

Onset Clusters in English: Simplified by Rule or Simple due to Constraint?

Frans Plank, vii 06

Modern English doesn't allow onset clusters such as /gn, gm, kn, km, ps/, even though they do not offend against sonority sequencing. The question this raises is: Is there a CONSTRAINT prohibiting such onset clusters (underlyingly or superficially) or a RULE actively simplifying them, should any turn up? Given that such clusters were permissible in earlier English, and are permissible in other languages from which English is a lexical borrower, it is conceivable that such clusters indeed do turn up.

The answer to this question bears on the question what the lexical representations are of words such as *gnat*, *gmo* (=genetically modified organism), *knot*, *Khmer*, *psalm*. Are they lexically /gnæt, gməʊ, knɒt, kməər, psɑ(l)m/, and are these clusters simplified by rule to yield pronunciation forms [næt, məʊ, nɒt, (k)məər, sɑm]? Or are the pronunciation forms themselves also the lexical forms, with no rules altering their onsets? (If the constraint against these clusters is assumed to only apply superficially, then lexical forms could have them, but wouldn't be pronounced. I wonder what the difference is then to the rule-based account.)

The following considerations would seem to me to argue for a constraint- and against a rule-based account.

Words like the following could be syllabified in line with the universal principle of onset maximisation, so far as is permitted by sonority sequencing (left column) or without onsets maximised (right column):

re.co.gnize	re.cog.nize
sta.gnant	stag.nant
ma.gma	mag.ma
fi.gment	fig.ment
a.cknow.ledge	ack.now.ledge
a.cne	ac.ne

a.cme	ac.me
har.psi.chord	harp.si.chord
ter.psi.cho.re	terp.si.cho.re
e(l).li.psis	e(l).lip.sis
so.li.psi.sm	so.lip.si.sm
au.to.psy	au.top.sy
pe.psi	pep.si
ti.psy	tip.sy
gy.psy	gyp.sy

In actual fact, they are invariably syllabified like in the right column, offending against onset maximisation. This offence is judicious: onsets are not maximised iff the resulting cluster would be language-particularly illicit.

Now, if such words **were** to be syllabified as in the left column, in line with onset maximisation, one could expect them to be pronounced with these onset clusters (illicit word-initially):

[re.kə.ɡnɑɪz, mæ.ɡmə, æ.knɪ, tɜ.psi.kə.rɪ, pe.psi] etc.

But they aren't. (One might assume that this is how such words are syllabified initially, only to be obligatorily resyllabified, so as to bring syllable onsets in line with a language-particular constraint on permissible clusters.)

One could also expect such words to be pronounced with the clusters simplified, assuming there is a RULE of cluster simplification rather than only a CONSTRAINT:

[re.kə.nɑɪz, mæ.mə, æ.nɪ, tɜ.si.kə.rɪ, pe.sɪ] etc.

Again, this is not how such words are ever pronounced.

I conclude there is no RULE of onset cluster simplification active in Modern English. It follows that the lexical representations of words such as *gnat*, *gmo*, *knot*, *Khmer*, *psalm*

are /næt, məʊ, nɒt, (k)mɛər, sɑ(l)m/, differing from the lexical representations and/or pronunciations of such words in earlier English or in the languages from which English borrowed them. The right pronunciations of words such as *recognize*, *magma*, *acne*, *Terpsichore*, *pepsi* are due to a judicious violation of onset maximisation (effected by resyllabification, if one assumes initial syllabifications as in the left column above), motivated by the (language-particular) constraint against onset clusters, holding word-initially as well as word-internally.