

Syllables: phonotactics

A little experiment

In most varieties of English there are 23 consonant phonemes:

/p, t, k, b, d, g, f, θ, s, ʃ, v, ð, ʒ, tʃ, dʒ, m, n, ŋ, l, ɹ, w, j/

English allows syllable onsets with two consonants:

tray, blood, snake, shrivel, ...

If all combinations of two consonants are possible in English (excluding geminates), there should be 506 ($23^2 - 23$) possible two-consonant onsets. Are all 506 really possible?

Try making a CCV syllable that *could* be a word in English – it may not actually exist, but it could be a name for a new product, or slang, etc. For starters, assume the vowel is *aɪ* and pick random consonants from above. What do you find?

Here are 30 combinations that I picked randomly:

sðʒaɪ	θwaɪ ☺	lhɑɪ	tʃgɑɪ	pŋɑɪ	ʃzaɪ
vjaɪ ☺	ŋʃɑɪ	vpaɪ	fgɑɪ	θtʃɑɪ	twaɪ ☺
ʃtaɪ (☺)	dʃɑɪ	bɹɑɪ ☺	klɑɪ ☺	ŋtaɪ	sðɑɪ
pnaɪ	vfaɪ	jwaɪ	hŋɑɪ	ltaɪ	jsɑɪ
θvaɪ	swaɪ ☺	vhaɪ	ktɑɪ	wɹɑɪ	wpaɪ

Out of 30, I would consider 6 to be possible English words and 1 to be borderline. The rest are most likely unpronounceable by an English speaker without phonetic training.

If we were to continue randomly picking two-consonant combinations, we would find that there are less than 40 acceptable English onset clusters.

What do the 6 acceptable clusters in the above list have in common?

vj θw sw bɹ kl tw

The first consonant is an obstruent; the second is an approximant. The following chart lists possible two-consonant onsets in English. Most are combinations of an obstruent and approximant.

English Onsets: two consonants

		[+son, +cont]				[+son, -cont]		[-son, -cont]		
		j	w	ɹ	l	m	n	p	t	k
[-son]	p	puke		prize	play					
	t	(tune)	twin	tree						
	k	cute	quit	creep	clay					
	b	beauty		breed	blink					
	d	(duty)	dwell	dream						
	g	ambiguity	Gwen	grip	glue					
	f	few		free	fly					
	θ	(enthuse)	thwart	three						
	ʃ			shrill	(schlep)	(schmuck)	(Schnapps)		(shtick)	
	s	(suit)	swell		sly	smile	sneer	spy	sty	sky

English Onsets: three consonants

	j	w	ɹ	l
sp	spew		spry	split
st	(stupid)		street	
sk	squeeze	skew	scream	sclerosis

The point of this exercise is that not just any combination of consonants is allowed in English. The set of possible sound combinations for a language is called the **phonotactics**.

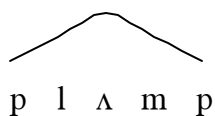
Phonotactics varies from language to language, but there are general tendencies across languages. Syllable structure:

- When a syllable onset has more than one consonant, the less sonorant sounds tend to be first, followed by more sonorant sounds.
- When a syllable coda has more than one consonant, the more sonorant sounds tend to be first, followed by less sonorant sounds.

In other words, the more sonorant sounds tend to be closer to the syllable nucleus.

obstruents < [+son] consonants < vowels

This is called the **Sonority Scale** or Sonority Sequencing Generalization.



If we exclude coda clusters formed from the suffixes *-s*, *-ed*, *-th*, we get the following coda combinations for English:

	m	n	l	p	t	k	b	d	g	f	θ	s	ʃ	tʃ	dʒ
m				ramp						lymph					
n					rant			land				sense		branch	
ŋ						rank									
l	helm			help	wilt	elk	bulb	build		elf	wealth	else		mulch	bulge
r	arm	warn	snarl	warp	wart	fork	orb	board	org	wharf	hearth	farce	harsh	arch	large
s				asp	last	ask									

What do you do with consonant clusters in the middle of a word?

fasting: [fæ.stɪŋ] [fæs.tɪŋ] or [fæst.ɪŋ] ?

Maximum Onset Principle : make syllable onsets as large as possible.

Turkish

'kalktuw s/he got up
 'kalkɑɣ s/he gets up
 'abla sister
 ak'ʃam evening
 ba'ba father

Armenian

marzpan governor
 gəzrɔts place to save stuff
 məkɔrtel to christen
 tɔrtʃnək a bird

Thai

mákn̄ã:ɪ
 sawàʔdi:
 kòpkun
 róʔtjon

Russian

bʌ'ɖrʲitʲ to invigorate
 ʦunʲ'ɖrə tundra
 ku'ɖrʲi curls
 kʌ'stʲak core, backbone

Farsi

xakestærha
 rægsi:dæn
 ræftæn
 amuzefgah
 keʃaværzi
 xosusijæt

Arabic

bustan
 bæʔtʲaman
 əstixʲɑdʒ