Week 1 – The Chomsky Hierarchy: A Formal Interlude

February 18 and 20, 2008

1 The phonetics-phonology interface

(1) Recall Figure 1 which shows the classical view of the phonetics-phonology interface.

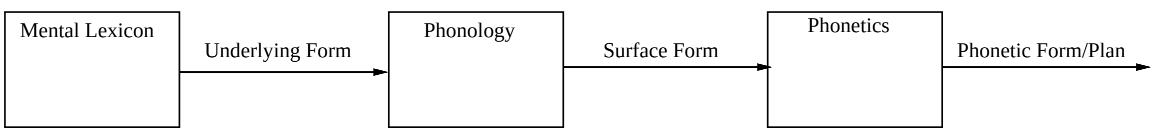

Figure 1: Classic Competence Model of the Phonology-Phonetics Interface

(2) We had noted the following:

a. Morpheme structure constraints are generalizations about the *set* of URs; they are constraints over the set of URs.
   (i) Ki-Rundi (Bantu): *[-long,+syl][+nasal][-syll]
   (ii) Tonkawa (Coahuiltecan): *[-syll,+glottalized][-syll]
   (iii) Russian: *[αvoice,-son][$-\alpha$voice,-son]

b. Surface constraints are generalizations about the *set* of SRs; they are constraints over the set of SRs.
   (i) English: *[-long,+syll][-son,+voice] (cf. [bɛt,bɛd])
   (ii) Japanese: *[-syll,-nasal] $]_\sigma$
   (iii) Russian: *[αvoice,-son][$-\alpha$voice,-son]

c. Phonology maps URs to SRs. We'll see some examples in a little bit.

d. We think that phonological rules themselves ought not be arbitrary but constrained.

(3) Today we make these ideas precise, by examining some of the foundations of linguistics and computer science.

(4) One emphasis is to make concrete the notions of *logically possible language* and *logically possible patterns*.

(5) There will be some notation as we go along; they are just handy abbreviations.

2 Being precise about Figure 1

2.1 Sets

(6) As you know, a *set* is a just some collection of objects. We write $a \in A$ iff object a is in the set A. With this simple idea, we can define a number of other handy abbreviations:

$$\begin{array}{ccccl} \text{union} & A \cup B & = & \{x : x \in A \text{ or } x \in B\} \\ \text{intersection} & A \cap B & = & \{x : x \in A \text{ and } x \in B\} \\ \text{difference} & A - B & = & \{x : x \in A \text{ and } x \notin B\} \\ \text{powerset} & 2^A & = & \{X : X \subseteq A\} \\ \text{product} & A \times B & = & \{(a, b) : a \in A \text{ and } b \in B\} \end{array}$$

★ Let $A = \{a, b, c\}$. Let $B = \{x, y, z\}$. What is $A \times B$? What about 2^A?

★ Possibly mind-bending math-type question. The empty set—the set with nothing in it—is denoted $\emptyset$. Can you see why for any set A, $\emptyset \subseteq A$?

(7) It is also useful to make clear a notion of equality between sets. We will say two sets A and B are equal if and only if A is a subset of B and B is a subset of A. With our handy abbreviations, we can write

$$A = B \text{ iff } A \subseteq B \text{ and } B \subseteq A$$

2.2 Words and Sentences

(8) Words and sentences are basically sequences formed from the alphabet. If we concatenate elements of the alphabet into a sequence we get a word, or sentence.

a. Suppose A={s,o,t,k,u,i}.
   (i) Then *soku* is a word.
   (ii) And *tui* is a word.

(9) In phonology, we are interested in sets of *phones* (I use phones to be deliberately ambiguous regarding *phonemes* or *allophones*), whose concatenation makes words. In syntax we are more interested in sets of *morphemes*, whose concatenation makes sentences. We will refer to both kinds of sets as an *alphabet*.

(10) For example, here are the phonemes of English, according to Hammond (1999). As you can see there are 36 of them (English diphthongs would be described as concatenating the vowels, e.g. eɪ).

{ p,t,k,b,d,g,f,θ,s,ʃ,v,ð,z,ʒ,tʃ,dʒ,m,n,ŋ,l,r,j,w,h,i,ɪ,e,ɛ,æ,u,ʊ,o,ʌ,ɔ,ɑ,ə}

(11) We obtain almost all the allophones of by taking the union of the set above with this next set[1]:

$$\{ \text{p}^\text{h}, \text{t}^\text{h}, \text{k}^\text{h}, \text{tʃ}^\text{h} \} \cup \{ \tilde{x} : x \text{ is a vowel}\}$$

★ Is the phoneme /p/ specified for aspiration? (There is more than one correct answer.)

(12) Let us call the set of allophones of English (the union of the two sets above) *Eng*.

(13) If we are interested in syntax and sentences, it is common to consider the alphabet to be morphemes. Here we are interested in words, and so we consider the alphabet to be phones.

(14) For now we consider words to be finitely-long sequences of phones and we address autosegmental representations of words later. If our alphabet is A, the set of all words with finite length formed by A is given by A^*. A^* is sometimes called the *Kleene closure* of A.

★ Given *Eng*, what are some examples of Eng^*?

★ Let *OED* denote the set of actual English words. Is $OED \subseteq Eng^*$??

★ Let *Blick* denote the set of possible English words. Is $Blick \subseteq Eng^*$??

★ Does $Eng^* = Blick$?

[1] I am excluding from this list the varieties of English approximants (e.g. light and dark [l] plus the varieties of [r], plus vowel length, see Ladefoged (2000) and also Ladefoged and Maddieson (1996))

(15) To include any possible word in any language, we may take the symbols of the IPA (multiplying out the possibilities with diacritics and so on) as the alphabet. Then the *logically possible human words* are given by IPA^*.

★ When we speak of phonotactic competence, which set most closely represents this? Why?

(16) In general, for some alphabet A, we consider a *language* or *language pattern* to be a subset of A^*. Thus given an alphabet, it is possible to consider the logically possible languages, i.e. all those subsets of A^*.

(17) Note it is common practice to use the Kleene closure in combination with concatenation to define sets of words.

a. Suppose A={s,o,t,k,u,i}.
b. Then *ku*ki* represents a set of words, with *kki, kuki, kuuki, kuuuki, ...* in this set.

(18) A few more mathematical notes:

a. The empty word is the unique word of length zero, we denote it ϵ.
b. The empty language is the language with no words in it, as we saw before this set is $\emptyset$.
c. Note the empty language is different from the language which contains only empty word! That language is given by the set $\{\epsilon\}$ and is of size 1.

2.3 Relations

(19) A *relation* R between two sets A and B is a subset of $A \times B$. If $(a,b) \in A \times B$, we sometimes write aRb.

(20) The business of linguistics is relations. Phonology relates URs to SRs, or members of a morphological paradigm to others. Syntax relates sentences to other sentences, and semantics relates sentences to meaning.

(21) For example, consider let $A = \{a, b, c\}$. Then

$$A \times A = \left\{ \begin{array}{ccc} (a,a) & (a,b) & (a,c) \\ (b,a) & (b,b) & (b,c) \\ (c,a) & (c,b) & (c,c) \end{array} \right\}$$

a. Let $R_1 = \{(a,a),(a,b),(a,c)\}$. (free variation)
b. Let $R_2 = \{(a,b),(b,c),(c,a)\}$. (one-to-one-ness)
c. Let $R_3 = \{(a,c),(b,c),(c,c)\}$. (neutralization)

★ Why do I label the examples above with the above terms?

★ How can we describe all the logically possible relations between two sets A and B?

(22) Terminology: For a relation $A \times B$, we sometimes call set A the *left-hand side* and B the *right-hand side*.

(23) Consider the following examples of relations (parentheses omitted for clarity).

a. Guinaang Kalinga — English gloss

dábo	,	d**in**ábo	(hypothetical)	,	(hypothetical)
dopá	,	d**im**pána	'fathom'	,	'he measured by fathom'
gobá	,	g**im**bána	'firing (pots)'	,	'she fired'
ʔomós	,	ʔ**im**m'osna	'bath'	,	'she bathed'
botáʔ	,	b**in**táʔna	'broken piece'	,	'she broke'
ʔodáw	,	ʔ**in**dáwna	'requesting'	,	'he requested'
bosát	,	b**in**sátna	'sudden break'	,	'he snapped'
ponú	,	p**in**núna	'filling'	,	'she filled'
toʔóp	,	t**in**ʔópna	'satisfaction'	,	'he satisfied'
sogób	,	s**iŋ**góbna	'burning'	,	'he burned'
doŋól	,	d**iŋ**ŋólna	'report'	,	'he heard'
ʔolót	,	ʔ**il**lótna	'tightening'	,	'he made tight'
ʔowá	,	ʔ**iŋ**wána	'doing, making'	,	'he made, did'

b. Palauan

X	,	*his/her/its X*	gloss
ɾákt^{h}	,	ɾəkt-ɛ́l	'sickness'
sɛ́səb	,	səsəb-ɛ́l	'fire'
bótk^{h}	,	bətk-ɛ́l	'operation'
ɾíŋəl	,	ɾəŋəl-ɛ́l	'pain'
kúk-	,	kəkú-l	'nail'
ɾɛ́ːk^{h}	,	ɾɛk-ɛ́l	'rustling sound'
ðəkóːl	,	ðəkol-ɛ́l	'cigarette'
ʔíːs	,	ʔis-ɛ́l	'escape'
búːʔə	,	buʔ-ɛ́l	'betel nut'

c. Polish

sg.		*pl.*	gloss
trup	,	trupi	'horse'
wuk	,	wuki	'bow'
snop	,	snopi	'sheaf'
kot	,	koti	'cat'
nos	,	nosi	'nose'
sok	,	soki	'juice'
klup	,	klubi	'club'
trut	,	trudi	'labor'
grus	,	gruzi	'rubble'
wuk	,	wugi	'lye'
žwup	,	žwobi	'crib'
lut	,	lodi	'ice'
vus	,	vozi	'cart'
ruk	,	rogi	'horn'

★ The tables above give examples of the relation; i.e. they are finite samples. To describe the full competence, what kind of relation would we be talking about?

(24) Originally the terms *feeding, bleeding, counterfeeding,* and *counterbleeding* were used to describe rule orderings. But over time, I think they have become *descriptive* terms used to describe relations.

★ Quick review. Which terms describe which of the above relations?

★ Are there any URs in the above relations? Why did we posit URs? See Adam Albright's work on identifying bases in morphological paradigms `http://web.mit.edu/albright/www/`.

(25) I think one reason people use URs is it overcomes the problems of which relations to care about. When there are many morphological categories (nominative singular, nominative plural, accusative singular, accusative plural, genitive etc.), we can form many relations among these sets. It appears to some extent redundant and how would we know which ones are absolutely necessary (again see Albright's work)?

(26) If we wanted to avail ourselves of URs, we might describe the Polish paradigm as follows, using instead of R the symbol $\rightarrow$. (I omit the brackets // and [] for clarity.)

	sg.			*pl.*		gloss
UR	→	SR	UR	→	SR	
trup	→	trup	trup+i	→	trupi	'horse'
wuk	→	wuk	wuk+i	→	wuki	'bow'
snop	→	snop	snop+i	→	snopi	'sheaf'
kot	→	kot	kot+i	→	koti	'cat'
nos	→	nos	nos+i	→	nosi	'nose'
sok	→	sok	sok+i	→	soki	'juice'
klub	→	klup	klub+i	→	klubi	'club'
trud	→	trut	trud+i	→	trudi	'labor'
gruz	→	grus	gruz+i	→	gruzi	'rubble'
wug	→	wuk	wug+i	→	wugi	'lye'
žwob	→	žwup	žwob+i	→	žwobi	'crib'
lod	→	lut	lod+i	→	lodi	'ice'
voz	→	vus	voz+i	→	vozi	'cart'
rog	→	ruk	rog+i	→	rogi	'horn'

(27) We will use the term $Phon - Polish$ to mean the relation which describes native Poles' knowledge from URs to SRs, (which of course extends beyond the small sample represented above).

(28) One problem for people who want to postulate URs is they have to explain how it is that people come to know them—since they never hear them.

(29) The relations above can be considered relations between words formed over the same alphabet, in the broadest sense, over $IPA^* \times IPA^*$.[2]

(30) Note all the logically possible relations is the powerset of $IPA^* \times IPA^*$; i.e. the subsets of $IPA^* \times IPA^*$.

a. Major Research Questions:
   (i) Is a phonology of a language best described with URs or not?
   (ii) Which of these subsets are phonologies of actual languages?
   (iii) Which of these subsets are possible phonologies of actual languages?

(31) There are a number of properties that relations may have, some are shown below.

[2] Actually, we ought to include the morpheme boundary and word boundary symbols so $(IPA \cup \{+, \#\})^* \times (IPA \cup \{+, \#\})^*$.

antisymmetry	Whenever xRy and yRx, it is the case that $x = y$.
reflexivity	For all $x \in S$, it is the case that xRx.
irreflexivity	For all $x \in S$, it is not the case that xRx.
symmetry	Whenever xRy, it is the case that yRx.
asymmetry	Whenever xRy, it is not the case that yRx.
transitivity	Whenever xRy and yRz, it is the case that xRz.

2.4 Functions

(32) A *function* f is a relation between two sets A and B such that if $(a, b) \in f$ then $(a, b') \in f$ iff $b = b'$. In other words, each a is related to only one b. We write $f(a) = b$ and call b the value of f at a. A is called the *domain* of f and B the *co-domain*. This is often indicated by writing $f : A \rightarrow B$.

a. $f(x) = x^2$ is a function.
b. $f(x) = \sqrt{x}$ is not a function.

(33) Functions have several nice properties and there are special kinds of function. If you are interested, check out this page: `http://en.wikipedia.org/wiki/Bijection\%2C_injection_and_surjection`

★ If we consider the relation from URs to SRs that have been posited in different languages, is this relation a function? Why or why not?

(34) We mention functions here for completeness only, as our main focus is on relations.

2.5 Concluding remarks

(35) An infinite set like *Blick* represents the phonotactic competence of a native speaker.

(36) Similarly an infinite relation like $Phon - Polish$ represents the phonological competence of the speaker.

★ Let *Schplick* denote the set of well formed words of Polish. How does *Schplick* relate to $Phon - Polish$?

(37) We are interested in the *characteristics* of sets like these which define human sound patterns.

(38) The learner's problem is easily stated in these terms—given a finite sample of words from *Blick*, how does the learner generalize to the infinite set *Blick*?

(39) Sets, relations, and functions are among the most useful ideas in science. They will come up in our discussion of Optimality Theory.

3 The Chomsky hierarchy

(40) The Chomsky hierarchy (Figure 2) provides a means to study language patterns. It divides the space of logically possible languages into different regions. These are, in order of increasing complexity:

a. Finite languages
b. Regular languages
c. Context-Free languages
d. Mildly Context-Sensitive languages
e. Context-Sensitive languages
f. Recursively Enumerable languages
g. Non-Recursively Enumerable languages

(41) In other words, points in the space in Figure 2 represent languages, or language patterns—that is, subsets of A^*.

(42) Languages that are in the same region share certain properties—namely the grammars that are necessary to generate such languages share certain properties, or are necessarily possess a certain level of complexity.

(43) You may be thinking that Chomsky Hierarchy depends on the alphabet and that the choice of alphabet A really matters, but as we will discuss later, one of the things that makes the Chomsky hierarchy so great is that it is largely independent of the choice of alphabet.

(44) You may also be thinking that languages are more than just sets of strings since languages have structure, which we may wish to represent with trees, thus making the Chomsky hierarchy irrelevant to linguists like ourselves.

(45) But this is something that can be handled by a carefully chosen alphabet, which as mentioned the Chomsky hierarchy is independent of.

a. Example: if we add the symbols *[,]* to our alphabet, we can add bracketing to our strings—and we can be interested in those languages with well-formed bracketing structures that correspond to trees.

(46) Additionally, many of these regions above have independently converging definitions! We will see one definition below. For others see for example, Kracht (2003:chap 2) or Vijay-Shanker and Weir (1994)).

(47) Also, within each of these regions, there are often additional hierarchies (see for example McNaughton and Papert (1971), Pullum and Rogers (2007)).

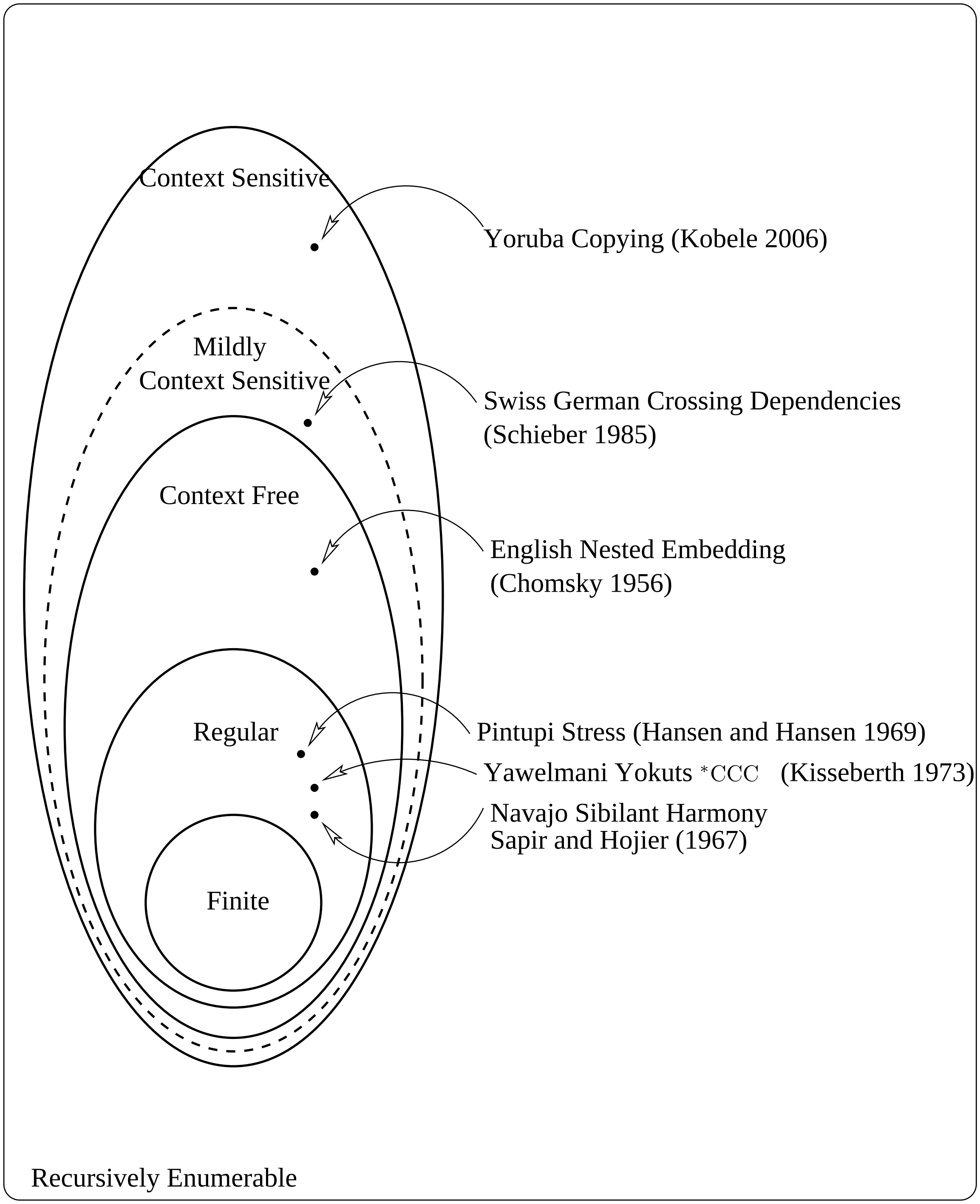

Figure 2: The Chomsky Hierarchy

(48) Finally, the Chomsky hierarchy can be generalized to include the study of relations, i.e. pairs of stringsets. Some properties of these classes change when we move to relations, but for the most part, the major divisions remain intact. So it can be used to study (UR,SR) mappings, or in syntax, (form, meaning) mappings.

3.1 Defining the Chomsky Hierarchy

(49) So far when we discuss the competence of a speaker, we have been discussing infinite sets, whether they are sets of words such as *Blick* or relations such as $Phon{-}Polish$.

(50) *Grammars* are finite devices that generate languages in some well-defined way.

(51) Today we look at a general kind of grammar, called re-write grammars. The phonological rules that we looked at last semester are an example of a re-write grammar.

3.2 Re-write grammars

(52) The ingredients of a rewrite grammar is the alphabet A (sometimes called the *terminals*) and a set of categories Cat (sometimes called the *non-terminals*). It is assumed Cat contains at least one element, which we call S; it is the *start* category.

(53) Rewrite grammars are a collection of rewrite rules.

(54) A *rewrite rule* has the following form:

$$(A \cup Cat)^*Cat(A \cup Cat)^* \rightarrow (A \cup Cat)^*$$

These rules are essentially a relation $\rightarrow$ with the left hand side being any finite sequence of alphabetic symbols and categories (but containing at least one category) and the right hand side being any finite sequence of alphabetic symbols and categories (possibly empty).

(55) Examples. Suppose X={s,o,t,k,u,i} and $Cat = \{S, X, Y\}$

a. $tSXkui \rightarrow YsoXXXSXXkuX$
b. $tttttkkkkkS \rightarrow o$
c. $S \rightarrow XY$
d. $S \rightarrow tui$.

(56) For all $w, z \in (A \cup Cat)^*$, we say w rewrites as z (written $w \Rightarrow z$) iff there are $x, y, u, v \in (A \cup Cat)^*$ such that

a. $w = uxv$
b. $z = uyv$, and
c. $x \rightarrow y$.

3.3 Languages of rewrite grammars

(57) The relation $\Rightarrow^*$ is the reflexive, transitive closure of $\Rightarrow$.

(58) Given a rewrite grammar G, the language of G is given here:

$$L(G) = \{w \in A^* : S \Rightarrow^* w$$

3.4 Major language classes

(59) A grammar is a *list grammar* iff all of its rewrite rules are like

$$S \rightarrow w \text{ where } w \in A^*$$

a. For example:

$$S \rightarrow dog$$

$$S \rightarrow cat$$

$$S \rightarrow mouse$$

★ What language does this grammar generate?

(60) A grammar is a *regular grammar* iff all of its rewrite rules are like

$$X \rightarrow aY$$

$$X \rightarrow a$$

where $X, Y \in Cat$ and $a \in A$. In other words, the left hand side only contains a category symbol and the right hand side contains

a. For example:

$$S \rightarrow kR$$

$$R \rightarrow aC$$

$$C \rightarrow tE$$

$$E \rightarrow \epsilon$$

$$C \rightarrow tS$$

★ What language does this grammar generate?

(61) A grammar is a *context-free grammar* iff all of its rewrite rules are of the form

$$X \rightarrow Z$$

where $Z \in (A \cup Cat)^*$. In other words, the left hand side may contain only one category symbol, but the right hand side can contain any sequence of categories and alphabetic symbols.

a. For example

$$S \rightarrow NP\,VP$$

$$S \rightarrow S \textit{ and } S$$

$$NP \rightarrow DP \textit{ dog}$$

$$VP \rightarrow \textit{bites } NP$$

$$DP \rightarrow \textit{the}$$

★ What sentences does this grammar produce?

(62) A grammar is a *context-sensitive grammar* as long as all of its rewrite rules do not have the empty string in the right hand side.

a. Example. Add the following rule to the above context-free grammar:

$$S \textit{ and } S \rightarrow S \textit{ or} S \textit{ or } S \textit{ and} S;$$

$$S \textit{ or } S \rightarrow S \textit{ or} S \textit{ or } S \textit{ or} S; S \textit{ or} S;$$

(63) *Recursively enumerable languages* are those that can be described with any rewrite grammar. *Non-Recursively enumerable languages* are languages that cannot be described with any rewrite grammar (and we know of no other grammar that can describe such languages).

(64) The Chomsky hierarchy specifies proper inclusion relations among collections of languages:

non r.e. language ⊂
r.e. languages ⊂
context-sensitive languages ⊂
context-free languages ⊂
regular languages ⊂
finite languages

3.5 Phonological rules

(65) When we consider phonological rules like the ones we studied last semester, we learned the following:

$$\text{A} \longrightarrow \text{B} \,/\, \text{X} \,\underline{\quad}\, \text{Y}$$

This means *XAY is rewritten as XBY* or, in other terms, *A is rewritten as B when preceded by X and followed by Y*

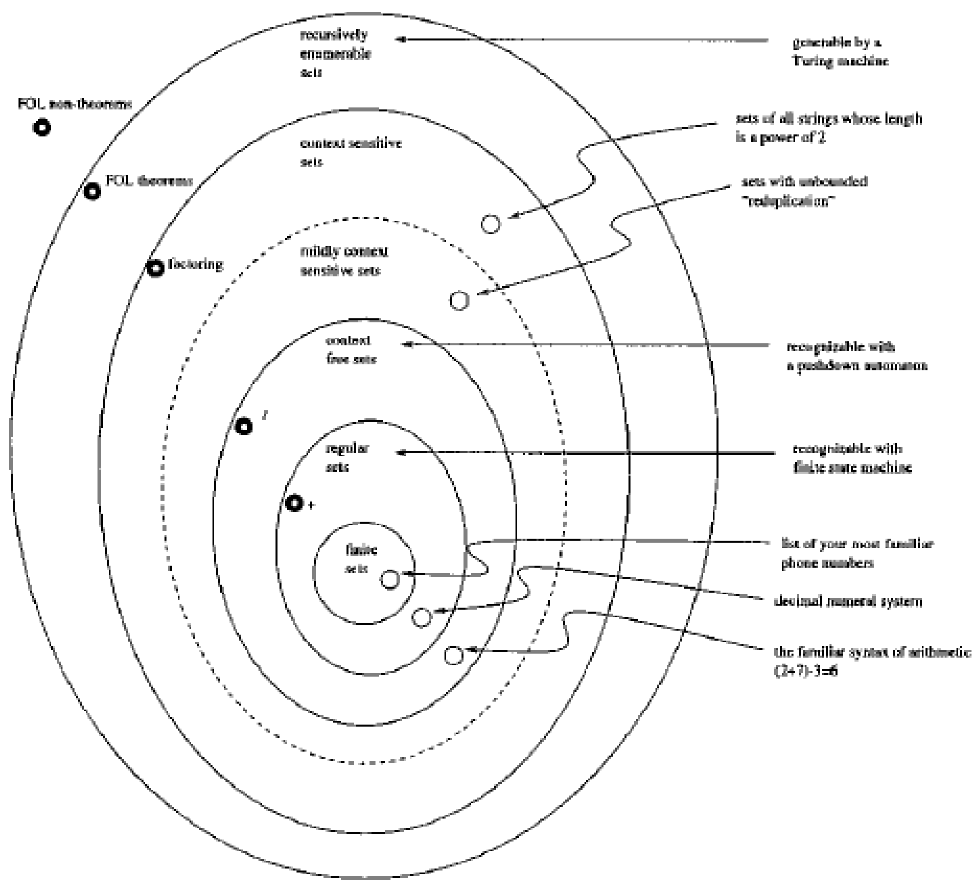

Figure 3: The Chomsky Hierarchy (by E. Stabler)

a. A is the affected **segment, focus,** or **target** of the rule.
b. B is the **structural change** that the rule requires.
c. X____ Y is the **context** of the rule.
d. XAY is the **structural description**.

★ What kind of rewrite rule is this?

4 Natural language and the Chomsky hierarchy

(66) Once we have established patterns in this way, we can ask about natural language patterns. Where do they belong? Figure 2 shows some answers to this question.

(67) Major research question:

a. What are the properties of human languages? The answer to this question will carve out some space in the Chomsky hierarchy since it will separate human languages from logically possible nonhuman ones.

(68) It is striking that almost all phonological phenomenon is known to be regular. This includes phonotactic patterns (Heinz 2007) as well as UR to SR mappings (Johnson 1972, Kaplan and Kay 1981, 1994, Koskenniemi 1983).

a. Individual MSCs and SCs like *CCC are also regular.
b. Their combination—due to closure properties of regular sets—are also regular. Thus *Blick* and *Schplick* are regular.
c. Phonological relations like $Phon - Polish$ are also regular.

(69) Note this does not mean any regular language is a possible phonological pattern!

a. It is much more likely that phonological patterns have additional properties which further characterize them: non-counting, locality, perceptual, articulatory, etc.

(70) The only challenges to the hypothesis that phonological patterns are regular that I know of have to do with reduplication (we'll look at this later this semester) and metathesis.

a. Metathesis in Kwara'ae (Heinz 2005a,b).

	Citation	Normal	
a.	ˈŋe.la	ˈŋe̯al	'child'
b.	li.ˌma.ku	ˈli.ˌmau̯k	'my hand'
c.	ˈke.ta.ˌla.ku	ˈke̯at.ˌlau̯k	'my height'
d.	da.ˈro.ʔa.ˌni.da	ˈdao̯r.ʔa.ˌni̯ɛd	'to share them'
e.	ˈra.ʔe.ˌra.ʔe.ˌna.ʔa	ˈrae̯ʔ.ˌrae̯ʔ.ˌnaˑʔ	'incline, slope'

(71) The metathesis problem may end up having some kind of solution (it may just be unbounded metathesis that's the problem, but this is a wide open area). Albro (2005) develops some restricted extensions to regular systems to handle reduplication (he provides a full analysis of phonology in Malagasy).

(72) Conclusions

a. The Chomsky Hierarchy is a powerful, flexible way to represent the complexity of patterns—both looking at sets of strings and relations between sets of strings—and natural language patterns can be studied within it.

b. The phonological rule systems we studied last semester essentially define regular relations.

(i) To be clear: with such rules, we could have described more more complex relations, like context-sensitive

(ii) But we didn't because the phenomenon we studied didn't actually need it.

c. Optimality theory is another way to define grammars. The kinds of languages generated by OT grammars depends in large part (but not entirely!) on the constraints chosen to be part of the OT grammar. Still, for the most part, the relations we study are actually regular!

References

Albro, Dan. 2005. A Large-Scale, LPM-OT Analysis of Malagasy. Ph.D. thesis, University of California, Los Angeles.

Chomsky, Noam. 1956. Three Models for the Description of Language. *IRE Transactions on Information Theory* IT-2.

Hammond, Michael. 1999. *The Phonology of English: A Prosodic Optimality-Theoretic Approach*. Oxford University Press.

Heinz, Jeffrey. 2005a. Optional Partial Metathesis in Kwara'ae. In *Proceedings of AFLA 12*. UCLA Working Papers, pages 91–102.

Heinz, Jeffrey. 2005b. Reconsidering Linearity: Evidence from CV Metathesis. In *Proceedings of WCCFL 24*. Cascillida Press.

Heinz, Jeffrey. 2007. The Inductive Learning of Phonotactic Patterns. Ph.D. thesis, University of California, Los Angeles.

Johnson, C. Douglas. 1972. *Formal Aspects of Phonological Description*. The Hague: Mouton.

Kaplan, Ronald and Martin Kay. 1981. Phonological Rules and Finite State Transducers. Paper presented at ACL/LSA Conference, New York.

Kaplan, Ronald and Martin Kay. 1994. Regular models of phonological rule systems. *Computational Linguistics* 20(3):331–378.

Kisseberth, Charles. 1970. On the Functional Unity of Phonological Rules. *Linguistic Inquiry* 1:291–306.

Kobele, Gregory. 2006. Generating Copies: An Investigation into Structural Identity in Language and Grammar. Ph.D. thesis, University of California, Los Angeles.

Koskenniemi, Kimmo. 1983. Two-level Morphology. Publication no. 11, Department of General Linguistics. Helsinki: University of Helsinki.

Kracht, Marcus. 2003. *The Mathematics of Language*. Mouton de Gruyter.

Ladefoged, Peter. 2000. *A Course in Phonetics*. Heinle, 2nd edition.

Ladefoged, Peter and Ian Maddieson. 1996. *Sounds of the World's Languages*. Blackwell Publishers.

McNaughton, R. and S. Papert. 1971. *Counter-Free Automata*. MIT Press.

Pullum, Geoffrey and James Rogers. 2007. Aural Pattern Recognition Experiments and The Subregular Hierarchy. In *Proceedings of 10th Mathematics of Language Conference*, edited by Marcus Kracht. pages 1–7. University of California, Los Angeles.

Shieber, Stuart. 1985. Evidence against the context-freeness of natural language. *Linguistics and Philosophy* 8:333–343.

Vijay-Shanker, K. and D.J. Weir. 1994. The equivalence of four extensions of context-free grammars. *Mathematical Systems Theory* 27:511–546.