Chapter 3
Phonological words: Calling all Scrabble players! /fɑnəlɑdʒɪkəl wʌɹdz: kɑlɪŋ ɑl skræbəl plejəɹz/

In this chapter, we learn about the language-specific restrictions that govern what sound sequences are possible in English phonological words (*phonotactics*), the regular processes that apply to produce different sounds in different contexts (*allophony*), and the rules according to which *stress* is assigned to English words, and how stress affects pronunciation. We'll look at how these three processes intersect to identify phonological words. We also consider all these properties in relation to the problem faced by babies: breaking the speech stream down into smaller parts so they can begin learning *listemes*.

3.1 Guessing at words: The Scrabble problem[24]

We are now finally able to consider the central problem of this section: what is a phonological word? We saw in the first chapter that our everyday use of the word "word" seems to pick out a kind of a phonological unit. According to our everyday way of thinking, *dogs* is one

[24] The empirical material in this section is drawn largely from Hammond (1999).

word, *-s* is not a word, and *works like a dog* is four words. We saw, however, that one common intuition about words—that they are minimal units of meaning—is faulty.

Not only do we have firm intuitions about how many of these phonological words are present in a given string of English (intuitions that are reflected in (and possibly affected by) the spacing conventions of English orthography)—we *also* have intuitions about strings of sounds we've never heard before. For instance, consider the following nine strings of sounds, written both in English orthography and in IPA, so you can get a precise idea of what they're intended to sound like:

(19) a.	timp	/tɪmp/
b.	rog	/rɑg/
c.	mbotto	/mbɑto/
d.	flezk	/flɛzk/
e.	spink	/spɪŋk/
f.	beh	/bɛ/
g.	bod	/bɑd/
h.	psore	/psɔɹ/

It is unlikely you have seen or heard most of these letter sequences before. Now, rate each of these "words" according to how confident you are that they are not English words. Give the words you are most confident are not English a score of 1 and the words you think are most likely to be English words a score of 5. Arrange the words in order according to the scores you assign, lowest to highest. No looking in the dictionary, of course!

Exercise 1: Give each of the strings of sounds in (8) a numerical rating, from 1-5, where 1 means the string is definitely *not* an English word, and 5 means that it definitely *is* an English word.

Here is a typical ranking and average score for each of these 'words', from a sample of 40 native English speakers:

(20)

Rank	Word	Score
1	bod	4.66
2	timp	4.30
3	rog	4.2
4	spink	4.17
5	beh	2.75
6	psore	2.02
7	flezk	1.69
8	mbotto	1.07

It's pretty likely that your ranking comes fairly close to this one. Yet you have never seen most of these strings before! Shouldn't they all seem equally unlikely as potential English words?

Scrabble players are familiar with this problem. In Scrabble, you get a rack of seven letters, such as, say, IOBUZRP, each worth a certain number of points. Your job is to arrange them in such a way that you can spell an English word with them. You will try to look for the English word that uses the largest possible number of letters, since the longer the word, the more points you will score. Since the "Z" is worth a very large number of points you'll certainly try to use it. You know that f you can make a four-letter word with the Z at the beginning, you can score double points for the Z! What are the possibilities you consider? You rearrange your letters, hoping to see a word. You might look at the combinations ZOIP, ZURP, ZOUB, ZOIB, ZIRB, ZORB, ZIRP, ZURB, ZUBI, ZOPI, ZORI… but you don't recognize any of them as words for sure. (In fact, "zori" is a word in the Scrabble dictionary; it's a name for a type of sandal. If you happen to know this, when you get to "zori" you will go "aha!" and put it down, rejoicing. But if not, you'll just keep scratching your head and rearranging.)

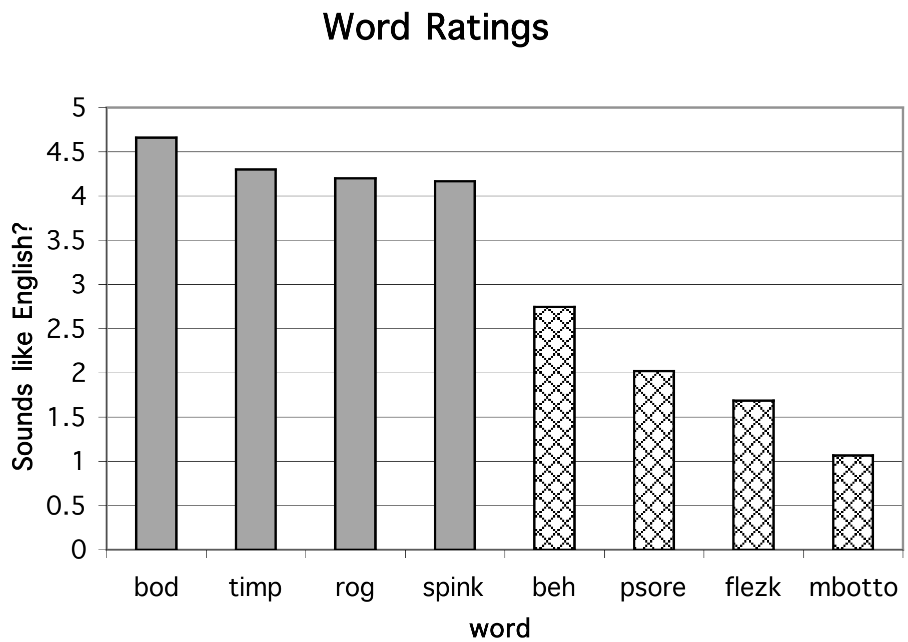
zori, *n.* Japanese thonged sandals with straw (or leather, wood, etc.) soles. From Japanese *sō* 'grass, (rice) straw' and *ri* 'footwear, sole.'

The point is, there are hundreds of arrangements of letters that you will never even consider as potential English words: ZPOI, ZROB, ZIPB, ZBRP, ZIUO, and so on. What is it that you know that makes you pause and wonder whether ZIRP might be a word of English, but makes you pass over ZIPB (unless you're Calvin)?

The answer is that you know English *phonotactics*, or the rules that describe possible sequences of sounds for forming English words. Languages can differ in their phonotactic rules, so that /mbɑto/ might be a possible word of Swahili, or /psɔɹ/ a possible word of Greek. These sequences, however, are not possible words in English.

Notice that a bar chart of the wordhood ratings given in (20) above shows a significant jump between *spink* and *beh*:

(21) *Bar chart of 'word-like' ratings*



From *bod* to *spink*, the bars decrease gradually, and similarly from *beh* to *mbotto*. But between *spink* and *beh* there's a big jump. The reason for this is that the four least-wordlike words violate rules of English phonotactics, while the four most-wordlike words do not. That is, these sound sequences fall into two groups: the last four are phonologically impossible words in English, while the first four are phonologically possible words. We're going to look at some of the constraints that determine what sound combinations can appear in possible English words.

3.2 Building Blocks III: The Syllable

One condition on a well-formed English word is that it has to be made up of at least one *syllable*. That's one reason why *-s*, as in *dogs*, can't be considered a phonological word. A syllable is, roughly, a phonological unit that contains at least a vowel. Syllables can start or end with one or more consonants, but even without any consonants, a vowel can be a syllable all by itself: the pronoun "I", /aj/, for instance, or the first syllable of the word *open*, /ow/, or *apart*, /ə/. Consequently, all English phonological word must contain at least a vowel.

> *Phonotactic Rule #1:* All phonological words must contain at least one syllable, and hence must contain at least one vowel.

How do we know that syllables are important units of speech? We can see that people pay attention to syllables in a number of ways. One very obvious one is in metered poetry. We know, for instance, that two lines of poetry that *scan*, i.e. that fall into a regular rhythmic pattern, usually have the same number of syllables. Consider this famous first verse of the nonsense poem *Jabberwocky*:

(22) 'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe.
All mimsy were the borogoves
And the mome raths outgrabe.

Count the syllables. You should find that the first three lines match, each containing eight syllables, and the last one is shorter, with six syllables. This pattern is repeated throughout the poem:

(23) One, two! One, two! And through and through
The vorpal blade went snicker-snack!
He left it dead, and with its head
He went galumphing back.

Again, eight, eight, eight and six. Considering that Lewis Carroll made up most of the words in the poem, he must have intended for the syllable counts to turn out this way—it can't just be a coincidence.

Sometimes poets will play with the intuition that syllable counting is an essential ingredient of verse. Consider the first two verses of *Poetical Economy* by Harry Graham:

(24) What hours I spent of precious time
What pints of ink I used to waste,
Attempting to secure a rhyme
To suit the public taste,
Until I found a simple plan
Which makes the lamest lyric scan!

When I've a syllable de trop
I cut it off, without apol.
This verbal sacrifice, I know,
May irritate the schol.;
But all must praise my dev'lish cunn.
Who realize that Time is Mon.

In an opposite direction, what about this *Rhyme for Remembering the Date of Easter*, by Justin Richardson?

(25) No need for confusion if we but recall
That Easter on the first Sunday after the full moon
following the date of Equinox doth fall.

This particularly unmemorable rhyme fails as a mnemonic (and succeeds as a joke) because it doesn't scan: trying to remember the rhyme is just as hard as trying to remember the plain prose fact. A *good* mnemonic rhyme scans, giving it a rhythm that helps you fit in the right individual words, as in the first two lines of this famous mnemonic for remembering the number of days in a month:

(26) Thirty days hath September
April, June, and November.

Another place where we see the notion of syllable at work is in hyphenation conventions in written English. When group of words won't exactly fit into a single line of text, one of the words has to be broken up and part of it placed on the next line, like this:[25]

> *When an orthographic word won't fit onto a single line of text, it is hyphenated at a syllable boundary.*

Since phonological words like *can't* and *caboodle* have to be made up of well-formed syllables, understanding what can be a well-formed English syllable will take us a long way towards understanding what is a possible well-formed English phonological word, and hence a long way towards understanding why, in your hypothetical Scrabble game, you wouldn't even consider *zpob* as a possibility.

Syllables can be made up of a simple vowel, even a reduced vowel, such as the intial /ə/ in "attempt", which has two syllables. They can be made up of a consonant and a vowel, such as /hij/ in "he". They can be made up of a consonant, a vowel and a consonant, like /sʌn/ in "Sunday". In fact, English syllables can have up to three consonants at the beginning (in the syllable's *onset*, as in the word /strɪŋ/, "string"), the vowel, and up to four consonants at the end (in the syllable's *coda*, as in the word /tɛksts/, "texts"). The anatomy of a syllable is diagrammed below; "C" stands for "consonant" and "V" for vowel; brackets indicate optionality:

(27) (C)(C)(C)V(C)(C)(C)(C)

onset coda

[25] Actually, the syllable boundary convention can be overridden by other considerations. If hyphenating at a syllable boundary would mean that there would only be one letter on a line, another breaking point is chosen—often an affix boundary. For instance, *unable* is hyphenated as *un-able*, rather than *u-nable* or *unab-le*.

It's the vowel plus the coda that make a rhyme. *Coast* and *toast* rhyme, as do *code* and *toad*, but *coast* and *code* don't rhyme, although they have the onset and vowel in common. (In poetic terms, *coast* and *code* are alliterative and assonant—but they don't rhyme).

Now, considering that we earlier identified 24 consonant sounds and 15 vowel sounds in English, and that a syllable can consist of anything from a single vowel to a vowel surrounded by seven consonants, in theory there should be 74,909,241,375 possible syllables in English. Obviously, this is far larger than the actual number of possible syllables. For one thing, that figure includes cases where repeated sequences of the same consonant in onset and coda are considered possible, as in /sssɪtttt/. In English, this does not occur.

> *Phonotactic Rule #2:* Sequences of repeated consonants are not possible.

Excluding sequences of identical sounds within an onset or coda we get 60,779,920,695 possible syllables, which is still orders of magnitude bigger than the actual number of possible English syllables. It includes such non-starters as /zɪpb/ and /msɑl/. What are the underlying principles that English speakers use to limit this combinatorial explosion?

Even at first glance, it's clearly wrong to assume that all 24 English consonants can occur freely in all the 7 available consonant positions in English syllables. For instance, although the velar nasal sound /ŋ/ frequently occurs at the end of syllables (*si**ng**, jumpi**ng**, ha**ng**man*), it never appears at the beginning of English syllables, and hence never at the beginnings of English words.

> *Phonotactic Rule #3:* The velar nasal /ŋ/ never occurs in the onset of a syllable.

Similarly, although /h/ is a well-formed syllable onset in English (as my own name attests), it is never found at the ends of syllables, and hence never at the ends of words. The symbol "h" is sometimes used at the end of a syllable in English spelling to indicate the pronunciation of a vowel—e.g. *ah* is the usual spelling of the syllable /ɑ/— but the /h/ sound itself is not pronounced.

> *Phonotactic Rule #4:* The glottal fricative /h/ never occurs in the coda of a syllable.

Excluding /h/ from codas and /ŋ/ from onsets, then, the number of possible English syllables is reduced to 44,881,090,380.

3.3 Phonotactic restrictions on complex syllable onsets

If we look for patterns in the formation of onset consonant *clusters* in English—when there is more than one consonant at the beginning of a syllable—they are easy to find. First, the affricates /t͡ʃ/ and /d͡ʒ/, and the glottal fricative /h/ can't occur in an onset with any other consonant: there are no English words like /**t͡ʃl**ɪŋk/, 'chlink,' /**d͡ʒl**æm/, 'jlam' or /**kh**ɑt/, 'khot,' so instead of 23 possible onset consonants, in clusters there are only 20.[26]

> *Phonotactic Rule #5:* The affricates /tʃ/ and /dʒ/, and the glottal fricative /h/, do not occur in complex onsets.

---

[26] Remember that we are talking about clusters of consonantal *sounds*, not spellings. The letter 'h' occurs as a letter in the spelling of many simple and complex syllable onsets, because of the English spelling convention of using 'h' to indicate certain fricatives and affricates. Words like *thin* and *chin* and *through* have more than one orthographic consonant in their spelling, but only the last one is really a complex onset when pronounced, and none of them involve the actual consonant sound /h/. What is the correct IPA transcription of these words?

Second, when a onset contains a sequence of two consonants, the first consonant of the sequence must be an *obstruent*—an oral stop or fricative. So while we find sequences like /**tr**ʌk/ 'truck', or /**d**rɑp/, 'drop' we never find sequences like /**r**tʌk/, 'rtuck', or /**rd**ɑp/, 'rdop'.

> *Phonotactic Rule #6:* The first consonant in a two-consonant onset must be an obstruent.

Since only 14 the 20 possible consonants are obstruents,[27] that whittles the possibilities down to 28,956,015,990.

The *second* consonant of a two-onset sequence can be anything except a voiced obstruent—that is, it can be a voiceless stop or fricative, or a nasal, liquid or glide, but not a voiced stop or voiced fricative. So while there are words like /s**n**ɪt/, 'snit', /s**w**ɪl/, 'swill', /s**t**ɪl/, 'still', /s**p**ɪl/ 'spill', and /s**f**ɪŋks/, 'sphinx', in English, there are no words like /s**g**ɪl/ 'sgill', /s**d**ɪl/ 'sdill' or /s**v**ɪŋks/ 'svinx'.[28]

> *Phonotactic Rule #7:* The second consonant in a two-consonant onset must not be a voiced obstruent.

That means that only 13 of the 20 possible syllable-initial English consonants can occur in second position in a two-obstruent consonant cluster in English, bringing our estimate of the number of possible English syllables down to 28,578,886,740.

When we consider combinations of consonants in two-consonant onsets in English, even more restrictions jump out at us. While it is true that any consonant except a voiced obstruent (13 sounds) can occur in second position, and any obstruent (14 sounds) can occur in first position, it's not true that they can pair up indiscriminately. Rather, when the first

---

[27] Note that nasals are not considered obstruents here, although they are stops.

[28] Even though the word *svelte* is spelled with an 'sv' sequence, it is pronounced with an /sf/ — /sfɛlt/.

consonant is anything other than an /s/, the second consonant has to be a liquid or a glide.[29] So while there are words like /**fl**ɪk/ 'flick' and /**dw**ɛl/, 'dwell', there are no words like /**fp**ɪk/ 'fpick' or /**dz**ɛl/, "dzell".[30]

> *Phonotactic Rule #8:* If the first consonant of a two-consonant onset is not an /s/, the second consonant must be a liquid or a glide—the second consonant must be /l/, /ɹ/, /w/, or /j/.

The only time the second consonant can be a voiceless obstruent or a nasal is when the first consonant is an /s/: we have words like /**sn**ijk/ 'sneak', /**st**ejk/ 'steak', /**sk**ejt/ 'skate', /**sp**ijk/, 'speak' and /s**f**ijr/, 'sphere'. (The only voiceless obstruent that cannot occur after /s/ is /ʃ/, probably because the /sʃ/ sequence is nearly impossible to make distinct.)

Taking this two-way restriction into consideration, our syllable possibilities are down to 28,117,067,940.

To summarize our observations so far about onsets in English:

(28) a. /ŋ/ is not a possible onset.
b. Complex onsets may not contain affricates or /h/
c. Two-consonant complex onsets may contain either
(i) First consonant: /s/
Second consonant: nasal, liquid, glide or voiceless obstruent (except /ʃ/).
(ii) First consonant: any obstruent other than /s/
Second consonant: liquid or glide

---

[29] In a few words borrowed from German or Yiddish, there can be an /ʃ/ instead of an /s/ before non-liquids, as in *shtick*.

[30] Again in a few borrowed words, we see the /ts/ sequence: tsunami, Tsetse. It is arguable, however, that this is not a sequence of two phonemes (a complex onset), but a new affricate

What about three-consonant onsets? Let's consider some examples of words that have three-consonant onsets:

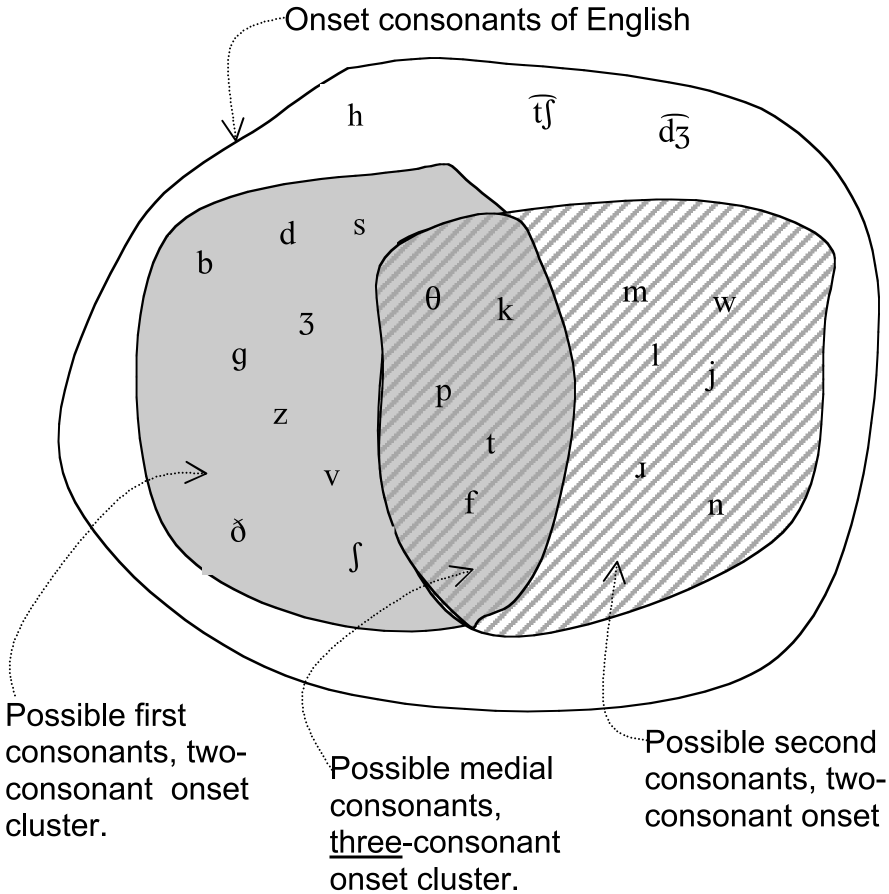
(29)	a.	/**spl**ijn/	'spleen'
	b.	/**spr**ɪŋ/	spring
	c.	/**str**ijm/	stream
	d.	/**skr**ijn/	screen
	e.	/**skl**r̩oʊsɪs/	sclerosis
	f.	/**skw**ijz/	squeeze
	g.	/**spj**uʊ/	spew

One thing that immediately leaps to the eye is that they all begin with /s/. It's also easy to see that they all end in a liquid or glide—that is, the third consonant is always either /l/, /ɹ/, /j/, or /w/. Considered in the light of our previous generalizations about two-consonant clusters, this doesn't seem like a coincidence. What's going on in three-consonant clusters is that each pair of consonants within the cluster must independently satisfy the restrictions on two-consonant clusters. We can call this the *Substring Rule*:

Phonotactic Rule #9: The Substring Rule: Every subsequence contained within a sequence of consonants must obey all the relevant phonotactic rules.

Let's call the three consonants of a three-consonant onset C_1-C_2-C_3. (In the word *string*, /stɹɪŋ/, C_1=/s/, C_2=/t/, and C_3=/ɹ/.) Such an onset has two subsequences: C_1-C_2 (/st/), and C_2-C_3 (/tɹ/). The substring rule says that the C_1-C_2 sequence must obey the two-consonant onset rules, and so must the C_2-C_3 sequence. That places some serious restrictions on what C_2 can be—it has to work as the *first* consonant of a two-consonant sequence (considered as part of the C_2-C_3 string), and it *also* has to work as the *second* consonant of a two-consonant sequence (considered as part of the C_1-C_2 string). So, for instance, in our example, the onset of *string*, /stɹ/, contains the two substrings /st/ and /tɹ/, each of which are themselves independently well-formed, as in *stick* /**st**ɪk/ and *try* /**tɹ**aj/. Recall that the first consonant of a two-consonant sequence can be any obstruent, but not a nasal, liquid or glide. The second consonant of a two-consonant sequence can be anything except a voiced obstruent. The consequence is diagrammed in Fig. 4 below. The only sounds that can possibly work both as first-consonants and second-consonants are the voiceless obstruents: /p/, /t/, /k/, /f/, and /θ/.

Fig. 4: The Substring Rule



Since the only potential legitimate C2 in a three-consonant cluster is a voiceless obstruent, it follows from the substring rule that the only possible C1 in a three-consonant cluster is /s/, since the only well-formed two-consonant onsets with a voiceless obstruent in the second position are ones which begin with /s/. It also follows that the third consonant must be a glide or a liquid, since the only well-formed two-consonant onsets with a voiceless obstruent in the first position have a glide or liquid in second position. If we include the effects of the Substring Rule on three-consonant onsets in our calculations, our inventory of possible English syllables is reduced still further, to 870,327,990.

This number still includes some unlikely onset combinations. Although /sθ/ is a very rare onset in English, it does occur in words like *sthenic* 'producing nervous energy, stimulating'; however, there are no examples of three-consonant onsets that begin with this sequence (a word like *sthrigal* or *sthlinky*). Similarly, we don't run into /spw/, /stw/, or /sfl/ onsets, despite the fact that /sp/, /pw/, /st/, /tw/, /sf/ and /fl/ do all occur independently as onsets of at least one or two English words. It's sometimes hard to tell if such gaps are the result of principles of English phonotactics or accidents of history. Try making up some words that begin with such sequences. How do they sound to you, compared to made-up words with other outlawed sequences? (E.g. compare /spwɛt/ or /sflɪŋ/ to /zbijd/ or /mræt/. There even is a technical word with an /sfr/ onset: *sphragistics*, the study of official seals or signet rings.)

3.4 Phonotactic restrictions on complex syllable codas

The codas of English syllables are, as you might suspect, subject to similar kinds of restrictions. We've already observed above that you never see /h/ in coda position. Similarly, /w/ and /j/ are not possible English codas. The only places where one might think they occur are in words like *cow,* /ka͡w/, and *lie,* /la͡j/. In such cases, though, they are part of the off-glide of a diphthong. The off-glide counts as part of the vowel in the nucleus of the syllable, not a truly separate and contrastive consonant in the coda.

> *Phonotactic Rule #10:* No glides in syllable codas.

Taking this into consideration, our inventory is reduced to 490,875,390.

In two-consonant codas, the first consonant can be pretty much anything (except /h/, which we saw above doesn't occur in codas at all). This is largely due to the fact that the two most common English suffixes—the past tense, written *-ed*, but often pronounced as just a single consonant /t/ or /d/, and the plural, written *-s*, often pronounced as /z/ or /s/, can be affixed to the end of almost any English noun or verb, creating a complex two-consonant coda.

In second position in a two-consonant sequence, there are a few more restrictions: /ŋ/ doesn't show up as the second consonant in a two-consonant coda, nor does /ɹ/, /ð/ or /ʒ/.

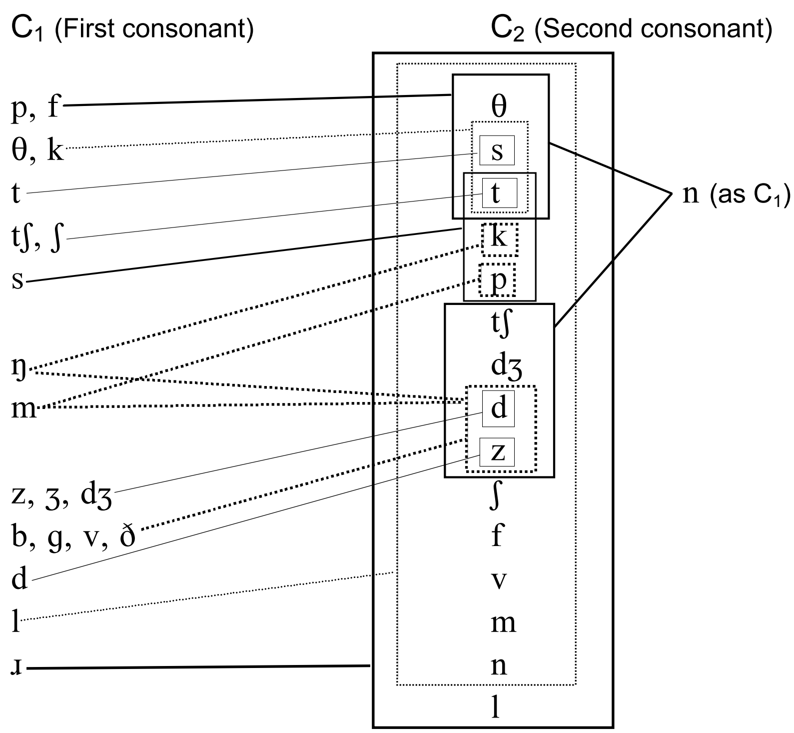
> *Phonotactic Rule #11:* The second consonant in a two-consonant coda cannot be /ŋ/, /ð/, /ɹ/, or /ʒ/.

These four sounds are consequently disallowed in third and fourth positions as well (by the Substring Rule applied to codas). These restrictions get us down to 330,467,370.

For the remaining 21x17 possibilities for two-consonant codas, we can't state absolute descriptions of availability in either first or second position, since just about any consonant can occur in either first or second position. Nonetheless, there are significant co-occurrence restrictions between the two. These are summarized in the diagram in Fig. 5. The consonants on the left are 20 of the the 21 possible first consonants (the *n* on the far right is the 21st). The consonants on the right are the 17 possible second consonants. The lines and boxes connecting the two groups indicate which first consonants may co-occur with which second consonants. For instance, /m/ on the right has a line connecting it to a (matching) square containing /p/, and another line connecting it to a (matching) square containing /d/ and /z/. This means that English contains two-consonant codas like /mp/, as in /læmp/, 'lamp', /md/, as in

/spæmd/, 'spammed', and /mz/ as in /dæmz/, 'dams'. Similarly, the right-hand consonants /z/, /ʒ/ and /dʒ/ have a line connecting them to a square on the left with only a /d/ in it, which means that there are two-consonant codas such as /zd/, as in /bʌzd/, 'buzzed', /ʒd/, as in /ɹuwʒd/, 'rouged', and /d͡ʒd/ as in /d͡ʒʌd͡ʒd/, 'judged'. Even though there's a lone /n/ is on the right, it's intended to be a first consonant. It's connected to two boxes, containing /θ/, /t/, /s/ and /d/, /z/, /tʃ/, /dʒ/, showing that it occurs as the first consonant in two-consonant codas like those in *tenth, tent, tense, tend, tens, stench,* and *lunge*.

Fig. 5: Two-consonant coda clusters of English



Some generalizations are immediately apparent. The only sound which can precede any other consonant in a two-consonant coda is /ɹ/, in words like *bird, heart, courage, purse, marsh, curl* etc. The other liquid /l/ is almost as flexible; it can precede everything except /ɹ/. (Notice that the bold /l/ and /ɹ/ in the bottom right hand corner of Fig. 5 have the biggest loops of any consonant.) Finally, the nasal /n/ can precede any consonant of English that is pronounced using the tip of the tongue (any *coronal* consonant), except /l/, /ɹ/, and /ʃ/.

coronal, *adj*. From Latin *corōnāl-is*, itself from *corōna*, 'crown', via French. Pronounced with the tip of the tongue.

Otherwise, clusters are very restricted. The only sounds that can occur after every consonant are the alveolar stops; any consonant can be followed by an alveolar stop with the same voicing except the alveolar stops themselves. Most consonants can be followed by an alveolar fricative with matching voicing as well, except for other alveolar and palatalalveolar fricatives. The nasals /m/ and /ŋ/, can also be followed by voiceless stops pronounced at the same place of articulation (*homorganic* stops), and /s/ can be followed by any voiceless stop. I won't list all the phonotactic rules summarized by Fig. 5, but here are a few:

Phonotactic Rule #12: If the second consonant in a complex coda is voiced, the first consonant in the coda must also be voiced.

Phonotactic Rule #13: When a non-alveolar nasal is in a coda together with a non-alveolar obstruent, they must have the same place of articulation, and the obstruent must be a voiceless stop.

Phonotactic Rule #14: Two obstruents in a coda together must have the same voicing.

Taking all these patterns into consideration, we find that there are 74 possible two-consonant codas in English, which reduces our number of possible syllables to 329,508,000.

homorganic, *adj*. From Greek *homos* 'same' + *organikos* 'of or pertaining to an organ'. Pronounced at the same place of articulation, i.e. pronounced using the same organ (teeth, alveolar ridge, velum, etc.)

Exercise 9: Try to think of a phonological word ending in each of the possible two-consonant codas represented in the diagram in Fig.5. Don't forget that consonantal suffixes like *-ed* and *-s* count!

The Substring Rule applies to codas as well as to onsets, as we mentioned earlier in connection with /ŋ/.To see how it affects the possible third consonant in a three-consonant string, examine Fig. 5 again. In order for a three-consonant cluster to be legitimate, the second consonant in the cluster must be both a legitimate second consonant and a legitimate first consonant. So, for instance, if the first consonant in the cluster is /ɹ/, and the second consonant is /p/, as in 'burp', the only possible third consonant is one that is a legitimate successor to /p/ in a two-consonant coda. The only possible successors to /p/ in a two-consonant coda are /θ/ (as in *depth*), /t/ (as in *clapped*), or /s/ (as in *lapse*). Consequently, they are the only possible extensions of a three-consonant coda cluster that begins with /ɹp/ (as in *burped*, /bʌɹpt/ or *burps* /bʌɹps/).

When an obstruent is in *first* position of a two-consonant cluster, it is generally the case that only a coronal obstruent can follow it. Since nearly all the consonants that work in second position are obstruents, we expect third consonants to nearly always be a coronal obstruent. There are only seven cases of three-consonant coda where the Substring Rule preditcs that a third consonant could be non-coronal.

> Exercise 10 (difficult): Looking at figure 5, identify the seven cases in which the Substring Rule predicts that a non-coronal third consonant should be possible in a three-consonant cluster.

Here's an example: /ɹm/ is a possible two-consonant cluster (as in *warm*), and /mp/ is too (as in *bump*). The Substring Rule, then, predicts that /ɹmp/ is a possible three-consonant coda. Since /p/ isn't coronal, /ɹmp/ is one of the seven cases we're looking for. A made-up word that ends in this cluster might be *termp*.

In fact, despite the fact that the Substring Rule predicts they could exist, none of the seven possibilities ever seem to occur in English codas. All third consonants in English codas are one of five coronal obstruents: /t/, /d/, /s/, /z/, or /θ/. Taking this into account, our possible-syllable count comes down to 22,901,710. Since the only way to get from three to four consonants in a coda is to add the suffix *-ed* (which can be pronounced as /t/ or /d/) or the suffix *-s* (which can be pronounced as /s/ or /z/) to a word, the only possible fourth-consonants are /t/, /d/, /s/, or /z/, which brings the total number of possible syllables down to 6,596,940.

There are many other lesser restrictions on phonotactic patterns in English which we will not consider at length or at all. For instance, no English syllables begin with a /dl/ or /tl/ onset, although both /dr/ and /tr/ are extremely common. As mentioned above, some vowel distinctions are neutralized before certain consonants, like /ɹ/, which reduces the possible inventory of syllables still further. Other restrictions apply only to particular kinds of words. For instance, we'll learn later about stressed, free morphemes in English. Such morphemes, if they are a single syllable, must either contain a coda, or they must contain a tense vowel (a grammar-school 'long' vowel). If they have no coda, their lonely vowel can't be a lax vowel. Consequently, for example, 'do' is pronounced /duw/, not /dʊ/ or /dɑ/; 'be' is /bij/, not /bɪ/, and 'so' is /sow/, not /sɑ/.

In any case, even without taking such subsidiary restrictions into account, our estimate of the possible syllable inventory of English is down to approximately 6,600,000. That's still quite a lot, but keep in mind that it's a *lot* smaller than our first raw calculation based just on the sound inventory of English. In fact, it's 4 orders of magnitude smaller than our first calculation. The additional constraints which we haven't considered here, of course, mean that the true figure in fact is probably less than 6,000,000. What we have learned here about English syllable structure, however, should give you a sense of the quantity and complexity of the tacit knowledge that you bring to bear every time you contemplate a Scrabble rack.[31]

3.5 From a stream of sound into words: Speech perception

As useful as phonotactic knowledge is in Scrabble, it is far more crucial to the problems faced by people listening to everyday speech. Although there are spaces between the words in the IPA transcriptions of phrases provided at the beginning of every chapter, in the actual speech stream that is heard by speakers of English, and children learning English, there are no breaks, except at the beginnings and ends of phrases. How do we identify the units in the utterance? How do we get from [wʊdʒəlajkfɹajzwɨðæʔ] to understanding the message "Would you like fries with that?"

Phonotactic rules are part of the answer. If you know that the consonant sequence [kf] is not a possible syllable onset or coda, then you

---

[31] Or play "Ghost", a good car game in which players take turns adding a letter to a string of letters. The letter you add must create the beginning of a correctly spelled word, but you don't want to be forced to add the letter which will complete the word. If you complete the word, you are assigned the first letter from the word "ghost". The winner begins a new word. Each player drops out after losing five rounds and becoming a "g-h-o-s-t"; the winner is the last player who is not a ghost. The advanced version of the game where you really see phonotactics in action is SuperGhost, where players may add letters to either end of a word under construction.

know that there must be a syllable break in that spot. From that, you also know it might be a word boundary. Similarly for the sequence [zw]: you know that no English syllable coda can contain that sequence of sounds, so there cannot be a syllable boundary after the [w]. There must be a syllable boundary either before the [w] or before the [z], which again lets you know that there could be a word boundary in one of those places. This is the beginning of a correct segmentation of this stream of sounds.

The process of analyzing the speech stream is called *parsing*, and the first step in parsing is identifying the phonological words in a stream of sound. This is the first problem faced by an infant or toddler trying to learn his parents' language. How can he, not knowing any language at all, take an unbroken string of sound and detect significant sub-units which might have meaning attached to them? For example, in the phrase /ajsijðədɑgij/, how does the child decide that /sij/ is an individual word with a particular meaning that it contributes to the meaning of the whole? Why not /sijð/? or /ajs/?

Part of the answer lies in the fact that babies are statistical supercomputers. As they hear speech directed to them and around them, their brains are detecting recurring patterns of repetition and tabulating the likelihood that particular sequences of sounds occur together as a unit. Studies have shown that babies can detect repeated sequences that occur in an unbroken string of CV syllables after only two minutes of exposure. To get a feel for what they're accomplishing, try reading the following aloud to yourself in an inflectionless, regular monotone:

> bidamodapamopanotabinopatapabinomodadapamobinopapanotadapa
> motapabinomodabidamodapamobinopapanotadapamotapabidapamo
> panotanomodabidamobinopapanotatapabidapamo…

After hearing two minutes of such stimuli, 8-month old babies have detected that the sequence is made up of six three-syllable "words", repeated in varying orders: *bidamo, panota, dapamo, binopa, nomoda,* and *tapabi*. Experimenters found this out by first playing them the string of syllables. Later, when presented with one of those six "words" in isolation, babies paid noticeably more attention to it than they did to a three-syllable sequence that didn't occur in the original 2-minute stream of syllables, even if the novel three-syllable word is made up individual syllables that were in the stream.[32] (Although *mobita,* for instance, is made up of individual syllables that appeared in the original string, the syllables in the original string never occurred in that order. Babies would play more attention to *dapamo* than to *mobita* after hearing the example stimulus string given above.) This shows that the babies have noticed particular sequences of syllables—the 'words'—not just the individual syllables—after just two minutes of exposure.[33] (The real speech that babies hear around them contains thousands of words, of course, not just six—which is why it takes longer than two minutes to learn your first words).

Babies' statistical engines also detect phonotactic probabilities, and babies use these phonotactic probabilities to extract good candidates for 'words' from the speech stream.[34] For instance, in English, within many words, the sequence of consonants /ŋgʔ appears, as in "finger", "tingly" or "anger". There are almost no words within which the sequence /ng/ appears, however. The only place where /n/ and /g/ tend to occur next to each other is when the /n/ is at the end of one word and the /g/ is at beginning of the next, as in phrases like "pine grosbeak" or "I win games" or "thin gruel". Because /ŋg/ is a sequence contained within many words, the probability of hearing a /g/ after you hear an /ŋ/ is pretty high, compared to other sound sequences. On the other hand, because /ng/ is a sequence that is *not* contained in many words, the probability of hearing a /g/ after you hear an /n/ is quite low—after all, there are usually thousands of words someone could pick to follow a word ending in /n/, and only a few of those words begin with a /g/. Probabilistically speaking, then, it's a

[32] To tell if a baby is paying attention to something in an experiment, psychologists either measure the rate at which it's sucking on a pacifier, which changes when the baby notices something new, or count the number of seconds it spends looking in the direction that the sound is coming from. Between stimuli, the babies are given a chance to revert to a neutral rate of sucking, or to bring their gaze back to the center of the room.
[33] See Saffran, Alsin and Newport 1996.
[34] See Mattys and Juscyk, 2001.

good bet if you hear /n/ and /g/ next to each other in the speech stream, the /g/ marks the beginning of a new word, but if you hear /ŋ/ and /g/ next to each other, the /g/ doesn't begin a new word, but is part of the same word.

To test whether 9-month old babies had discovered this phonotactic generalization, experimenters chose a CVC word beginning with "g", 'gaffe' (/gæf/), that the babies were very unlikely to be familiar with, and embedded it in 'sentences' of other CVC words (the sentences didn't mean anything really, they were just strings of nonsense words). In some 'sentences', the word immediately before "gaffe" ended in an /n/, as in "…bean gaffe hold…". In other 'sentences', the word immediately before 'gaffe' ended in an /ŋ/, as in "…fang gaffe tine…". (Of course, there were no gaps in pronunciation between the words that might signal their separation independently). A group of 9-month-old babies was divided into two, groups A and B. Group A was played the sentences where 'gaffe' followed a word ending in /ŋ/, and Group B was played the sentences where 'gaffe' followed a word ending in /n/. Both groups then heard just the word 'gaffe' by itself. The experimenters knew that if the babies had noticed that 'gaffe' was a word on its own, then they would pay attention to it significantly longer than to a control word, like 'fooz', that hadn't been in the sentence. Lo and behold, the Group B babies, who heard the sentences where 'gaffe' followed a word ending in /n/, paid a lot more attention to the word 'gaffe' played alone afterwards than they did to 'fooz'. In contrast, the Group A babies, who heard the sentences where 'gaffe' followed a word ending in /ŋ/ , didn't pay any more attention to 'gaffe' than to 'fooz'. The conclusion is that the group B babies already knew that an /ng/ sequence is a pretty improbable sequence, and hence that it's unlikely to be contained in single English word. That's why they picked up on the fact that the 'gaffe' syllable had to begin its own word. Similarly, the group A babies knew that an /ŋg/ sequence a relatively *probable* sequence, and hence that it's likely to occur within a single word. They assumed that the 'gaffe' syllable in the sentences they heard was not a word on its own, but formed part of a word with the previous syllable ("fanggaffe" in the example above). Consequently, they were as unfamiliar with "gaffe" on its own as they were with "fooz", which they'd never heard.

So phonotactic knowledge is not only very important to the adult listening to the speech stream, but it is one of the first kinds of knowledge acquired by very young children who are trying to find clues which will enable them to break up the speech stream into segments to which they can attach meaning—when they are trying to detect words. The phonotactic rules of English mean that certain kinds of consonant sequences are very likely to occur at the beginnings and ends of words, and hence at the beginnings and ends of utterances. If babies pay attention to the strings of sounds at the beginnings and ends of the utterances they hear, they will learn these high-probability, edge-marking sequences, and be able to use that knowledge to discern the edges of syllables, and hence potential word boundaries, *within* the speech stream.

3.6 Syllables, rhythm, and stress

There is one major segmenting feature of the English speech stream that we haven't discussed yet. Syllables are an important phonological subunit, certainly, as demonstrated by the metered poetry examples given above, but that's not all. Read the limerick below aloud to yourself and count the syllables in each line:

(30) A foolish young hunter named Shepherd
was eaten for lunch by a leopard.
Said the leopard, "Egad!
You'd be tastier, lad
If you had been salted and peppered!"

You should have come up with 9 syllables in the first line, 9 syllables in the second, 6 syllables each in the third and fourth lines, and 9 in the last. Indeed, this is a very common syllable pattern for a limerick. But it's not just syllable count that makes for a limerick. Read the following 'lame-erick' to yourself, and count the syllables:

(31) A foolish farmer chased elephants.
Elephants were squashing his best plants.

The foolish farmer was
Mad as a bee abuzz.
His angry hopping seemed like a dance.

Correct syllable count, right? And the rhymes are in all the right places. So why doesn't this seem like much of a limerick?

The answer is that the *stress* is in all the wrong places. Go back and read the first limerick to yourself, but don't say the words. Instead, substitute 'la' for every syllable, so you can hear where the stress goes.

You should have ended up saying something like this:

(32) la LA la la LA la la LA la
la LA la la LA la la LA la
la la LA la la LA
la la LA la la LA
la LA la la LA la la LA la

In the first, second and third lines, there is a regular pattern of weaker and stronger stress:

weak STRONG weak weak STRONG weak weak STRONG weak

and similarly, there's a particular pattern for the third and fourth lines:

weak weak STRONG weak weak STRONG

Now look at the lame-erick in (31). What is the stress pattern in the first line, if you just read it normally, forgetting that it's trying to be a limerick? It's something like this (There are dots between syllables within a word. S" and "w" stand for "STRONG" and "weak"):

(33) A FOO.lish FAR.mer chased EL.e.phants
w S w S w w S w w

But of course, we know the first line of a true limerick is supposed to have the rhythm *wSwwSwwSw*, not *wSwSwwSww*. Similarly, the third and fourth lines of the lame-erick, read naturally, have the following rhythm:

(34) The FOO.lish FAR.mer was MAD as a BEE a.BUZZ.
w S w S w w S w w S w S

but the third and fourth lines of the true limerick have this rhythm:

(35) Said the LEO.pard, "e.GAD! You'd be TAS.ti.er, LAD..."
w w S w w S w w S w w S

Getting the hang of it? Every content word of English (nouns, adjectives, verbs, and adverbs) has its own particular pattern of stress. Less contentful words (prepositions, articles, auxiliary verbs, pronouns) are generally not stressed when they're part of a longer utterance (though they can be stressed for emphasis). To make a good limerick, you have to string together content words and function words that add up to the limerick stress pattern, or something fairly close to it.

Exercise 11: To practice your ear for stress, sort the 25 words listed in (36) below into five groups of five, according to stress pattern. (Hint: start by sorting according to number of syllables, since two words with same stress pattern have to have the same number of syllables).

(36) arrest, arrogant, beautiful, belittle, beware, building, compute, computer, data, defeat, disbelief, donation, hammer, inferno, inspire, intervene, misbehave, national, paragraph, printed, redefine, shadow, telephone, umbrella, underfed.

In IPA transcription, main stress is indicated by placing a vertical tick — ˈ — *before* the syllable that receives stress. So, for instance, the word ‘mother’, which has stress on the first syllable, is transcribed /ˈmʌðəɹ/, while the word ‘appear’, which has stress on the second syllable, is transcribed /əˈpijɹ/. In transcription, we will use the IPA tick to indicate main stress. However, when indicating main stress in regular English orthography, I’ll just write the stressed syllable in captial letters, like this: ‘MOther’, ‘apPEAR’.

The different stress patterns you discovered in the exercise above have names, which you may be familiar with from English class. A group of syllables containing one strong stress is called a *foot*. Two-syllable groups can have either a Sw pattern, like ‘MOther’, or a wS pattern, like ‘apPEAR’. The Sw pattern is called a *trochee*, or a *trochaic foot*. The wS pattern is called an *iamb*, or an *iambic foot*. (Shakespeare wrote in *iambic pentameter*—that is, most of his lines consisted of five iambic feet: “Shall I / comPARE / thee TO / a SUM / mer’s DAY?”). Three syllable groups can have a Sww pattern, like ‘TElephone’, a wSw pattern like ‘comPUter’, or a wwS pattern, like ‘redeFINE’. Sww feet are called *dactyls*, and wwS feet are called *anapests*. (Lines 3 and 4 of a good limerick are made up of two anapests). Feet with wSw stress are called ‘amphibrachs,’ but in normal speech this footing occurs only rarely.

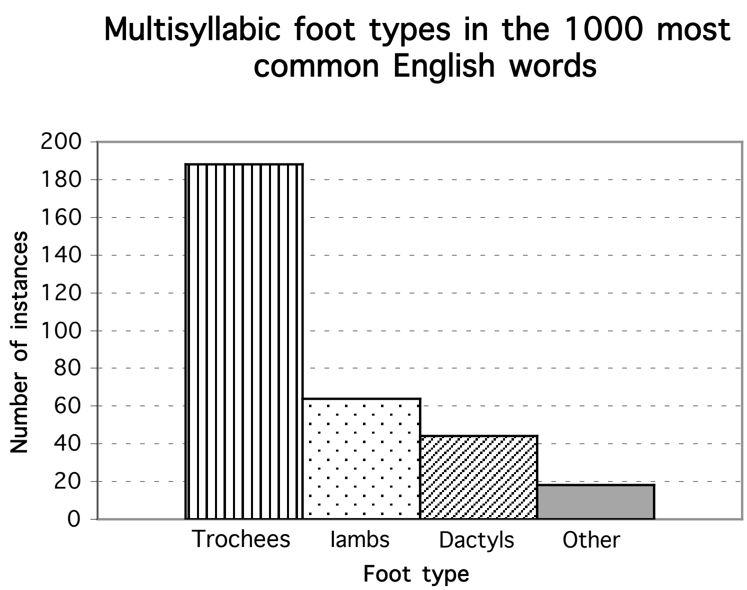
© 2003 Darby Conley. Dist. by UFS, Inc.

Because every content word in English has to have a stress on it somewhere, stress is also an important clue for babies trying to parse the speech stream into units. A sequence of unstressed syllables can’t contain a content word. So, for instance, in the sentence “John is arrogant”, whose transcription is /ˈdʒɑnɪzˈæɹəgənt/, we know that the last two syllables of the utterance, /ɹəgənt/, can't be a content word on their own, or individually, because neither of them bears stress.

We know that babies pay attention to stress as a cue to word boundaries because of another series of 'Did they think it's a word?' experiments.[35]

First, it's important to know that stress can be a pretty good indication of a word boundary in English. English has a tendency to initial stress—if a word has two syllables, it's more likely to have a Sw pattern than a wS pattern. Similarly, common trisyllabic words have a Sww pattern, not a wwS pattern. A bar chart showing the number of trochees, iambs, dactyls and 'other' patterns among the 1000 most common (printed) English words is given in Fig. 6:

Fig. 6



[35] Jusczyk 2001.

Most of the most frequent multisyllabic words have initial stress. In infant-directed speech, or "motherese", this tendency is even more exaggerated.

Do infants make use of this Sw tendency when parsing the speech stream? In an experiment designed to investigate this question, 7.5 month-old infants were divided into two groups. Two trochaic words (e.g. 'doctor' and 'candle') were played to Group A until they were familiar with them. Two iambic words (e.g. 'guitar' and 'device') were played to Group B until they were familiar with them. Then, two readings were played to each group. The first reading contained the words they'd just heard, while the second reading didn't contain those words. The experimenters hypothesized that if the babies recognized the words they'd just heard, they would listen longer to the reading than if they didn't recognize them. If the Group A babies recognized their words, but the Group B babies didn't, it would show that babies learning English are paying attention to initial stress as a cue to wordhood.

Sure enough, the Group A babies listened longer to the passages containing their familiar trochaic words than they did to passages containing unfamiliar words. The Group B babies, however, didn't listen any longer to the passages containing their familiar iambic words than they did to the passages containing words they were unfamiliar with.

The experimenters thought that the Group B babies were having trouble recognizing their iambic words because they were using initial stress as a cue to word boundaries. That is, they thought that babies might be guessing that every stressed syllable began a new word. When they heard a sentence like /ˈbɪlɪzˈbajɪŋəgɪˈtɑɹ/, 'Bill is buying a guitar," they made a mistake and parsed the stressed syllable /ˈtɑɹ/ as the beginning of a new word.

To confirm that that was what was happening, the researchers conducted a second experiment. They familiarized babies with *single*-syllable words that were the same as the stressed syllable in the iambic words from the passage. For instance, instead of familiarizing the babies with 'guitar' and 'device', they familiarized them with 'tar' and 'vice'. Then they played the original passages containing the iambic words 'guitar' and 'device' to those babies, as well as a control passage containing other words that they hadn't been familiarized with. Sure enough, this time the babies did listen longer to the passages containing the iambic words than to the control passages. The babies were using stress as a clue to word boundaries—even though in these passages it was leading them to make a mistake!

3.7 Misparsing the speech stream, mondegreens and allophones

We adults sometimes get a taste of the problem faced by babies when we misparse the speech stream. This kind of mistake is one of the main ingredients of a particular type of joke, usually a pun. Remember this bit of doggerel?

> Mairzey dotesen
> Dozey dotesen
> Liddle lamzey divy
> A kiddley divy too,
> Wooden shoe?[36]

Bart Simpson of the TV show *The Simpsons* specializes in this type of joke. He phones Moe, the dimwitted barman, and asks to speak to someone with an improbable name, presented in the order last name, first name. Moe shouts out the requested name, in the right order, to the assembled barflies, who respond with hoots of laughter, to Moe's consternation. Here are some typical exchanges:

Bart: I want to talk to a guy there named Coholic, first name Al.
Moe: Phone call for Al...Al Coholic...is there an Al Coholic here?

Bart: I want to talk to a guy there named Butz, first name Seymour.
Moe: Hey, is there a Butz here? Seymour Butz? Hey, everybody, I wanna Seymour Butz!

Bart: I'm looking for a Miss Huggenkiss, first name Amanda.

[36] Mares eat oats, and does eat oats, and little lambs eat ivy. A kid'll eat ivy too, wouldn't you?

Moe: Uh, Amanda Huggenkiss? Hey, I'm looking for Amanda Huggenkiss! Ah, why can't I find Amanda Huggenkiss?

Another place where this phenomenon can inspire a few chuckles is in "Mondegreens." A mondegreen is a misheard song lyric. One famous one comes from the Jimi Hendrix song *Purple Haze*. He sings "Excuse me, while I kiss the sky," but many people hear "Excuse me, while I kiss this guy." Children are particularly susceptible to this problem: they sing about "Olive, the other reindeer"[37], say prayers to "Our father, who art in Heaven, Howard be thy name,"[38] and think "Donuts make your brown eyes blue".[39]

© NEA, Inc.

What happens when we misparse speech in this way? Obviously our phonotactic system is fooled by such examples. How does this happen?

Sometimes it's just that one phrase is homophonous with another phrase, and the only way to detect which parse is the correct one is by context. This is true of "Olive, the other reindeer," and "All of the other reindeer." In any situation, "Olive" and "All of" are just homophonous, and context is the only thing which will allow you to figure out which is intended.

---

[37] Really "All of the other reindeer," from the Christmas carol *Rudolf the red-nosed reindeer*.

[38] Really "Our Father, who art in Heaven, hallowed be thy name," from the Lord's Prayer.

[39] Really "Don't it make your brown eyes blue," from the Crystal Gayle song.

In other contexts, more complicated phenomena come into play. Sounds often vary a little bit in pronunciation, depending on what other sounds precede or follow them. For instance, when the alveolar stop /t/ is followed by the palatal glide /j/, the stop quality of the /t/ is softened, and the combination is pronounced more like the in-between alveopalatal affricate /t͡ʃ/ (for some speakers of English, this happens within words like "Tuesday" and "tune"). When you're speaking precisely, the phrases 'Wouldn't you?" and "Wooden shoe" sound quite distinct: /ˈwʊdəntˈjuw/ vs. /ˈwʊdənˈʃuw/. However, when you're speaking quickly, the /t/ at the end of "Wouldn't" and the /j/ at the beginning of "you" get smeared together into the alveopalatal affricate, and you end up saying /ˈwʊdənˈt͡ʃuw/—which is essentially indistinguishable from "wooden shoe." Normally you'd never confuse a /t/ and a /ʃ/—but if the /t/ is followed by a /j/, it can happen easily.

Expectations about stress assignment can facilitate misparses as well. In a song by Queen, *Bohemian Rhapsody*, one lyric goes, /ˈbejɛlzəbʌbhæzəˈdɛvɨlpʌtəˈsaɪdfɔɹmij/, "Beelzebub has a devil put aside for me." This line has a very odd, but nonetheless common, mondegreen: "Beelzebub has a devil for a sideboard(eeee)." Consider the word boundaries and stress assignments in the last phrase of the correct parse and the mondegreen:

(37) a. ...a DEvil put aSIDE for...
b. ...a DEvil for a SIDEboard

In the correct lyric, there's an iambic word—a word whose main stress does *not* fall on the first syllable. The mondegreen doesn't contain an iamb—it only contains single-syllable words and trochees. The person hearing this mondegreen has reanalyzed the phrase so that it conforms to the normal English expectation that a stressed syllable begins the word. They're falling prey to the same expectation that the babies in the stress experiment did when they behaved as if the phrase /ˈbɪlɪzˈbajɪŋəgɪˈtɑɹ/, contained the word /ˈtɑr/.

Here's a slightly more complicated example. Consider the transcriptions of the last part of the Hendrix line and its mondegreen:

(38) Excuse me while I…
 a. /kɪsðɨskaj/
 b. /kɪsðɨsgaj/

Evidently, we mishear the /k/ as a /g/. This doesn't happen most of the time. Is anyone likely to mishear the phrase "the coat" as "the goat"? Try saying them aloud to yourself, and comparing them to "the sky" and "this guy". It turns out that the difference in syllabification is playing a role here. In "the sky", the /k/ is the *second* sound in a syllable, following an /s/ in the same syllable—*sky* is a syllable with a complex onset, made up of two consonants. By contrast, in "the coat" and "the goat", the /k/ is a syllable onset all by itself. Is there anything about the pronunciation of /k/ that is different when it occurs as a syllable onset by itself vs. when it occurs as the second consonant in an onset? It turns out there is.

Remember, early in the last chapter, we observed that voiceless stops were often *aspirated* in English. Words like *pot, tab,* and *kill* are pronounced with an extra puff of air right after the consonants /p/, /t/, and /k/. If you compare these words with *bought, dab* and *gill*, by placing your hand close to your mouth, you will detect that voiced stops are not pronounced with aspiration—no extra puff of air. But remember, in some cases, we noticed that the voiceless stops are *not* aspirated. When they're part of a complex onset in words like *spot, stab* and *skill*, the voiceless stops lose that extra aspiration.

It is this difference in pronunciation of /k/ which makes it easy to mistakenly hear *the sky* as *this guy*: the /k/ in *sky* is unaspirated, because it's the second element in a complex onset, which makes it sound much more like /g/ than a /k/ in a simple onset does.

We've just seen that the same basic sound—/k/—has two slightly different pronunciations, depending on where in the syllable it is. This is an example of a phenomenon called *allophony*. Allophones are different pronunciations of the same sound that arise in different phonological environments. Compare your pronunciation of the vowels in the following two lists of words:

(39)

write	ride
trite	tried
height	hide
tripe	tribe
rice	rise
trice	tries
lice	lies
lout	loud
bout	bowed
house (*n.*)	house (*v.*)
mouth (*n.*)	mouth (*v.*)
sat	sad
bat	bad
cat	cad

You perhaps have noticed that the vowels in the column on the left sound different than the vowels in the column on the right. This is again allophony in action. There are two different pronunciations of each of these vowels, a shorter one, and a longer one. (This is a genuine difference in length of pronunciation, not a difference in sound quality). Which column has the longer vowels in it, and which has the shorter ones? What aspect of the surrounding phonological environment seems to determine whether the vowel is long or short?

Exercise 11: What is the conditioning environment which dictates when to produce a longer vowel and when to produce a shorter vowel in the words listed in (18)?

In modern English, certain vowels are longer when they appear before voiced consonants, and shorter when they appear before voiceless consonants. These different pronunciations are allophones of the same basic vowel sound, in the same way that aspirated and unaspirated /k/ are allophones of the same basic consonant.

In the two cases of allophony we have so far considered, the difference in pronunciation usually helps the listener to distinguish an otherwise potentially difficult-to-hear contrast. Aspirating voiceless stops at the beginnings of stressed syllables heightens the difference between them and voiceed stops—it makes the difference between 'pat' and 'bat' easier to detect. Changing the vowel length before voiceless stops at the *end* of the syllable does the same thing. The length of the vowel is a clue that helps the listener to tell the difference between voiced and voiceless stops, this time in an environment—the end of the syllable—where aspiration can't provide the extra hint.

The differences in pronunciation make the distinction easier for listeners to hear, and makes misparses less likely to occur. Only in rare cases, like *Excuse me while I kiss the sky,* do they have the opposite effect. Allophony, then, is another tool that is useful to the parser in dividing the speech stream into discrete phonological words.

3.8 *What we know about phonological words*

We've covered a broad range of material in this section on phonological words, and assimilated a lot of new information. In the last chapter, we learned a precise system for representing actual pronunciation, the International Phonetic Alphabet. The architecture of the vocal apparatus provides the means for producing families of sounds that differ by only one feature: Place, Manner and Voicing, in the case of consonants, Height, Tenseness and Backness in the case of vowels. These families of sounds tend to behave alike with respect to particular phonological patterns.

We then moved on to consider what kind of patterns of sounds define a possible phonological word in English. Words are divided into syllables, which have very significant restrictions on the ordering of the elements that make them up. These restrictions constitute the *phonotactics* of English words. Understanding phonotactic patterns enables us to understand what kinds of sound sequences can be legitimate phonological words of English, and also how we can parse the speech stream into discrete units. We saw evidence that babies use phonotactic probabilities to identify the strings of sounds that are potential words of their language. Next, we learned about the requirement in English that content words receive stress, and saw that the most common pattern in English words is for stress to fall on the first syllable of the word. Finally, we considered cases where misparses of speech strings arise, even in adult speech, and saw how allophony can help speakers avoid misparses, providing additional clues to speakers about how to divide up the speech stream into phonological words.

Having identified some of the factors at work which create the English phonological word—the factors that our orthographic spacing conventions are based on—we'll now move on to consider the mismatch between phonological words and listemes in some detail. (Remember listemes? They are the *minimal meaningful* units of sound). We're moving into the wild world of word-formation, full of prefixes and suffixes: *morphology*.

XX to come: problem sets, further reading XX