OCP Effects in Catalan Cliticization*

Eulàlia Bonet
Filologia Catalana
Edifici B
Universitat Autònoma de Barcelona
08193 Bellaterra
eulalia.bonet@uab.es

Maria-Rosa Lloret
Filologia Catalana
Gran Via de les Corts Catalanes, 585
Universitat de Barcelona
08071 Barcelona
lloret@lincat.ub.es

Abstract
In Catalan, sequences of sibilants are never pronounced as such. In most contexts all varieties coincide in the "strategy" used to avoid such sequences, namely epenthesis or deletion. Variation is only found in the domain of pronominal clitics (but not with other types of clitics). One source of variation is accounted for by decomposing a general constraint into two specific ones, which implies partial constraint reranking. The other source of variation, which involves a case of apparent opacity, is explained through an Output-Output constraint that makes reference to paradigmatic relations.

Key words: OCP, clitics, epenthesis, deletion, fusion, opacity, Optimality Theory

Resum
En català les seqüències de sibilants no es pronuncien mai com a tals. En la majoria de contextos totes les varietats coincideixen pel que fa a les "estratègies" que utilitzen per evitar aquestes seqüències, que poden ser l'epèntesi i l'elisió. La variació només es troba amb els clítics pronominals (però no amb altres menes de clítics). Una de les fonts de variació s'explica mitjançant la divisió d'una restricció general en dues restriccions més específiques, la qual cosa implica una reordenació parcial de restriccions. L'altra font de variació, que inclou un cas d'opacitat aparent, s'explica mitjançant una restricció del tipus output-output que fa referència a relacions paradigmàtiques.

Paraules clau: OCP, clítics, epèntesi, elisió, fusió, opacitat, teoria de l'optimitat

1. Introduction

In Catalan, sequences of sibilants are never pronounced as such (in normal speech). In the cases where two adjacent sibilants would potentially occur, two ways of avoding this adjacency can be found: (a) reduction of the two sibilants to one, and (b) insertion of an epenthetic vowel between the two sibilants.[1]

All varieties of Central Catalan coincide in inserting an epenthetic schwa between two adjacent sibilants when they belong to the same word and the second *s* is a suffix, as is the case in *cuses* /kuz+z/, pronounced [ku@.z´s] '(you) sew'. In the other contexts where all varieties of Central Catalan coincide, only reduction of the two sibilants to one (fusion or deletion) is found. Reduction is found, for instance, when a definite determiner in the plural, *els* (masc.) or *les* (fem.) (both being clitics), is adjacent to a word starting with a sibilant (e.g., *els sons* [´l.sç@ns] 'the sounds' or *les sopes* [l´.so@.p´s] 'the soups') or when the sibilants belong to different words (e.g., *portes sacs* [pç$r.t´.sa@ks] '(you) bring sacks', or *pis senzill* [p"›.s´n.z"€¥] 'simple apartment'). Within words, deletion/fusion of sibilants is found in compounds (e.g., *dos-cents* [do.se@ns] 'two hundred') and in prefixed words (e.g., *dessalar* [d´.s´.la@] 'to desalt', with the prefix *des* and the verb *salar*).[2]

Within Central Catalan (more concretely, in the Barcelona area), variation with respect to the realization of adjacent sibilants is only found in pronominal clitic - verb sequences.[3] One of the varieties, let us call it variety A, systematically inserts a schwa between a pronominal clitic and the verb whenever the clitic ends in a sibilant and the verb starts with a sibilant; this schwa is not found when the verb starts with other consonants. The pronominal clitics that appear in (1) and (2) are all the clitics that end in *s* in Catalan**:** *es* /s/ 'third person reflexive/impersonal clitic', *ens* /nz/ 'first person plural', *us* /uz/ 'second person plural', *els* /l+z/ 'third person accusative masculine plural', and *les* /l+a+z/ 'third person accusative feminine plural'. In (1) these clitics appear without a final schwa because the verb starts with a non-sibilant consonant. In (2) the same clitics appear with a final schwa because the verb starts

* ACKNOWLEDGEMENTS.

[1] In Majorcan and Minorcan Catalan, a sequence of sibilants is, in most contexts, pronounced as an affricate (a sequence like *dos sons* is pronounced [do.tÉsç@ns]). In this paper we disregard these data and concentrate on different varieties belonging to Central Catalan.

[2] Even though we refer simply to sequences of sibilants, all the cases we contemplate in this paper are sequences of voiced or voiceless anterior coronal sibilants. We ignore sequences of sibilants involving a different place of articulation (/S/ or /Z/) because they are fairly rare and because their behavior implies complications that escape the scope of this paper.

[3] It is impossible to know what the facts would be with enclisis: the only clitic starting with *s* is the third person reflexive/impersonal clitic *es* (/s/), and this clitic can never cooccur with a verbal form ending in a sibilant (only the imperative second person forms *vés* 'go!' and *fes* 'do!' end in a sibilant, and they cannot combine with the clitic /s/).

with a sibilant; the presence of this schwa has its reflex in the orthography only in the case of the third person reflexive/impersonal clitic *es* (i.e., *se*), illustrated in (2a). In addition, some clitic forms show initial epenthetic schwas for syllabification purposes; for instance, the clitic *els*, with an underlying form /l+z/, needs an initial epenthetic vowel in proclitic position because the sequence [lz] does not constitute a possible onset. (From now on, we underline all occurences of epenthetic schwas for expository reasons.)

(1)a. Es trenca. [́s.tRE@N.k ́]
itself breaks
'It breaks.'
b. Ens parla. [́ns.pa@r.l ́]
to-us talks
'(S/he) talks to us.'
c. Us creu. [us.kRE@w]
you (pl.) believes
'(S/he) believes you all.'
d. Els porta. [́ls.pç@r.t ́]
them (masc.) brings
'(S/he) brings them.'
e. Les compra. [l ́s.ko@m.pR ́]
them (acc. fem.) buys
'(S/he) buys them.'

(2)a. Se sap. [s ́.sa@p]
Impers. knows
'It is known.'
b. Ens sent. [́n.z ́.se@n]
us hears
'(S/he) hears us.'
c. Us citarà. [u.z ́.si.t ́.Ra@]
you (pl.) will-quote
'(S/he) will quote you all.'
d. Els sé. [́l.z ́.se@]
them (acc. masc.) know
'(I) know them.'
e. Les supera. [l ́.z ́.su.pe@.R ́]
them (acc. fem.) overcomes

'(S/he) overcomes them.'

The clitics that end in a consonant other than *s* do not cause the appearence of a schwa after the clitic even when the verb does start with a sibilant**.** If vowel insertion is required for syllabification purposes, a schwa appears, as mentioned, in initial position (cf. *em sent* /m#sent/, [ˊm.se@n] '(s/he) hears me'; *el simula* /l#simul+ˊ/, [ˊl.si.mu@.lˊ] '(s/he) simulates it (masc.)'). The epenthetic schwa between the pronominal clitic and the verb appears only to break the contact between two sibilants.[4]

A different variety, variety B, also belonging to Central Catalan, inserts a schwa after a clitic only when the first sibilant belongs to the third person reflexive/impersonal clitic (shown in (3)). With all other clitics ending in a sibilant, fusion/deletion is found when the verb starts with a sibilant (as shown in (4)).

(3) Se sap. [sˊ.sa@p]
Impers. knows
'It is known.'

(4)a. Ens sent. [ˊn.se@n]
us hears
'(S/he) hears us.'
b. Us citarà. [u.si.tˊ.Ra@]
you (pl.) will-quote
'(S/he) will quote you all.'
c. Els sé. [ˊl.se@]
them (acc. masc.) know
'(I) know them.'
d. Les supera. [lˊ.su.pe@.Rˊ]
them (acc. fem.) overcomes
'(S/he) overcomes them.'

[4] It is interesting to recall that, in variety A, the behavior of the third person accusative plural pronominal clitic is very different from the otherwise identical definite determiner. Before a host starting with a segment other than a non-sibilant consonant they both surface, in their masculine forms, with an initial schwa, inserted for syllabification purposes (cf. *els portes* '(you) bring them': [ˊls.pç@r.tˊs], from an underlying form /l+z#pçRt+ˊ+z/, and *els porcs* 'the pigs': [ˊls.pç@rks], from an underlying form /l+z#pçRk+z/). Before a sibilant, the behavior of both clitics differs. In the case of the pronominal clitic, as illustrated in (2d,e), epenthesis takes place (cf. *els sé* '(I) know them': [ˊl.zˊ.se@], from an underlying form /l+z#se/), while in the case of the definite determiner, the "strategy" used to avoid the contact of sibilants is deletion/fusion (cf. *els sons* 'the sounds': [ˊl.sç@ns], from an underlying form /l+z#sçn+z/.

Finally, what we can call variety C systematically presents fusion/deletion when a clitic ending in *s* is adjacent to a verb starting with an *s*. This is shown in (5).

(5)a. Se sap. [ˍ.sa@p]
Impers. knows
'It is known.'
b. Ens sent. [ˍn.se@n]
us hears
'(S/he) hears us.'
c. Us citarà. [u.si.t´.Ra@]
you (pl.) will-quote
'(S/he) will quote you all.'
d. Els sé. [ˍl.se@]
them (acc. masc.) know
'(I) know them.'
e. Les supera. [l´.su.pe@.R´]
them (acc. fem.) overcomes
'(S/he) overcomes them.'

The table in (6) summarizes all the facts concerning the realization of underlying sequences of sibilants in the varieties A, B, and C of Central Catalan. In (6) all epenthetic vowels appear underlined. (As said, those include not only the schwas that break sequences of sibilants but also the initial schwas that are needed for syllabification purposes.)

(6)

		Epenthesis	*Reduction*	*Epenthesis & Reduction*
X+suffix	*cuses* /kuz+z/	**A, B, C** [ku.zˍs]		
cl # Verb	*se sap* /s#sab/	**A, B** [sˍ.sa@p]		**C** [ˍ.sa@p]
	els sé /l+z#se/	**A** [ˍl.zˍ.se@]		**B, C** [ˍl.se@]
cl # [+N]	*els sons* /l+z#sçn+z/			**A, B, C** [ˍl.sç@ns]

Others: in prefixed words in compounds between words	 *dessalar* /d´s##sal+a+R/ *dos-cents* /doz##sent+z/ *portes sacs* /pçRt+´+z##sak+z/		**A, B, C** [d´.s´.la@] [do.se@ns] [pç$r.t´.sa@ks]	

Epenthesis and reduction (deletion or fusion) are two ways of avoiding the contact of sibilants, which would constitute a violation of the Obligatory Contour Principle (OCP), a principle originally proposed in Leben (1973) for tones, but later extended to other phonological and morphological domains. Myers (1997) argues convincingly for an Optimality-theoretic account of the OCP and the whole range of "repair strategies" to avoid it in tone phonology. Our goal in this paper is to account for the choice between epenthesis and deletion/fusion in the different contexts where a potential OCP violation occurs in Catalan and to explain the variation found in clitic–verb sequences, but not in other contexts.

In the next section, we first provide an analysis of variety A, including all the cases in which this variety does not differ from the others. Section 2.2 is devoted to variety B, and section 2.3 deals with variety C. Most of our assumptions are based on Bonet and Lloret (2001), which provides a detailed analysis of the phonology of pronominal clitics, both in contact with the verb and in clitic sequences. As previously mentioned, a crucial claim in that paper, assumed here, is that the schwa that appears associated to some clitics and in different positions is the product of epenthesis, and that it is not necessary to resort to allomorphy to account for all the shapes a clitic might surface with. Our analysis is framed in Optimality Theory, a framework that has proved to be more adequate than others in accounting for the phonology of Catalan clitics.

2. The analysis

2.1. Variety A

The fact that sibilant sequences are systematically avoided in Catalan shows that the constraint OCP is very highly ranked in Catalan, in the version of it that makes reference to sibilants. We give an informal formulation of the constraint in (7); for simplification, we call it simply OCP.

(7) OCP: adjacent sibilant segments are forbidden.

The high ranking of the Markedness constraint OCP in Catalan forces outputs to be less faithful to their inputs. Both the addition of an epenthetic vowel and the deletion of a consonant constitute violations of Faithfulness constraints. In the case of epenthesis, the constraint that punishes it is DEP-IO.

(8) DEP-IO (DEP): "Every segment of the output has a correspondent in the input"; i.e., epenthesis is prohibited (McCarthy and Prince 1995: 264).

The appearence of a single sibilant (represented by [s] in (9), ignoring voicing phenomena to which we refer below) instead of two adjacent ones can, in principle, be interpreted in one of three ways: as deletion of the first *s*, shown in (9a), as deletion of the second *s*, shown in (9b), and as fusion, in which an output [s] corresponds to two input identical segments /s s/, shown in (9c).

(9)a.	/s_1	s_2/	b.	/s_1	s_2/	c.	/s_1	s_2/	input
	|			|		\	/	
	[s_1]			[s_2]			[$s_{1,2}$]		output

The representations in (9a,b) constitute violations of the constraint MAX-IO, which bans the deletion of a segment.

(10) MAX-IO (MAX): "Every segment of the input has a correspondent in the output"; i.e., deletion is prohibited (McCarthy and Prince 1995: 264).

In (9c), the case of fusion, MAX is not violated because the two instances of /s/ do have a correspondent in the output; it just happens to be the same one. The constraint that is violated in (9c) is UNIFORMITY-IO:

(11) UNIFORMITY-IO (UNIF): No segment of the output has multiple correspondents in the input; i.e., fusion is prohibited (see McCarthy and Prince 1995).

With the constraints given so far, the appearence of a schwa in clitic-verb sequences in the context of two input sibilants could be obtained with the following ranking:

(12) Provisional ranking: OCP » UNIFORMITY, MAX-IO » DEP

The tableau in (13) shows how the correct output [ˈl.zˈ.se@], from an input /l+z#se/, is obtained through this ranking.

(13) /l+z#se/: [ˈl.zˈ.se@] '(I) know them (masc.)'

/l+z_1#s_2e/	OCP	UNIF	MAX	DEP
a. ˈls_1.s_2e@	*!			*
b. ˈl.$s_{1,2}$e@		*!		*
c. ˈl.s_2e@			*!	*
d. ˈl.z_1e@			*!	*
☞ e. ˈl.z_1ˈ.s_2e@				**

All the candidates in (13) violate the constraint DEP because all of them have at least one epenthetic vowel (the first one being needed for syllabification purposes). The output [ˈl.zˈ.se@] is the optimal candidate in spite of the fact that it is the only one with a double violation of DEP. This is so because the other candidates all violate more highly ranked constraints. All the candidates in (13), as in the rest of the tableaux in this paper, appear with the voicing specifications (voiced or voiceless) of the relevant sibilants according to what would be expected. For instance, the candidate in (13a) appears with a first voiceless sibilant because in that context voicing assimilation takes place; the candidate in (13d) appears with a voiced sibilant because we assume, with Mascaró (1986) and others, that (pro)clitics keep their voicing specification intact when the host starts with a vowel, and the first sibilant in (13d) belongs to the clitic.

If nothing else is said, it is easy to see that the ranking given in (12) would cause the wrong candidate to win in an example like *els sons* 'the sounds', with a clitic (here a determiner) followed by a noun. In the same way as in *els sé* in (13), the ungrammatical output *[ˈl.zˈ.sç@ns], instead of the grammatical [ˈl.sç@ns], would be obtained from the input /l+z#sçn+z/. The solution to this difference cannot be related to prosodic domains (like the clitic group, as in Nespor and Vogel (1986) and later work), because both the determiner *els* and the pronominal clitic *els* are clitics, more specifically proclitics; therefore they belong to the same type of prosodic domain.

In Bonet and Lloret (2001) it is argued that what motivates the presence of peripheral epenthesis in clitic-verb sequences (cf. *en tira*: [ˈn.t"€.Rˈ] *vs.* *[nˈ.t"€.Rˈ], from an underlying form /n#tiR+ˈ/, '(s/he) throws some') is the constraint ALIGN(CL-V), defined in (14).[5]

[5] A parallel constraint, ALIGN(V-CL), accounts for enclisis, and the presence of peripheral epenthesis in that context, that is, the presence of a schwa following the clitic (as in *tirem-ne* 'let us throw some'

(14) ALIGN(CL-V) (AL(CL-V)): Align the left edge of V(erb)[+tense] with the right edge of a pronominal clitic.

This constraint is violated in candidates like *[n_´.t"€.R´] due to the presence of the epenthetic vowel between the clitic and the verb (it is not violated when epenthesis is peripheral, that is, when the schwa precedes the proclitic, as in the grammatical form [_´n.t"€.R´]).

Parallel to ALIGN(CL-V), the constraint ALIGN(CL-[+N]) tries to ensure the adjacency between a determiner and its host. This constraint is informally stated in (15).

(15) ALIGN(CL-[+N]) (AL(CL-N)): Align the left edge of [+N] with the right edge of a determiner.

This constraint is violated whenever the configuration $_{CL})(_{N}$ is not obtained. The ranking of ALIGN(CL-[+N]) right below OCP causes a candidate *[_´l.z_´.sç@ns], for *els sons* 'the sounds', to be discarded as the optimal candidate, as shown in (16).

(16) /l+z#sçn+z/: [_´l.sç@ns] 'the sounds'

/l+z_1#$s_2$çn+z/	OCP	AL(CL-N)	UNIF	MAX	DEP
a. _´ls_1.s_2ç@ns	*!				*
b. _´l.$s_{1,2}$ç@ns		*!	*		*
☞c. _´l.s_2ç@ns				*	*
☞d. _´l.z_1ç@ns				*	*
e. _´l.z_1_´.s_2ç@ns		*!			**

The candidate in (16e), with an epenthetic vowel between the determiner and the noun —and with a configuration $_{CL})´(_{N}$—, is not the only one to violate ALIGN(CL-[+N]). Fusion also causes a violation of this constraint, given that the left edge of the noun is *before* $s_{1,2}$, while the right edge of the determiner is *after* $s_{1,2}$ (the configuration being [_´l($_{N}$ $s_{1,2}$ $_{CL}$)ç@ns]). In (16), then, both the candidate with epenthesis and the candidate with fusion are ruled out. However, two ouputs fare even, (16c), with deletion of the first sibilant, and (16d) with deletion of the second sibilant and a faithfully voiced first sibilant, which does not correspond to the

/tiR+E+m#n/, [ti.RE@m.n_´]). For variety A, there is no evidence for a different ranking of the two constraints; for this reason they can be collapsed under the term ALIGN(CL/V).

grammatical phonetic form, [˝l.sç@ns], with a voiceless sibilant. An additional constraint, justified below, is needed to undo the tie.

Words like *esport* 'sport' or *especificar* 'to specify' surface, in Catalan, with an initial epenthetic vowel [´]: [˝s.pç@rt], [˝s.p´.si.fi.ka@]. When a proclitic is added to these words, this epenthetic vowel remains: /l#spçRt/ 'the sport' surfaces as [l´s.pç@rt] (cf. *[˝ls.pç@rt], also with a well-formed syllabification in Catalan), and /u#sp´sifik´/ '(s/he) specifies it' surfaces as [w˝s.p´.si.f"€.k´]. In the latter case, a sequence like *[us.p´.si.f"€.k´] would be more faithful to the input and would not cause any syllabification problems. Traditionally it has been assumed that epenthesis takes place first at the word level, and clitics are added later to the epenthesized word. Leaving aside some problems such a cyclic account would have to face in dealing with the phonology of clitics, Optimality Theory offers several alternatives to cyclicity that avoid having to resort to levels, one of them being Output-Output correspondence constraints (see, for instance, Benua 1995 or Kenstowicz 1996). In this type of Faithfulness constraints a correspondence relation is established between the base form (which has to be a possible free standing word) and an affixed (or cliticized) form. The Output-Output constraint stated in (17) makes reference to the initial segment of the base. The final segment of the base is more unstable (for instance, the final schwa of a verbal form like *passa* 'pass!', [pa@.s´], might be deleted before the neuter clitic *ho* (/u/) in some of the varieties discussed in this paper (cf. *passa-ho* 'pass it!': [pa@.su]). A parallel constraint, OUTPUT-OUTPUT$_{FINAL}$, punishes the deletion in a candidate like [pa@.su], for *passa-ho*, and can be ranked differently with respect to OUTPUT-OUTPUT$_{INITIAL}$ depending on the variety.

(17) OUTPUT-OUTPUT$_{INITIAL}$ (OO$_{IN}$): The initial segment of a base has a correspondent in the affixed or cliticized word.

For an input form like /u#sp´sifik´/, the highly ranked constraint OUTPUT-OUTPUT$_{INITIAL}$ favors the candidate [w˝s.p´.si.f"€.k´] over *[us.p´.si.f"€.k´], because the former, but not the latter, keeps the first segment of the base [˝s.p´.si.f"€.k´]. The corresponding tableau is given in (18). (The constraint ONSET (ONS), which demands that syllable have onsets, has to be ranked below ALIGN(CL-V) to ensure initial epenthesis in examples like *en sap* /n#sab/, [˝n.sa@p], instead of *[n˝.sa@p] '(s/he) knows some'; cf. Bonet and Lloret 2001.)

(18) /u#sp´sifik´/: [w˝s.p´.si.f"€.k´]

/u#sp´sifik´/	OO$_{IN}$	AL(CL-V)	ONS	DEP
a. us.p´.si.f"€.k´	*!		*	

☞b.w´s.p´.si.f"€.k´		*		*

Going back to the problem posed by the tableau in (16), corresponding to *els sons* 'the sounds', the tie between the two candidates [´l.s$_2$ç@ns] and *[´l.z$_1$ç@ns] can be resolved with the inclusion of OUTPUT-OUTPUT$_{INITIAL}$, ranked between OCP and ALIGN(CL-[+N]), as shown in (19).

(19) /l+z#sçn+z/: [´l.sç@ns] 'the sounds'

/l+z$_1$#s$_2$çn+z/	OCP	OO$_{IN}$	AL(CL-N)	UNIF	MAX	DEP
a. ´ls$_1$.s$_2$ç@ns	*!					*
b. ´l.s$_{1,2}$ç@ns			*!	*		*
☞c. ´l.s$_2$ç@ns					*	*
d. ´l.z$_1$ç@ns		*!			*	*
e. ´l.z$_1$´.s$_2$ç@ns			*!			**

The candidate *[´l.z$_1$ç@ns] is ruled out because the initial segment of the base [sç@ns] has not been kept, while the initial sibilant survives in the optimal candidate [´l.s$_2$ç@ns].

The first tableau for the clitic-verb sequence *els sé* '(I) know them' (realized [´l.z´.se@]), in (13), took into account only the constraints OCP, UNIFORMITY, MAX, and DEP. The inclusion of the constraint ALIGN(CL-V), which has to be ranked above MAX (as argued for in Bonet and Lloret 2001), and OUTPUT-OUTPUT$_{INITIAL}$ would rule out the grammatical candidate [´l.z´.se@]; the optimal candidate would be *[´l.s$_2$e@].

(20) /l+z#se/: [´l.z´.se@] '(I) know them (masc.)'

/l+z$_1$#s$_2$e/	OCP	OO$_{IN}$	AL(CL-V)	UNIF	MAX	DEP
a. ´ls$_1$.s$_2$e@	*!					*
b. ´l.s$_{1,2}$e@			*!	*		*
☞ c. ´l.s$_2$e@					*	*
d. ´l.z$_1$e@		*!			*	*
e. ´l.z$_1$´.s$_2$e@			*!			**

The decisive constraint that is missing from (20) is the well established constraint REALIZE-μ, defined below.[6]

(21) REALIZE-μ (REAL$_{\mu}$): A morpheme must have some phonological exponent in the output (Walker 1998).

The ranking of REALIZE-μ, between OUTPUT-OUTPUT$_{\text{INITIAL}}$ and ALIGN(CL-V) gives the grammatical output [ˈl.zˈ.se@] as the optimal candidate, as shown in (22).

(22) /l+z#se/: [ˈl.zˈ.se@] '(I) know them (masc.)'

/l+z$_1$#s$_2$e/	OCP	OO$_{\text{IN}}$	REAL$_{\mu}$	AL(CL-V)	UNIF	MAX	DEP
a. ˈls$_1$.s$_2$e@	*!						*
b. ˈl.s$_{1,2}$e@				*	*!		*
c. ˈl.s$_2$e@			*!			*	*
d. ˈl.z$_1$e@		*!				*	*
☞ e. ˈl.z$_1$ˈ.s$_2$e@				*			**

REALIZE-μ rules out the candidate *[ˈl.s$_2$e@] because the deletion of the first sibilant implies the lack of realization of the plural morpheme of the clitic. UNIFORMITY ends up being the decisive constraint for the two surviving candidates.

There is an additional context in which epenthesis is used to avoid the OCP problem posed by sequences of sibilants: all varieties, not only variety A, insert a schwa after a stem-final sibilant and a sibilant morph. The example *cuses* '(you) sew', [ku@.zˈs], from an underlying form /kuz+z/, illustrated such a case.[7] The Alignment constraints introduced so far, ALIGN(CL-[+N]) and ALIGN(CL-V), are irrelevant in this type of case, but a similar kind of constraint needs to be invoked, one that ensures the adjacency relation between the stem and the suffix. We give a general formulation of this constraint in (23).

(23) ALIGN(μ-μ) (AL(μ-μ)): Align the left edge of a morph X with the right edge of morph Y.

---

[6] An equivalent constraint, within the Containment model of OT, can be found, with the label PARSE-MORPH, in Akinlabi (1996), for instance.

[7] This type of example is discussed by Colina (1995) and Jiménez (1997), but they do not consider candidates with fusion or deletion of one of the segments.

For an input like /kuz+z/, ALIGN(μ-μ), unranked with respect to REALIZE$_{\mu}$, is violated by the candidate with epenthesis (actually the only grammatical candidate for all the varieties considered in this paper) and the candidate with fusion. ALIGN(μ-μ) together with OCP and REALIZE$_{\mu}$ would wrongly give *[ku@s$_2$] as the optimal candidate. A possible solution to this puzzle can be related to the constraint MAX, which has, and must have, a fairly low ranking. So far we have considered MAX to be a constraint that punishes the deletion of *any* segment. However, this constraint can be "broken" into more specific constraints, by distinguishing, for instance, a version of it that makes reference to vowels and another one that makes reference to consonants (see, e.g., McCarthy 2000 for arguments in favor of this possibility). Although a complete analysis of the phonology of Catalan might give arguments for a fairly detailed specification of the different MAX-constraints, for the purposes of this paper it is enough to distinguish the general MAX constraint (with the same definition and ranking assumed so far), from a specific version of it that makes specific reference to sibilants. This constraint is stated in (24).

(24) MAX-SIBILANT (MAX$_S$): Every sibilant segment of the input has a correspondent in the output; i.e., deletion of a sibilant consonant is prohibited.

This specific version of MAX receives support from general facts related to deletion in Catalan. In coda consonant clusters, for instance, deletion of the second consonant is fairly common, especially in dialects like Valencian (words like *augment* 'augmentation' are commonly pronounced [ˊw.me@n]; cf. *[ˊg.me@n]); this is so except when the second consonant is *s*, in which case the first consonant is deleted (a word like *monstre* 'monster' is often pronounced [mç@s.tRˊ]; cf. *[mç@n.tRˊ]). Ultimately, the fact that sibilants are reluctant to deletion could be related to their perceptual prominence. The more specific MAX-SIBILANT must universally be ranked above the more general MAX, as is shown in the tableau corresponding to [ku@.zˊs]. We exclude from the tableau all the constraints (like OO$_{\text{INITIAL}}$) that are irrelevant for this example.

(25) /kuz+z/: [ku@.zˊs] '(you) sew'

/kuz$_1$+z$_2$/	OCP	REAL$_{\mu}$	AL(μ-μ)	MAX$_S$	UNIF	MAX	DEP
a. ku@s$_1$s$_2$	*!						
b. ku@s$_{1,2}$			*		*!		
c. ku@s$_2$				*		*!	
d. ku@s$_1$		*		*!		*	

☞ e. ku@.z$_1$˙s$_2$			*				*

The unordered status of REALIZE-μ, ALIGN(μ-μ), and MAX-SIBILANT causes a tie between the candidates in (25b), (25c), and (25e); UNIFORMITY and MAX become, then, the decisive constraints.[8]

For an example like *els sons*, [˙l.sç@ns], 'the sounds' (see the tableau in (19)) the inclusion of MAX-SIBILANT does not alter the results already obtained, given that MAX-SIBILANT is ranked lower than ALIGN (CL-[+N]), the lowest decisive constraint. The results are not different, either, for an example like *els sé*, [˙l.z˙.se@], '(I) know them' (see the tableau in (22)): MAX-SIBILANT punishes two candidates, (22c) and (22d), which also violate REALIZE-μ and $OO_{INITIAL}$, respectively. The definite tableau corresponding to [˙l.z˙.se@] is almost identical to the one for *se sap*, [s˙.sa@p], 'it is known', given in (26).

[8] Following Kenstowicz (2001), one could attribute the ungrammaticality of an output *[ku@s] corresponding to the input /kuz+z/ (second person singular) to a Contrast constraint, which would rule out deletion or fusion because the output would become identical to another form of the verbal paradigm, namely the third person singular of the same tense, *cus* [ku@s]. However, pursuing this type of approach could have many consequences for other paradigms not too easy to foresee.

(26) /s#sab/: [s_´.sa@p] 'it is known'

/s_1#$s_2$ab/	OCP	OO_{IN}	$REAL_{\mu}$	MAX_S	AL(CL-V)	UNIF	MAX	DEP
a. s_1s_2a@p	*!							
b. $s_{1,2}$a@p					*	*!		
c. s_2a@p			*	*!			*	
d. s_1a@p		*!		*			*	
☞ e. s_1_´.s_2a@p					*			*
f. _´s_1.s_2a@p	*!							*

So far we have accounted for most of the cases in which a potential OCP violation might occur in variety A (even though some of the results are shared by varieties B and C). Before examining the behavior of the other varieties in clitic-verb sequences, let us briefly see how the rest of the cases (under *Others* in (6)) could be dealt with, even though many of the issues that arise are of a very general nature and fall beyond the goals of this paper. These cases include prefixed words (like *dessalar* [d´.s´.la@] 'to desalt'), compounds (like *dos-cents* [do.se@ns] 'two hundred') and adjacent independent words (like *portes sacs* [pç$r.t´.sa@ks] '(you) bring sacks', or *pis senzill* [p"›.s´n.z"€¥] 'simple apartment'). Given the analysis put forward so far, it is not possible to account for the lack of epenthesis in prefixed words, like *dessalar* 'to desalt'. The solution to this problem might be related to whatever properties cause prefixes to behave phonologically as independent words (in many respects), like the other cases to be dealt with. Compounds like *dos-cents* 'two hundred' are made out of independent words; therefore, at least for the time being, they can be treated like word sequences, like *portes sacs* '(you) bring sacks' or *pis senzill* 'simple flat'. Sequences of words can readily be dealt with, if two considerations are made. An Alignment constraint has to ensure that adjacent words are in fact adjacent (as was the case with clitic-verb sequences, morphemes, etc.). Let us call this constraint ALIGN(WORD-WORD) (in the following tableaux, abbreviated as AL(W-W)). Moreover, although we defined $OO_{INITIAL}$ as a constraint that related bases to their affixed or cliticized counterparts, it can be reformulated in such a way that it establishes a correspondence relation between a base and all occurrences of that base. Assuming this modification of the constraint, the candidate (27d), below (corresponding to the input /piz##s´nzi¥/) violates $OO_{INITIAL}$ because the first segment of the second word, [´] (the /s/ having been deleted), does not correspond to the first segment of the base [s´n.z"€¥].

(27) /piz##s´nzi¥/: [p"›.s´n.z"€¥] 'simple apartment'

/piz$_1$##s$_2$´nzi¥/	OCP	OO$_{IN}$	AL(W-W)	REAL$_\mu$	MAX$_S$	UNIF	MAX	DEP
a. p"›s$_1$.s$_2$´n…	*!							
b. p"›.s$_{1,2}$´n…			*!			*		
☞c. p"›.s$_2$´n…					*		*	
d. p"›.z$_1$´n…		*!			*		*	
e. p"›.z$_1$´.s$_2$´n…			*!					*

In (28) we give a slightly different example with the same result, *portes sacs* '(you) bring sacks', in which the first sibilant (i.e., the last sement of *portes*) constitutes the second person singular morph.

(28) /pçRt+´+z##sak+z/: [pç$r.t´.sa@ks] '(you) bring sacks'

/pçRt+´+z$_1$##s$_2$ak+z/	OCP	OO$_{IN}$	AL(W-W)	REAL$_\mu$	MAX$_S$	UNIF	MAX	DEP
a. …t´s$_1$.s$_2$a@ks	*!							
b. …t´.s$_{1,2}$a@ks			*!			*		
☞c. …t´.s$_2$a@ks				*	*		*	
d. …t´.z$_1$a@ks		*!			*		*	
e. …t´.z$_1$´.s$_2$a@ks			*!					*

2.2. Variety B

In variety B most pronominal clitics behave like determiners do; that is, when a proclitic ending in a sibilant is adjacent to a verb starting with a sibilant, a single [s] is present in the surface form. In this variety, then, no distinction is made between the two types of clitics. For this reason there is no need for the existence of a constraint ALIGN(CL-V) different from ALIGN(CL-[+N]); the two constraints can be collapsed into a more general one, ALIGN(CL-LEX).[9]

[9] In footnote 5 it was mentioned that ALIGN(CL-V) and ALIGN(V-CL) have an identical ranking in variety A; for this reason they can be collapsed into a single constraint ALIGN(CL/V). In a similar fashion, one could wonder whether in variety B it is possible to collapse all proclisis and enclisis into a constraint ALIGN(CL/LEX). This might be the case, but a detailed study of enclisis in this variety is needed before jumping to such conclusions.

(29) ALIGN(CL-LEX) (AL(CL-LEX)): Align the left edge of a lexical word with the right edge of a clitic.

ALIGN(CL-LEX) occupies the same position as ALIGN(CL-[+N]) in variety A, unordered with respect to $OO_{INITIAL}$. The tableau in (30) shows how the output [ˈl.se@] is obtained for *els sé* '(I) know them'.

(30) /l+z#se/: [ˈl.se@] '(I) know them'

/l+z_1#s_2e/	OCP	OO_{IN}	AL(CL-LEX)	$REAL_{\mu}$	MAX_S	UNIF	MAX	DEP
a. ˈls_1.s_2e@	*!							*
b. ˈl.$s_{1,2}$e@			*!			*		*
☞ c. ˈl.s_2e@				*	*		*	*
d. ˈl.z_1e@		*!			*		*	*
e. ˈl.z_1ˈ.s_2e@			*!					**

Not all the pronominal clitics behave like *els*. The impersonal/reflexive clitic surfaces with an epenthetic vowel instead of deletion or fusion (cf. *se sap* [sˈ.sa@p]). At first sight this different behavior among the clitics might seem surprising, but there is a crucial difference between the reflexive clitic and all the other clitics that end in a sibilant: the reflexive clitic is underlyingly a single segment, /s/, while the other clitics have more than one segment (cf. third person masculine plural, /l+z/; third person feminine plural, /l+a+z/; first person plural, /nz/; second person plural, /wz/ or /uz/. Therefore, if the sibilant of the reflexive clitic is deleted, the whole clitic is deleted, while this is not the case for any of the other clitics. The constraint that punishes the deletion of a whole clitic is REALIZE-CLITIC.

(31) REALIZE-CLITIC ($REAL_{CL}$): A clitic must have some phonological exponent in the output.

As mentioned, this constraint will only be relevant when the clitic is a single segment (of course, in any complete tableau for other clitics there will be a candidate with deletion of all the consonants, thus violating REALIZE-CLITIC; but this candidate will also violate many other constraints and will not have a chance to survive). The tableau in (32) shows how *se sap* is obtained in variety B.

(32) /s#sab/: [s˙.sa@p] 'it is known'

/s_1#$s_2$ab/	OCP	OO_{IN}	$REAL_{CL}$	AL(CL-LEX)	$REAL_{\mu}$	MAX_S	UNIF	MAX	DEP
a. s_1s_2a@p	*!								
b. $s_{1,2}$a@p				*			*!		
c. s_2a@p			*		*	*!		*	
d. s_1a@p		*				*!		*	
☞ e. s_1˙.s_2a@p				*					*

REALIZE-CLITIC is a constraint also present in variety A, although it did not appear in the corresponding tableau in (26) for expository reasons. In a way parallel to what we see in (32), REALIZE-CLITIC fatally punishes the candidate with deletion of the clitic (*[s_2a@p]), a candidate that also violates REALIZE-μ and MAX-SIBILANT, as can also be seen in (32).

2.3. Variety C

Variety C is almost identical to variety B.[10] The lack of epenthesis between the clitic and the verb in cases like *els sé* [˙l.se@] '(I) know them', which are treated like sequences with determiners (like *els sons* [˙l.sç@ns]), indicates that the relevant Alignment constraint for these cases is ALIGN(CL-LEX). We can assume the same ranking it has in variety B.

The only difference between variety B and variety C lies in the behavior of the reflexive/impersonal clitic, the only clitic of the language that consists of a single sibilant. A sequence like *se sap* 'it is known' is realized, in variety C, as a sequence with sibilant deletion/fusion *and* initial epenthesis: [˙.sa@p] (from an underlying form /s#sab/). This case raises one of the most difficult problems to solve in Optimality Theory, namely the problem of opacity. The sequence [˙.sa@p] is opaque because there is no apparent need for the initial epenthetic vowel: the OCP problem is solved via deletion or fusion, and the initial schwa is not needed for syllabification (*[sa@p] would be fine). Given our claim that the schwa is not present in the underlying form, the presence of an epenthetic vowel in the grammatical output cannot be explained straightforwardly. With the constraints we have presented so far, *[sa@p], in any of the interpretations for the appearence of a single [s], will always violate a subset of the constraints violated by [˙.sa@p], which violates DEP in

[10] Variety C, in comparison with varieties A and B, is spoken by few people in the Barcelona area. Most of the speakers of this variety have Catalan as a second language.

addition to others. In a serial model of phonology one could stipulate that the rule inserting the epenthetic vowel applies before the rule that deletes one of the sibilants or fuses them (an instance of extrinsic ordering), a possibility not available here.

Intuitively, the absence of the initial schwa (a form like *[sa@p]) would cause the sequence to become phonetically identical to the simple verb, [sa@p] (the presence of the clitic could not be perceived). The schwa in [́.sa@p] is the clue for the presence of the clitic, even if the clitic itself cannot be identified (at least, apparently). Moreover, the presence of the schwa between the proclitic and the verb (like in [s ́.sa@p] for varieties A and B) would constitute the only instance of epenthesis between a clitic and its host, all other proclitics having initial epenthesis when needed (cf. *en sap* [́n.sa@p] '(s/he) knows some', *em veu* [́m.bE@w] '(s/he) sees me'). On the contrary, and leaving aside the OCP problem, initial epenthesis in [́.sa@p] causes the clitic /s/ to have exactly the same phonological behavior as all the other consonantal clitics, which drives them to have a final VC(C) shape in proclitic position (cf. [́m], [́t], [́n], [́l], [́ls], [́ns], *and* [́s]). Even though capturing these intuitions within the framework adopted here is not easy and the issue needs further investigation, we can assume, following the lines of Kenstowicz (2001) for paradigmatic uniformity phenomena, an Output-Output constraint between members of a paradigm, in the present case pronominal clitics.[11] For the clitic *es*, this Output-Output constraint, let us call it OUTPUT-OUTPUT$_{\text{PARADIGM}}$ (OO$_{\text{PAR}}$), will favor an output with a VC shape, namely [́s], parallel to the other consonantal proclitics of the language. This constraint can be ranked fairly low, just above UNIFORMITY. The tableau corresponding to [́.sa@p] is given in (33).[12] Notice that in this case, but not in other cases of deletion /fusion (in which deletion of the first consonant was the "strategy" chosen), the optimal candidate is realized with fusion of the sibilants; therefore the clitic does have a final VC shape ([́s]) without causing, at the same time, a violation of OUTPUT-OUTPUT$_{\text{INITIAL}}$ in the verb ([sa@p]): [́.s$_{1,2}$a@p].

[11] Kenstowicz (2001) considers paradigmatic uniformity phenomena the cases where "the grammar strives to maintain the same output shape for pairs of inflected words that the regular phonology should drive apart". In this paper, we extend this notion to the set of pronominal clitics.

[12] Another, obvious, way of avoiding the opacity problem raised in this section is to assume that the initial schwa in [́.sa@p] is not epenthetic but part of the clitic. If the underlying form of the clitic (at least in certain contexts) is /́s/, the problem disappears, and the realization [́.sa@p] for an underlying sequence /́s#sab/ is explained exactly like the other cases. However, this would be the only case in which an underlying schwa would have to be posited for pronominal clitics.

(33) /s#sab/: [ˊ_.sa@p] 'it is known'

/s_1#$s_2$ab/	OCP	OO_{IN}	$REAL_{CL}$	AL(CL-LEX)	$REAL_{\mu}$	MAX_S	OO_{PAR}	UNIF	MAX	DEP
a. s_1s_2a@p	*!						*			
b. $s_{1,2}$a@p				*			*!	*		
c. s_2a@p			*		*	*	*!		*	
d. s_1a@p		*				*	*!		*	
e. s_1ˊ_.s_2a@p				*			*!			*
f. ˊ_s_1.s_2a@p	*!									*
☞ g. ˊ_.$s_{1,2}$a@p				*				*		*
h. ˊ_.s_2a@p			*		*	*	*!		*	*
i. ˊ_.s_1a@p		*				*!			*	*

The $OUTPUT\text{-}OUTPUT_{PARADIGM}$ constraint is obviously also present in the other varieties, although its effects are not noticeable because it is ranked lower (at least below UNIFORMITY and MAX).

3. Conclusions

In this paper we have examined the effects that a highly ranked OCP constraint for sibilants has in different varieties of Catalan, especially in environments involving clitics. As we saw, the main difference among the three varieties under discussion is whether all clitics behave alike (as in varieties B and C) or a distinction is made between pronominal clitics and other types of clitics (as in variety A). In OT, dialectal and language variation has commonly been accounted for by constraint reranking. In our analysis, the different behavior of variety A with respect to varieties B and C is captured by the decomposition of a general constraint, ALIGN(CL-LEX), into more specific constraints, ALIGN(CL-[+N]) and ALIGN(CL-V), which are to be considered members of the same constraint family. The fact that there is no subset relation between ALIGN(CL-[+N]) and ALIGN(CL-V) implies that they are not universally ranked with respect to each other. In addition, the decomposition of a general constraint into more specific ones implies that at least one constraint has to intervene between them (in the case at hand, three unordered constraints intervene: REALIZE-μ, ALIGN(μ-μ), and MAX-SIBILANT); otherwise there would be no evidence

for such decomposition. Decomposition necessarily involves partial constraint reranking, as shown in (34).

The other source of variation we have seen is restricted to the behavior of a specific clitic, namely the third person reflexive/impersonal clitic. In variety C, as opposed to varieties A and B, the output form of this clitic constitutes an apparent opacity case. As an alternative to levels or extrinsic rule ordering, we have proposed an approach in terms of an Output-Output constraint, $\text{OUTPUT-OUTPUT}_{\text{PARADIGM}}$, which forces it to adopt the same output shape (a VC structure) as all the other consonantal proclitics, establishing thus a kind of analogical relation.

To conclude, in (34) we give the complete hierarchy for varieties A, B, and C. The constraints that constitute a source of variation appear in bold face.

(34)a. *Variety A*

OCP » OO_{IN}, REAL_{CL}, AL(W-W), **AL(CL-N)** » REAL_{μ}, AL(μ-μ), MAX_{S} » **AL(CL-V)** » UNIF, MAX » DEP, $\mathbf{OO_{PAR}}$

b. *Variety B*

OCP » OO_{IN}, REAL_{CL}, AL(W-W), **AL(CL-LEX)** » REAL_{μ}, AL(μ-μ), MAX_{S} » UNIF, MAX » DEP, $\mathbf{OO_{PAR}}$

c. *Variety C*

OCP » OO_{IN}, REAL_{CL}, AL(W-W), **AL(CL-LEX)** » REAL_{μ}, AL(μ-μ), MAX_{S}, $\mathbf{OO_{PAR}}$ » UNIF, MAX » DEP

References

Akinlabi, A. (1996). "Featural Affixation". *Journal of Linguistic* 32: 239-289.

Benua, L. (1995). "Identity Effects in Morphological Truncation". In: Beckman, J.; Walsh Dickey, L.; Urbanczyk, S. (eds.). *Papers in Optimality Theory*. Amherst, Mass.: Graduate Linguistic Student Association (University of Massachusetts Occasional Papers in Linguistics 18), pp. 77-136.

Bonet, E. and Lloret, M.-R. (2001). "More on Alignment as an Alternative to Domains: the Syllabification of Catalan Clitics", ms. Universitat Autònoma de Barcelona and Universitat de Barcelona.

Colina, S. (1995). *A Constraint-based Analysis of Syllabification in Spanish, Catalan, and Galician*. Doctoral dissertation, University of Illinois, Urbana-Champaign.

Jiménez, J. (1997). *L'estructura sil·làbica del dialecte valencià*. Doctoral dissertation, Universitat de València.

Kenstowicz, M. (1996). "Base-Identity and Uniform Exponence: Alternatives to Cyclicity". In: Durand, J.: Lacks, B. (eds.). *Current Trends in Phonology Models and Methods*, vol. 1. Salford: European Studies Research Institute (ESRI), University of Salford, pp. 363-393.

Kenstowicz, M. (2001). "Paradigmatic Uniformity and Contrast". Handout, *3èmes Journées Internationales du GDR Phonologie*, Université de Nantes, June 1, 2001.

Leben, W. (1973). *Suprasegmental phonology*, doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

Mascaró, J. (1986). *Morfologia*, Barcelona: Enciclopèdia Catalana.

McCarthy, J. (2000). "The Prosody of Phase in Rotuman". *Natural Language and Linguistic Theory* 18: 147-197.

McCarthy, J. and A. Prince (1995). "Faithfulness and reduplicative identity", in J. Beckman, L. Walsh Dickey and S. Urbanczyk (eds.), *Papers in Optimality Theory*, University of Massachusetts Occasional Papers in Linguistics 18, University of Massachusetts, Amherst, pp. 249-384.

Myers, S. (1997). "OCP Effects in Optimality Theory". *Natural Language and Linguistic Theory* 15: 847-892.

Nespor, M. and Irene Vogel (1986). *Prosodic Phonology*, Dordrecht: Foris.

Walker, R. (1998). *Nasalization, Neutral Segments, and Opacity Effects*. Doctoral dissertation, UC, Santa Cruz. To be published, New York: Garland.