

LING 106: Homework 2

1. LEXICAL ANALYSIS: WHAT LANGUAGES ARE

- a. Sentences that are descriptively ungrammatical, in any dialect of English.
- b. Sentences that are descriptively grammatical (in some dialect) but prescriptively ungrammatical
- c. Sentences that are syntactically well-formed but semantically anomalous.
- d. Sentences that are syntactically and semantically well-formed, but which are hard to understand for processing reasons.

Answers here were somewhat varied; I won't go into any detail here.

2. FORMAL DEFINITIONS OF LANGUAGE

2.1. *Deriving a language from its grammar*

Lexicon = { <a, VOW, ...>, <o, VOW, ...>, <i, VOW, ...>, <p, CON, ...>, <t, CON, ...>, <n, CON, ...> }

Syntax: a string x is a sentence of W if x has the form CON VOW CON, where the first and last lexical items are not the same.

$W = \{ \text{pat, pan, pot, pon, pit, pin, tap, tan, top, ton, tip, tin, nap, nat, nop, not, nip, nit} \}$

2.2. *Deriving a grammar from a language*

chan yana tho,
chan yana yana tho,
chan yana yana yana tho,
chan yana yana yana yana tho,
...
chan yana yana ... yana yana tho, ...

Lexicon: { chan , yana , tho }
Syntax: <i>S</i> is a sentence of <i>D</i> iff it consists of chan followed by one or more instances of yana followed by tho .

That's the easiest way to do it; if you want to phrase it in terms of syntactic categories:

Lexicon: {< chan , START, ...>, < yana , MIDDLE, ...>, < tho , END, ...>}
Syntax: <i>S</i> is a sentence of <i>D</i> iff it consists of a START followed by one or more instances of MIDDLE followed by END.

3. THE DISTRIBUTIONAL METHOD

3.1. *Welsh*

- (1) **a charnedawr tra gogyhwc gwyr trin**
“And there was raising large cairns over the heroes of toil.”
[etc.]

This is a straightforward “sound/meaning” correspondence question. The word gwyr appears in all six lines, and the only thing the translations have in common is the word “heroes”, or “warriors” or “men”. Thus, gwyr probably means “heroes” or “men”.
--

Ideally: one should also observe the problem of using a 150-year-old poetic translation, which makes it vague as to whether Williams interprets a word meaning “men” to refer to heroes, or whether he poetically understates a word meaning “heroes” as referring merely to men. (Remember, your analysis of the data can only ever be as good as your data.)
--

3.2. Finnish

3.2.1. Distributional information

In each pair, there's a string of six words that differ only by a single word somewhere in the middle. Those are bolded and underlined below. The apparent interchangeability of the italicized words in each pair suggests that they're words in the same syntactic category.

- (7) a. ...istumaan ja halaili äitiä iloissaan, **vaikka ei tiennyt syytä minkä tähden** hän halaili.
b. Vaan sen puheen johdosta se kuitenkin tuntui olevan, **vaikka ei tiennyt oikein minkä tähden** se oli.
- (8) a. Ja jos niihen nyt muutaman tai parin satamarkkasta panee, **niin se ei ole mitään, kun** niissä on ijäksi ja vielä pojan pojallekin kalua, kun ne hyvät laittaa.
b. Menkään se vaikka Matin pellon ojaan, **niin se ei merkitse mitään, kun** se kerran on näiksi lähtenyt.
- (9) Samoin on laita, **jos toisen paha sana vastataan rakkauden** sanalla, niin sen viha muuttuu rakkaudeksi, **jos toisen paha teko vastataan rakkauden** työllä, niin se vaikuttaa rakkauden ja viha sammuu.
- (10) ...että lähetän hänelle rakkaan tervehdyksen ja että kiitän häntä **kaikesta siitä hyvästä, mitä hän minulle** on tehnyt, ja **kaikesta siitä kunniasta, mitä hän minulle** on osoittanut lapsuudesta pitäin...

Note that for (10), the following is not correct:

...että lähetän hänelle rakkaan tervehdyksen ja että kiitän häntä kaikesta siitä hyvästä, **mitä hän minulle on tehnyt**, ja kaikesta siitä kunniasta, **mitä hän minulle on osoittanut** lapsuudesta pitäin...

Harris's condition requires an *X* and a *Y*, that is, it requires text on *both* sides of the words you're testing. Using two words that can both *follow* the same string would give you cases like:

- (10') a. You may imagine the young people brushed up **after the labours of the day,** ***and*** making this novelty...
- b. ...working men were gardening **after the labours of the day,** ***children*** were being put to bed...
- (10'') a. He went to bed a little after midnight, and was awakened from lurid dreams **in the small hours of Monday** ***by*** the sound of door knockers...
- b. London, which had gone to bed on Sunday night oblivious and inert, was awakened, **in the small hours of Monday** ***morning***, to a vivid sense of danger.

Thus we would conclude that *and* and *children* have the same syntactic category, as do *by* and *morning*.

3.2.2. Using Harris's Conditions

- (11) a. ...ajatteli alottaa jotakin puhetta sen kirjeen johdosta, **mutta ei tiennyt mitä oikein sanoisi.**
- b. Hannan mielessä tuntui, että pitäisi siihen jotakin sanoa, **mutta ei selvinnyt, mitä oikein sanoisi.**

As above, the identical contexts suggests that *tiennyt* and *selvinnyt* are the same category. The pair that helps prove that these are non-monomorphic is (b) and (d), which contain the words *tiessä* and *selvissä*. This suggests that *tie-* and *selvi-* are part of a class that can come before a variety of suffixes such as *-nnyt* and *-ssä*.

It's by no means a perfect Harris context—we'd want to see *tiessä* and *selvissä* in the same context as *tiennyt* and *selvinnyt*, or at least in the same context as each other—but there is at least the ABCD of ***tie-* / *-ssä* // *selvi-* / *-nnyt***, which can be swapped.

Note that the other four sentences contain (what may be) other forms of these words:

selvi-ämättömään, selvi-ttäen, tie-tä, tie-lle

These sentences aren't *as* helpful, but they can be invoked in a discussion of a Harris-style class of things that the roots can come before, i.e. {***-ssä, -nnyt, -mättömään, -ttäen, -tä, -lle***}, and how one would now want to find evidence of ***tiettäen, selvitä***, etc.

But giving, as an answer, that ***selvissä*** and ***selvittäen*** are helpful isn't really enough here. After all, the words ***preacher*** and ***pretty*** start with the same string as ***precook***, but they're not instances of the same morpheme. You need a class of things that *both* ***tiennyt*** and ***selvinnyt*** can appear before.