ACADEMIA
Accelerating the world's research.

Do you believe in magic? The story of s+C sequences.

Jonathan Kaye

Related papers

Download a PDF Pack of the best related papers

Auschwitz im Bild. Zur kritischen Analyse der Auschwitz-Alben
Tal Bruttmann, Stefan Hördler, Christoph Kreutzmueller

DIE KUNST DES INTAVOLIERENS: GEBUNDENHEIT UND FREIHEIT
Dr. Johannes Ring

Verstreute Editionen in Zeitschriften
Klaus Amann

Do you believe in magic?
The story of *s+C* sequences*

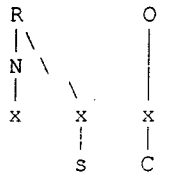
Jonathan Kaye

This file is the scanned faksimile of an unpublished version of the following reference. Its text, but not its pagination, is identical to the published version. [A scanned faksimile of the published version soon in this theatre...]
Kaye, Jonathan 1992. Do you believe in magic? The story of s+C sequences. SOAS Working Papers in Linguistics and Phonetics 2, 293-313. [Reprinted in A Festschrift for Edmund Gussmann, edited by Henryk Kardela & Bogdan Szymanek, 155-176. Lublin: Lublin University Press.]

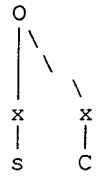
0. Introduction

In this article I will discuss true sequences of *s+C*, where "C" represents a non-nuclear phonological position along with its associated segmental material. English examples of these sequences are words like ***stop***, ***skip***, ***rescue***, ***western***, ***wasp*** and ***mask***. I will argue that such sequences are always heterosyllabic and never tautosyllabic. Concretely, *s+C* clusters will be shown to be rhymal complement-onset sequences as in (1a) and not branching onsets as in (1b).[1]

(1) a.



b.



The most controversial aspect of this claim involves word initial *s+C* sequences as in ***stop*** and ***skip***. Accordingly, I will spend most time justifying the structure (1a) for these cases.

In section 1, I will give the theoretical arguments for rejecting the tautosyllabic structure (1b) as the structure for these sequences. In the following sections I will provide evidence from Italian, (European) Portuguese, (Ancient) Greek and English in support of this claim.

1. Theoretical Considerations

This discussion is based on the principles of *Government Phonology*.[2] In this theory conditions of strict locality and strict directionality allow us to derive the *binarity theorem* presented in (2) below.[3]

(2) **The binarity theorem**
All constituents are maximally binary.

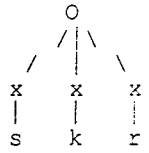
Applying (2) to English onset is not problematic except for a very specific set of cases. Putative ternary English onset occur in form like ***scream***, ***spray***, ***splatter*** and ***stress***. The superficial initial clusters of these forms contain sequences of three consonants. These examples could

prove problematic for (2) if we assume something like (3).

(3) Syllabify all initial consonant sequences up to the first vowel into the onset.

If (3) is correct then a form like ***scream*** must have an initial onset as in (4) in clear violation of the binarity theorem.

(4)



If (2) is correct then (4) could not be the structure of the initial onset of ***scream***. Indeed, the only sequences that even give the appearance of violating the binarity theorem with respect to onset are all of the form *s+C*$_1$C$_2$ where C$_1$C$_2$ is a well formed onset. If it can be shown that *s+C* sequences in general are heterosyllabic, then we will have eliminated all potential counterexamples to the binarity theorem with respect to English onset.

There is another reason to doubt that *s+C* sequences form branching onset. A branching constituent be it onset, rhyme or nucleus constitutes a head initial governing domain. Constituents have the further requirement that their governors are charmed and their governees are both neutral and simplex.[4] *s* is normally assumed to be neutral (R° · h°) and thus is not a potential constituent governor. Further, in obvious branching onset cases consisting of charmed governors, there are severe constraints on what may constitute the governed member of the constituent. *t* may only govern *r* within an onset; *b* may only govern *r* and *l*. *s*, in contrast occurs (word initially) before the set {p,t,k,m,n,l,w}.[5] It is clear that *s* is not behaving like a typical onset governor.

There are two theory-internal reasons to doubt the tautosyllabic status of *s+C* sequences: (1) they are involved in binarity theorem violations and (2) the putative constituent head does not behave in a fashion typical to this status. What evidence suggests that *s+C* sequences are branching onset? Only the assumption in (3) would cause us to analyze the sequences in this way. We then can ask, what is the evidence for (3)? Does it follow from any principle of UG? Is it empirically motivated? The answers to these questions appear to be negative. I am unaware of any evidence supporting (3). This leaves open the possibility that (3) does not represent an accurate generalisation about phonological systems. If this can be shown to be so, then all support for the tautosyllabic analysis of *s+C* clusters disappears. In contrast, there is empirical

evidence to support the claim that these clusters are heterosyllabic and that they indeed have the structure shown in (1a). Let us turn to the evidence now.

2. The Italian Evidence

The argumentation found here first appeared in KLV, 1990:204ff.[6] I will summarise the arguments rather briefly. The reader is referred to the original article for a more complete discussion.

2.1 Italian Metrical Lengthening

In Italian stressed vowels must occur in rhymes containing two positions. This requirement may be satisfied either with a branching rhyme or by a branching nucleus. Since branching nuclei are not lexically present in Italian, "open syllables" (i.e. rhymes with no post-nuclear rhymal position) acquire an additional point resulting in the lengthening of the tonic vowel. The examples in (5a) show cases of such metrical lengthening. (5b) shows failure of metrical lengthening in closed syllable cases. The branching rhyme constraint on tonic rhymes is already satisfied and the binarity theorem predicts that branching nuclei could not appear within branching rhymes.[7] The examples (5c) show that the sequences *pr*, *kr*, and *tr* are indeed analyzed as branching onset (as they are in English). As such their preceding rhyme remains non-branching and accordingly, subject to metrical lengthening.

(5) a. fá:to 'fate' b. fátto 'fact'
mé:ro 'pure' mánto 'coat'
pé:lo 'hair' párko 'park'

c. ká:pra 'goat'
sá:kre 'sacred'
ré:tre 'behind'

Let us now consider Italian *s+C* sequences. If they behaved like branching onset (structure (1b) then we would expect metrical lengthening exactly as in (5c). If, on the other hand, they are heterosyllabic (structure (1a) then we expect no metrical lengthening. The data in (6) show that *s+C* clusters behave heterosyllabically and **not** like branching onset.

(6) pásta 'pasta' véspa 'wasp' fyásko 'flask'
pésta 'trail' móska 'fly' kréspo 'pancake'

Have we then established our claim? Do the data in (6) suffice to show that (3) is false? This would be the case if we accepted some version of a *Uniformity Principle* as in (7).

(7) **The Uniformity Principle**
a. The weak version
In a given language, sequences of contiguous positions that are in a governing relation and contain the same phonological material have the same constituent structure.
b. The strong version
For all languages, sequences of contiguous positions that are in a governing relation and contain the same phonological material, have the same constituent structure.[8]

The Uniformity Principle states that if we find some sequences, say, *st* that is analyzed as having some structure, then every sequence of *st* must have the same structure. In its weak form, once we have established the constituent structure in one position in a language, we are required to maintain that structure **in all positions** for that language. The strong version states that once we have established the constituent structure in one position in a language, we are required to maintain that structure **in all positions** for all languages.

Applying this to the case at hand, even the weak version of the Uniformity Principle requires that word-initial *s+C* sequences be analyzed heterosyllabically in Italian. Concretely, the fact that *st* in *pasta* is heterosyllabic means that it is heterosyllabic in *studente*. This is so, whatever version of (7) is used. If the strong version of (7) were accepted, then this would mean that the initial cluster of English *student* is also heterosyllabic. It is a fact of life that most phonological theories do not accept either version of (7) and so having proved that **word-internal** *s+C* sequences are heterosyllabic does not imply that they have the same status word-initially for those theories. Let us now investigate the Italian evidence concerning the structure of these word-initial sequences.

2.2 The Masculine Definite Article

Italian selects the form of the masculine definite article (henceforth mda) according to phonological criteria. If the initial onset of the noun is filled (i.e. contains a position) then *il* is selected. If the initial onset is empty (i.e. contains no position) then *lo* is selected.[9]

(8) a. il costo 'the price' b. l'arco 'the arch'
il lato 'the side' l'elenco 'the list'

c. il piombo 'the lead' d. il solco 'the wake'
il treno 'the train' il sale 'the salt'

In (8a) we see nouns with filled initial onset. The form of the mda is *il*. In (8b) the nouns appear without an initial onset and as noted above the mda form is *l'* (< *lo*). The examples in (8c) show that initial branching onset behave exactly like non-branching ones. As in (8a) the form *il* is selected. Finally, (8d) shows that word-initial *s* has no special property. There is no "surface phonotactic constraint" that would prevent surface *ls* sequences from occurring. The selection of the mda appears to be a reliable indicator of word-initial constituent structure.

Let us now consider word-initial *s+C* sequences. We have clear expectations as to the form of their mda. If (3) is correct then we expect *il studente*. If some version of (7) is correct, then we expect *lo studente*. In fact (3) is incorrect. The correct form is *lo studente* and **il studente* is ungrammatical. Further examples are given in (9) below.

(9)	lo straccio	'the cloth'	lo sprezzo	'the scorn'
	lo scuro	'the darkness'	lo slancio	'the élan'

2.3 Raddoppiamento Sintattico

A second argument involving the process of raddoppiamento sintattico (RS) is discussed in KLV, 1990.[10] Word-initial onset consonants will geminate when following a final tonic vowel. In (10a) the onset is non-branching and geminates under the stated conditions. (10b) shows cases of branching onset. It is important to note that the initial member of a branching onset geminates in exactly the same way as a non-branching onset does.

(10) a.	paltó pulíto	[paltóppulíto]	'clean coat'
	é caríno	[ékkaríno]	'it is pretty'
b.	čittà tríste	[čittáttríste]	'sad city'
	caffè fréddo	[kaféffréddo]	'cold coffee'

This brings us to the *s+C* cases. If *s+C* were an onset, we would predict that it should geminate exactly as in (10b). Thus, if, say, *spésso* had the same constituent structure as *fréddo*, we would expect the same behaviour under RS conditions. The initial *s* as the first member of a putative branching onset should geminate exactly as the initial *f* of *fréddo* does. In fact, the RS behaviour of *s+C* sequences differs markedly from that of true branching onset. The relevant data appear in (11) below.

(11)	città straniéra [čittástraniéra]	*[čittásstraniéra]	'foreign city'
	caffè spésso [kaféspésso]	*[kafésspésso]	'thick coffee'

Let us summarise. The Italian data shows quite clearly that (3)

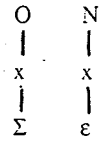
cannot be maintained. It is simply false to claim that all initial consonant sequences up to the first vowel must be syllabified into the onset. In Italian the initial consonant sequence up to the first vowel of *studente* is *st* and now we know that *st* is not an onset. Where does that leave us with respect to English. It could be claimed that only some languages (like Italian) violate (3) while others (say, English) respect it. But how are we to know. We have at least one language where (3) is false. Therefore, it could be false in English. One cannot analyze English *s+C* sequences as onset simply on the basis of (3). Independent evidence must be found.

I now turn to Ancient Greek for another language where word-initial *s+C* sequences can be shown to be non-onset.

3. The Ancient Greek Evidence[11]

Ancient Greek (unlike Modern Greek) formed the perfect tenses by a process of reduplication. In essence, reduplicative prefix took the following form:

(12)



where Σ is a variable unified with the initial onset of the verb stem. ε is a constant, i.e. a lexical property of the reduplicative prefix.

(13)	Verb	Perfect Active	Gloss	
a.	λύω	λέ-λυκα	'loose'	le-l
	παιδεύω	πε-παίδευμαι	'educate'	pe-p
b.	γράφω	γέ-γραφα	'write'	ge-gr
	βλάπτω	βέ-βλαφα	'harm'	be-bl
c.	στρατεύω	ἐ-στράτευκα	'take the field'	e-st

The Ancient Greek data are shown in (13) above. In (13a) non-branching onset are given. The prefix initial onset is a copy of the initial onset of the stem. (13b) gives the branching onset cases. Note that only the head (first member) of the branching onset appears as the reduplicative initial onset. We now turn to the initial *s+C* case. If Ancient Greek initial *s+C* *sequences* behaved like branching onset, then we would expect the

behaviour of a stem like (13c) to behave exactly like those of (13b). Concretely, given the stem στρατεύω, an onset analysis predicts the Perfect Active form to be σε-στράτευκα, just as βλάπτω yields βέ-βλαφα.

On the other hand, if the heterosyllabic analysis of Ancient Greek *s+C* sequences is correct then the structure must be that of (1a). But (1a) has an empty initial onset. If this is indeed the correct structure then the reduplicative prefix must also have an initial empty onset. This yields the prediction that the form of the reduplicative prefix of the stem στρατεύω should be ἐ-, that is, the vowel ε preceded by the first onset of the verb stem which is ∅. A glance at (13c) shows that this prediction is correct.[12]

In sum, the analysis of Ancient Greek reduplication confirms our Italian results, to wit, *s+C* sequences are not branching onset but rather have the structure of (1a). I now turn to further evidence for this conclusion. Let us consider European Portuguese.

4. The European Portuguese Evidence[13]

Vowel nasalisation is a well-known phenomenon in Portuguese. Stated informally, a nasal consonant may not close a syllable in Portuguese. In cases where this could occur, a nasal vowel appears. Consider the following data:

(14)

a.	admissível	'admissible'	(in]admissível	'inadmissible'
	aplicável	'applicable'	[in]aplicável	'inapplicable'
	elegível	'eligible'	[in]elegível	'ineligible'
	obediente	'obedient'	[in]obediente	'disobedient'
b.	pureza	'purity'	[ĩ]pureza	'impurity'
	capaz	'capable'	[ĩ]capaz	'incapable'
	decente	'decent'	[ĩ]decente	'indecent'
	fecundo	'fertile'	[ĩ]fecundo	'infertile'
c.	frangível	'breakable'	[ĩ]frangível	'unbreakable'
	grato	'grateful'	[ĩ]grato	'ungrateful'
	criticável	'criticizable'	[ĩ]criticável	'uncriticizable'
	tratável	'sociable'	[ĩ]tratável	'unsociable'
d.	sonhável	'conceivable'	[ĩ]sonhável	'inconceivable'
	satisfeito	'satisfied'	[ĩ]satisfeito	'dissatisfied'
e.	[šk]apável	'escapable'	in[šk]]apável	'inescapable'
	[šp]erado	'expected'	in[šp]erado	'unexpected'
	[št]imável	'estimable'	in[št]imável	'inestimable'
	[šk]rupuloso	'scrupulous'	in[šk]rupuloso	'unscrupulous'
			*[ĩšk]apável	
			*[ĩšp]erado	
			*[ĩšt]imável	
			*[ĩšk]rupuloso	

In (14a) we see a series of stems beginning with an empty onset. When the negative prefix *in-* is added, notice that no nasalisation takes place. Stems with filled onset are given in (14b). In these cases the negative prefix is realised with a nasal vowel, [ĩ]. This establishes the pattern, [in] when followed by an empty onset; [ĩ] when followed by a filled onset. (14c) shows that branching onsets behave exactly like non-branching onsets - the prefix is realised as [ĩ]. (14d) demonstrates that sibilant-initial stems do not manifest any abnormality. This brings us to the *s+C* cases. If this sequence were an onset we would expect behaviour similar to that of (14c). A stem like [šp]erado would yield *[ĩšp]erado, just as tratável yields [ĩ]tratável. If *s+C* cases have the structure of (1a) then [šp]erado should yield in[šp]erado just as aplicável yields [in]aplicável. The data in (14e) show this to be the case. European Portuguese behaves exactly like Italian and Ancient Greek.[14]

Once again we have seen that apparent word-initial *s+C* sequences

do not pattern with word-initial filled onset. Rather they display behaviour similar to that of forms with word-initial empty onset. European Portuguese augments the earlier cases from Italian and Ancient Greek which indicate that *s+C* sequences, even word-initially, are not onset. Our final example comes from English. It involves the behaviour of postconsonantal *y* (the palatal glide) in some dialects of British English.

5. The British English Evidence

English dialects vary with respect to the occurrence of postconsonantal *y* after coronals. Some American English dialects such as New York City, lose postconsonantal *y* after all coronals tonically and pretonically. In these contexts *y* is conserved in many Southern British English dialects as well as some Canadian and Australian English dialects. The following table gives an indication of the occurrence of *y* in a variety of forms.[15]

Word	New York City	London
few	+	+
cute	+	+
mute	+	+
beauty	+	+
putrid	+	+
huge	+[16]	+
tune	-	+
dew	-	+
news	-	+
suit	-	+
lurid	-	+[17]
rude	-	-

If we restrict ourselves in English dialects that display *ly* as in *l[y]urid* we note that postconsonantal *y* **never** occurs after a branching onset. The sequences *ply, bly, fly, kly, gly* simply do not occur.[18] The following data illustrate this point:

(15) a.	l[y]ute	b.	pluvial	*pl[y]uvial
	all[y]uvial		plumage	*pl[y]umage
	l[y]urid		fluid	*fl[y]uid
	all[y]ure		occlusion	*occl[y]usion
			glue	*gl[y]ue[19]
			blew	*bl[y]ue

In (15a) we see that *l* in a non-branching onset can be followed by *y*. In (15b) possible sources for *y* following a branching onset are given. These forms are deemed impossible by speakers who otherwise produce *ly* sequences.

We are now in a position to test the status of initial *s+C* sequences in English, at least for *ly* speakers. We have seen that postconsonantal *y* is not found following a branching onset. If *s+C* sequences are branching onsets, then we expect sequences of the type *s+Cy* to be impossible. On the other hand, if English aligns itself with Italian, Portuguese and Greek with respect to these sequences, then *s+Cy* should be well formed in English. In fact these sequences are well-formed as the data in (16) show.

(16) st[y]upid sp[y]ew sk[y]ewer

If our observation, to with, English has no *y* following branching onset is correct, then it must be the case that *st, sp* and *sk* are not branching onset. One could argue, however that the absence of forms like *gl[y]ucose and the presence of forms like *st[y]upid* does not reflect differences in constituent structure. Rather it might be claimed that this effect is due to the so-called "sonority hierarchy". If one were bent on maintaining the claim that *s+C* sequences are tautosyllabic in English, one could formulate the distributional constraint discussed above in the following way:

(17) Postconsonantal *y* cannot occur in sequences with rising sonority curves.

If it can be shown that (17) is correct then the presence of *sty* as in *st[y]upid* and the absence of *ply* as in **pl[y]uvial* could not be taken as evidence for the non-onset status of *st*. One could claim that *t* is less sonorous than *s* and therefore *st* has a falling, not a rising sonority curve. On the other hand, *l* is more sonorous than *p* giving it a rising curve. Under this interpretation (17) would exclude *ply* but admit *sty*. No conclusions could be drawn about onset status for these sequences.

In fact there is a crucial test that can decide whether it is onsethood or the sonority hierarchy that underlies these distributional

facts. Consider the sequence *sl*. This sequence must be analyzed, as indeed are all *s+C* sequences as the heterosyllabic (1a). If the distributional facts are based on onsethood then we expect that *sly* should be possible since *sl* is not an onset. On the other hand if (17) is correct then *sly* should be impossible. *s* is less sonorous than *l* so *sl* has a rising sonority curve. Postconsonantal *y* would be excluded by (17) in such a case.

sly, though rare is possible in English. English *ly* speakers pronounce the verb *slew*, a nautical term meaning to turn around a fixed point, as *sl[y]ew*.[20] (17) turns out to be false. One could still try to salvage the situation by modifying (17) into something like

(18) Postconsonantal *y* cannot occur in sequences with rising sonority curves, unless the initial segment is *s*.

It hardly seems worth taking (18) seriously. Its only function would be to escape the conclusions that the distribution of postconsonantal *y* is based on constituent structure and that *s+C* sequences have the structure (1a) and not (1b). In addition the *OED* lists alternative pronunciations for less widely used words in *pneum-* such as *pneumatology*, *pneumatosis* and *pneumectomy*. Both "pniu..." and "niu" are given. If speakers still use the former pronunciation then (18) would have to be modified still further. On the sonority scale, *n* is more sonorous than *p*.

It seems safe to conclude that the behaviour of postconsonantal *y* is unrelated to the sonority hierarchy but rather strictly a function of constituent structure. It may freely occur following a transconstituent sequence of a rhymal consonant followed by an onset. It cannot occur after a branching onset. We conclude then that *s+C* sequences in English are not onsets, even in word-initial position. Given the convergence of evidence from Italian, Greek, Portuguese and English, the most prudent assumption would be that *s+C* sequences are not analyzable as onset in any language. The empirical record supports the theoretical claim of Government Phonology. For this reason one would have to look at theories of constituent structure which allow for *s+C* sequences in an onset position with some degree of scepticism.

5.1 Psycholinguistic Evidence concerning *s+C* sequences

Treiman, Gross and Glavin, 1991 (TGG) have conducted a series of psycholinguistic experiments concerning the status of *s+C* sequences. First, they have compared these sequences with true branching onset. Secondly, they have sought to compare *s+C* sequences with a "falling sonority curve" (such as *st*) with those exhibiting a rising curve (such as *sl*). While their experiments do not treat word-initial *s+C* sequences, they do related to our previous discussion of the "sonority hierarchy" and

support the negative conclusions with respect to its relevance to the issues at hand. I briefly summarise the experiments and the conclusions that TGG draw from them.

Experiment 1 was a written syllabification task. It involved three types of clusters: type 1 was an *s+C* cluster with a falling sonority curve (e.g. *st*), type 2 was an *s+C* cluster with a rising sonority curve (e.g. *sl*) and type 3 was a true branching onset (e.g. *fl*). TGG conclude that "The present results suggest that all /s/ clusters are unusual with respect to syllabification. This is true whether their sonority profile is normal, as with /sw/ or /sl/, or unusual, as with /st/. These results support the claim of Kaye et al. [KLV, 1990/JK] that cluster onset may not begin with /s/. Because onset like /sp/ and /sw/ are illegal, people usually place the two phonemes in separate syllable when these sequences occur in the middles of words." (p. 11).

Experiment 2 was similar to experiment 1 except that it was an oral task rather than a written one. The conclusions of experiment 2 were quite similar to those of experiment one. In general, the subjects treated all *s+C* sequences, whether of type 1 or type 2, heterosyllabically.

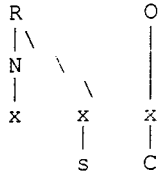
In experiment 3 subjects were asked to read nonsense words identical to those used in experiment 1. TGG used forms some of which could be construed as having a prefix (such as be, co, etc.). The idea was to see if the initial vowel of these nonsense forms was produced as a reduced vowel or not in cases where stress fell on a later syllable. In a word like *wospeem* if the first syllable is open we expect vowel reduction to schwa; if it is closed the vowel might remain unreduced. TGG used words like *wospeem* (type 1), *wosweem* (type 2) and *wofleem* (type 3) to test vowel reduction behaviour. They concluded, "When the biasing worked and the first vowel of the nonword was unstressed, the vowel was reduced a small majority of the time for non-s clusters. The vowel reduction implies that the first consonant of the cluster belongs to the second syllable, making the first syllable weak and subjecting it to reduction. For s clusters, in contrast , unstressed first vowels were usually full. This result implies that the s is syllabified with the first syllable, making the syllable strong and protecting it from reduction." (page 19).

Thus, the three experiments performed by TGG lead to the same conclusion: word-internal *s+C* sequences are **not** syllabified like onset. They conclude, "Thus, our results do not support the usual interpretations of either the maximum onset principle or the sonority hierarchy principle." (page 20). Neither exist within the framework of Government Phonology and it seems reasonable to conclude that the theoretical claims of this framework are borne out by the empirical record - both by linguistic and psycholinguistic evidence.

6. Theoretical Excursus

In the preceding sections I have presented evidence from several languages as well as from a number of psycholinguistic experiments to show that *s+C* sequences are not onset. Specifically, given the assumptions of constituent structure adopted in this paper, these sequences when found in word-initial position must have the structure of (1a) repeated as (19) below.

(19)



Notice that *s* occurs in a rhyme whose nuclear head is empty. Furthermore the nucleus receives no phonetic interpretation. Government phonology provides very precise conditions under which an empty nucleus is licensed. It is now our task to see if the empty nucleus of (19) falls under one of these conditions. These are presented in (20).

(20) The Phonological ECP: *A p-licensed (empty) category receives no phonetic interpretation.*

P-licensing: 1. Domain-final (empty) categories are p-licensed (parameter: **true** German Polish Arabic, **false** Italian Japanese Vata

2. Properly governed (empty) nuclei are p-licensed.

Proper government:[21]

α properly governs ß if

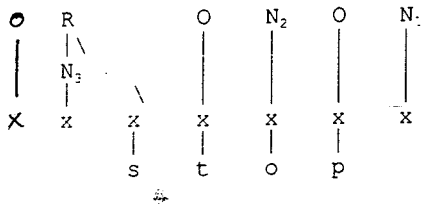
1. α and ß are adjacent on the relevant projection,
2. α is **not itself** licensed, and
3. No governing domain separates α from ß.

The phonological ECP states the conditions under which an empty nucleus receives no phonetic interpretation. Two types of P-licensing are then listed. Type 1 refers to domain-final empty nuclei. Languages like English, French and Arabic P-license all domain-final empty nuclei. Such is not the case in Japanese or Vata.[22]

The second case of P-licensing involves proper government. A properly governed empty nucleus is P-licensed and following the ECP receives no phonetic interpretation. Let us now apply the ECP to a word-

initial *s+C* sequence using the constituent structure given in (19) above.

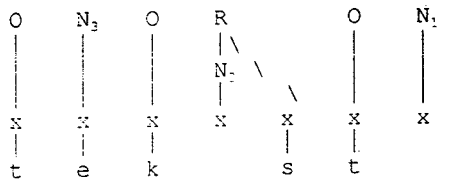
(21)



(21) above represents the constituent structure of *stop*. There are two empty nuclei in this form neither of which receive phonetic interpretation. According to the ECP both should be P-licensed. N_1 is domain-final and English P-licenses all domain-final empty nuclei. It should therefore be inaudible, which it is. What about N_3? N_3 is not domain-final. Could N_3 be properly governed? Three conditions must be satisfied in order for this to take place. The proper governor and governee must be adjacent on the relevant projection. Assuming that this projection is the nuclear projection then N_3 and N_2 would be adjacent. The second condition requires that the putative proper governor not be licensed itself. This is the case in (21). N_2 is not licensed. It is a potential proper governor for N_3. The final condition states that no governing domain intervene between governor and governee. This condition is not satisfied in (21). A transconstituent governing domain, to wit *st*, lies between N_3 and N_2. Therefore, N_3 cannot be P-licensed by proper government.

Softening the restriction on proper government across a governing domain is not a viable solution to this problem. Besides an unwelcome weakening of the theory, it runs afoul of forms like *text*. This word must have the structure in (22) below.

(22)



As in the word-initial *s+C* cases, a P-licensed empty nucleus precedes the cluster in this example. A weakened Proper Government definition would allow N_3 to govern across the transconstituent governing domain *st*. However N_1 is itself licensed, being in domain-final position. If N_2 is P-licensed by proper government then N_1 must be the proper governor. There are no reported cases of licensed domain-final nuclei serving as proper governors. The most prudent move, under the circumstances would be to maintain the stricture against proper government across an

intervening governing domain and to reject the idea that proper government P-licenses the internal empty nuclei in *Østop0* and *tekØstØ*.

It is not obvious how these internal empty nuclei are licensed. For the moment we can only observe that a *s+C* sequence serves as a P-licenser in some languages. I will call this type of P-licensing "Magic Licensing" as a constant reminder that it is a pure stipulation in need of an explanation. At the moment the theory offers no reason why *s+C* sequences should have this "magical" property. This gives us three contexts in which P-licensing takes place.

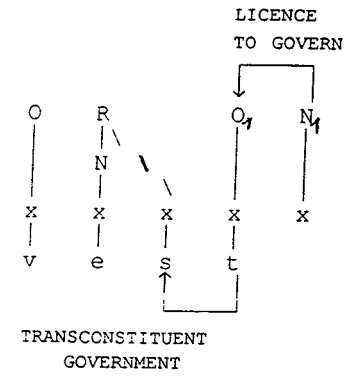
(23) P-Licensing Contexts
- i. Domain-finally (Parameterised)
- ii. Proper Government
- iii. Magic Licensing (Parameterised)

In (23) I have indicated that Magic Licensing is parameterised. This is obvious from a comparison of European Portuguese with Brazilian Portuguese, or Italian with Spanish. The former members of each pair display magic licensing one of whose manifestations is initial *s+C* sequences. The latter member of each pair does not display magic licensing. A prothetic vowel occurs before what are initial *s+C* sequences in the magic licensing cases.

Aside from P-licensing empty nuclei, *s+C* sequences have another magical property. In section 1, I noted that within C_1C_2 sequences, those beginning with *s* permit much more freedom concerning the segment that may occupy the C_2 position. Given their transconstituent nature, we expect that C_2 should govern C_1 in such cases. Following the assumption that *s* has neutral charm, C_2 requires either a charmed segment or a neutral segment greater in complexity (i.e. containing more elements) than *s*. These conditions are satisfied in the case of *sp*, *st*, *sk*. Let us call them "natural (transconstituent) sequences". Consider now *sw*, *sl*, *sn* and *sm*. The set {w,l,n,m} are all neutral segments and all are equal or less than *s* in complexity. Consider *sw* as in *sweep*, *swoon*, etc. As we have mentioned above, *s* is $(R^\circ \cdot h^\circ)^\circ$ and *w* is $(U^\circ)^\circ$. In other words, *w* is even less complex than *s* - hardly a fitting characteristic for a neutral governor. Let us call these sequences "unnatural (transconstituent) sequences". It appears then that not only does *s* have the property of combining with its "governor" to license a preceding empty nucleus, it also has the power to confer on the following onset, the ability to govern it (the *s*). This is true in the case of the unnatural sequences.

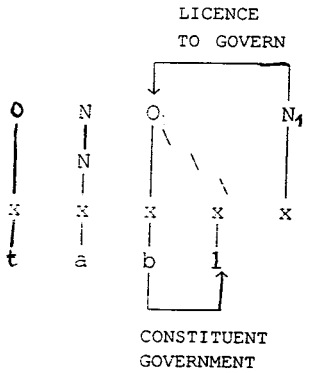
Comparison of Polish and French, on the one hand, with English and Dutch, on the other, give us an indication of how *s* acquires this power. All four of these languages license final empty nuclei. The empty nucleus in all four languages is a "government-licensor"[23], i.e. it licenses its preceding onset to govern another position. Consider a form like *vest* which is well-formed in all for languages.

(24)



In (24) we see that the onset O_1 governs the preceding rhymal consonant. It is licensed to govern by the following final empty nucleus, N_1. Compare (24) with a structure whose final onset is branching as in (25) below.

(25)

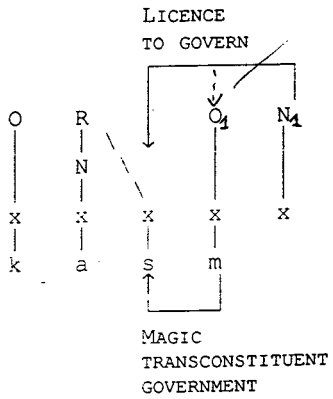


With (25) we encounter a parametric division between English and Dutch vs. Polish and French. (25) is ill formed in the former two languages and well formed in the latter two. Comparing (24) with (25), we note that in the former case the government licensor, N_1 is strictly adjacent to the licensee, the head of O_1. This is not the case in (25) where N_1 must government license across the governed position in the branching onset. We can now say that government licensing must be strictly local (in the usual sense of the term) in English and Dutch but need only satisfy some

locality requirement (say, at the head projection) in French and Polish.

The point of this discussion is that the unnatural *s+C* sequences pattern like branching onsets. We have seen that natural sequences such as *st* are well formed in languages requiring strictly local government licensing such as Dutch and English. *sm*, an unnatural sequence, is excluded in these languages before final empty nuclei. It is permitted in Polish and French. So [plasm] is possible in Polish and French but excluded in English and Dutch.[24] This leads us to believe that the mysterious property of conferring governing powers to a following onset, **independently of its segmental makeup**, comes ultimately from the following nucleus as shown in (26) below.

(26)



Suffice it to say that the magical properties of *s* remain an interesting topic for further research.

In this paper I hope to have shown that *s+C* cannot be onset. This conclusion follows from both theoretical considerations, in particular the binarity theorem , and finds ample support in the empirical record drawn from a number of different languages. It is no longer possible to maintain that any word-initial sequence of consonants is to be placed in the onset position. While not all the properties of *s* are completely understood, it is to be hoped that this segment will soon lose its "magical" nature and join the ranks of those phenomena that are better understood.

Notes

*. This is indeed a "working paper." I apologise in advance for any typos, spelling errors and other silly things that have not yet been weeded out of the manuscript. This notwithstanding, thanks to Monik Charette for pointing out some of the more salient bloopers. She is in no

way responsible for any that remain.

1. Note the presence of the onset position *O* in this structure. Its presence is required by the condition that states that there must be an onset position for every rhyme and *vice versa*. Its presence will be important in the discussion of the Greek data below.

2. See Kaye, Lowenstamm and Vergnaud, 1985, 1990, Charette, 1991, Kaye, 1990a, 1990b as well as other references contained therein for a detailed exposition of this theory.

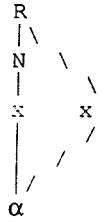
3. See KLV, 1990 for discussion.

4. Branching onset whose governed member is "l" do not conform to the simplex requirement. "l" contains two elements as shown in the expression (R° · ?°). It is interesting to note that *Cl* onset display instability in a number of languages (e.g. Italian) while *Cr* onset do not.

5. We can add "r" to the list if we allow for the fact that "s" is realised as "sh" in such cases: e.g. shrink, shriek, etc.

6. Similar arguments for Italian are found in Davis, 1990. Davis, however, makes no claim about the universal impossibility of *s+C* sequences as onset. He also makes the claim that *pn* is an onset in Italian - a claim which I reject. He resorts to modifying the "sonority hierarchy" in his analysis. It will be argued below that the sonority hierarchy (whose very existence is problematical) is unrelated to the question at hand.

7. In fact that structure of such metrically lengthened rhymes is really



Where α represents some nuclear phonological expression. This hybrid structure has been called the "Johnsen vowel" after Lars Johnsen of the University of Trondheim, who first proposed this structures. Exploring the implications of his proposal would take us beyond the scope of this paper. The existence or non-existence of the Johnsen vowel is irrelevant to the discussion here.

8. Technically speaking, this should be called "The Uniformity Theorem" rather than "The Uniformity Principle". This strong version can be derived from the fact that both phonological expressions and constituent structure receive the same interpretation across linguistic systems. Strict directionality states that headedness is not a parameter. Finally, charmed expressions cannot be governed and only charmless ones can be governed. Uniformity follows as a consequence of these theoretical primitives.

9. As we shall see, the "o" of "lo" is truncated when immediately followed by a nucleus.

10. See that article for the relevant references concerning RS data and analysis.

11. Ancient Greek Source: Jannaris, 1968.

12. A number of authors have noted that apparently aberrant behaviour of *s+C* sequences and suggested that the "s" be treated as an appendix or assigned an "extrasyllabic" status. These proposals run afoul of the Ancient Greek facts. If the initial "s" of στρατεύω were an appendix or extrasyllabic, then the first onset of this form would contain τρ. We would expect the Perfect Active form to be *τὲ-στράτευκα just as βλάπτω yields βέ-βλαφα. In fact this form is incorrect. This demonstrates that the σ is neither an appendix nor extrasyllabic. It is in the rhymal position of the first Onset-Rhyme pair. The stem initial onset is empty.

13. This analysis is due to work on European Portuguese done by Augusta Cavaco M. (Cavaco, in preparation).

14. Brazilian Portuguese manifests an audible vowel before *s+C* clusters. [šp]erado is realised as i[šp]erado (or i[sp]erado in some Brazilian dialects. In this sense, European Portuguese behaves like (standard) Italian and Brazilian Portuguese behaves like Spanish.

15. A + in a cell indicates that the yod is pronounced; a - indicates that it is not pronounced.

16. In New York City English some speakers do not pronounce the initial "h".

17. Southern British speakers differ in their treatment of "y" following "l". Many speakers do not produce "y" in this context but a number still do.

18. There is no point in discussing branching onset in "r" since yod does not follow "r" even in non-branching onset, cf. "rude" *r[y]ude.

19. In general these data correspond to those listed in the OED (Compact Edition, 1972) which gives alternant pronunciations involving [glyu...] ~ [glu...] in words like "glue" and "glucose". It gives [...glyu...] as the only pronunciation of the word "agglutinate". Wells, 1990 gives only [...glu...] for the words cited above. He does give many words where [...ly...] is at least one pronunciation.

British English speakers who pronounce "ly" in "lurid" react negatively to "gly..." forms. "I wouldn't say that in a million years" was a typical reaction to "gl[y]ue".

Roberston & Cassidy, 1954:398 state:

> Normally *l* and *r* are followed by the simple vowel, in England as well as in America; *lucid, lure, grew, true,* and *rule* are typical instances. This is probably more generally true, however, in respect to British pronunciation, for *r* than for *l*. ...there is no doubt about the [u] when *l* is preceded by another consonant: *blue, clue, flute,* and *glue* are examples.

20. Interestingly, the past tense of the verb to slay is pronounced *slew* and not **sl[y]ew*. Speakers distinguish the two "slews" in

(i) Saint George *slew* (**sl[y]ew*) the dragon.

(ii) The yacht will have to *sl[y]ew* (**slew*) sharply to win the race.

Three southern British English speakers immediately accepted [slyu] as the only pronunciation of this verb. One of the three speakers also used [slyuθ] for *sleuth*; the other two rendered this word as [sluθ]. I have not been able to find other English *sly...* words.

21. This version of proper government differs slightly from that found in Charette, 1991:83. The difference in formulation has no bearing on the discussion at hand.

22. Domain-final licensing is discussed and motivated in Kaye, 1990b.

23. See Charette, 1990 for details of government licensing.

24. In English word-internal unnatural sequences are exceedingly rare. They appear to be limited to word-initial position. Word-internal *s* spellings are generally pronounced [z] in unnatural sequences, e.g. *plasma* [zm], *grisly* [zl]. This may be a crucial clue to the solution of the magical *s* mystery.