



Linguistics Challenge: Solutions

Note: Solutions to problems from the Linguistics Challenge website (<http://webscript.princeton.edu/~ahesterb/puzzles.php>) are not posted.

Problem 1: Agta (2.5 points (0.5 points/answer))

This problem is found on the Linguistics Challenge website:
<http://webscript.princeton.edu/~ahesterb/puzzles.php>.

Problem 2: Tokyo Japanese (4 points (1 point/answer))

- a. yaé + sákura → yae-sákura
“many” “cherry tree” “a fully-blooming cherry tree”
- b. kúmo + kakuré → kumo-kákure
“cloud” “hide” “hidden behind a cloud”
- c. kúi + jidóusha → kei-jidóusha
“light” “car” “compact car”
- d. oó + kámakiri → oo-kámakiri
“big” “praying mantis” “large praying mantis”

Notes on the solution: For compounds where the second word is at least three syllables: if the second word’s accent is on the last two syllables, accent the first syllable of the second word, else keep the second word’s accent. The first word’s accent position doesn’t matter.

This problem was written by Mitcho Erlewine.

Problem 3: German (5 points (0.5 points/answer))

- | | | | | | |
|----------|----------|-------------|------------|----------|-----------|
| a. sechs | <u>c</u> | “six” | f. Elch | <u>c</u> | “elk” |
| b. Docht | <u>x</u> | “wick” | g. Kelche | <u>c</u> | “goblets” |
| c. Stich | <u>c</u> | “stab” | h. Suche | <u>x</u> | “search” |
| d. Sucht | <u>x</u> | “addiction” | i. Pflicht | <u>c</u> | “duty” |
| e. Ochse | <u>x</u> | “ox” | j. sprach | <u>x</u> | “said” |

Notes on the solution: *Ch* is pronounced as *x* after *a/o/u* (called *back* vowels) and as *c* after *e/i/l* (i.e. elsewhere).

This problem was written by Peter Graff.

Problem 4: Constituents (5 points)

$$\frac{2N-1}{\binom{N}{2}+N} = \frac{2N-1}{\frac{N(N-1)}{2}+N} = \frac{2(2N-1)}{N(N-1)+2N} = \frac{4N-2}{N^2+N}$$

Notes on the solution: With a binary tree, there are N terminal vertices and $N - 1$ branching vertices, so $2N - 1$ vertices total and therefore $2N - 1$ unique constituents. The total number of subsentences is N (the number of single-word subsentences) + $\binom{N}{2}$ (the number of multi-word subsentences = choosing a first word and last word).

This problem was written by Mitcho Erlewine.

Problem 5: Turkish (8 points (2 points/answer))

This problem is found on the Linguistics Challenge website:

<http://webscript.princeton.edu/~ahesterb/puzzles.php>.

Problem 6: Maasai (10 points (1 point/answer))

This problem is found on the Linguistics Challenge website:

<http://webscript.princeton.edu/~ahesterb/puzzles.php>.

Problem 7: Nahuatl (Aztec) (12 points (2 pts/answer))

This problem is found on the Linguistics Challenge website:

<http://webscript.princeton.edu/~ahesterb/puzzles.php>.

Problem 8: Niuean (15 points (3 points/answer))

- a. John swam. Ne koukou a Sione.
- b. You will eat a dog. To kai e koe e kuli.
- c. Pule is abandoning you. Kua liti e Pule a koe.
- d. A bird will see a boy. To kitia he manu e tama.
- e. A dog is flying. Kua lele e kuli.

Notes on the solution: The first word in each sentence marks tense: *ne* = past, *kua* = present, *to* = future. The tense is followed by the verb, then the subject, and then the object (if any). Each noun is preceded by *a*, *e* or *he*. In sentences with both a subject and an object, *e* precedes subjects that are proper names or pronouns, and *he* precedes all other subjects. An object is preceded by *a* if it is a proper name or pronoun, and *e* otherwise. In sentences with only a subject, the subject behaves like an *object* (*a* for names and pronouns; *e* otherwise). Such a correspondence between objects and subjects is called *ergativity*.

This problem was written by V. I. Belkov for the Moscow Linguistics Olympiad.

Problem 9: Hmong Daw (17 points (4pts/4pts/4pts/5pts))

- a. noog “bird”
- b. cw “lobster”
- c. hais lus “to speak”
- d. qhov muag kiv “prone to dizziness”

Notes on the solution: In Shong Lue Yang’s writing system, syllables are written left-to-right, but letters within a syllable are written right-to-left. Tones are marked with symbols above the vowel:

| Shong Lue Yang | missionaries |
|----------------|------------------|
| line | s |
| two circles | g |
| one circle | <i>no letter</i> |
| <i>no mark</i> | v |

Combinations of consonants and combinations of vowels each correspond to a single special symbol. The consonant *k* is not marked in any way.

This problem appeared in the 2002 Moscow Linguistics Olympiad.

Problem 10: Egyptian Arabic (24 points (2 pts/cell))

| n | 1/n (singular) | 1/n (plural) |
|----|----------------|--------------|
| 3 | tult | itlāt |
| 4 | rubī | irbāī |
| 5 | xums | ixmās |
| 6 | suds | isdās |
| 7 | subī | isbāī |
| 8 | tumn | itmān |
| 9 | tusī | itsāī |
| 10 | īušr | īšār |

Notes on the solution: Egyptian Arabic nouns come in three forms that vary with number: singular, dual, plural. The dual form, used when you have exactly two of a certain item, is just the singular form with the suffix $-ēn$, but the plural form is quite different. Bizarrely, the plural form is used only for things of indeterminate number (e.g. the bananas) and for three to ten (three bananas, four bananas...). After ten, they go back to using the singular form again (eleven banana, twelve banana...)!

The difference in forms between the singular and plural nouns illustrates another interesting fact about Arabic and related languages such as Hebrew. Arabic morphology is said to be based on roots and patterns. The roots generally consist of three consonants and indicate a broad class of meaning, for example tlt for ‘three’ and $rbī$ for ‘four’. Words with the root ktb are all related to writing. For example, *kataba* means ‘he wrote’, *maktūba* means ‘it was written’, *kitāb* means ‘book’, *maktab* means ‘office’, etc. The pattern is what is left after taking out the three consonants. For example, the pattern for forming a word meaning the place where some action usually takes place is $maCCaC$.

All Arabic words in this problem are constructed according to one of the following patterns, where $1, 2, 3$ correspond to consonants: $1u23$, $1u23ēn$, $i12ā3$, $1a2a3t$. Each particular combination of three consonants corresponds to a number n between 3 and 10. The pattern $1u23$ is used for $\frac{1}{n}$; $1u23ēn$ indicates $\frac{2}{n}$; $i12ā3$ is used for $\frac{1}{n}$ in the plural, with a numerator > 2 ; $1a2a3t$ represents n .

Adapted from a problem for the First International Olympiad in Theoretical, Mathematical and Applied Linguistics written by Ivan Derzhanski.