Department: LING

Course No. : 110

Credits : 3

Title : The Science of Linguistics

Contact Person : Harry van der Hulst

Content Area : CA3 Science and Technoology

WQ: Q

Catalog Copy : LING 2010Q. The Science of Linguistics
(110Q) Either semester. Three credits.
An introduction to the methods and major findings of linguistic research as applied to the sound systems of languages and the structure and meaning of words and sentences. CA 3.

Course Information: Course goals and objectives. The course provides a 100-level introduction to formal linguistics. Emphasis is placed on data collection, theory construction, and hypothesis testing.

1. Course requirements. In-class group exercises. Weekly homework assignments. Three (short-answer) examinations, including the final.
2. Topics covered. The course covers syntax, semantics, phonology and morphology.

Meets Goals of Gen Ed: The course places emphasis on evaluating competing hypotheses by devising and conducting simple experiments. These experiments often take the form of systematically eliciting judgments of grammaticality from a native-speaker consultant. The course thus satisfies at least two of the goals of general education. First, students are required to practice critical judgment (Goal 2) as they evaluate competing hypotheses about the grammar of a particular language. Second, the students acquire a working understanding of a major process (scientific hypothesis testing) by which they can continue to acquire knowledge (Goal 7).

CA3 Criteria: 1. *Contemporary science and technology* . The course is an introduction to the theory of grammar, as developed in modern linguistics since Chomsky's groundbreaking 1957 work, Syntactic Structures. Following Chomsky's proposals, languages are understood and studied as a formal system in which utterances are 'propositions' that need to be validated as true (i.e. grammatical), given an underlying formal expression (the premise) and a system of derivational rules. Thus, grammatical descriptions are technically 'formal proofs'. The various steps in this proof (the linguistic derivation) are formalized and their interaction or ordering is studied in detail. A "grammar" in this sense is a model of the native speaker's tacit linguistic knowledge, and is explicit enough to serve, for example, as the basis for natural language processing algorithms. Indeed, an understanding of what information a grammar provides is the basis for modern language technology, such as speech recognition, automatic translation tools, and information retrieval.

2. *Scientific methods* . The course provides extensive training in the scientific method, using language as an example. Linguistics is particularly suitable for a first introduction to scientific investigation, because data collection and experimentation can be relatively simple, with the reward of immediate results. It is therefore possible to focus on the reasoning process underlying the formulation of a theory. The course emphasizes data collection and evaluation, developing and testing a hypothesis, and presenting that hypothesis formally. Given what was said under 1, hypothesis formation and testing follows very explicit procedures that allow the student to clearly see what kind of data would falsify his or her proposal. It is due to the formalized and explicit nature of linguistic theorizing that the students, in doing linguistic analysis, have a true hands-on ‘laboratory experience'.

3. *Presenting unresolved questions* . Since the 1950's, new theoretical proposals in linguistics have been arriving at a breathtaking pace. The empirical evaluation of these proposals is very much an on-going process. The students in this course will learn about many points of present-day controversy, and will see how data that they themselves discover may favor one or another side in these debates.

4. *Promote continued learning* . Students, we hope, will be caught up in the excitement of scientific research on human language, and will continue to follow new developments in both linguistics and language technology.

Detailed Rationale for Placement in Group III: Language can be (and has been) studied from many different perspectives, including social, cultural, historical, psychological and philosophical perspectives. The Chomskyan approach added to this a (neuro)biological perspective. The proposed placement of LING 110 in Group III, Science and Technology, rather than Group I, II or IV, therefore merits discussion. The particular perspective on human language that is taken in LING 110 is a (neuro)biological one. Language is examined as a distinctive physical characteristic of the human species, i.e. as a mental ‘organ' that is part of (or implemented in) the human brain. Even though different languages of the world vary greatly in superficial aspects such as vocabulary, linguists have discovered astonishing commonalities in the grammatical principles that they obey. The current consensus is that the space of possible grammars for human language is tightly constrained by the genetic make-up of the human species. A linguistic theory, in this perspective is understood as an abstract model of the human language capacity or ‘the language organ' and, indirectly, of that part of the genome that underlies this organ. While the native speaker of a language automatically obeys a richly articulated system of grammatical principles, the principles are not available to simple introspection, and must instead be discovered by scientific induction. Contemporary scientific investigations that follow this line of research into human language are providing remarkable new insights into the computational architecture of the human brain, with implications at the genetic level. (For recent discussion, see W.T. Fitch, M.D. Hauser, *Science* 303, 377-380 (2004).) The human capacity for language has important consequences for social and cultural interaction in our species, and mental processing, but this particular course take a perspective that is logically independent from these consequences. The biological approach to language is relatively new and perhaps not widely understood, but it is a valid, expanding and promising take on the study of language. Students in the course receive hands-on experience in the scientific investigation of the individual human's capacity for language. Accordingly, the course is best placed in Group III.

Q Criteria:

This course meets all three of the criteria:

- 1. *Mathematics above the basic algebra level*. Mathematical tools are a necessary part of formulating a grammar for natural language, and of applying grammatical analyses to particular examples. Analysis of datasets requires the detecting of **recurrent patterns/structures**, which students have to express in formulaic expressions, using **variables**. Linguists call such formulaic expressions **rules**. Combinations of rules form **algorithms**, comparable to a computer program and, accordingly, grammar (as collections of rules) are **functions** that map an input representation intro an output representations. For example (in phonology), the word ‘electricity’ is derived from the smaller parts ‘electric’ (ending in the sound ‘k’) and the part ‘ity’. When ‘ity’ is attached to ‘electric’ the final ‘k’ must be changed into ‘s’. This is done by a rule. This is a simple example. In many languages, several rules have to applied in some specific **order**. Often rules apply to more than one of sound, such that students need to establish **sets** of sounds that undergo some operation. A further aspect of linguistic analyses is that the structure of **strings** (like sounds, or words) is represented in the form of **trees** (graph-theoretic objects) which reflect **grouping** and **hierarchical organization**. Student learn some of the formal properties of such tree structures such as ‘root node’, ‘terminal node’, ‘binary branching nodes’ etc. In the areas of syntax (sentence structure) and semantics (meaning), we encounter similar formal mechanisms.
- 2. *Mathematical concepts*. The concepts used in this course include sets, operations on sets, and relations between sets; functions and partial functions; formal logic; and informal discussion of algebraic structures.
- 3. *Mathematical manipulations*. Exercises. The weekly homework assignments will require students to apply the mathematical concepts listed in (1) and (2) as they solve novel problems. The exercises will be challenging, and will require the student to have mastered the formal tools. Please see the selection of sample exercises that has been submitted separately by Prof. van der Hulst.

Role of Grad Students : Each semester, one section will be taught by a faculty member, and the other sections will be taught by qualified doctoral students from our department. The homework exercises and in-class activities will be created by the faculty member and shared with the doctoral students for use in their sections. At least three additional graduate students will be needed to help with the following: Creation and grading of homework assignments; the creation of in-class activities; and the creation and grading of examinations.

Supplementary Information : For LING 110Q, which was submitted for inclusion in Group III (Science and Technology), the committee requested additional motivation. I believe that we have strengthened this motivation in the revised proposal that is attached.

LING 11Q, LING 205Q and 206Q were all three rejected for inclusion in the category of Q courses. Here we have not undertaken further action at the moment. That is, we have not refomulated the proposals. However, we do request reconsideration of the decision on the basis of the considerations below.

It seems to us that the criteria for accepting a course as a Q course under the new GenEd format are very strict and AS FORMULATED hard to meet in all details by the above mentioned courses. It almost seems as if only math courses can meet the criteria.

Here I wish to ask the committee to reconsider the rejection on the basis of the following considerations.

All three courses are generally experienced by students as involving a lot of 'difficult' formal metalanguage, i.e. the explicit formal tools that linguistS use to characterize linguistic representations. Especially, LING 205Q and LING 206Q offer advanced and in-depth formalizations, while LING 110Q uses these formalizations at a more basic level. The concepts and reasoning processes used differ somewhat from course to course, because they cover different areas, but we STRONGLY feel that each course crucially relies on the application of mathematical reasoning. Following the work of Chomsky in the late fifties, languages are understood and studied as a formal system in which utterances are 'propositions' that need to be validated as true (i.e. grammatical), given an underlying formal expression (the premise) and a system of derivational rules. Thus, grammatical descriptions are technically 'formal proofs'. The various steps in this proof (the linguistic derivation) are formalized and their interaction or ordering is studied in detail.

In addition, the representation of linguistic expressions themselves is heavily formal, relying crucially on tree-structures and feature systems whose properties are studied. In all areas of the grammar, students learn to see representations as expressions of a formal language in which a finite set of primes and a finite set of rules or constraints characterize an infinite set of expressions.

Even though there is no extensive graph-theoretical component in any of the courses, the use and study linguistic trees certainly involves important formal notions from this mathematical domain.

In all three courses, basic notions of set theory are used all the time (even though the student may not always be fully aware of it because it plays a role as part of the linguistic analysis, while not forming a goal in itself). In syntax and semantics (206), explicit formal notations that come directly from predicate and propositional logic play a major role. In phonology, on the other hand, a lot of insight is based on statistical results concerning the study of phoneme inventories. With a component of phonetics (in both 205 and 110) students are confronted with an elementary study of the physics of speech which involves basic insight into understanding spectrograms and various other graphic representations.

In short, these courses are what they are. They are not explicit math courses, abut, in our mind,qualify as involving a lot of 'applied math'. It is up to the committee, of course, to set the standards, while we on the other hand do not wish to artificially inflate the math content of the courses. However, we do believe that our courses have justifially and fruitfully been Q courses under the old 'regime'. Best regards, Harry van der Hulst