

Toward the Logical Description of Languages in their Phonemic Aspect

Written at Harvard in 1952 with a couple of mathematicians. Read quickly for general notions, not the mathematics and communications theory as it stood then.

452) notion of Markoff process. This is a stochastic process in which the probabilities depend on the previous events. This fits phonology, as the successive chain of elements determines what may or may not follow.

453) Fig 1. If you have 8 objects – a power of 2 – the number of yes/no answers needed to identify each of the objects in the set is $\log_2 8$. Look at the figure. First you divide them into two halves, 4 yes and 4 no. Then each half subdivides to yes and no, then each half after that. The three questions distinguish each of the 8 as a chain of three yes/no decisions.

Natural languages of course do not work this way; they have redundancies and pose some questions that only a few of the ‘objects’ need answer.

454) Phonemes of Russian need 11 features. Some do not need to be specified at all. Look at the *j*, for example, at the very end. The two minuses already distinguish it from all the others. It does not matter what the other features are.

Look at the inventory. Left to right, compact C’s start off the columns. The comma means palatalized. The fifth from left, *c*, is a weird symbol for tch or č. The next two are š and ž. Note that none of the compact C’s need be specified for diffuse (vs. vowels). None need nasal, as there is no nasal, nor strident, as some are only redundantly strident, nor stress, as only V’s are stressed. All need continuous, and the stops need sharp specification. *x* and č do not need voice, as there is no opposition here. They all have to have the tonality opposition grave-acute.

Here is how they divide up. The five velars are a group as grave, the three palatals as acute (šč is not considered, as a double consonant). Of the grave ones, only *x* is continuous, so it is complete. *g* and *g*, are voiced and oppose each other in sharpness, while *k* *k*, are voiceless and oppose each other in sharpness.

Try to work out the dentals and labials, and the vowels, by yourself. Note that the fifth phoneme from the right on the top is actually what we usually write *c*, the dental affricate (ts). What is interesting about this affricate in its specification?

456) Removal of non-specified features. A recoding. See bottom of page and top of 457. Why are some of the chains shorter than others? Note the dental affricate *c*, or ts, labeled *s* with a circumflex. It is fully specified.

461) frequency tables here are interesting.

