Computation, Information, and Intelligence (COMS/ENGRI/INFO/COGST 172), Fall 2005

11/11/05: **Lecture 32 aid — Nearly blank-slate segmentation of Japanese**

Agenda: Consider whether simple statistics drawn from massive amounts of not immediately-obviously helpful data can be used to accomplish a difficult NLP task.

Announcements: You might find it useful to employ the following "sanity checks" upon your solutions to Homework Five. You are not required to include the results in your writeups, but if doing so would make your solution clearer, you are encouraged to put them in.

- Question 2: we gave some example sentences that your grammar should or shouldn't generate. You might want to explicitly check your grammar against those examples.

- Question 4: various normalizations guarantee that certain quantities should sum to one. Verifying that such sums are correct might help in proof-reading your answers. Also, you might want to check that at least some of the probabilities and weights you arrive at make intuitive sense.

- Question 5: since you should be able to figure out what the correct parse(s) for the given sentence are just by looking at the grammar, you might want to check that at least some of the parse states you derive are consistent with the parse(s) you know about.

Follow-ups:

- A ten-second sound clip from the Saffran, Aslin, and Newport (1996) training data is available here: http://oak.psych.gatech.edu/~spieler/intro/stream.wav . If you download it and replay it in repeat or loop mode, you should hear "words" emerging.

- A movie clip of an infant undergoing a very similar experiment to what we described is available here: http://www.waisman.wisc.edu/infantlearning/media/Language.mpg . The duration of fixation is very clearly demarcated.

- The "two cups" experiment was introduced by Jean Piaget, and is known as the "A not B" task. Diamond pointed out the flaw in using lack of object permanence as an explanation. (References from Eliot (1999), *What's going on in there? How the brain and mind develop in the first five years of life.*)