

Informatics 1 Cognitive Science – Tutorial 1 Solutions

Frank Keller, Carina Silberer, Frank Mollica

Week 2

The goal of this tutorial is to familiarise yourself with terms and ideas about vocabulary and grammar and to start thinking about language.

1 Vocabulary and Grammar

Linguistics (the scientific study of human language) is an important part of cognitive science, and understanding language can help us understand more general phenomena in cognition. The following tasks will address the aspects of language you heard about in class and will give you the opportunity to do some practical exercises to get a better understanding of them.

1.1 Words and Rules

Many cognitive scientists believe that human linguistic knowledge consists of two different kinds of ‘mental tissue’:

- a **lexicon**, containing **words**
- a **grammar**, containing **rules**

Here is an example lexicon:

<i>John</i>	is a	proper noun
<i>Mary</i>	is a	proper noun
<i>held</i>	is a	past tense verb
<i>caught</i>	is a	past tense verb

And here is an example grammar:

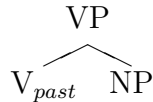
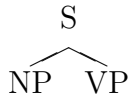
a sentence	can consist of a noun phrase followed by a verb phrase
	$S \rightarrow NP VP$
a verb phrase	can consist of a past tense verb followed by a noun phrase
	$VP \rightarrow V_{past} NP$
a noun phrase	can consist of a proper noun on its own
	$NP \rightarrow PN$

Exercise 1

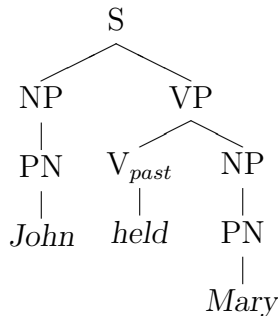
1. Draw out the lexicon as trees, like the ones in lectures.

PN	PN	V	V	V_{past}	V_{past}
<i>John</i>	<i>Mary</i>	<i>hold</i>	<i>catch</i>	<i>held</i>	<i>caught</i>

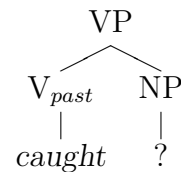
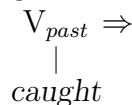
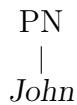
2. Likewise, draw out the grammar rules as trees.



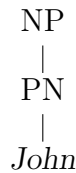
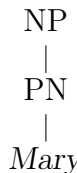
3. Use the grammar and lexicon to show that *John held Mary* is a grammatical sentence.



4. Show that **John caught* is **not** a grammatical sentence.



5. Show that **Mary catch John* is **not** a grammatical sentence.



\Rightarrow There is no rule in the grammar covering base verbs (V).

6. How many sentences are grammatical, according to this grammar and lexicon?

Mary caught John. John caught Mary. Mary caught Mary. John caught John. Mary held John. John held Mary. Mary held Mary. John held John.

\Rightarrow 8 sentences

1.2 Productivity & Reuse

Words and Rules theory embodies two broader principles that we discussed in class: Productivity and Reuse. Lexical entries are memorized chunks that we consistently reuse. They are productively combined using rules to create phrases and sentences. In English, it's fairly straightforward to distinguish lexical entries as they are often separated by spaces when written and there are not too many important prefixes and suffixes.

For this exercise, let's consider Swahili, a language spoken in East Africa. In Swahili, a single typed "word" can translate to a full sentence. However, these sentences are still the result of productivity and reuse. Take a look at the following data:

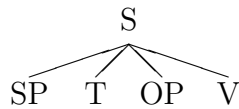
Exercise 2

1. Identify the parts of the words that are frequently reused—i.e., the lexicon.
See table below.
2. For each lexical entry, identify its "part of speech" (Verb/Subject/Object pronoun/Tense).
Verb: lipa, penda, sumbua, piga
SubjectPronoun (SP): ni, tu, u, wa, a
ObjectPronoun (OP): ku, m, wa, tu, ni

atanipenda 's/he will like me'	atanipiga 's/he will beat me'
atakupenda 's/he will like you'	atakupiga 's/he will beat you'
atampenda 's/he will like him/her'	atampiga 's/he will beat him/her'
atatupenda 's/he will like us'	ananipiga 's/he is beating me'
atawapenda 's/he will like them'	anakupiga 's/he is beating you'
nitakupenda 'I will like you'	anampiga 's/he is beating him/her'
nitampenda 'I will like him/her'	amekupiga 's/he has beaten you'
nitawapenda 'I will like them'	amenipiga 's/he has beaten me'
utanipenda 'you will like me'	amempiga 's/he has beaten him/her'
utampenda 'you will like him/her'	alinipiga 's/he beat me'
tutampenda 'we will like him/her'	alikipiga 's/he beat you'
watampenda 'they will like him/her'	alimpiga 's/he beat him/her'
wametulipa 'they have paid us'	atakusumbua 's/he will annoy you'
tulikulipa 'we paid you'	unamsumbua 'you are annoying him/her'

Tense (T): me, na, ta, li

3. The data can all be explained by a single rule. Draw out that rule as a tree.



4. Use the grammar and lexicon to provide the Swahili words for *You are annoying me* and *I am paying him*.

You are annoying me: u-na-ni-sumbua

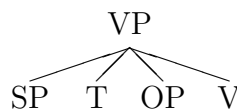
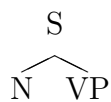
I am paying him: ni-na-m-lipa

5. How many sentences are grammatical, according to this grammar and lexicon?
 $5SP \cdot 4T \cdot 5OP \cdot 4V = 400$ sentences.

6. In Swahili, *the lion has chased me* translates to *simba alinifukuza*. Expand the grammar and lexicon to account for this new data.

Simba is a noun.

-fukuza is a verb.



ni- I	tu- we	-me- past tense
-lipa pay	-penda like	-na- present tense
a- s/he	-sumbua annoy	-ta- future tense
-ni- me	-m- him/her	-li- present perfect tense
-piga beat	wa- they	u- you (subject)
-tu- us	-wa- them	-ku- you (object)

Let's take a look at another language to get some more leverage on productivity and reuse. While it often seems that we have an English word for every animal in the world, there was a point when we first encountered an animal and did not have a word for it. In English, we typically either borrow or create an arbitrary and conventionalized word (e.g., axolotl). In scientific communities, there are conventionalized rules for how to name a new animal.

Take a look at this sentence in Mohawk, an indigenous language to Northeastern US/Canada (Figure 1) which shows how Mohawk speakers refer to porcupines.

<i>Nek</i>	<i>tsi</i>	<i>akitshé:nen</i>	<i>aneñ:taks</i>	<i>è:rhar</i>	<i>raotíkhwa?</i>
<i>just</i>	<i>that</i>	<i>my-pet</i>	<i>it-evergreen-eat-s</i>	<i>dog</i>	<i>their-food</i>
	<i>wahí:nonte?</i>	<i>tsi</i>	<i>wahatehià:ron.</i>		
	<i>past-I/him feed</i>	<i>that</i>	<i>past-he-grow</i>		

I raised my pet porcupine on dog food.

Figure 1: The top line is the Mohawk sentence in Latin script. The bottom line reflects a word-by-word gloss. The translation is given at the end. Taken from Mithun (1984).

Exercise 3

- Based on Figure 1, does Mithun analyse *porcupine* as a morphological object or a psychological one?
Morphological object because it's a linguistic description, not a processing claim.
- The Mohawk word for *porcupine* provides insight about the productive process involved in naming the animal. Do you think Mohawk speakers store the full word or productively construct the word as needed? Why?

There is no data on the frequency with which Mohawk speakers deal with porcupines so there is no answer but it's a useful discussion. Here are some relevant points:

Productively construct:

- Porcupines may be really low frequency. Why store a word that you'll probably use once in a lifetime.
- If porcupines are rarely in the environment being productive allows you to adapt your reference. For example, would a porcupine still be named after eating evergreens when there are no evergreens around because of seasons?

Reuse:

- It takes effort to productively name something. If porcupines are frequent, why would you want to constantly repeat the computation?
- If we store porcupine as a lexical entry, we also fix the conventionalized expression. This way we can talk about porcupines even in winter when there are no evergreens and when porcupines aren't physically present to provide context.