understood as a channel by which speakers convey, among other things, newsworthy and informative messages (i.e., content that is otherwise unpredictable to the comprehender). We therefore might expect comprehenders to show a preference for such messages. However, psycholinguistic studies on language comprehension often point to a preference for the opposite (i.e., processing ease for real-world predictable content). Several decades of research have shown that comprehenders can deploy knowledge about situation plausibility to generate fine-grained context-driven predictions about upcoming words. This course reviews this work and then considers how to test whether comprehenders can also deploy this knowledge in favor of newsworthy content. The goal is to draw a distinction between our guess of what a speaker is likely to encounter in the world and what they are likely to **say** about it. We will read psycholinguistic papers on predictability and formal pragmatic papers on informativity. We will discuss how to design experiments that test language users' awareness of these two components during processing. Assignments will consist of crafting

two pieces that constitute important components of experimental psycholinguistic papers: a Methods section (assignment 1) and a Results & Discussion section (assignment 2).

Prosody and Meaning (Catherine Lai)

This class will explore how prosody affects how speech is understood, and how this fits into formal semantic and pragmatic frameworks. As a specific case study, we will explore the relationship between prosody and information structure, and how this affects presupposition projection. For example, people usually perceive the embedded clause 'Alex ate the cake' to be presupposed in 'Jean didn't REGRET that Alex ate the cake' but not in 'Jean didn't regret that ALEX ate the cake'. However, the relationship between presupposition projection and prosodic focus is less clear when we look at more complex and realistic discourses. To build up the tools to understand why we observe the data we do, we will look at the following:

1. What is prosody? How do we analyze it?
2. What is information structure? How does this affect prosody?
3. What sort of discourse structures do we need to understand how prosody affects pragmatic phenomena like presupposition projection? e.g. Questions Under Discussion.
4. What sort of experimental methods can we use to test theories of prosody and meaning?

A major component of this class will be critically evaluating experimental approaches to understanding these problems. The current syllabus focuses mainly on issues in English prosody, but students are welcome to look at similar issues in other languages. Students will get the most out of the class if they have some experience in formal semantics, pragmatics, and phonetics.

Quantifiers and numbers (Chris Cummins)

Language provides us with many ways to talk about quantity – for instance, we can use numbers, we can use proportions (such as *half*), we can use less precise quantifying expressions (*few, some, many, most*) and we can relate quantities to fixed reference points (*more than half, about 100*). Expressions of quantity are interesting to researchers in semantics and pragmatics for several reasons. From a formal semantic perspective, quantifiers such as *some* and *most* appear to invite a straightforward analysis based on mathematical intuitions – but as applied to natural language, these analyses turn out to be problematic in subtle ways. For instance, though it's true that more than 50% of people are female, hearers do not endorse the assertion *most people are female*. From a theoretical pragmatic viewpoint, the wide choice of possible quantity expressions for a given situation makes this area an interesting testbed for theories about the role of alternatives in pragmatic meaning. And, in terms of real-life communication, the accurate description of quantity is crucial in many important domains, whether we're talking about finance, medical risk, or indeed scientific research itself – and much of this is underexplored within linguistic semantics/pragmatics. This guided research course will look at some of the major themes in the study of quantifiers and numbers, from a theoretical and experimental perspective, and explore their significance to linguistics more broadly.

What's Where When Why: Linguistic Diversity, Typology, and Historical Linguistics (Pavel Iosad)

How do we approach the study of diversity and variation across languages? Why are some structures rare, and others common? Why do certain features tend to occur together? Explaining such patterns, which are pervasive across language families and linguistic subfields, is the aim of linguistic typology. A venerable tradition, going back to Wilhelm von Humboldt in the early 19th century, gives a role to structural factors, by identifying 'types' that languages belong to, and the properties associated with them. Linguistic structures, however, do not exist in a vacuum, and patterns of variation are often related to

contingencies of history. Closely related languages, for instance, are often typologically similar. In the early 20th century, the notion of 'language areas' was crystallized and thereby introduced geography into the mix. A better understanding of the mechanics of language contact also led to deeper insights into how it influences patterns of diversity. As a result, linguistic typology has increasingly abandoned its focus on structural properties of languages, moving towards a research programme that aims to explain how linguistic diversity patterns in space and time: from 'what' to 'where', 'when', and of course 'why'. In this seminar we will look at the range of diversity in a number of domains of grammar, consider some key concepts of linguistic typology, and analyse the ways in which theories of grammatical structure and the study of language change, history, anthropology, and geography all contribute to understanding the diversity of human language.