<1> Get Edumacated![1] (1/2)   (30 points)

"Homeric infixation is a morphological construction that has recently gained currency in Vernacular American English. People who are familiar with this construction invariably credit the TV animation series, *The Simpsons,* particularly the speech of the main character Homer Simpson, for popularizing this construction." (Yu, A.C.L. 2004. Reduplication in English Homeric infixation. *NELS 34*)

Many speakers of American English, particularly younger generations, can insert the syllable "*ma*" into a word (like "edumacation" or "saxomaphone") to produce a humorous variant. For many words, everyone agrees on how the "edumacated" variant should be formed, but there's some disagreement, too.

Below, three people give what they feel are the correct "edumacated" versions of twelve words. We've capitalized the stressed syllables of the respondent's answers. You should likewise indicate stress with capitalization in your answers.

	Alan	**Barbara**	**Chris**
Alabama	*ALamaBAma*	*ALamaBAma*	*ALamaBAma*
capital	*CApimaTAL*	*CApimaTAL*	*CApimaTAL*
captain	*CApamaTAIN*	*CAPtamaTAIN*	*Uh... I'm not sure.*
congratulations	*conGRAtumaLAtions*	*conGRAtumaLAtions*	*conGRAtumaLAtions*
hypothermia	*HYpomaTHERmia*	*HYpomaTHERmia*	*HYpomaTHERmia*
oboe	*ObamaBOE*	*OboemaBOE*	*OOOmaBOE*
octagon	*OCtamaGON*	*OCtamaGON*	*OCtamaGON*
octet			*I dunno...*
purple			
tuba		*TUbamaBA*	
wonder	*WONdamaDER*	*WONdermaDER*	*WONNNmaDER?*
wonderful	*WONdermaFUL*	*WONdermaFUL*	*WONdermaFUL*

Task 1. We've left out some of their responses. Fill in the blanks with the appropriate words from the list below.

PURpamaPLE	*OCtemaTET*	*TUbamaBA*
TUUUmaBA	*PURplemaPLE*	*OcamaTET*
	PURRRmaPLE	

---

[1] Created by Patrick Littell.

<1> Get Edumacated! (2/2)

Task 2. How would each respondent say the following words? We've given you a few to get started.

	Alan	Barbara	Chris
antiseptic	_____	_____	_____
Canada	_____	_____	_____
feudalism	_____	FEUdamaLISm	_____
optics	_____	_____	_____
party	PARtamaTY	_____	_____
table	_____	_____	_____
water	_____	_____	WAAAmaTER

Task 3. What do YOU feel is the correct "edumacated" version of the following words:

multiplication	_____
anecdotal	_____
graduation	_____
hyperactivity	_____
Kalamazoo	_____

Task 4. Respondents usually hesitate before two-syllable words, and are less sure that their answers feel "correct". Why, and what motivates Alan's, Barbara's, and Chris's eventual answers?

_____

_____

_____

_____

_____

_____

<2> How's your Hittite?[2] (1/2)   (15 points)

Hittite is an extinct language that belongs to the Anatolian branch of the Indo-European language family. It was spoken in the ancient Hittite Empire in second millenium BCE. Hittite was written using cuneiform script.

The excerpt below is a (simplified) phonetic transliteration of a cuneiform passage found on a tablet. The simple capitalized words denote so called Sumerograms – signs that were written using Sumerian cuneiform, and the capitalized words in italics denote Akkadograms – signs that were written using Akkadian cuneiform.

Našta illuyankan
ḫantešnaz šarā kallišta
kāšawa EZENan iyami
nuwa adanna akuwanna eḫu

našta illuyankaš *QADU* [DUMUMEŠ-*ŠU*]
šarā úēr nuza eter ekuer
našta palḫan ḫūmandan ekuer
neza ninkēr

ne namma ḫattešnaš kattanta
nūmān pānzi ḫupašiyašša úit
nu illuyankan išḫimanta kalēliēt

IM-aš úit nukán illuyankan
kuenta DINGIRMEŠ-ša kattišši ešer

Here is its translation into English:

And he called up the snake from the hole: "Behold the feast I'm making! Come to eat and drink!" And the snake came up with his sons. And they ate and drank. And they could no longer go down into the hole again. And Hupasiyas came and tied the snake with a rope. The Stormgod came and killed the snake; and the gods were with him."

Task 1: Match the following Hittite word forms with their English translations by writing the appropriate letter (a...g) to the right of the Hittite word in the table below.

A.	eter		a.	*snake*	
B.	**ḫanteššar**		b.	*hole*	
C.	úit		c.	*came*	
D.	illuyankaš		d.	*and*	
E.	našta		e.	*up*	
F.	**šarā**		f.	*drank*	
G.	ekuer		g.	*ate*	

---

[2] Created by Dorottya Demszky.

<2> How's your Hittite? (2/2)

Task 2: Match the following suffixes (A to G) with their grammatical roles (a to g). Answer by placing the appropriate letter (a-g) to the right of the suffix:

A.	–aš	
B.	–ša	
C.	MEŠ	
D.	-er/ēr	
E.	-an	
F.	-anna	
G.	–it	

a. marker of the infinitive (e.g., in English: *to* sleep, *to* walk)

b. plural marker (e.g., in English: apple*s*)

c. marker of 3rd person plural, past tense (e.g. in English: *they* walk*ed*)

d. marker of the direct object (in English only personal pronouns have distinct 'object' forms which contrast with their 'subject' form: e.g., 'she hit *him*' and not 'she hit *he*')

e. marker of the subject (in English the subject is unmarked, but it is the entity performing the action: e.g., in '*The dog* chased the cat', 'the dog' is the subject while 'the cat' is the object.)

f. marker for 3rd person singular, past tense (e.g., in English: *he* walk*ed*)

g. a marker denoting 'and'.

Task 3: As you may have noticed, ends of lines do not always coincide with the ends of clauses or sentences. What is the part of speech of the word that a Hittite clause (or sentence) ends with? Select from the following choices by placing a tick (✓) beside your choice.

adjective	
adverb	
noun	
preposition	
verb	

<3> FAN FICTION[3] (1/2)   (18 points)

MARY SU.0 is a fan-fiction writing robot. Unfortunately, she's not very good at what she does. MARY writes fan-fiction by reading in the text of a book (or series of books) and randomly generating new sentences based on the text. Her latest effort is fan-fiction based on the Harry Potter book series.

MARY SU.0 has a few different methods that she's able to use for generating sentences. The first class of methods are called *ngram* methods. The simplest of these methods is the *unigram* method. In the unigram method, MARY chooses each token of the sentence completely randomly from the entire vocabulary of the book she read. (A token can also be a punctuation mark.) An example of a sentence generated using this method might look like this:

```
gave spiral the truly poisoned , Neville the shoulder Invisibility
```
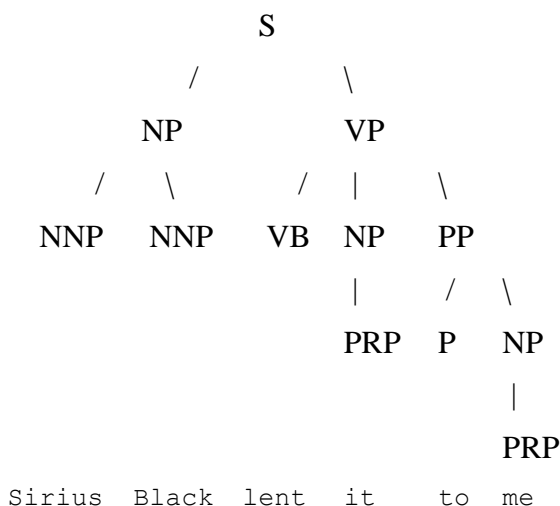
A second method is the *bigram* method. In this method, MARY first finds all the tokens that were used to start a sentence in the text and randomly chooses one of these to start the sentence. Then she builds the rest of the sentence by looking at the most recent token generated, finding all tokens that occur immediately after that token in the text, and randomly choosing one of these. For example, if the most recently generated token was "red", MARY would find all the tokens in the text that immediately follow red, {"hair", "curtains", "as" …} and randomly choose one of these to be the next word. A sentence generated using the bigram method might look like this:

```
Face your nose noisily after you saying stuff  .
```

A third method is called the *trigram* method. This method is very similar to the bigram method, but uses the previous two tokens (instead of the previous one) to decide what the next token will be. A sentence generated using the trigram method might look like this:

```
But Harry hardly noticed that six extra chairs . "
```

The last method that MARY can use to generate sentences is called the *Context Free* method. This method starts by taking each sentence in the text and generating a grammar tree, like the one below, for it.

```
                    S
                  /     \
               NP          VP
             /   \       /  |   \
          NNP    NNP    VB  NP   PP
                            |   /  \
                           PRP  P   NP
                                    |
                                   PRP
      Sirius   Black   lent   it   to   me
```

---

[3] Created by Ben King.

<3> FAN FICTION (2/2)

To generate a new sentence, she first generates an "S" which represents a sentence. Then she looks through her collection of grammar trees for all the sets of symbols ([NP VP .] for example) that occur immediately under an "S". She then repeats this process recursively for each of the new items generated until the tree has no more nodes that can be expanded (once a token is generated, it cannot be expanded). (You don't have to concern yourself with these symbols.) A sentence generated by this method might look like this:

> The next question will cast by Ron .

Task: Below is a collection of sentences. Two of them are real sentences from the Harry Potter series. The rest were generated using one of the methods above. Next to each sentence, write either "U" (for unigram), "B" (for bigram), "T" (for trigram), or "CF" (for context free) to indicate the method that most likely generated that sentence, or if you think the sentence was not automatically generated, write "R" (for real).

	Headmaster uninjured could that was Malfoy that badges		"Sorry!" he said," said Mr. Malfoy's eyes.
	He bent over top of the water blushing furiously.		Harry wasn't," said Dumbledore went slightly surprised .
	There were crouching in your bedroom.		years beginning at to annoyance spider!" just months Harry
	He lived about a hundred wizards were closing.		You might have been an impostor.
	Ron spooned iron bolts, keyholes, and a heavy wooden breadboard onto her back and picked up a fistful.		They'll be the first to rise up in the Invisibility Cloak on," said Professor Flitwick pressed a box into his bag.
	"What?" said Harry.		The broom gave them an enormous wink.

<4> Best to deny it all in Warlpiri[4] (1/1)   (30 points)

Warlpiri is the first language of over 2,000 people who live in communities on or near their traditional lands in the Tanami Desert area of the Northern Territory of Australia. Since 1973 children in these communities have been educated in both Warlpiri and English, so that many Warlpiri people both speak and write their language. National and local news is broadcast in Warlpiri on the national ABC (Australian Broadcasting Commission) network.

Study these Warlpiri sentences and their English translations:

	Warlpiri	**English**
1.	Nganangkungku pakarnu?	*Who hit you?*
2.	Kulaju nganangku pakarnu.	*Nobody hit me.*
3.	Ngananpa nyangu?	*Who did you see?*
4.	Kularna ngana nyangu.	*I didn't see anyone.*
5.	Ngana wantija watiyangurlu?	*Who fell from the tree?*
6.	Nganamayi wantija.	*I don't know who fell.*
7.	Kula ngana wantija.	*No one fell.*
8.	Nyiyarlunpaju pakarnu?	*What did you hit me with?*
9.	Kularnangku nyiyarlu pakarnu.	*I didn't hit you with anything.*
10.	Kularna ngana pakarnu.	*I didn't hit anyone.*

Using the knowledge of Warlpiri gained from studying the sentences above and those below, fill in the missing translations in the appropriate language.

	Warlpiri	**English**
11.	Nyarrparakurra yanu?	*Where did he go to?*
12.	Nyarrparakurramayi yanu.	
13.		*He didn't go anywhere.*
14.		*Who hit me?*
15.	Nganangkumayingki pakarnu.	
16.	Nyarrpararlanpa nyangu?	*Where did you see him?*
17.	Kularna nyarrpararla nyangu.	
18.	Nyiyanpa nyangu?	
19.		*I didn't see anything.*
20.	Nyarrparangurlunpa yanu?	
21.		*I didn't go from anywhere.*
22.		*I don't know what you hit me with.*

---

[4] Created by Mary Laughren.

<5> Don't Sell The House!⁵ (1/1)　　(35 points)

This problem involves the Nung language of north-eastern Vietnam, spoken by about a million people and related to the Thai, Lao, Isan, Shan, and Zhuang languages of Southeast Asia in the Tai-Kadai family. It is not related to Chinese, Vietnamese, Khmer, Hmong, Malay, or Burmese, so far as we know. In this problem, the Nùng Phạn Slinh variety of Nung will be used. In Nùng Phạn Slinh, as seen here, word order is fixed: that is, for every sentence containing certain words, *there is only **one** way to properly order those words*.

Task: Here is a list of sentences in Nung and their English translations. Find the sentences without English or Nung equivalents and write down the missing translation.

Note: The marks above vowels indicate tone and the length of the vowel. đ and sl are consonants. You do not need to know how to pronounce Nung in order to solve the problem.

1.	Cáu ca vửhn nhahng kíhn.	I was about to continue to eat it.
2.	Cáu cháhn slồng páy mi?	Do I truly want to go?
3.	Cáu mi slày kíhn.	I don't have to eat it.
4.	Cáu ngám hẻht pehn tế.	I did it like that just now.
5.	Cáu tan đohc hảhn mưhng.	I only saw you.
6.	Cáu vửhn nhahng bô sạhm tảhng hẻht hơn.	I also continue to build the house alone.
7.	Da kíhn!	Don't eat it!
8.	Da khải hơn!	Don't sell the house!
9.	Mưhn chóng ca cháhn fải khải.	Then she truly was about to have to sell it.
10.	Mưhn mi cháhn đày non.	She truly can't sleep.
11.	Mưhn náhc-thày chóng bô sạhm kíhn.	Then she also just previously ate it.
12.	Mưhng náhc-thày slồng tảhng páy.	You wanted to go alone just previously.
13.	Cáu cháhn đày non.	
14.	Da páy non!	
15.	Mưhn bô sạhm mi slồng hẻht hơn mi?	
16.	Mưhn ngám bô sạhm páy hơn.	
17.		I wasn't about to eat it just previously.
18.		She didn't have to eat it alone like that just now.
19.		The house truly can't eat you.
20.		Then were you also about to go just previously?

---

⁵ Created by Alex Wade.

<6> Tunnelling back through Tocharian prehistory (1/2)   (22 points)

The Tocharian languages were an extinct branch of the Indo-European language family (including English, French, German, Greek, and many others in Europe). Linguists have reconstructed the ancestor language, called Proto-Indo-European, from which all the daughter branches descended.

A major part of language change is sound change, where a language's phonemes shift around over time. Sound change is importantly regular, and can be encapsulated neatly by writing down rules to describe how one stage of the language proceeds to the next. For example, a rule like:

   t > d /_r

means that all instances of 't' change to 'd' before 'r', so *tree* would become *dree*, while:

   p > ∅ / _ #

means that all instances of 'p' disappear (change to 'zero' represented by 'ø') at the end of a word (represented as the hash #), so *stop* would become *sto*.[6] Sound changes apply to all sounds, in all words, that fit their criteria (the stipulation after '/' in the rules).

Because many ancient languages were never written down until recent millenia, linguists have to rely on clever deductions to work out the details of their early history. Our only records of Tocharian are some 9th century manuscripts around the Tarim Basin in western China, so our knowledge of its development comes from inferences of this type.

Here are some Tocharian words with their English equivalents. These groups of words represent seven stages in the very early history of the language, in a random order:

share	*row of teeth*	*knee*	*war*	*hundred*	*dog*	*prop*
pákos	kómos	kónu	kóro-	kṃtóm	kuṓ	stema-
págos	gómos	gónu	kóro-	kṃtóm	kuṓ	stema-
bʰágos	jómbʰos	jónu	kóro-	kṃtóm	kuṓ	stembʰa-
bʰágos	jómbʰos	jónu	kóro-	cṃtóm	cuṓ	stembʰa-
páko	kómo	kónu	kóro-	kṃtóm	kuṓ	stema-
bʰágos	gómos	gónu	kóro-	kṃtóm	kuṓ	stema-
bʰágos	gómbʰos	gónu	kóro-	kṃtóm	kuṓ	stembʰa-

Note: The dot under the first 'm' in the word for 'hundred' signals a breathy unvoiced nasal sound. (Not relevant for solving this puzzle.)

As we can see, between these stages of Tocharian, some sound changes have occurred.

---

[6] Hash (#) can also mark the beginning of a word.

<6> Tunnelling back through Tocharian prehistory (2/2)

Task: Put the stages in historical order, and write down rules describing the sound changes that happened in between each stage. If you can find different orders, explain which you think is the most likely. (The accent ´ on a vowel can be ignored.) Answer by rewriting the words in the chart below placing the oldest forms at the top and the most recent at the bottom. Write the relevant sound change in the rightmost column. The one ordered set of sound changes must account for all forms of each word.

	share	*row of teeth*	*knee*	*war*	*hundred*	*dog*	*prop*	**Sound Change**
1								
2								
3								
4								
5								
6								
7								