8 Phonology

ABIGAIL COHN

1 Introduction

Consider the “words” shown in (1):

(1)		I	II	III	
	a.	xoda [x] as in German **ch**.	poda	poda[z]	(cf. coda, codas)
	b.	rudih	hurid	hurid[z]	(cf. hoard, hoards)
		ngatus = [ŋ]	matus	matus[ɪz]	(cf. mattress, mattresses)
	c.	bnick	blick	blick[s]	(cf. block, blocks)

Fluent speakers of English would agree that none of these are actual words of English, yet most speakers would also agree that those in column I are not possible words, while those in column II are. In addition, most speakers would agree that the plurals of the would-be words in column II would be pronounced as indicated in column III. How do we know this? Our knowledge of the sound patterns of our native language(s) comes not through memorizing a list of words, but, rather, by internalizing information about the allowed and disallowed sound patterns of that language. As fluent speakers of English, we know which sounds, or segments, occur in our language and which don't. For example, in (la), the [x] sound of German (written **ch** in borrowings from German, as in the German pronunciation of *Ba**ch***) just doesn't occur in English. In addition, some sounds which are sounds of English are nevertheless restricted in the position where they occur within the word. For example, as shown in (1b), the sound represented by the spelling sequence **ng** [ŋ] can occur in the middle (si**ng**er) or end (si**ng**) of a word, but not the beginning, and **h** occurs at the beginning (**h**ot) or middle (a**h**ead), but not the end of a word. We also know which sounds can be combined into a sequence. Thus in

(1c), **bl** is an allowable sequence at the beginning of a word (**bl**ue), while **bn** is not. Finally, we also know how to manipulate alternating sound patterns. For example, in the regular formation of the plural in English, what is written as **s** or **es** is pronounced [s], [z], or [ɪz] depending on certain properties of the last sound of the word; as native speakers, we automatically produce the expected forms (block[s], hoard[z], mattress[ɪz]). It is this knowledge about sound structure – which sounds occur, what their distribution is, how they can be combined and how they might be realized differently in different positions in a word or phrase, that constitutes the study of *phonology*.

Central to the study of phonology is observing ways in which languages differ in terms of their sound structure, as well as what the full range of attested possibilities or options are within each facet of the phonology. In this chapter, we explore some of the central cross-linguistic generalizations about sounds, using some of the theories and tools that allow us to insightfully analyze these patterns. We will focus on three areas: sound inventories and contrasts (section 2), structure above the level of the sound unit or segment (section 3), and structure internal to the segment (section 4). Finally we conclude (section 5) with a brief discussion of phonology as a system.

2 Inventories and Contrasts

2.1 *Inventories*

All languages have consonants and vowels. Consonants are sounds with a constriction in the vocal tract, while vowels lack such a constriction. Vowels can serve as the core of a syllable (see below in section 3), while consonants generally cannot. Consonants must co-occur with vowels to produce forms which are pronounceable. Both consonants and vowels can be defined in terms of where in the mouth they are produced and how they are produced. For consonants, this is characterized in terms of *place* and *manner* of articulation. Place of articulation indicates where the obstruction occurs. The places relevant in English, as we'll see below, include the lips (labial), the tongue tip approaching the teeth (dental), the tongue tip approaching or contacting the ridge behind the teeth (alveolar), or a bit farther back (palato-alveolar), the body of the tongue approaching or contacting the hard palate (palatal) or the soft palate (velar), and finally the position of the vocal cords, or the glottis (glottal). The manner of articulation indicates the degree of constriction: complete closure (stops), noticeable obstruction (fricatives) or a combination of closure and obstruction (affricates), closure in the mouth with air escaping through the nose (nasals), or only slight approximation (liquids and glides). Vowels are generally characterized in terms of the height of the tongue or jaw (high, mid, low) and the relative backness of the tongue (front, central, back). In addition, other properties play a role, such as whether the vocal cords are close together and vibrating (voiced) or farther apart, allowing freer passage of air from the lungs.[1]

So far we have presented examples using English spelling, with some additional pronunciation information provided in [ ]'s. English spelling is sorely inadequate for describing the sounds of current American English accurately. The 26 symbols of the Roman alphabet are not sufficient to represent all of the consonant and vowel sounds of English (as we'll see below there are 39), and so in some cases two symbols are used to represent a single sound. But this isn't the only problem. In order to describe sounds reliably, we need a completely systematic relationship between sound and symbol, something which English spelling doesn't provide, since in the English spelling system there are far too many correspondences of sound to symbol. Take for example the sound [k], which can be represented by several different symbols or symbol combinations (as shown in (2a)) and the letter **c** which can represent various different sounds (in (2b)).

(2) a. symbols used to represent the sound [k]
cat
kite
khan
quite (**qu** = [kw])
e**ch**o
pa**ck**
bo**x** (**x** = [ks])
b. sounds represented by the letter **c**
(not including two-symbol combinations, such as **ch**)
[k] **c**at
[s] **c**ite
[tʃ] **c**ello

In addition we often need to be able to include more pronunciation detail. (The need for greater detail is true even of those languages which have much better spelling systems than English.) We need what is called *phonetic transcription*. The International Phonetic Alphabet (IPA) is a system of phonetic transcription which allows us to systematically represent the sounds of any language. This system, developed by the International Phonetic Association (founded in 1886) is periodically updated, to reflect changes in general thinking on transcription and to include new speech sounds which have been "discovered." In 1989, the International Phonetic Association had a congress to address such questions and fine-tune the system in a number of ways. The common systems of phonetic transcription used in the United States differ in a few small ways from the standard IPA, but still most such systems are quite close to the IPA.

A sound inventory is the selection of sounds occurring in a particular language. Looking across the inventories of the languages of the world, we find that the number of consonants and vowels, as well as the specific selection of sounds, varies enormously from one language to another. In his study of the sound inventories of 317 languages, Maddieson (1984) found that the number of consonants in a language ranged from 6 to 95, with a mean of 22.8; while the number of vowels ranged from 3 to 46 with a mean of 8.7; and 62.1 percent of the languages in his sample have between 5 and 9 vowels.

Considering this range of sound inventory size, let's see how the sound inventory of American English compares, shown in (3). For the consonants, the places of articulation are the column headings and the manners of articulation are the labels for the rows. When two sounds appear within a single cell in the table, the one on the left is voiceless (without vocal cord vibration) and the one on the right is voiced (with vocal cord vibration). For the vowels, in addition to tongue backness (marking the columns) and height (marking the rows), the adjacent pairs within a category differ in "tenseness" vs. "laxness." (C = consonant, V = vowel.)

(3) Sound inventory of English

C's	labial	dental	alveolar	palato-alv.	palatal	velar	glottal
stop	p b		t d			k g	
fricative	f v	θ ð	s z	ʃ ʒ			h
affricate				ʧ ʤ			
nasal	m		n			ŋ	
liquid			r l				
glide					j	w	

V's	front	central	back
high	i ɪ		ʊ u
mid	e ɛ	ɚ ʌ	ɔ o
low	æ	a	

Diphthongs: aj, aw, oj

There is some variation in the number of sounds argued to occur in English (for example should the affricates, [ʧ] (***church***) and [ʤ] (***judge***), be treated as single units or as sequences of sounds?); however, the characterization of American English in (3) with 24 consonants, 12 vowels and 3 diphthongs (vowel glide combinations that function as a single unit) is fairly common. Thus, English has an average-sized consonant inventory, though notable in its rich array of fricatives. There are whole classes of other consonants that English doesn't exemplify, such as clicks, found in some languages of Southern Africa. With 12 vowels, English has a relatively rich vowel inventory, especially considering that the distinctions are all made using only the two dimensions of tongue height and backness. (In the inventory above, we haven't included schwa [ə], which occurs only in unstressed position.) Some languages make additional, or different, vowel contrasts. For example, in English the front vowels have an unrounded lip position and the non-low back vowels have a rounded lip position, but in many other languages, there are both unrounded and rounded front and / or back vowels (e.g. French *riz* [ri] "rice," with a high front unrounded vowel, vs. *rue* [ry] "street," with a high front rounded vowel and *roux* [ru] "red" (of hair), with a high back rounded vowel).

Compare the English inventory with that found in Arabic (Modern Literary), as shown in (4):

(4) Sound inventory of Arabic

C's	labial	dental	dental pharyn-gealized	alveolar	alveolar pharyn-gealized	palato-alveolar	palatal	velar	uvular	pharyngeal	glottal
stop	b bː	t d tː dː	tˤ dˤ tːˤ dːˤ					k g kː gː	q qː		ʔ ʔː
fricative	f fː	θ ð ðː	ðˤ ðːˤ	s sː z zː	sˤ sːˤ	ʃ ʃː			χ ʁ χː ʁː	ħ ʕ ħː ʕː	h hː
affricate						ʤ					
nasal	m mː			n nː							
liquid		l lː	lˤ lːˤ	r rː							
glide							j jː	w wː			

V's	front	central	back
high	i iː		u uː
low		a aː	

In Modern Literary Arabic, we find a very small vowel inventory, only three distinct vowel qualities (though length differences (indicated by ː for a long vowel or consonant) also result in differences in meaning, e.g. [dur] "turn!" vs. [duːr] "houses"), but a very rich consonant inventory. Not only are most of the consonants seen in English found here, but there are additional places of articulation, notably at the back of the mouth (uvular – the back of the soft palate, and pharyngeal – the throat). In addition, there is a contrast between plain consonants and those with a superimposition of a back tongue position (pharyngealized) and finally consonants also contrast for length ([ba**r**a] "sharpen" vs. [baɾːa] "acquit"). Including all these contrasting dimensions, there are 48 consonants in this dialect, though there is some variation in the consonant inventory of different dialects of Arabic.

While there is a tendency for languages with large consonant inventories to have correspondingly small vowel inventories and vice versa, this is not necessarily the case. Consider for example Rotokas, spoken in Papua New Guinea (following Maddieson 1984, the smallest inventory found in his database), with a very common 5 vowel inventory, but only 6 consonants for a total of only 11 segments.

(5) Sound inventory of Rotokas

C's	labial	alveolar	velar
stops	p	t	k g
fricatives	β		
liquids		ɾ	

V's	front	central	back
high	i		u
mid	e		o
low		a	

While there is great variation in the segments that occur in particular languages – Maddieson identifies over 800 in his study – strong predictions can nevertheless be made about which sounds will occur. Some sounds and categories of sounds are just more common than others. For example, all languages have stops, but not all languages have fricatives. Beyond these basic observations, there are also many cases where the presence of one property implies the presence of something else in the same system; such generalizations are called implicational language universals. For example, if a language has the mid vowels [e, o] (as in English, *bait* [bet] and *boat* [bot]), it can be predicted that it will also have the high vowels [i, u] (English *beat* [bit] and *boot* [but] and the low vowel [a] (English *pot* [pat]); but the converse doesn't hold, as we've seen in Arabic which has [i, u, a], but lacks [e, o].

2.2 *Contrast*

When we characterize the inventory of sounds of a language, we need to draw an important distinction between those sounds that can be used to make meaningful contrasts in a language vs. those that occur, but are predictable in their distribution. The description of the inventories of English, Arabic, and Rotokas, provided above, present those sounds argued to be distinctive in the language (though, as we discuss below in section 4, the status of [ŋ] in English is debatable).

In order to determine the status of such sounds, we use a simple test to determine if two sounds are distinct by looking to see if there are *minimal pairs*. Minimal pairs (or sets) are words with distinct meanings differing only in one sound. Thus we can show that [m] and [n] (differing only in place of articulation) are distinct sounds in English, since the substitution of these sounds alone is enough to change the meaning of a word:

(6) **m**eat vs. **n**eat
si**mm**er vs. si**nn**er
ra**m** vs. ra**n**

In (6) we see that the presence of [m] vs. [n] at the beginning, middle, or end of a word results in different words.

If a sound is used distinctively in a particular language, it is what we call a *phoneme* in that language (and is represented in / /'s). Phonemes are argued to be the level of representation at which segments are encoded in lexical entries (the forms in our mental dictionaries) and the level at which speakers judge "sameness" and "differentness." However, phonemes can vary in their actual realization or pronunciation, depending on the context of the neighboring sounds, the structure of the utterance, and so forth.

Two languages may have the same sounds or *phones* (the actual phonetic events, represented in [ ]'s), but their grouping into phonemes or contrastive units might be different. In English, for example, the sounds [b, p, p^{h}] all occur (that is, voiced, voiceless unaspirated, and voiceless aspirated); while [p^{h}] and [b] contrast, whether [p] or [p^{h}] will appear is predictable from the context, as exemplified in (7). *Buy* [baj] contrasts with *pie* [p^{h}aj], but the realization of a voiceless stop as aspirated (*pie* [p^{h}aj]) or unaspirated (*spy* [spaj]) is predictable and there are no minimal pairs for [p] and [p^{h}]. (We use an asterisk to indicate something non-occurring or "ungrammatical.") Thus these three phonetic categories are mapped to only two abstract phonological categories. Yet in Thai, all three sounds occur and can produce contrasts in meaning, as shown by the minimal set in (7).

(7)

	Thai						*English*	
		phonemes		*phones*		*phonemes*		
[baa]	"crazy"	/b/	—	**[b]**	—	/b/	buy	[baj]
[paa]	"aunt"	/p/	—	**[p]**		/p/	pie	[p^{h}aj]
[p^{h}aa]	"cloth"	/p^{h}/	—	**[p^{h}]**			spy	[spaj]
						but no *[paj] or *[sp^{h}aj]		

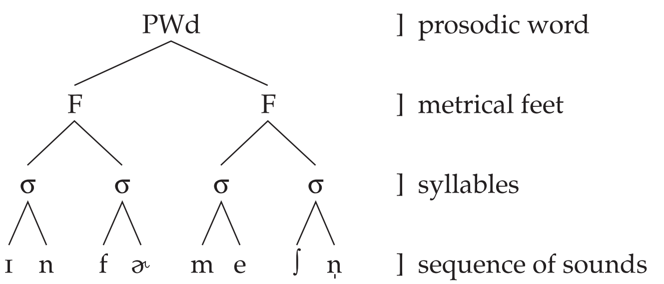
To summarize, these three phones [b, p, p^{h}] constitute three separate abstract sounds or phonemes in Thai, but only two in English.

In English [p, p^{h}] are phones which stand in a special relationship to each other, since they are part of the same phoneme (usually taken to be /p/). Such sounds are called *allophones*. We can capture this relationship by describing the distribution, e.g. [p^h] occurs at the beginning of words and [p] occurs after [s]. (There is a lot more to this pattern, but we won't pursue it here.) Or we can go a step further and argue that the phoneme /p/ occurs at an abstract or *underlying* level and account for the observed surface distribution with a rule (typically of the form a → b / c__d, which says that "a becomes b in the environment following c and preceding d"). This general approach is fundamental to the view of generative phonology (see Chomsky and Halle 1968, Kenstowicz and Kisseberth 1979) where the goal is to develop a theory which accurately models a speaker's knowledge of his or her language; we return to the issue of rules in section 4.

3 Structure above the Level of the Segment

The sound structure of a word (a unit which can be defined on several linguistic levels, including morphologically and phonologically) includes not only the sequence of sounds (made up in turn of bundles of distinctive features, as discussed in section 4), but also entails the hierarchical grouping of these sounds. Let's take the English word *information* as an example which we can use as a reference point:

(8)



This word consists of a sequence of sounds **ɪ-n-f-ɚ-m-e-ʃ-n̩**. These sounds are grouped into sequences of consonants and vowels, known as syllables (σ). Most speakers of English would agree that this form consists of four syllables broken up as **ɪn-fɚ-me-ʃn̩**. Consonants and vowels are grouped into syllables in non-arbitrary ways, with a vowel forming the core or **nucleus** (such as [m**e**], and consonant or consonants preceding (**onset**, such as [**m**e]) or following (**coda**, such as [ɪ**n**]). In the final syllable [**ʃn̩**], the nucleus is **n̩**, which is a syllabic nasal, serving the role of a vowel. These syllables are in turn organized into stress groupings (**ɪ̀n-fɚ**) (**mé-ʃn̩**). The third syllable is the most prominent (primary stress, indicated with a ´) and the first also has some prominence (secondary stress, indicated with a `). These patterns of prominence can be accounted for by grouping the syllables together into units known as metrical feet (F). Finally the feet are grouped together into the Prosodic Word (PWd).

The Prosodic Word often has the same shape as what we would define morphologically as a word, but not necessarily. There are, for example, grammatical words, which we take to be words morphologically, but which can't stand on their own phonologically, such as *a*, or *the*. The syllables, feet, and prosodic words are together the prosodic structure of a word. Words in turn can be grouped into higher levels of prosodic structure as well.

We can focus on the structure at the level of the segment and above, how segments are combined, how syllables, metrical feet, and prosodic words are constituted; and we can in turn examine the subsegmental structure, how distinctive properties of sounds are organized into segments. In the remainder of this section, we examine syllable structure as an example of the nature of structure above the segment and then turn to the question of subsegmental structure in section 4.

3.1 Syllable structure

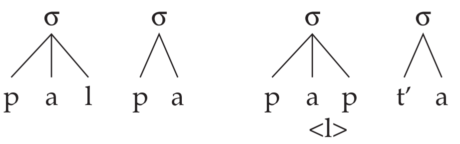
Many processes result in the insertion or deletion of a segment. This is often due to the influence of syllable structure. Consider an example from Korean, shown in (9) where we observe that sometimes a cluster of consonants occurs and sometimes one of the members of the cluster is deleted. This is an example of what we call an *alternation* where the same morpheme varies in its realization, conditioned by some aspect of the sound system (in this case the allowable syllable structure). The result is an alternation between the presence of a consonant and zero in morphologically related forms. ([t'] represents a voiceless alveolar stop with a stronger articulation than a plain voiceless stop.)

(9) Consonant ~ Zero alternations in Korean clusters

root	*+ vowel initial suffix*	*+ consonant initial suffix*
	-**a** nominalizing suffix	-**t'a** infinitive
/palp/ "tread on"	palp + a "treading on"	pap + t'a "to tread on"
/salm/ "boil"	salm + a "boiling"	sam + t'a "to boil"

The basic syllable structure in Korean is (C)V(C). The underlying clusters (/lp/ and /lm/ are allowed to surface before a vowel initial suffix, since the second member of the cluster can be syllabified as the onset of the second syllable, producing *palpa* and *salma*. But when the root occurs before a consonant initial suffix (verbs cannot occur without some kind of suffix), the first consonant of the cluster, in the cases illustrated here /l/, is deleted, producing *papt'a* and *samt'a*. (In other cases, it is the second consonant which is deleted.) The syllabification of forms with vowel initial and consonant initial suffixes respectively is shown in (10) for /palp/ (where < > indicates a segment not incorporated into the syllabic structure):

(10)



Here we can see that this deletion is directly driven by the allowable syllable structure.

As noted in section 1, restrictions also exist on possible sequences of sounds. For example in English, *[bn] can't occur at the beginning of a word (11a) or at the end of a word (11b), but it is not the case that the sequence [bn] is always bad in English.

(11) a. **bnick*
b. **kibn*
c. *lab-network*
d. *drabness*
e. *Abner*

In (11c), this sequence is fine, but the word is a compound and we might argue that it consists of two prosodic words grouped together (into a structure such as $[[\text{lab}]_{pwd}[\text{network}]_{pwd}]_{pwd}$) and therefore it is not held to the same restrictions. The fact that (11d) is allowable might be attributed to the sounds belonging to different morphemes (*drab* and *-ness*). But in (11e) there aren't two words or two morphemes. So what is the difference between [bn] in (11a and 11b) and in (11d)? In the latter case, the [b] and [n] are in different syllables, while in the former they are in the same syllable.[2] The restriction holds of a sequence within a syllable and seems to be due to the fact that [b] and [n] are too similar in terms of *sonority*. Sonority can be defined loosely as the degree of constriction in the mouth during the production of a particular sound. Most important for our purposes here is the observation that there is a hierarchy of how *sonorous* sounds are. Vowels are more sonorous than consonants; and within the consonants, further divisions can be made. Stops, which have complete closure, and fricatives, which have enough of a constriction to create frication or noise (as well as affricates), together are known as *obstruents*, since there is a significant obstruction in each of these cases. These are less sonorous than the nasals, liquids, and glides, together known as *sonorants*. Thus we find the following strong cross-linguistic pattern:

(12) Sonority hierarchy

more sonorous				*less sonorous*
vowels	>	sonorants	>	obstruents

The sonority hierarchy characterizes the behavior of sounds in syllable structure and many other aspects of phonological patterning. Whether finer grained distinctions of the sonority hierarchy are required is a question open to much debate, though we will see some evidence for some additional distinctions below.

As mentioned above, syllables are organized around vowels, sometimes preceded and / or followed by consonants. All of the examples in (13) are well-formed English syllables (and in these cases independent words too).

(13)

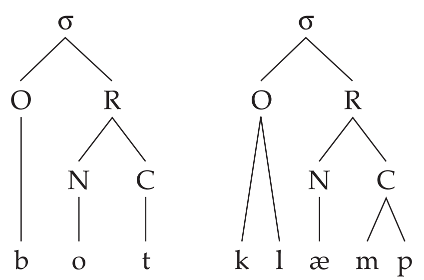
coda: / *onset*:	∅		C		CC		CCC	
∅	oh	[o]	ode	[od]	old amp	[old] [æmp]	amps	[æmps]
C	bow	[bo]	boat	[bot]	bolt	[bolt]	bolts	[bolts]
CC	blow	[blo]	bloat clam	[blot] [klæm]	clamp	[klæmp]	clamps	[klæmps]
CCC	spree	[spri]	split	[splɪt]	splint	[splɪnt]	splints	[splɪnts]

In English, anything from a single consonant to a complex structure of up to three consonants preceding and four following may constitute a well-formed syllable. (Four consonants following the vowel are not included in (13); an example is *texts* [tɛksts].) Many restrictions hold, however, on possible combinations of consonants preceding or following the vowel and only a small subset of the logically possible combinations occur. For example, in three-consonant clusters starting a syllable ($C_1C_2C_3$), the first sound (C_1) must be [s], followed by a voiceless stop ([p, t, k]), followed by a liquid ([r, l]) or glide ([j, w]). Many of the occurring patterns can be characterized with reference to the sonority hierarchy (12), though other factors also come into play. Thus in CCC clusters the pattern of C_2 and C_3 follows the sonority hierarchy, with the beginning of syllables showing a rise in sonority going from C_2 to C_3: stops followed by the more sonorous liquids and glides. Some evidence for the fact that a more fine-grained sonority hierarchy is required comes from the fact that stops (voiced or voiceless) followed by liquids or glides are well formed (e.g. **bl**oat, **cl**am), but stops followed by nasals are not (*bn, *kn). Yet nasals are also members of the class of sonorant consonants. This suggests that the sonorant consonants should be further divided into the oral sonorants (the liquids and glides) and the nasals, with the oral ones being more sonorous than the nasals. But the occurrence of [s] preceding such clusters is not predicted even with further modification of the sonority hierarchy, since [s] is not less sonorous than the stops, and therefore requires a distinct explanation.

Similarly in characterizing what coda clusters (the sequences of consonants following a vowel) can occur in English, sonority also plays an important role. In general, the first member of a two member coda cluster must be of the same or greater sonority than the second member (e.g. le**nt**, be**lt**, li**ft**, mi**st**, a**pt**). In most monosyllabic words with more than two consonants following the vowel, these forms are morphologically complex, usually involving the [s] or [z] of the plural marker or third person singular or the [t] or [d] of the past tense (though there are some three consonants clusters which occur as codas in the same morpheme, such as [kst] in *text*). Such patterns can be characterized simply if we make reference to the syllable, but are much harder to characterize if we only refer to the string of segments.

Good evidence thus exists for making formal reference to the syllable as part of the hierarchical structure of the phonological system to account for observed alternations and also to be able to capture consonant sequencing restrictions. In addition, the syllable is often argued to be divided into subparts. Evidence for this comes from the fact that co-occurrence restrictions hold on the consonants preceding the core of a syllable, as well as following, but not generally across the subparts of the syllable. One general approach to the internal organization of the syllable is as shown in (14), where the substructure of *boat* and *clamp* are illustrated:

(14)



O = onset
R = rime
N = nucleus
C = coda

Based on a wide range of evidence, there is argued to be a major break in the syllable between the onset and the rime constituents. The division into onset and rime allows us to capture various consonant sequencing restrictions and is also relevant for other aspects of the phonology, as well as language games and poetry. The rime corresponds to the unit which rhymes, e.g. **oat**, b**oat**, bl**oat**; and the onset is the unit shared in poetic patterns of alliteration, e.g. **bl**ue, **bl**ow, **bl**ithe, **bl**oat. The rime is then further divided into the nucleus, the core of the syllable which contains the vowel or vocalic elements(s), and the coda, which contains any following consonant(s). In English, the only required element of the syllable is the nucleus (e.g. **oh** [o], **I** [aj]), although in many languages the onset is also an obligatory part of the syllable.

How much explicit or formal internal structure to the syllable is warranted and how it should be encoded is a much debated question, which we won't pursue here, but reference to some degree of substructure of the syllable is useful in capturing insightful generalizations about allowable sequencing restrictions and other aspects of sound distribution. Indeed in English, we can capture the pattern presented in (11) by observing that the sequence [bn] cannot occur together as part of an onset or coda. In addition, reference to syllable subconstituency allows us to capture the broader distribution of sounds in many cases. For example, as noted in (1), the distribution of /h/ in English is limited: it can occur only in the onset of a syllable (and if it is not word-initial, only if the syllable is stressed, e.g. *vehicle* [véɪkl̩] vs. vehicular [ve**h**íkjəlɚ]).

While it is relatively straightforward to count the number of syllables in a word, it is often trickier to decide where to divide syllables in words of two or more syllables. Typically in the case of (C)VCV, the division is before the medial C, (C)V$CV (where $ is used to indicate a syllable break). In English, the situation is additionally complicated by the stress pattern. In words such as those in (15a), it is widely agreed that the syllable divisions are as shown, characteristic of the strong cross-linguistic tendency.

(15) a. attáck [ə$tǽk]
belów [bə$ló]
b. áttic [ǽtɪk]
béllow [bɛ́lo]

However, many researchers have argued that in the cases such as (15b), the medial consonant either belongs to the first syllable or is shared by the two syllables in order to account for otherwise systematic observations about the relationship between syllable structure and stress in English. (Even though the middle consonants in the forms in (15), except for *below*, are written with a doubled consonant (tt, ll), they are just single consonants. The doubling of a consonant in English spelling usually indicates something about the pronunciation of the preceding vowel, not the pronunciation of the consonant itself (compare *tapper* [tæpɚ], *taper* [tepɚ]).

In the case of (C)VCCV(C), the syllabification depends on the specific sequence of consonants. In English, if the CC is an allowable word onset (and therefore an allowable syllable onset) the syllable division will be before both consonants (16a), but otherwise it will be between the two consonants (16b).

(16) a. apply [ə$**pl**aj] cf. plea [pli]
abrupt [ə$**br**ʌpt] cf. brush [brʌʃ]
b. Adler [æd$lɚ] *[dli]
Abner [æb$nɚ] *[bni]
ardent [ar$dn̩t] *[rdi]

Some other languages show much greater restrictions on syllable structure than English does. Consider some examples from Japanese in (17).

(17) Allowable syllables in Japanese: CV, V, CVN, CVC

a. CV, V

[ki] "tree"
[kokoro] "heart"
[mado] "window"
[tegami] "letter"

[ito]	"string"
[origami]	"paper folding"

b. N$C

[tombo]	"dragonfly"
[hantai]	"opposite"
[neŋkin]	"pension"

c. C$C

[kitte]	"stamp"
[onna]	"woman"
[hakka]	"peppermint"
[kaʃʃa]	"pulley"

As illustrated in (17), only (C)V and (C)VC occur in Japanese (as well as some limited cases of long vowels (C)VV(C)). CV syllables can occur in any position in the word (17a). But CVCs are allowed only if the coda consonant is a nasal (17b), or part of a geminate (long consonant) (17c), and in these cases usually followed by another syllable. Thus, [tom] is a well-formed syllable when followed by [bo], but it would not be an allowable syllable, if it occurred on its own or as the final syllable in a word. A final alveolar nasal (as in [neŋki**n**] above in (17b)) is well formed, but other nasals and other consonants in this position are not allowed.

Additional evidence for the allowable patterns can be seen by looking at the ways foreign words are modified when they are borrowed into Japanese. Let's consider what happens to some words borrowed from English, as shown in (18).

(18) Borrowings from English into Japanese:

	word	*English*	*Japanese*
a.	pin	[pɪn]	[pin]
	pie	[paj]	[paj]
	Chicago	[ʃɪkago]	[ʃikago]
b.	million	[mɪljən]	[mirion]
	avocado	[avəkado]	[abokado]
	rally	[ræli]	[rarii]

Some words are borrowed as is shown in (18a) (with slight modifications of vowel quality in some cases), or with modifications to any non-occurring segments, with these being substituted by a similar sound which does occur in Japanese (18b). (Some of the vowels of English (e.g. [i, e, o, u]) are perceived to be long in Japanese, indicated here with the doubling of the vowel symbol.)

Of particular interest are cases where non-allowable consonant clusters occur; in such cases, Japanese uses the strategy of adding extra vowels, as illustrated in (19):

(19) More borrowings from English into Japanese

	word	*English*	*Japanese*
a.	free	[fri]	[fUrii]
	spray	[spre]	[sUpUree]
b.	peak	[pik]	[piikU]
	kiss	[kɪs]	[kisU]
	Bill	[bɪl]	[birU]
	beat	[bit]	[biitO]
c.	speed	[spid]	[sUpiidO]
	cross	[krɔs]	[kUrosU]
	test	[tɛst]	[tesUtO]
	street	[strit]	[sUtOriitO]
	contrast	[kantræst]	[kontOrasUtO]
	baseball	[besbɔl]	[basUbarU]

Consider first cases with onset clusters as shown in (19a). The inserted vowels are indicated in upper case symbols. (The vowel which is inserted in these cases is usually [u] (U), except after alveolar stops, where an [o] (O) is inserted.) (19b) shows cases of either monosyllables or final syllables of the shape CVC. These too are modified, since a consonant can occur in coda position only if it is followed by an appropriate consonant in the next syllable in the same word. Finally cases with both onset clusters, final consonants and final clusters are shown in (19c). All of these clusters are broken up into many more syllables in Japanese than found in the original English source.

In the case of non-allowable clusters in borrowed words, other languages delete segments. Consider what happens to final consonant clusters in Indonesian in words borrowed from English or Dutch. In Indonesian, in general the allowable syllable structure is (C)V(C), so final clusters in borrowed words pose a problem. As shown in (20), the final clusters are simplified by deleting the final consonant (similar to the pattern seen for Korean above in (9), although for those examples, it was the first member of the cluster which was deleted).

(20)	*word*	*English / Dutch*	*Indonesian*
	sport	[spɔrt]	spor
	aqueduct	[ækwədəkt]	akuaduk
	tolerant	[taləˑrn̩t]	toleran
	test	[tɛst]	tes

To account for such systematic syllable patterns, researchers have proposed various devices including rules, templates, and well-formedness conditions. A current approach, *Optimality Theory*, involves the idea of competing constraints, which can be ranked in importance with respect to each other. Due to such ranking, a less important constraint can sometimes be violated in order to obey a more important constraint (see Prince and Smolensky 1993, McCarthy and Prince 1993 inter alia). Languages differ in how they rank particular constraints. If we have correctly identified the relevant constraints (a major research agenda in itself), then the set of logically possible rankings of those constraints should match up with the range of sound patterns seen across languages. Optimality Theory offers an insightful account of syllable patterns and makes strong predictions about allowable syllable types cross-linguistically, and it also accounts for certain implicational universals such as the fact that if a language allows CVC syllables it will also allow CV syllables and if it allows V syllables, again it will also allow CV ones.

As discussed by a wide range of scholars, the ideal syllable is CV. Syllables minimally consist of a vowel; onsets are preferred; and codas are dispreferred. To account at the same time for the preference for CV syllables and the range of cross-linguistic variation observed in syllable structure, two general sorts of constraints interact. First there are markedness constraints – constraints which capture systematic cross-linguistic generalizations. In the case of the preference for CV syllables, this has been argued to emerge from three constraints, stated here informally:

(21) Syllable structure markedness constraints:

	constraint	*informal definition*
a.	Nuc	Syllables must have a nucleus
b.	Onset	Syllables must have an onset
c.	NoCoda	Codas are not allowed

If this were all there were to the story, all languages would have only CV syllables, but this is clearly not the case. There are also constraints that mediate between the underlying representation or abstract form (the input to the constraints) and the actual realization of the form, or the output of the constraints. The two constraints relevant for our purposes, again stated informally, limit how different the input and output can be. (* = Don't)

(22) Input / output constraints

	constraint	*informal definition*
a.	*Add	Only the material of the input should appear in the output; don't add material to the input
b.	*Delete	Underlying material should be incorporated in the output; don't delete material from the input

There are other constraints that can also affect syllable structure, but these five constraints are sufficient for our discussion here. To test constraint rankings, we compare the input of a form and a list of possible (expected) outputs (placed in the leftmost column in what is termed a "tableau") with respect to a particular ranking of the relevant constraints (placed in columns, going from higher to lower ranking as we go from left to right). No matter what the relative ranking of these five constraints in a particular language, if we have an input or underlying form of the shape CV(CV)(CV), then all of the above constraints, those affecting syllable structure and those affecting input / output relations, can be satisfied. This is true in both English and Japanese, as shown in (23a) for English *banana* and (23b) for Japanese [kokoro] "heart"; even though as we've seen above they have very different syllable patterns. In these tableaux, the constraints are all unranked, indicated by the dashed vertical lines, in contrast to solid vertical lines that we'll see in the tableaux below.

(23) a. English *banana*

/bənænə/	Nuc	Onset	NoCoda	*Add	*Delete
[bə\$næ\$nə]	√, √, √	√, √, √	√, √, √	√, √, √	√, √, √

b. Japanese [kokoro] "heart"

/kokoro/	Nuc	Onset	NoCoda	*Add	*Delete
[ko\$ko\$ro]	√, √, √	√, √, √	√, √, √	√, √, √	√, √, √

Here a checkmark in the relevant cell indicates that the constraint is met; there are three checkmarks in each cell referring to each of the three syllables in these cases. It is the combination of Onset and NoCoda (no matter what their ranking) that ensures that an intervocalic consonant (VCV) will be syllabified with the consonant as the onset of the second syllable (V\$CV).

Let's now consider some cases where the same input or underlying form results in different outputs in different languages. Consider the English word *test*, which as we saw above is realized as [tesuto] in Japanese and [tɛs] in Indonesian. I leave Nuc and Onset out of the following discussion, as they are met by all of the cases we are considering. This particular case doesn't provide evidence for the ranking of Onset, but the abundance of vowel initial forms in all three languages shows that Onset can be violated under certain circumstances. On the other hand, Nuc is very high ranking, and therefore unviolated, in each of the three languages.

In English, the input [tɛst] matches the output, even though it violates NoCoda twice. This provides evidence that NoCoda is lower ranked than both *Add and *Delete. In other words, meeting the requirements of the input / output constraints is more important in English than adhering to the markedness constraints.

(24) English *test* [tɛst]

/tɛst/	*Add	*Delete	*NoCoda
a. ☞ [tɛst]	√	√	**
b. [tɛs]	√	*!	*
c. [tɛ]	√	*!*	√
d. [tɛstV]	*!	√	*
e. [tɛsVtV]	*!*	√	√

The optimal or best formed candidate in this case is [tɛst], indicated by ☞. An ! indicates an insurmountable violation. This is followed by shading of the successive cells in the same row, indicating that the adherence to these lower ranked constraints isn't relevant to the outcome. (23a) is the optimal candidate in this case, even though this form violates NoCoda twice. This is still preferable to a violation of either *Add (24d and e) or *Delete (24b and c), providing evidence that both of these constraints outrank NoCoda (hence NoCoda is positioned to the right, separated by a solid vertical line). Since both *Add and *Delete have to be met, we don't have evidence for their relative ranking in English.

The pattern in Japanese is very different. In Japanese, priority is given to the markedness constraints over the input / output constraints. In order to meet the high ranking NoCoda constraint, vowels are inserted, providing evidence that *Delete outranks *Add, as shown in the tableau in (25):

(25) Japanese [tesuto] "test"

/tɛst/	NoCoda	*Delete	*Add
a. [test]	*!*	√	√
b. [tes]	*!	*	√
c. [te]	√	*!*	√
d. [testV]	*!	√	√
e. ☞ [tesVtV]	√	√	**

We see here that (25e) [tesuto], which respects both NoCoda and *Delete, is the optimal candidate. We use V to represent an inserted vowel and assume that it is a language-specific question what the actual quality of the inserted vowel will be. We also leave aside the additional question of the /ɛ/ being realized as [e]. As we saw above in (17b and c), some limited violations of NoCoda are tolerated. The intuition is that coda consonants cannot have their own place specification, rather, they must share it with the following onset consonant, either as part of a geminate or as part of a nasal-stop cluster agreeing in place of articulation.

Finally in Indonesian, we find a case where deletion is tolerated, indicated by the relatively low ranking of *Delete, though this is balanced with a violation of NoCoda, since the optimal form involves one violation of each NoCoda and *Delete (in contrast to English which violates NoCoda twice and Japanese which violates *Add twice).

(26) Indonesian [tɛs] "test"

/tɛst/	*Add	NoCoda	*Delete
a. [tɛst]	√	*!*	√
b. ☞ [tɛs]	√	*	*
c. [tɛ]	√	√	*!*
d. [tɛstV]	*!	*	√
e. [tɛsVtV]	*!*	√	√

The optimal candidate in Indonesian is (26b). Our analysis accounts for the fact that both (26d and e) are eliminated, but more needs to be said about why the optimal outcome is (26b) rather than (26a or c). An additional constraint must be involved; while I won't formalize it here, the intuition is that on one hand a single consonant in coda position is more acceptable than a cluster and on the other, there is a limit to how much deletion the system will tolerate. There is more to the story in Indonesian, since in the case of onset clusters, vowels are inserted rather than consonants being deleted, for example [sətasion] from Dutch *station*, but we leave aside these additional details in our current discussion.

There are clearly additional complexities, since all three languages allow vowel initial words (hence limited violations of ONSET) and more needs to be said about why, in Japanese, a final syllable such as [kin] is allowed but one such as [tom] is not. Finally, additional constraints are needed to account for the division of medial consonant clusters into codas and onsets, e.g. English *abrupt* [ə$brʌpt] vs. *Abner* [æb$nɚ]. In many languages, VCCV will surface as V$CCV if CC is an allowable onset (clearly additional constraints are required to define which consonant clusters are and are not allowable). If CC is not an allowable onset, the VC$CV syllabification would result in a minimal violation of NOCODA.

While I haven't provided a complete account of any of these three cases, we can see that the relative ranking of this limited set of constraints allows us to capture these different strategies of syllabification. Other languages are predicted to show different outputs. For example, the form [tɛstV] would result in a language that had some tolerance of single consonant codas (like Indonesian), but ranked *DELETE over *ADD.

In this section we have seen that reference to syllables as well as subsyllabic constituents offers a more insightful account than one where only reference to the segment can be made. In addition we have looked briefly at how a constraint-based approach, where minimal violation of constraints is tolerated, allows us to account for some of the cross-linguistic variation observed in syllable structure.

4 Subsegmental structure

4.1 Features and segmenthood

Up until this point in our discussion, we have focussed on segments (and larger units). Good evidence for the psychological reality of segments exists, including speaker intuition, alphabetical writing systems, speech errors, and the fact that phonological processes manipulate such units. But there is also good evidence that segments are made up of smaller units and that a more insightful discussion of sound patterning is possible, if we make reference to these smaller units. We have an intuition that [p, b] are more similar than [l, b]. This is because the former share more sound properties than the latter. These sound properties are called *distinctive features*. The notion of distinctive features grows out of the work of Trubetzkoy, Jakobson, and others (see Anderson 1985 for an excellent survey of the history of phonology). While numerous specific systems have been proposed, most current systems have evolved from that proposed by Chomsky and Halle (1968). Most approaches to phonology assume some kind of feature system and take the features to be the smallest building blocks of phonology. Segments thus consist of bundles of features, or feature matrices, as exemplified in (27):

(27) feature matrices

$$\begin{array}{ccc} \text{b} & \text{ɪ} & \text{l} \\ \begin{bmatrix} \text{+consonantal} \\ \text{–continuant} \\ \text{–sonorant} \\ \text{–nasal} \\ \text{labial} \\ \end{bmatrix} & \begin{bmatrix} \text{–consonantal} \\ \text{+high} \\ \text{–back} \\ \text{–tense} \\ \\ \end{bmatrix} & \begin{bmatrix} \text{+consonantal} \\ \text{–continuant} \\ \text{+sonorant} \\ \text{–nasal} \\ \text{coronal} \\ \text{lateral} \end{bmatrix} \end{array}$$

There are many interesting and important issues about the status of features. First there is much debate about an adequate specific set of features which can account for all the occurring sounds in the languages of the world. Additionally there are issues such as the number of values that characterize particular features. There are some features which clearly define two classes, for example [±sonorant], where [+sonorant] defines the class of sonorants and [–sonorant] defines the class of obstruents. Such features are appropriately characterized as two-valued or binary. In the case of other features, their presence or absence seems sufficient, that is, they are single-valued or privative; for example this is argued to be the case for [nasal]. Finally other parameters, such as vowel height or sonority seem to have multiple values. Such dimensions are often treated with two or more binary features (e.g. [±high] and [±low] to capture three vowel heights,

$$[\text{i}] = \begin{bmatrix} \text{+high} \\ \text{–low} \end{bmatrix}, \quad [\text{e}] = \begin{bmatrix} \text{–high} \\ \text{–low} \end{bmatrix}, \quad [\text{æ}] = \begin{bmatrix} \text{–high} \\ \text{+low} \end{bmatrix}),$$

but some researchers argue that multivalued features should be incorporated directly into the system. While some have argued that place of articulation might also be multivalued, there is good evidence that the specific categories are grouped together into broader categories, e.g. those sounds involving contact with the front part of the tongue, the dentals, alveolars, and alveo-palatals sometimes pattern as a group and are referred to by the cover term *coronal*. I will not provide a systematic discussion of distinctive features, since a number of good overviews are available (see, for example, Keating 1987, Clements and Hume 1995) and I will refer somewhat informally to specific features here. Leaving aside finer differences between specific proposals, a striking result about the nature of most feature systems is that the features themselves are not arbitrary classificatory elements, but rather are closely linked to phonetic structure. Thus we find a convergence of phonetic events and the sounds that are found to pattern together in the phonologies of language after language.

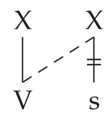
Evidence for specific feature proposals comes from their adequacy in capturing the recurrent cross-linguistic grouping of sounds, referred to as *natural classes*. The same groupings of sounds are found in a wide range of phonological patterns. Take for example the feature [±sonorant]. [+sonorant] defines the class of sounds for which spontaneous vocal cord vibration (or voicing) is possible. This includes those sounds for which there is not a close obstruction of the vocal tract (nasals, liquids, glides, vowels). In the typical case the sonorants are voiced and do not show a contrast between voiced and voiceless. For the obstruents – the stops, fricatives, and affricates – on the other hand, which are [–sonorant], voicing involves certain articulatory adjustments to maintain air pressure and keep the vocal cords vibrating. For the obstruents, the least marked category is voiceless, but the obstruents often show a contrast between [+voice] and [–voice]. A strong implicational universal is that if there is a voicing contrast in the sonorants (as found, for example, in Burmese where there are both voiced and voiceless nasals and other sonorants), then there is also a voicing contrast in the obstruents. Additional examples of reference to the natural class defined by [±sonorant] include syllabic consonants in English (the nasals and liquids in the final syllable of such forms as *bottle* [l̩] and *button* [n̩]) and the division between the sonorants and obstruents crucial to the sonority hierarchy discussed earlier.

Sometimes the patterning of sounds is characterized in terms of the specific featural content of segments, but other times the presence or absence of segments themselves accounts for the observed pattern. Thus sometimes it is appropriate to refer to the segment as a unit independent of its featural content. To incorporate the notion of the segment as such, some approaches include so-called "timing units," and others propose an internal hierarchical grouping of features within the segment, including a "root node," which, in effect, identifies a bundle of features as a segment. Such approaches allow us to account for the changes in timing which are independent of segment content.

Sometimes a segment might be deleted without leaving any evidence behind (such as the Korean consonant deletion case illustrated above in (9)), but in other cases, the timing of a deleted segment "stays behind." This is the case of what is called *compensatory lengthening*. Consider the widely discussed case from Latin illustrated in (28).

(28)	/kosmis/	[koːmis]	"courteous"
	/kasnus/	[kaːnus]	"gray"
	/fideslia/	[fideːlia]	"pot"

We see in (28) that an /s/ is deleted before another consonant. (The relevant consonants are labial and coronal nasals and /l/. Not all /s/'s disappear, as we can see by the fact that final /s/'s still surface.) But the /s/ doesn't completely disappear; rather, it leaves its timing unit (indicated here by an X) behind, resulting in a lengthening of the preceding vowel, hence the term compensatory lengthening. We can capture this change as follows (where I am informally representing the bundle of features which make up the content of relevant segments as **V** and **s**).

(29)


The feature bundle of /s/ is deleted but its timing unit is reassociated with the preceding vowel. Direct reference to the timing aspect of a segment allows us to capture this straightforwardly.

The facts of /s/ deletion in Latin are actually more complex, as there are cases where /s/ deletes, again before a nasal or /l/, but no compensatory lengthening occurs:

(30)	/smereoː/	[merioː]	"deserve" (present)
	/snurus/	[nurus]	"daughter-in-law"

Once again syllable structure plays a role: the /s/ in these cases is in the onset of the syllable, while in the cases in (28) above it is in the coda. A strong cross-linguistic observation is that consonants deleted from coda position may result in compensatory lengthening, while those in onset position almost never do. (There are alternative proposals besides "timing units" which capture this asymmetry.)

4.2 *Alternations*

With these further refinements of the representation of phonological units – features organized into segments and timing units, in turn grouped into larger units – we are ready to consider one of the central observations in phonology. Often phonemes are realized in different ways in different contexts – position in the word, next to certain sounds, in stressed or unstressed position, and so forth. Such differences in the realization of a phoneme, and as a result alternations in the shape of a morpheme, are the clearest evidence of the effects of phonology. As already seen above, alternations can result from aspects of the higher level organization (as we saw, for example, in the consonant ~ zero alternations in Korean due to syllable structure). But effects are also found due to the quality of neighboring segments. To take a simple example from English, the prefix /ɪn-/ changes its shape depending on the following consonant:

(31) [ɪn]
inappropriate
intolerant
indecent

[ɪm]
impossible
imbalance

[ɪŋ]
incoherent
inglorious

Here the nasal is becoming more similar to the following consonant by sharing the place of articulation, with a coronal nasal [n] before coronals (and also vowels), a bilabial nasal [m] before bilabial stops, and a velar nasal [ŋ] before velars. The morpheme /-ɪn/ has three allomorphs: [ɪn-, ɪm-, ɪŋ-]. This is an example of *assimilation*, whereby a sound becomes more similar to its neighbor(s). While such patterns of nasal place assimilation are very common cross-linguistically, this pattern is not as systematic in English as in some other languages, since a nasal consonant doesn't always share the place of articulation of the following consonant. For example, in forms compounded with the particle /ɪn-/, for some speakers, assimilation doesn't take place: cf. *input*, [n-p] *income* [n-k]. (There are systematic explanations of these differences, but considering these would take us beyond the scope of the present discussion.)

It is also assimilation, in this case, of voicing, which accounts (in part) for the alternation in the shape of the regular plural marker in English that we saw above in (1). As we observed above, what is spelled as *s* or *es* is pronounced as [s], [z], or [ɪz]. The distribution of these three variant shapes or allomorphs of the plural morpheme is not arbitrary. Rather, the distribution is systematically determined by the voicing and place of articulation of the final sound of the stem:

(32)	a.	[s]		b.	[z]		c.	[ɪz]	
		cap	[p]		cab	[b]		match	[ʧ]
		cat	[t]		fad	[d]		judge	[ʤ]
		book	[k]		dog	[g]		mess	[s]
					can	[n]		buzz	[z]
					file	[l]		wish	[ʃ]
					bow	[o]		garage	[ʒ]

If the final sound of the stem is voiceless, as shown in (32a), then the shape of the plural marker is [s]. (This holds systematically for the stops, but the situation with voiceless fricatives is more complicated: sometimes the voiceless fricative itself becomes voiced and then takes the voiced allomorph [z], such as *leaf* [f], *leaves* [vz], but sometimes the same pattern for the stops is found, *chef* [f] *chefs* [fs].) As shown in (32b), if the final sound of the stem, whether an obstruent, sonorant consonant, or vowel, is voiced, then the shape of the plural marker will be [z]. Thus the voicing of the final sounds in the stem conditions the shape of the plural marker, which agrees in voicing with that sound, another example of assimilation. But there is a systematic exception to the pattern seen in (32a and b), as illustrated in (32c). If the final sound is either an affricate [ʧ, ʤ], or an alveolar or palato-alveolar fricative [s, z, ʃ, ʒ], then the shape of the plural marker is [ɪz]. The intuition here is that [s] or [z] added to stems ending in these sounds would be too similar to be perceptually distinct and so a vowel is inserted to break up the cluster. While some limited exceptions exist, such as *mouse-mice*, *sheep-sheep*, *child-children*, there is good evidence for the fact that speakers intuitively know the rule that is responsible for the correct phonetic shape of the plural marker. Such evidence comes from the fact that both children acquiring English and adults when faced with new words added to the language apply these rules in forming the plural, for example *macs* [s] and *pentiums* [z] and some people even say *mouses* [ɪz].

We can see by comparing these two examples from English that assimilation can result from a preceding segment being affected by a following one as in the case of nasal place assimilation or vice versa as in the case of voicing assimilation of the plural marker.

Such patterns of assimilation are very common across the languages of the world. Again this is an area where we see a close parallel between phonology and phonetics. It is a common property of speech that neighboring sounds are coarticulated, that is, that the articulation of adjacent sounds overlaps. Such phonetic effects can become exaggerated and over time result in phonological assimilation. Let's consider another example, the case of vowel nasalization in Sundanese (a regional language of Indonesia).

(33) Sundanese vowel nasalization

a.	[atur]	"arrange"	[ŋãtur]	"arrange" (active)
	[obah]	"change"	[ŋõbah]	"change" (active)
	[parios]	"examine"	[mãrios]	"examine" (active)
b.	[tiis]	"relax in a cool place"	[nĩĩs]	"relax in a cool place" (active)
	[saur]	"say"	[ɲãũr]	"say" (active)

In Sundanese, an initial vowel or one following an oral consonant is oral, while one following a nasal consonant is nasalized. This alternation between nasalized and oral vowels can be seen in corresponding bare stems and active forms, since the active is formed by adding [ŋ] or [+nasal] to the initial consonant of the root, as shown in (33a). Not only is a single vowel following a nasal consonant affected, but a sequence of vowels will become nasalized, as shown in (33b). Such examples illustrate the importance of distinctive features for an adequate description of such alternations. If we couldn't make reference to a single feature (e.g. [voice] or [nasal]) or set of features (needed, for example, to account for nasal place assimilation), we would be missing a fundamental insight into what is going on in such cases. Within the generative framework, following the seminal work of Chomsky and Halle (1968), *The Sound Pattern of English* (SPE), such patterns are accounted for by rules of the following form:

(34) a. general rule schema: a → b / c__d
"a becomes b in the environment following c and preceding d"

b. Sundanese Vowel Nasalization: V → [+nasal] / [+nasal] ____
Condition: applies iteratively
"A vowel becomes [+nasal] when it is in the environment following a sound which is [+nasal]"

c.

Underlying representation	/tiis/	/[+nasal] + tiis/
Vowel Nasalization	–	niis
iterative		nĩĩs
Surface representation	[tiis]	[nĩĩs]

The general rule schema offers a formalism for accounting for observed phonological alternations. Rather than just describing the distribution of the differing allophones (or allomorphs as the case may be), this rule formalism incorporates the fundamental idea that one of the variants is basic, or underlying, and that the other variant(s) are derived by rule. Such rules are an attempt to capture the knowledge that a speaker has about the sound patterns in his or her language. Following this approach, the pattern of nasalization in Sundanese can be represented as shown in (34b), with an example of the application of the rule or "derivation" in (34c).

Such formalism, central to the view that phonology is about capturing the speaker's knowledge about language, indeed offers an explicit account of phonological patterns. However there are also some serious limitations for which alternative proposals have been developed. First, this approach does not formally account for the fact that some kinds of assimilation are so common. For example, there is nothing in the notation itself that accounts for the fact that the [nasal] specification changes by its proximity to [+nasal] as opposed to some specification for a different feature. More recent work has suggested that a more accurate account follows from the idea of assimilation as "feature spreading," rather than the changing of feature values (see Goldsmith 1976). This is part of a more general approach termed *autosegmental phonology*, where specific features can function independently of segments. Following this approach, we could characterize vowel nasalization in Sundanese as follows:

(35) a. X V

[+nasal]

b. n i i s

[+nasal]

The autosegmental rule in (35a) indicates that the [+nasal] feature specification spreads to the right to a following vowel, resulting in structures such as that illustrated in (35b). Here the pattern of assimilation is captured directly through the sharing of a single feature specification. This has the added advantage of allowing us a straightforward account of the iterative nature of this process.

We also saw an example of spreading in our characterization of compensatory lengthening above in (29), where the whole feature matrix specifying the vowel is shared between the vowel's timing unit and the following timing unit, freed up by the loss of the feature matrix of the /s/. Viewed in this way, this too can be seen as a sort of assimilation, in this case total assimilation. Within the formalism of SPE, such patterns of compensatory lengthening were represented as transformational rule as illustrated in (36).

(36) compensatory lengthening

$$\begin{array}{cccc} \mathrm{V} & \mathrm{s} & \begin{bmatrix} \text{+sonorant} \\ \text{−anterior} \end{bmatrix} & \\ 1 & 2 & 3 & \rightarrow \\ 1 & 1 & 3 & \end{array}$$

"The string 1, 2, 3, where 2 = /s/ and 3 is $\begin{bmatrix} \text{+sonorant} \\ \text{−anterior} \end{bmatrix}$ is rewritten as 1, 1, 3 where 1 is the preceding vowel"

Use of transformational rules has generally been rejected now in both phonology and syntax due to their excessive power, since there are no predictions about what are allowable structures formally. There is also no insight resulting from such formalism as to why particularly these sorts of patterns occur in language after language.

In addition to assimilation of a single feature (e.g. vowel nasalization) and total assimilation (e.g. compensatory lengthening), there are cases where two or more features systematically pattern together, such as the case of nasal place assimilation, as exemplified above in English for the prefix /-ɪn/. In SPE notation, where place of articulation is represented with the two features [coronal] and [anterior], this would be represented as shown in (37).

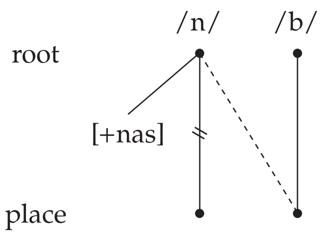
(37) $$\begin{bmatrix} \text{−continuant} \\ \text{+nasal} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha\ \text{anterior} \\ \beta\ \text{coronal} \end{bmatrix} / \underline{\qquad} \begin{bmatrix} \text{−continuant} \\ \text{−sonorant} \\ \alpha\ \text{anterior} \\ \beta\ \text{coronal} \end{bmatrix}$$

"A nasal consonant takes on the place specification (same values for [anterior] and [coronal]) as a following stop"

Here "alpha notation" is used to show that the resulting feature values are dependent on those elsewhere in the rule, in this case the values for both [anterior] and [coronal]. We see similar formal problems as in the case of single feature spreading and in addition, there is no explanation why certain features are seen to group together in language after language. In cases of nasal place assimilation, it is precisely the set of features that define place of articulation that pattern together.

Cases where a particular set of features pattern together in assimilation and other phonological processes provide strong evidence for the grouping of features (see McCarthy 1988 and Clements and Hume 1995 and work cited therein). This general approach, termed *feature geometry*, not only captures the notion of the segment as a unit independent from its featural content (represented by a *root node*), but it also offers an explicit proposal of hierarchical structure or subgrouping of features, making direct reference to elements such as the *place node*. An account of nasal place assimilation following this approach is schematized in (38).

(38)



Most recently such patterns of feature "spreading" have also been characterized in Optimality Theory in terms of competing constraints.

Segments can influence each other in a wide variety of ways. There is a rich array of patterns of assimilation, including cases where the segments affecting each other are not adjacent, such as *vowel harmony* where vowels agree in a certain property (e.g. height or rounding) irrespective of the quality of the intervening consonants. We also find that segments can become less like each other; this is termed *dissimilation*. The contrast between segments might be lost in a particular environment. This is known as *neutralization*. Feature changes may be brought about due to the segmental context (that is, influence for neighboring segments), but it is also the case that the influence of prosodic structure can drive such effects. It is quite common that the range of contrasts which occurs in syllable onsets is reduced in syllable codas. One very common pattern of neutralization is what is known as Final Devoicing. Consider the following example from Polish:

(39) Polish Final Devoicing

a.	klup	"club" sg.	klubi	"club" pl.
	trut	"labor" sg.	trudi	"labor" pl.
b.	trup	"corpse" sg.	trupi	"corpse" pl.
	kot	"cat" sg.	koti	"cat" pl.

We see the alternation in the voicing of the final consonant of the stem. Just looking at the forms in (39a), we might think that either the voiceless stops are underlying and become voiced in a particular environment (between vowels) or that the voiced stops are underlying and become voiceless in a particular environment (at the end of the word, though additional data suggest that it is actually syllable-final position more generally). But looking at the data in (39b), we see that not all cases show the same alternation; in these cases a voiceless stop surfaces in both forms. This makes it clear that it must be the voiced stops becoming voiceless. We also note that this pattern seems to be applying not to a random set of sounds, but to a natural class of sounds, in this case the stop consonants. We would predict that if we found forms ending in velar consonants, similar alternations would be observed. As we see in (40a), this pattern actually applies not just to stops including velar ones, but also to fricatives, that is to the class of obstruents or [–sonorant]. We can capture this pattern by positing underlying forms as shown in (40a) and applying a rule of Final Devoicing. This rule can be characterized in SPE terms as shown in (40c). Or we can account for such patterns in an autosegmental notation with the delinking of the relevant feature specification, in this case [+voice] (40d). In either case, we can see that the rule works by looking at sample derivations in (40e).

(40) Polish Final Devoicing

a.

wuk	"lye" sg.	wugi	"lye" pl.
grus	"rubble" sg.	gruzi	"rubble" pl.

b.

/klub/	"club"
/trud/	"labor"
/trup/	"corpse"
/kot/	"cat"
/wug/	"lye"
/gruz/	"rubble"
/-Ø/	singular
/-i/	plural

c. [–sonorant] → [–voice] / ____ # (# = word boundary)
"A member of the class of [–sonorant] becomes voiceless in word final position."

d.

root		[–son] #
laryngeal	[+voice]	

e.

Underlying representation	/klub + Ø/	/klub + i/
Final Devoicing	klup	–
Surface representation	[klup]	[klubi]

Polish also provides a nice example of how one phonological process can interact with another. Consider the additional data presented in (41).

(41) Polish Vowel Raising

a.	bur	"forest" sg.	bori	"forest" pl.
	sul	"salt" sg.	soli	"salt" pl.
	ʒur	"soup" sg.	ʒuri	"soup" pl.
b.	sok	"juice" sg.	soki	"juice" pl.
	nos	"nose" sg.	nosi	"nose" pl.
c.	ruk	"horn" sg.	rogi	"horn" pl.
	vus	"cart" sg.	vozi	"cart" pl.

In (41a), we see that before liquids (actually sonorants more generally), there is an alternation between [u] and [o]. We might think that /u/ becomes [o] in some environment or that /o/ becomes [u]. Since, as we see in the third pair of forms ("soup"), some [u]'s correspond to [u], it must be that /o/ → [u], what we might term Vowel Raising. Here the relevant environment seems to be before sonorants. An important question is whether this process of Vowel Raising happens more generally. Above in (39), in the form /kot/ "cat" there was no such alternation. There is also no alternation seen in the cases in (41b), but consider the cases in (41c), where the forms include a following underlying voiced obstruent. In these cases we also see the [o] ~ [u] alternation. This suggests that the environment for this rule is more general, not just [+sonorant], but rather [+voice], grouping the sonorants (which are voiced) together with the voiced obstruents, as stated in (42a) (expressed in SPE terminology for ease of exposition), with the additional underlying representations as shown in (42b).

(42) a. $\begin{bmatrix} +\text{syllabic} \\ +\text{back} \\ -\text{high} \end{bmatrix} \rightarrow [+\text{high}] \;/\; \underline{\quad\quad} \; [+\text{voice}] \; \#$

"A back non-high vowel becomes high in the environment before a voiced sound in word final position."

b.

/bor/	"forest"
/sol/	"salt"
/ʒur/	"soup"
/sok/	"juice"
/nos/	"nose"
/rog/	"horn"
/voz/	"cart"

Since part of the trigger of the Vowel Raising rule, the following voiced sound, is the target of the Final Devoicing rule, an obvious question is how these two rules interact. It could be that Final Devoicing applies first or that Vowel Raising applies first. Consider the two possible orderings shown in the derivations in (43):

(43) a.

Underlying representation	/rog + Ø/	/rog + i/
Final Devoicing	rok	–
Vowel Raising	–	–
Surface representation	*[rok]	[rogi]

b.

Underlying representation	/rog + Ø/	/rog + i/
Vowel Raising	rug	–
Final Devoicing	ruk	–
Surface representation	[ruk]	[rogi]

It is clear comparing the two derivations that the Vowel Raising rule must apply before Final Devoicing, otherwise Final Devoicing would in effect rob relevant cases from Vowel Raising. Such cases show that the ordering of rules may be crucial. We have characterized these patterns of alternation following a rule-based approach. We could equally well pursue a constraint-based approach, but in either case, we need to be able to account for the ways in which phonological processes might interact with each other.

Before concluding this section, let's return to the question raised above about the status of [ŋ] in English. While we included [ŋ] in the chart of the sound inventory in English presented in (3) above, we also noted in (1) that [ŋ] has a defective distribution. One approach to this would be to say that /ŋ/ just has a defective distribution, period – parallel to /h/. Yet this would leave a number of distributional observations unaccounted for. Consider the distributions of the three nasals of English, [m, n, ŋ] in (44):

(44)

	initial	medial	final	N−Vstop	N+Vstop	N−Vstop	N+Vstop
m	map	dimmer	dim, bomb	camper	amber	camp	–
n	nap	sinner	sin	canter	candor	can't	land
ŋ	–	singer[3] [ŋ]	sing [ŋ]	canker [ŋk]	anger, finger [ŋg]	bank [nk]	–

[m, n] can occur in word-initial position, as well as medially and finally. They can also occur before an oral stop, either medially or finally (except that [mb] doesn't occur as a cluster within a syllable coda, hence *bomb* [bam], but *bombardment* [bəmbardmn̩t]). [ŋ], on the other hand, doesn't occur in word-initial position. Basically [ŋ] only occurs in the syllable coda, not in the onset. This generalization accounts for why it can't occur word initially and accounts for all the cases except *singer*. The important observation here is that *singer* consists of the root *sing* plus the suffix *-er* and so the [ŋ] is, in effect, in syllable-final position until the suffix is added (assumed to cause resyllabification). This generalization accounts for the distribution, but doesn't explain why it

should be so. As noted above, sometimes sounds are limited in their distribution, but cross-linguistically we find if a consonant is limited, the more restricted distribution occurs in the coda, not in the onset. In other words, neutralization (such as the case of Final Devoicing) tends to occur in codas, not onsets. If we take the spelling as a cue in the cases of [ŋ], a solution presents itself. We might argue that [ŋ] is not part of the underlying inventory of English, but rather that it is derived from /ng/ or /nk/ sequences. Very briefly the analysis would work as follows. The underlying nasal consonants in English are /m, n/. As noted above in our discussion of the prefix /-ɪn/, English has a rule of Nasal Place Assimilation whereby a nasal assimilates to a following stop (schematized above in (38)). Based on the evidence from the lack of word-final [mb] clusters we might also posit a rule of Voiced Stop Deletion which applies to non-coronals, whereby a voiced stop following a nasal consonant is deleted word finally (45a). Given the underlying representations presented in (45b), the rules of Nasal Place Assimilation and Voiced Stop Deletion together (as well as some understanding of the interaction of phonology and morphology for cases like *singer* which we won't develop here) account for the observed patterns, as shown in the derivations in (45c). As in the Polish case, these rules must be crucially ordered, otherwise the deletion of the voiced stop would have removed the information about place specification needed for the Nasal Place Assimilation rule.

(45) a. Voiced Stop Deletion

$$\begin{bmatrix} -\text{sonorant} \\ -\text{continuant} \\ +\text{voice} \\ -\text{coronal} \end{bmatrix} \rightarrow \varnothing \, / \begin{bmatrix} +\text{consonantal} \\ +\text{nasal} \end{bmatrix} \_\_\_\_ \#$$

"A voiced non-coronal stop is deleted word finally following a nasal consonant."

b. /dɪm/
/bamb/ or /banb/
/bænk/
/sɪng/
/fɪngɚ/

c.

Underlying representation	/banb/	/bænk/	/sɪng/	/fɪngɚ/
Nasal Place Assimilation	bamb	bæŋk	sɪŋg	fɪŋgɚ
Voiced Stop Deletion	bam	–	sɪŋ	–
Surface representation	[bam]	[bæŋk]	[sɪŋ]	[fɪŋgɚ]

In the case of *bomb*, we might assume an underlying /n/ or /m/ or even a nasal consonant which is unspecified for place of articulation. /n/ when it occurs before a velar consonant assimilates in place of articulation and then in the case of a following voiced stop, this is deleted. The restricted distribution of [ŋ] in English follows directly from this approach without our having to posit an underlying phoneme with a defective distribution.

In this section we have seen a number of ways in which segments might affect each other and evidence for reference to distinctive features, as well as their grouping. We have also seen that the division we made between structure above the level of the segment and subsegmental structure is somewhat artificial, since syllable structure can affect feature specification and so forth.

5 Phonology as a System

In concluding this introduction to phonology, it is useful to step back and consider how all these aspects of phonology that we have discussed fit together.

Most basically a phonology consists of a set of representations – an inventory of sounds, in turn defined by distinctive features matrices – and a system of rules or constraints which act on the representations. Fundamental to the generative approach is the idea that the idiosyncratic and predictable information of the phonology are treated separately: the idiosyncratic information is part of the underlying representations and the predictable patterns arise through the systematic manipulation of these sounds through rules or constraints. Consider the following schematic figure:

(46)	underlying representations:	input
		constraints / rules
	surface representations:	output

The underlying representation includes the abstract sounds or phonemes for each morpheme in the language and the surface representation incorporates the phonetic variations or allophones, seen in the systematic alternations of the language, introduced as a result of the applications of a system of rules or constraints. The phonological representation includes not only the sequence of sounds, made up of timing units and featural content, but also the hierarchical grouping of these sounds into syllables and higher level prosodic units.

A phonology of a language consists of the whole system taken together. A complete phonology consists of dozens and dozens of rules (or constraints) often with complex interactions. To illustrate both the nature of phonological patterns and the mechanisms involved in accounting for these patterns, we have considered a number of examples of phonological patterns, but only by studying the whole phonology of a language can we understand its full complexity.

NOTES

Thanks to Beverley Goodman, Lisa Lavoie, Ayako Tsuchida, and Draga Zec for providing helpful input on earlier drafts of this chapter.

1 The description of possible sounds used in language is part of the purview of (linguistic) phonetics and so I will not provide a full discussion here (for an introduction, see Ladefoged 1993, also chapter 7, in this handbook).

2 The situation is actually a bit more complex, since if we have a syllabic nasal, such as in *gibbon* [gɪbn̩], then the sequence is allowable, but here the [n] is functioning as a vowel and is in a different subconstituent of the syllable from the [b].

3 For some speakers, the author included, this is pronounced [sɪŋgɚ], rhyming with *finger* [fɪŋgɚ].