



**LUND**  
UNIVERSITY

Faculties of Humanities and Theology

## **HTXN07, Linguistics and Cognitive Science: Creating a Language Corpus, 7.5 credits**

*Språk- och kognitionsvetenskap: Att skapa en språkkorpus, 7,5 högskolepoäng*

**Second Cycle / Avancerad nivå**

---

### **Details of approval**

The syllabus was approved by the programmes director by delegation from the pro-dean for first- and second-cycle studies on 2013-01-14 to be valid from 2013-01-14, spring semester 2013.

### **General Information**

Subject: Language and Linguistics, Cognitive Science

The course is offered as an elective component of Master's programmes in Psychology, Cognitive Science or Language and Linguistics. It can normally be included in a first- or second-cycle degree.

*Language of instruction:* English

The course is normally taught in English, but if both the students and the lecturer agree, it may be taught in Swedish. Such an agreement would depend on both the lecturer and the students having a good knowledge of Swedish.

*Main field of studies*

-

*Depth of study relative to the degree requirements*

A1N, Second cycle, has only first-cycle course/s as entry requirements

### **Learning outcomes**

On completion of the course the students shall

#### **Knowledge and understanding**

- be able to explain what a corpus is and how it can be created
- be able to explain how different aims, research questions and data may affect the structure of a corpus

- be able to understand and account for the possibilities and limitations of corpus analyses

### **Competence and skills**

- be able to master and apply basic skills in transcription and coding according to at least one transcription standard
- be able to carry out a simple audio or video recording and link the material to language data in at least one transcription program
- be able to execute a few elementary linguistic analyses of the corpus material
- be able to independently plan one or several parts of the structure, coding or analysis of a language corpus
- be able to independently formulate research questions typical for treatment by corpus analysis

### **Judgement and approach**

- be able to evaluate whether a certain coding and analyse system can resolve a certain research question
- be able to discuss in speech and writing and take a position on ethical aspects of data collection, data storage and coding.

### **Course content**

The aim of the course is to provide students with practical and theoretical knowledge of how a language corpus is created. The course deals with questions of how to collect, transcribe, code and analyse language data. It also includes exercises in fundamental corpus analysis (e.g. counting the number of words, establishing predetermined search strings, identifying morphological and syntactic patterns, examining lexical measures). Furthermore, the course addresses a selection of texts and ethical aspects of data collection and coding. The emphasis is on spoken language data but other corpus types (written data, sign language) are also considered. The course provides students with a practical and theoretical introduction to different tools for transcribing and analysing the material.

### **Course design**

The teaching includes lectures and 4-6 compulsory laboratory exercises.

### **Assessment**

The assessment is based on an oral assignment during the course and a final written assignment.

*Subcourses that are part of this course can be found in an appendix at the end of this document.*

### **Grades**

Marking scale: Fail, Pass, Pass with distinction.

### **Entry requirements**

To be admitted to the course, students must have a Bachelor's degree including 60 credits in Linguistics, a language subject or Cognitive Science, or the equivalent.

### **Further information**

1. The course is offered at the Humanities Laboratory, Lund University.
2. The credits allocated for course content that in whole or in part is commensurate with another course can only be credited once for a degree. For further details see the current registration information and other relevant documentation.

## Subcourses in HTXN07, Linguistics and Cognitive Science: Creating a Language Corpus

Applies from V13

1301 Creating a Language Corpus, 7,5 hp  
Grading scale: Fail, Pass, Pass with distinction