3

Phonological Words: Calling All Scrabble Players!

/fanəˈladʒɪkəl ˈwʌɹdz: ˈkalɪŋ ˈal ˈskɹæbəl ˈplejəɹz/

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*), 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

We are now finally able to consider the central problem of this first part of the book: what is a phonological word? Why do we have intuitions about where we should put spaces in written English text? 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 word, *-s* is not a word, and *works like a dog* is four words. These units that we're counting are not minimal units of meaning, though – they are some sort of unit of *sound*.

Not only do we have firm intuitions about how many of these phonological words are present in a given sentence 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:

(1)	a.	timp	/tɪmp/
	b.	rog	/ɹɑg/
	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.

Exercise 3.1 Give each of the strings of sounds in (1) 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. Arrange the words in order according to the scores you assign, lowest to highest.

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

(2)	*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. *Bod* gets a high score because it actually is a word of English (short for *body*); *mbotto* gets a low score, and it is clearly not a word of English. But why do *timp*, *rog* and *spink* get higher scores than *mbotto*? They are

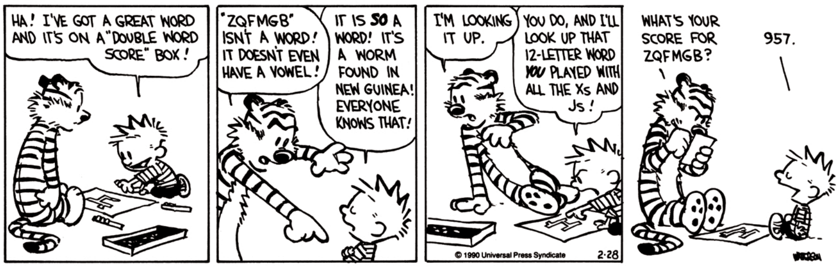


Figure 3.1 Calvin and Hobbes. © 1990 Watterson. Dist. By Universal Press Syndicate. Reprinted with permission. All rights reserved

not English words either – the people rating the words had probably never seen these strings before! Shouldn't they all seem as unlikely to be English as *mbotto*?

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. Let's imagine the board layout at this point in the game is such that if 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 (Figure 3.1). (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 say "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 *so‹* "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,

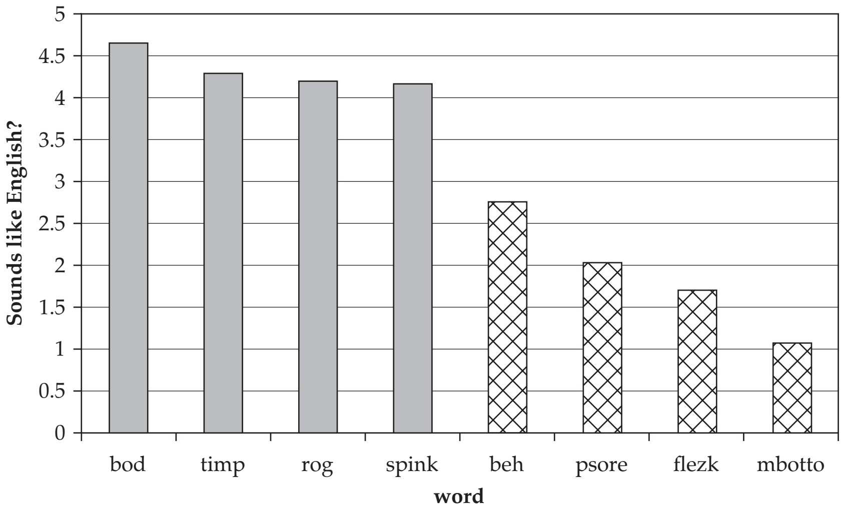


Figure 3.2 Word-like ratings for strings of English sounds

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 (2) above shows a significant jump between *spink* and *beh* (Figure 3.2).

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. You might consider playing the first four in an imaginary game of Scrabble, but you wouldn't even wonder about the last four. In this chapter, we're going to look at some of the constraints that determine what sound combinations make for a possible English word.

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 stand on its own as a phonological word, although it has its own meaning. 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, is a syllable. The first syllable of the word *open*, /ow/, is made up of just one vowel, and so is the first syllable of *apart*, /ə/. Consequently, all English phonological words must contain at least a vowel (as Hobbes, in Figure 3.1, knows very well).

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*:

(3) '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:

(4) 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:

(5) 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 the opposite direction, what about this *Rhyme for Remembering the Date of Easter*, by Justin Richardson?

(6) 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:

(7) 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 a 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:

When an ortho-
graphic word
won't fit onto

a single line of text, it is hyphenated at a syllable boundary.

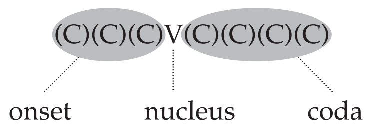
The convention of hyphenating at a syllable boundary can be overridden by other considerations. If putting a hyphen at a syllable boundary would result in there being only one letter on a line, another breaking point is chosen – often an affix boundary. For instance, *unable* is hyphenated as *un-able*, rather than *u-nable* or *unab-le*.

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, as in the word /**stɹ**ɪŋ/, "string," and up to four consonants at the end, as in the word /tɛ**ksts**/, "texts."

Syllables can be divided into three parts: the beginning, or *onset* of the syllable, made up of one to three consonants, the required middle, or *nucleus* of the syllable, made up of the vowel, and the end, or *coda*, of the syllable, again made up of one or more consonants, up to four. The onset and coda are optional, as we saw above; syllables can be made up of just a vowel. Most syllables, though, also have an onset, and many also have a coda. The anatomy of a syllable is shown in (8); "C" stands for "consonant" and "V" for vowel; brackets indicate optionality:

(8)



It's the nucleus 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 nucleus in common. (In poetic terms, *coast* and *code* are alliterative and assonant – but they don't rhyme.)

3.3 Phonotactic Restrictions on English Syllables

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, i.e., about 75 billion. Obviously, this is far larger than the actual number of possible syllables. For one thing, that figure includes cases where contiguous 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.

If we incorporate the effects of phonotactic rule 2 into our computation, excluding sequences of identical sounds within an onset or coda, we reduce the number of possible syllables by about 20 percent, to a mere 60,779,920,695 possible syllables. This 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 of English that 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 seven available consonant positions in English syllables. For instance, although the velar nasal sound /ŋ/ frequently occurs at the end of syllables (*s**ing***, *jump**ing***, *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.

Taking into account rule 4, the number of possible English syllables is reduced by another 25 percent to 44,881,090,380. Still too much, and we still haven't found out what's wrong with /zɪpb/.

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/ cannot 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.

> Always keep in mind that we are talking about clusters of consonantal *sounds*, not spellings. Although 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, the *sound* /h/ never occurs in a complex onset. Words like *thin* and *chin* and *through* have more than one orthographic consonant in their spelling, but only *through* 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?

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

Second, when an 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 /**tɹ**ʌk/ "truck," or /**dɹ**ɑp/, "drop" we never find sequences like /ɹtʌk/, "rtuck," or /ɹdɑp/, "rdop."

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

Only 14 of the 20 possible onset consonants are obstruents. Between them, generalizations 5 and 6 whittle the number of possibilities down another 35 percent, to 28,956,015,990.

The *second* consonant of a two-consonant 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." Note that even though the word *svelte* is spelled with the letters "sv," it is pronounced with an /sf/ – /sfɛlt/.

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 a measly 1 percent to 28,578,886,740.

When we consider other combinations of consonants in two-consonant onsets, 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, they cannot pair up indiscriminately. Rather, when the first consonant is anything besides an /s/, the second consonant has to be a liquid or a glide.[1] 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."[2]

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 – it 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 /**sf**ijɹ/, "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 another 1 percent to 28,117,067,940.

Here's a summary so far of our observations about onsets in English:

(9) 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.

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

(10) a. /**spl**ijn/ spleen
  b. /**spɹ**ɪŋ/ spring
  c. /**stɹ**ijm/ stream
  d. /**skɹ**ijn/ screen
  e. /**sklɹ̩**oʊsɪs/ 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 of consonants contained within a bigger sequence must itself obey all the 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/.) This 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

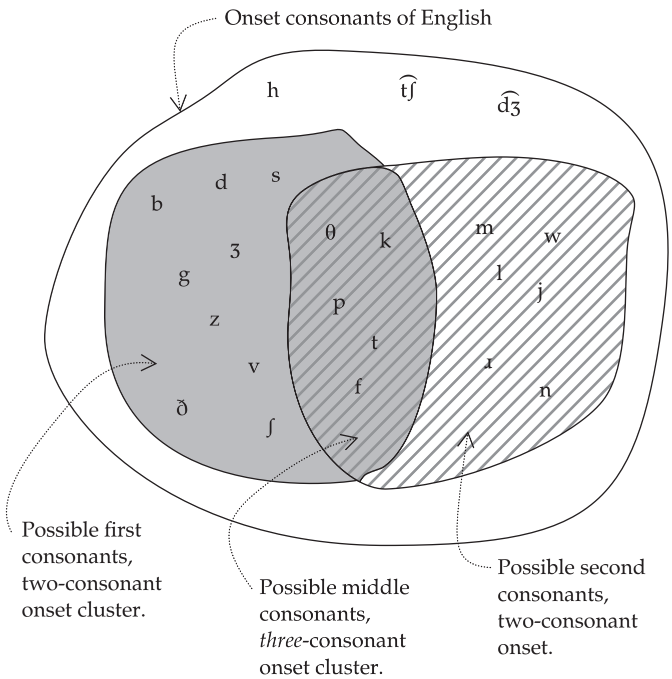


Figure 3.3 The Substring Rule

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 shown in Figure 3.3. The only sounds that can possibly work both as first-consonants and second-consonants in a three-consonant onset are the voiceless obstruents: /p/, /t/, /k/, /f/, and /θ/.

Since the only potential legitimate C_2 in a three-consonant cluster is a voiceless obstruent, it follows from the substring rule that the only possible C_1 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 has to 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 by a huge amount – about 96 percent – all the way to 870,327,990. Now we're really starting to help our Scrabble player out.

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*), though it's technically possible. Similarly, we don't run into /spw/, /stw/, or /sfl/ onsets, despite the fact that /sp/, /pw/, /st/, /tw/, /sf/ and /fl/ all do 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? Try comparing /spwɛt/ or /sflɪŋ/ to /zbijd/ or /mɹæt/. (There does exist a technical word with an /sfɹ/ onset: *sphragistics*, the study of official seals or signet rings.)

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 about another 44 percent to 490,875,390.

In two-consonant codas, the first consonant can be pretty much anything (except /h/, which 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*, pronounced as /z/ or /s/, can be suffixed to the end of almost any English noun or verb, creating a complex two-consonant coda.

In the second position in a two-consonant sequence, though, 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 another 32 percent to 330,467,370.

For the remaining 21 × 17 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 (fairly complicated!) Figure 3.4. The consonants on the left are 20 of the 21

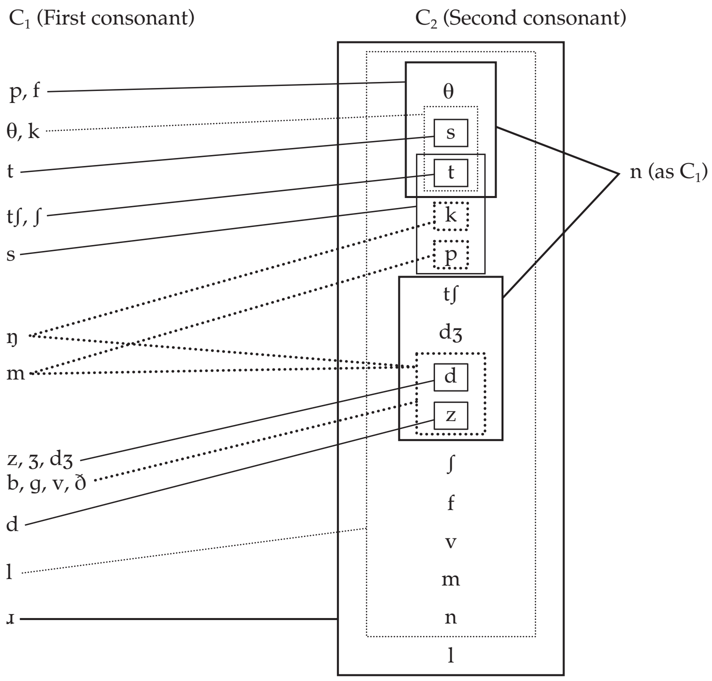


Figure 3.4 Two-consonant coda clusters of English

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 left 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." (That lone /n/ on the right, rather than the left, is still intended to be a first consonant. It was just difficult to fit in neatly on the left side where it belonged. 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*.)

Some generalizations are immediately apparent. The only sound which can precede all the other consonants 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 /l/ and /ɹ/ in the bottom left-hand corner of Figure 3.4 are connected to the biggest squares of any consonants.) Finally, the nasal /n/ can precede any consonant of English that is pronounced using the tip of the tongue (any **coronal** consonant, which includes all the alveolars and interdentals), 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 Figure 3.4, 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 just a little bit more, 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 3.2 Try to think of a phonological word ending in each of the possible two-consonant codas represented in Figure 3.4. 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 Figure 3.4 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* /bʌ**ɹpt**/ or *burps* /bʌ**ɹps**/).

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 codas where the Substring Rule predicts that a third consonant could be non-coronal.

Exercise 3.3 (difficult): Looking at Figure 3.4, figure out four of the coda sequences where 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 also (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 possible 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 by 93 percent 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 another 70 percent to 6,596,940.

There are many other lesser restrictions on phonotactic patterns in English which we will not consider in detail. For instance, no English syllables begin with a /dl/ or /tl/ onset, although both /dr/ and /tr/ are extremely common. As mentioned in the last chapter, 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 (i.e. a "long" vowel) – if they have no coda, their lonely vowel can't be lax (short). Consequently, for example, *do* is pronounced /duw/, not /dʊ/; *be* is /bij/, not /bɪ/, and *so* is /sow/. (This is why "beh" scored low in our original "Does it sound like an English word?" test.)

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 four orders of magnitude smaller than our first calculation – we've reduced the total number of syllables by 99.99 percent. The additional constraints which we haven't considered here, of course, mean that the true figure in fact is 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.

Another game in which your phonotactic instincts are brought to bear is "Ghost," a good car game for spelling geeks. In Ghost, players take turns adding a letter to a string of letters. The letter added must create the beginning of a correctly spelled word, but you don't want to add a letter which will complete the word. If you do complete the word, you are penalized by getting the first letter from the word "ghost." The winner begins a new word. Players who lose five rounds become a "g-h-o-s-t" and drop out; 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.

3.4 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 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:

bidamodapamopanotabinopatapabinomodadapamobinopapanotadapam
otapabinomodabidamodapamobinopapanotadapamotapabidapamopa
notanomodabidamobinopapanotatapabidapamo . . .

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*.

How can one tell what an 8-month-old baby has detected? Experimenters have found that they can measure the amount of time a baby pays attention to a sound by how long they keep their head turned toward it. When the baby gets bored, he'll turn his head away. In this experiment, the baby and his mother sit in a booth with speakers in it. At first, they just hear the two-minute string of unbroken syllables given above. (This is the "training" phase.) Later, the babies are presented with just some three-syllable "words" in isolation, e.g. *bidamo*, *bidamo*, *bidamo*. When the isolated "word" was one they had heard during the training sequence, babies paid noticeably more attention to it than they did to isolated "words" that hadn't occured in the original 2-minute stream of syllables – even if the isolated "word" is made up individual syllables that were in the stream, but in a different order. (Although *mobida*, 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 keep their heads turned toward a speaker playing *bidamo*, *bidamo*, *bidamo* than to one playing *mobida*, *mobida*, *mobida* after hearing the training string given above.) This shows that the babies noticed particular sequences of syllables – the "words" and not just the individual syllables – after just two minutes of exposure. (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 for them to learn their first words.)

Babies' statistical engines don't just pay attention to syllable sequences. They also detect phonotactic probabilities, and then use these probabilities to help them notice good candidates for "words" from the speech stream. For instance, in English, within many words, the sequence of consonants /ŋg/ appears, as in *finger* /fɪŋgəɹ/, *tingly* /tɪŋglij/, *anger* /æŋgəɹ/, or even *English* /ɪŋglɪʃ/. There are almost no words within which the sequence /**n**g/ 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," /pajngɹowsbijk/ or "I win games" /ajwɪngejmz/ or "thin gruel" /θɪngɹuwl/. Because /ŋg/ is a sequence contained within many words, the likelihood of hearing a /g/ after you hear an /ŋ/ is pretty high. 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 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 . . . ,"/ . . . bijngæfhowld . . . /. In other "sentences," the word immediately before "gaffe" ended in an /ŋ/, as in ". . . fang gaffe time . . ."/ . . . fæŋgæftajm . . . /. (Of course, there were no pauses in pronunciation between the words that might signal their separation independently.) A group of 9-month-old babies from English-speaking homes was divided into two groups, A and B. Group A was played the sentences where "gaffe" followed a word ending in /ŋ/, (as in ". . . fang gaffe time . . ."), and Group B was played the sentences where "gaffe" followed a word ending in /n/ (as in ". . . bean gaffe hold . . ."). Both groups were then played 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 than they did to "fooz." In contrast, the Group A babies, who heard the sentences where "gaffe" had followed a word ending in /ŋ/, didn't pay any more attention to "gaffe" than to "fooz." This showed that the Group B babies, living around English speakers all the time, already knew that an /ng/ sequence is a pretty improbable sequence, and hence that it's unlikely to be contained in single word. That's why they picked up on the fact that the "gaffe" syllable had to be a word on its own. Similarly, the Group A babies knew that an /ŋg/ sequence is a relatively *probable* sequence, and hence that it's likely to occur within a single word. Those babies assumed that the "gaffe" syllable in the sentences they heard was *not* a word on its own – they probably thought it 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 information acquired by (the brains of) very young children, which are trying to find clues which will enable them to break up the speech stream into smaller segments to which they can attach meaning. 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 notice these high-probability, edge-marking sequences, and then they can use that knowledge to discern the edges of syllables, and hence potential word boundaries, *within* the speech stream.

3.5 Syllables, Rhythm, and Stress

There is one major segmenting feature of the English speech stream that we haven't yet discussed. 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:

(11) 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:

(12) 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.

Exercise 3.4 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:

(13) 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 (12). 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.")

(14) 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:

(15) 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:

(16) Said the LEO.pard, "e.GAD! / You'd be TAS.ti.er, LAD . . ."
w w S w wS 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 3.5 To practice your ear for stress, sort the 25 words listed 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.)

arrest, arrogant, atrocious, beautiful, belittle, beware, building, compute, computer, data, defeat, disbelief, donation, hammer, inferno, inspire, intervene, misbehave, national, paragraph, printed, redefine, shadow, telephone, 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 capital letters, like this: "MOther," "apPEAR."

The different stress patterns you discovered in Exercise 3.5 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 feet 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*.

3.6 Using Stress to Parse the Speech Stream into Words

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 /ˈdʒɑnəzˈæɹəgənt/, we know tha the last two syllables of the utterance, /ɹəgənt/, cannot 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 "Do they think it's a word?" experiments.

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 Figure 3.5. 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 words have 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 sentences were played to each group. The first sentence 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 sentence that contained them. If the Group A babies recognized their words, but the Group B babies didn't, it would show

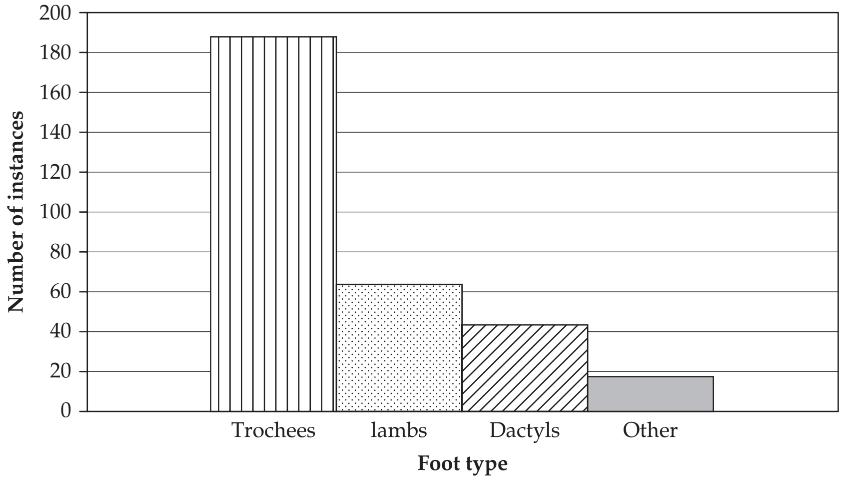


Figure 3.5 Foot types in the 1,000 most common English words

that babies learning English are paying extra attention to trochaic stress as a cue to wordhood.

Sure enough, the Group A babies listened longer to the sentences 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 playing the babies "guitar" and "device," they played them "tar" and "vice." Then they played the original sentences containing the iambic words "guitar" and "device" to those babies, as well as a control sentence containing other words that they hadn't been familiarized with. Sure enough, this time the babies did listen longer to the sentences containing the iambic words than to the control sentenecs. The babies were using stress as a clue to word boundaries – even though in these cases it was leading them to make a mistake – they thought "tar" was a word in the sentence, even though it was actually a subpart of the bigger word "guitar."

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.

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, which Bart gives 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.
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" (from the line "All of the other reindeer," in "Rudolph the red-nosed reindeer"), say prayers to "Our father, who art in Heaven, Howard be thy name" (from the line ". . . hallowed be thy name" in the Lord's Prayer) and think "Donuts make my brown eyes blue" (from "Don't it make my brown eyes blue," a Crystal Gale song).

Mondegreens are named after a mondegreen that arose when people mis-heard the lyric of an old English ballad: "They had slain the Earl of Moray, and laid him on the green," heard as ". . . the Earl of Moray and Lady Mondegreen."

This doesn't just happen with song lyrics or set phrases; it often happens in speech too, and is only revealed when the misparser writes down what she *thought* she heard, for example, *old wise tale* for "old wives" tale," or *take for granite* for "take for granted." (These are often called *eggcorns* – from someone who misunderstood "acorn". A collection of eggcorns can be found at http://eggcorns.lascribe.net/.)

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." For most people, "Olive" and "All of" are just homophonous, and context is the only thing which will allow you to figure out which is intended.

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,

/beˈjɛlzəbʌbhæzəˈdɛvɨlpʊtəˈsaɪdfɔɹˈmij/, "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. Eeeeee . . ." Consider the word boundaries and stress assignments in the last phrase of the correct parse and also in the mondegreen:

(17) 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 /ˈtɑɹ/.

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

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

Evidently, people hearing the mondegreen 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.

3.8 Allophony

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:

(19)	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. (In some dialects of English, the difference is just in the length of the vowels. For others, there's also a difference in the way some of the vowels are pronounced.) 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 3.6 What is the conditioning environment which dictates when to produce a longer vowel and when to produce a shorter vowel in the words listed above?

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 voiced 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** does not 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.

3.9 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 Chapter 2, 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 in creating the English phonological word – the factors that our orthographic spacing conventions are based on – we'll now move on to think about the relationship between phonological words and listemes. (Remember listemes? They are the *minimal meaningful* units of sound.) First, we'll consider where new listemes come from. We've already seen one mechanism that can generate new listemes – misparses of the speech string by learners of the language. In the next chapter, we'll think about that some more, and also look at other ways that new words can enter the language. We'll gradually move into the wild world of word-formation, full of prefixes and suffixes and other exotic beasts: **morphology**.

Study Problems

1. In the text, phonotactic rules are given describing possible combinations of sounds in the onsets of English syllables. Here are some generalizations about possible combinations of sounds in the codas of English syllables. (These generalizations only apply in words with no suffixes attached.)

a. If a syllable ends in two consonants:
   (i) If one of the last two consonants is a *nasal*, then
      (a) both the consonants have the same place of articulation;
      (b) the non-nasal consonant always comes second;
      (c) the non-nasal consonant is always either a stop or a fricative;
      (d) if the non-nasal consonant is bilabial or velar, it must be voiceless;
   (ii) If the first of the last two consonants is /s/, the second consonant must be a voiceless stop.
   (iii) If the first of the last two consonants is an obstruent, then both consonants must have the same voicing.

b. If a syllable ends in *three* consonants:
   (i) The last consonant must be dental, alveolar, or palatal.

Below are some invented "words" written in IPA, each of which violates one of the above generalizations, and hence are not actual or possible words of English. For each of the words below, try pronouncing it, and then state which of the generalizations above it violates:

a. kɪɹlf
b. pɛðəsθ
c. zeibk
d. flɔɪsh
e. slʌgθb
f. ʃɹænk
g. sɪŋg
h. əlɑɪsz

2. Using a dictionary that gives information about word origins, e.g. the Oxford English Dictionary, Merriam-Webster's, etc., look up the following words:

a. psychiatry
b. psilocybin
c. psoriasis

   (i) What language do these words trace their origins to?
   (ii) Transcribe each word in IPA as it is pronounced in your dialect of English. Indicate main stress with a tick before

the stressed syllable, and indicate syllable boundaries by placing a dot between syllables.

(iii) How is the spelling of these words misleading? Given their spelling, how do you suppose the onset of the first syllable of each of these words was pronounced in the original language (show it in IPA)? What phonotactic rule of English did these words most likely violate?

3. In several dialects of English, including American English, there are subtle variations in the pronunciation of the /l/ sound. It occurs in two variants, one that is transcribed as [l] and the other as [ɫ]. Here is a list of words containing /l/, with transcriptions showing its varying pronunciation in these dialects of English.

leaf	[lijf]	filter	[fɪɫtɹ̩]
feel	[fijɫ]	sold	[soɫd]
collect	[kəlɛkt]	lap	[læp]
inhale	[ənhejɫ]	milking	[mɪɫkɪŋ]
police	[pəlijs]	letter	[lɛtɹ̩]

a. Some of these words have more than one syllable in them. Which ones are they? Re-transcribe them showing where the syllable boundary is by placing a dot between the two syllables, and indicate which syllable receives main stress by putting a tick before it, as described in the chapter.

b. The two different kinds of /l/ are allophones. Which one you get depends on where in the word the /l/ appears. Describe where each variant of /l/ occurs in terms of syllable structure by filling in the blanks:

(i) /l/ is pronounced as [l] when it occurs in the ______________ of a syllable.

(ii) /l/ is pronounced as [ɫ] when it occurs in the ______________ of a syllable.

c. The [ɫ] sound is often called a "dark l," while the [l] sound is called "clear l." Dark l is produced by pulling the body of the tongue back toward the velum while making the /l/ sound; in the clear l, the tongue body is more towards the front of the mouth. In some dialects of English, particularly in Southeastern Britain, the dark l allophone has altered still further, becoming the glide [w] or the vowel [u]. In these dialects, “feel” is pronounced [fijw], “table” is pronounced [tejbu], and so on. Why do you think this sound change happens to the dark l allophone but not to the clear l allophone? (Consult the vowel diagrams given in Chapter 20. Your answer should mention the position of the tongue.)

4. We have seen that /t/ has an aspirated allophone [tʰ] and an unaspirated allophone [t]. We learned that the aspirated one occurs when the /t/ is by itself at the beginning of a stressed syllable. In American English /t/ also has other allophones. One of them is written [t˺]; this allophone is called “unreleased t,” and involves a glottal stop, Another allophone of /t/ is written [ɾ]; this one is called a **flap**, and involves quickly touching the tongue to the alveolar ridge without creating a full closure like a regular stop. Below, some words containing these other two allophones of /t/ are given, with transcriptions showing which allophone of /t/ occurs in each:

pat	[pʰæt˺]	patted	[pʰæɾəd]
repeating	[ɹəpʰijɾɪŋ]	Batman	[bæt˺mæn]
repeat	[ɹəpʰijt˺]	bucket	[bʌkət˺]
atom	[æɾəm]	analytic	[ænəlɪɾək]
quiet	[kwajət˺]	quietly	[kwajət˺lij]
quieted	[kwajəɾəd]	footman	[fʊt˺mən]

a. Seven of these are multisyllabic. Retranscribe the last six, showing where the main stress is with a tick in front of the stressed syllable. The first one is done for you:

   ['pæɾəd]

b. Describe the distribution of the [ɾ] allophone of /t/ by filling in the blank:

   The [ɾ] allophone of /t/ is produced when the syllable before the /t/ is __________ , the sound immediately following the /t/ is __________ and that following syllable is ____________.

c. Describe the distribution of the [t˺] allophone by filling in the blank:

   The [t˺] allophone of /t/ is produced when the /t/ occurs in the ___________ of a syllable.

Further Reading

On English phonotactics (technical!):
Hammond, M. (1999) *The Phonology of English*. Oxford: Oxford University Press.

On child language acquisition and speech perception:
Golinkoff, Roberta M. and Hirsh-Pasek, Kathy (1999) *How Babies Talk*. New York: Penguin.
Werker, Janet F. (1995) "Exploring developmental changes in cross-language speech perception," in L. Gleitman (ed.), *An Invitation to Cognitive Science*, Vol. 1: *Language*. Cambridge, MA: The MIT Press.

On poetry and meter:
Attridge, Derek (1995) *Poetic Rhythm: An Introduction*. Cambridge: Cambridge University Press.

Notes

1 In a few words borrowed from German or Yiddish, there can be a /ʃ/ instead of an /s/ before non-liquids, as in *shtick*.
2 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.