



Department of Linguistics

Prof. W. Detmar Meurers

222 Oxley Hall  
1712 Neil Avenue  
Columbus, OH 43210-1298

Phone (614) 292-0461  
FAX (614) 292-8833  
Email dm@ling.osu.edu

Kathleen Hallihan  
Director, Curriculum and Assessment  
Colleges of the Arts and Sciences

October 24, 2007

Dear Kate,

thank you for the feedback from the Humanities College and Curricular Committee regarding our Linguistics 484 "Code Making and Code Breaking" course proposal. We're happy to clarify the issues raised below.

- *Please expand on rationale for course: How does this course fit into Linguistic department? (i.e. Why is Linguistics the right place for such a course)*

This is a course which will be taught by members of one of the strongest computational linguistics programs in the country. The perspective of language as a code originates in information theory, which is at the heart of modern computational linguistics. One of the first traces of this in the literature is in a memo written by Warren Weaver in July 1949, spelling out a position which led to the first steps towards automatic machine translation, and which actually arose directly from cryptographic efforts in the Second World War. While just one of many perspectives on language, Weaver's view is among the most influential in shaping the direction of computational linguistics, and is having increasing impact in other areas of linguistics.

- *Writing assignment – why Korean (as opposed to other languages)? Because it is part of the textbook? Please clarify. There was some concern expressed about teaching a "language" as a "code" to be deciphered.*

Korean is distinguished by a particularly elegant writing system, designed on mathematical and phonetic principles by a group of scholars around and possibly in-

cluding King Sejong the Great (1397–1450). When seen as a code, Korean writing has three virtues:

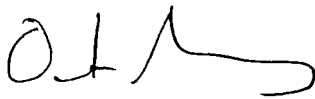
1. It challenges Anglocentric assumptions about how letters correspond to sounds: hangul signs stand for a combination of consonants and vowels;
2. it is regular and systematic enough to be the basis of a fair code breaking assignment;
3. it prepares the way for the reasoning needed by possible later courses in phonology, which require flexible thinking about how sounds can be combined together into larger units.

This module does not actually teach Korean, or claim that Korean really is more or less like a code than any other language. It uses the nice properties of the writing system to teach the students about how to approach linguistic problems by thinking about patterns and regularities. These skills get re-used over and over in the rest of the course. For example, the fact that consonants and vowels tend to alternate can provide a way in to many of the simple ciphers that we study in the course.

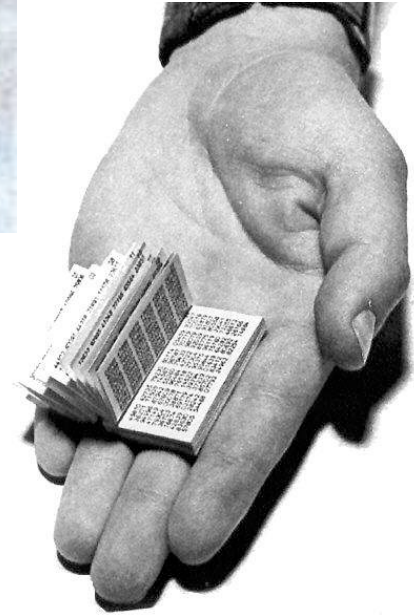
- An updated syllabus, including clarification of the grading scale, and the detailed final project description is attached to this letter.

I hope the above information addresses the relevant issues. Thank you for considering this updated new course request.

Sincerely,



Detmar Meurers  
Undergraduate Studies Director, Linguistics



## Linguistics 484

This course has two main aims: it introduces some of the old and new technology associated with codes and code-breaking and it discusses ways in which codes have made, are making and might make a difference to peoples' lives.

## Course Book

The textbook is *The Code Book*, by Simon Singh. You should buy this and expect to read all of it as background to the course. There will be some overlap between the technical material of the course and that presented in the book, but there will be material presented in class that is not covered at all in the book.

## Course Objectives

Students in Linguistics 484 will have an opportunity to:

- Acquire a thorough knowledge of the fundamental terminology, concepts and techniques of cryptology.
- Learn some of the history of codes, and their importance, both from the point of view of the code user and the code breaker.
- Develop an understanding of what a cryptanalyst looks for when trying to break a code.
- Gain experience in problem solving, in synthesizing ideas, and in writing reports.

## Instructor Details

Chris Brew  
200A Oxley Hall  
1712 Neil Ave  
Columbus Ohio 43210-1298  
Email: will be disclosed in class  
Office hours: 4pm-5pm TWR

## Class location

Bolz Hall 314: 10:30-12:18 Monday Wednesday

## Topics

### Codes

- Monoalphabetic ciphers: Caesar cipher, keywords
- Polyalphabetic ciphers: Vigenère cipher
  - Transposition ciphers
  - Polygraphic ciphers: Playfair; Hill Cipher
  - Perfect ciphers: The one-time pad.
  - Enigma: the technology

### Linguistic Codes

- Linear-B: Decoding Ancient Texts
- Hangul: Korean Writing

### Codes and Intelligence in War

- Enigma: the intelligence
- Exploiting Intelligence from Cryptography

## Assessment

There will be regular short code-breaking assignments. To succeed on these you need to attend the classes, and make a serious attempt to solve the codes. There will also be in-class quizzes on the readings from *The Code Book*. There will be a mid-term exam testing technical material and a final project that will involve a 5-page write up of a piece of independent work. There may be occasional extra credit opportunities

### The Final Project

The final project is group-based and requires you to do four things:

1. Design a cipher system that strikes a good compromise between usability and security.
2. Prove to me that you can use it,
3. Make a serious attempt to break the system created by one of the other groups.
4. As a group, write a well-organized and clear report on the things that you did.

Component	Score
Weekly assignments	50 points (10 at 5 points each)
Quizzes	5 points (5 at 1 point each)
Mid-term	20 points
Final project	20 points
Class participation	5 points
Available extra credit	5 points

The table below shows the connection between grades and point scores. If you make the point score, you can count on getting **at least** the grade listed. I reserve the right to give a higher grade if this is warranted by something you do in the course, but I will never give a lower grade.

As a rough guide, you can get a B by simply learning and understanding the material I teach in the course. To get an A you need to do that, but also show some originality.

Grade	Point Range
A	94-100
A-	90-93
B+	86-89
B	83-85
B-	80-82
C+	77-79
C	74-76
C -	70 – 73
D+	64-69
D	60-63
E	0-59

### **Your responsibilities**

All class members are responsible for

- Keeping up with the assignments and reading
- Monitoring your own progress and understanding of the material. If there is something you don't understand, please do ask, preferably in class.
- Contributing to class discussion.
- Helping to form a "course community". This includes responding appropriately and helpfully to other class members.

### **Academic Misconduct**

It is the responsibility of the Committee on Academic Misconduct to investigate or establish procedures for the investigation of all reported cases of student academic misconduct. The term "Academic misconduct" includes all forms of student academic misconduct wherever committed; illustrated by, but not limited to, cases of plagiarism and dishonest practices in connection with examinations. Instructors shall report all instances of alleged academic misconduct to the committee (Faculty Rule 3335-5-487). For additional information, see the Code of Student Conduct ([http://studentaffairs.osu.edu/resource\\_csc.asp](http://studentaffairs.osu.edu/resource_csc.asp)).

Cheating is wrong, wastes your time and ours, and will not be tolerated. Working together to find the answer is fine, but talking to someone who has already figured out the answer is cheating. You must also do your homework by yourself unless it is specifically designated as group work. We will assume that you are honest, but if we are confronted with clear evidence of cheating, it is our duty to take action.

### **Students with Disabilities**

Ohio State is committed to extending access and opportunity to those who are disabled. Any student who feels s/he may need an accommodation based on the impact of a disability should contact me privately to discuss your specific needs. You may also contact the Office for Disability Services at 614-292-3307 in room 150 Pomerene Hall.

### Sample schedule (from Spring 2007)

The table below indicates roughly when which piece of the course will be covered. Things may change as the course develops.

Week	Dates	Topic	Notes
1	Mar 26, 28	Monoalphabetic ciphers	Singh ch 1
2	Apr 2, 4	Polyalphabetic ciphers	Singh ch 2
3	Apr 9, 11	Decoding ancient languages	Singh ch 5
4	Apr 16, 18	Polygraphic ciphers	Sinngh ch 3
5	Apr 23, 25	Enigma: the intelligence	Singh ch 6
6	Apr 30, May 2	Transposition ciphers	Midterm May 2nd
7	May 7, May 9	Korean writing	Singh ch 7
8	May 14, May 16	Perfect ciphers	Singh ch 8
9	May 21, May 23	Enigma, the technology	Singh ch 4
10	May 30	Exploiting intelligence	Memorial Day
Exam	Jun 6	Final project due	



Student's Names:

Course Name: Linguistics 294L

Due in class M 21,W 23 May 2007

Teacher's Name: Chris Brew

---

## Code design challenge

This is the final project for this class. Your task is to design a cipher system that strikes a good practical balance between security and usability. Since this is a complex multi-stage assignment, I first lay out the steps. The I give some advice about how to approach this. The overall goal is to develop your skills in executing and analysing cryptographic processes

### The assignment

The set-up is the following:

- You will be working in groups of four or five. This is the smallest possible size, because I will need to split the group in two, and I do not want anyone working alone.
- Your first task is to design a cryptosystem that is reasonably secure. See below for a quick summary of the tools you have and the design considerations you will want to consider. The system must use some kind of simple shared secret (a key-word, key phrase or key number). I'll call this shared secret the key from now on.
- As a group, prepare a document that clearly describes the cryptosystem, giving enough detail that it will be possible for another group to understand and use the system, without further help. You should make two versions of the document. The first version not only describes the system but also gives two keys that can be used with it. Label the keys A and B. Version two is the same, except you don't give the keywords.
- On 21 May, we will trade system descriptions between groups. You will give feedback to the other group on (a) whether the write-up is clear enough for you to use (b) what you think of the cryptosystem itself. If the answer to (a) is "Not clear", then you may not be able to say much about (b).
- On 23 May, at the start of the class, I will split the groups in two and hand each half a message to be encoded. This will be about 200 characters in length. You will have 30 minutes to encode this message according to your cryptosystem. You need to check ahead of time that your system is efficient enough to allow you do this in the available time. When you are coding and decoding, the only external material you are allowed to refer to is version two of document that you created in the previous step. That is, you have to remember the keyword.
- Next, you will trade messages with the other half of your group, and decode. Again, you will have 30 minutes. Do not write directly on the cryptotext, because this is going to be photocopied and handed to another group after the class period is over.
- Next, I will collect in the coded and decoded messages, and you should use the remaining class time to reflect on what, worked, what didn't, and what you want to change. Maybe the system isn't as usable as you hoped. How will you fix it? The first part of the assignment for May 30th is to create a revised version of your document that fixes any problems that have emerged with the first go-round. Again you need two copies. One that just describes the system and one that additionally includes information about the shared secret

- Finally, I will give each group two 1000 character messages. Do not share this message with the other half of your group. Use your revised cryptosystem to encode the message and bring to class three sealed envelopes. The first envelope contains the cryptotext and the revised system description, but no keywords. This is for the other half of your group, The second envelope should be similar to the first, except that you must give two seven to ten-letter words from the message as a clue. The choice of these words is up to you, but you must give them, This envelope will be given to another group to try to break. The third envelope contains the cryptotext, and the revised system description, but this time the key information is provided. This is also for another group, but this time that group's task is to do an authorized decryption. This tests how usable the system really is.
- Each group now has three messages to work with. One is from the other half of your group. Decode it. A second is from another group, but with known key and system. Decode it. The third has a known system but unknown key. Try as hard as you can to break it. The final project report is a four page write-up of what you did with these three messages. It is due on June 6 (the Wednesday of exam week).

## Design Considerations

You know about several kinds of cipher and several possible attacks on these ciphers. You also might be able to use ideas from the work on Korean, Linear B and Hieroglyphics, such as encoding words in syllables, or something similar. As long as you can explain it clearly in your system description, anything you choose to do is fair game.

### Cipher types

Shift- Monoalphabetic - Polyalphabetic (e.g. Vigenere) - Polygraphic (e.g. Playfair, Beaufort) - Transposition.

You could use any of these, or a combination of one or more. The more complicated you make it, the harder it is going to be to use effectively under time pressure and the greater the risk of error. But if you can describe it, and you have tested it to see if you can use it, do what you want.

### Attacks

Simple frequency analysis - frequency analysis on letter pairs or triples - use of possible words - use of tables of words and their patterns (as in the Army field manual approach to foursquare, or as in use of possible words in polyalphabetic).

You should say in your system description which features of your system are designed to defend against which types of attack, and explain why you think the system is going to be effective.

### Step by step

You **MUST** include a detailed, step-by-step description of how the processes of encryption and decryption work. This needs to be good enough that another class member can do what you say.