

# Lexical conservatism and its analysis

Donca Steriade, UCLA

"Un mot quelconque peut toujours évoquer tout ce qui est susceptible de lui être associé d'une manière ou d'une autre. ..Un terme donné est comme le centre d'une constellation, le point où convergent d'autres termes coordonnés, dont la somme est indéfinie."  
Saussure

"There is, I believe, a certain economy in language; new forms are not created just because the formal mechanism is there, if there is no need for them, and their creation would serve merely to crowd an already-existing and perfectly usable form." Cowgill

## 1. Listedness

The first studies of generative morphology, Halle's (1973) "Prolegomena" and Aronoff's (1976) "Word Formation", have identified the phenomenon of blocking: a pre-existing, listed word blocks productive word formation processes from creating potential synonyms to it. Because of *fury*, *\*furiousness* is blocked: it cannot be used in any of the senses known to be already covered by *fury*. Blocking reflects the speakers' preference to use known words, a phenomenon referred to here as lexical conservatism.

This paper identifies the phonological side of lexical conservatism: phonological processes are, under certain circumstances, blocked from creating novel phonological variants to a listed stem. Rather than generating new allomorphs, speakers recycle already existing ones, even when none of the listed allomorphs gives full satisfaction to the applicable phonological and morphosyntactic conditions. The general interest of this variety of lexical conservatism is that, in order to provide it with an explicit description, it will be necessary to revise some of our basic assumptions about the relation between bases and their derivatives.

The notions of listed word and listed allomorph will be essential to the analysis. I borrow these, with some extension, from Halle (1973), who notes that speakers are aware of the difference between potential and actual results of the word formation system of their language. Correspondingly, the term "listed" denotes here a degree of familiarity with a word, sufficient to give a speaker the confidence that the word has been sanctioned by past linguistic usage. A listed word is a non-hapax, a non-nonce form. A listed word, in the sense adopted here, may be a word whose morphological and phonological properties are fully predictable, given knowledge of the grammar and lexicon of the language: thus

*happiness, demonstrative, demonstrable, readable* are listed words for most speakers of English. In contrast, I expect that a form like *matchability* is a clear hapax for most speakers. Since listedness is a matter of individual linguistic experience, the listed status of a word may vary from speaker to speaker: thus I expect that words like *nouniness* or *pronounceable* may register as listed with some English speakers but not with others. I have no proposal on what causes erstwhile hapaxes to become listed words<sup>1</sup>, nor any clear criteria to distinguish listed from unlisted items, but it seems clear that the intuition of listedness exists, as Halle had noted. This will be the only assumption we will need in what follows. I should also note that the term *listed* is used with a restricted meaning by morphologists like Aronoff (1976, 1995) and Lieber (1981): a listed item, for these writers, is a form entitled to a lexical entry in virtue of possessing properties that cannot be derived via productive rules of phonology or word formation from those of other lexical items. In this sense, *readable, nouniness* or *happiness* are not listed words. For the moment it will suffice to say that this restricted sense is not the one adopted here.

Related to the notion of listedness adopted here is Kurylowicz's (1949) concept of *sphère d'emploi* (range or domain of use) of a given expression: this refers to the set of linguistic contexts in which the expression can be used. The phenomena discussed here involve surface analogy. And surface analogy, according to Kurylowicz, takes place when one linguistic expression with a broader *sphère d'emploi* determines the shape of another one, whose own domain of use is narrower, a subset of the former. Thus the pronunciation of *cycle* [sɪkl̩] may influence that of the derivative *cyclicity*, which is increasingly being pronounced [sɪkl̩ɪsɪti] (as against the earlier, and still standard, [sɪkl̩ɪsɪti]). This is a *sphère d'emploi* effect, the domain of use of *cyclicity* being a proper subset of that of *cycle*. As Kurylowicz would predict, the pronunciation of the derivative *cyclicity* is unlikely to ever affect that of the base, because of this necessary subset relation between the range of uses of the two expressions. I believe that the notion of *sphère d'emploi* is related to and perhaps reduces to that of listedness: the narrower the range of uses of an expression, the lower its token frequency, hence the greater the likelihood that it will be unfamiliar, not listed, for many speakers. *Cyclicity* is influenced by *cycle* especially in the pronunciation of those speakers for whom the latter is a listed word and the former is not.

I draw this connection between listedness and *sphère d'emploi* because it is not always possible to establish, in comparing two expressions, whether they differ in their listed status for some population

---

<sup>1</sup>Clearly frequency of occurrence is not the only factor here: the term *matchability* was repeatedly used by the participants in the seminar on which this study is based, yet its status has not progressed from nonce to listed for any one of these speakers.

of speakers: but it is easier to compare their sphères d'emploi, given elementary knowledge of the linguistic system they belong to. If the two notions of listedness and range of uses are indeed related, then information about one can occasionally serve as substitute for the other.

## 2. Lexical relations: vertical and lateral

The focus in phonology and morphology has until recently been on a class of lexical relations that one may dub *vertical*: these are relations between pairs of terms in which one is viewed as the derivational antecedent of the other. The relation between underlying and surface form, that between base and derivative (*happy -happiness*) or that between base and reduplicant (*ge -grapha* in Greek *gegrapha* 'I have written') fit into this category. But it is at least conceivable that *lateral* relations are also linguistically significant: these are relations obtaining between terms such that neither can be viewed as the derivational antecedent of the other. An example are co-derivatives like *demonstrable* and *demonstrative*. As we explore the phenomenon of lexical conservatism we will observe that it is possible for co-derivatives to entertain direct relations, unmediated by their common underlying representation or by their common base word.

## 3. An instance of phonological conservatism: French liaison

### 3.1. Basic data

Effects of lexical conservatism can be observed in the paradigm of French liaison (Dell 1973, Selkirk 1974, Tranel 1981, 1987, Encrevé 1988, and further references in Tranel's and Encrevé's works). French adjectives possess so-called liaison allomorphs, to be used when preposed to a vowel initial noun or adjective, as in (1): *beau* [bo] (1.b) is the basic singular masculine form of the adjective, while *bel* [bEl] is its liaison allomorph, used to avoid hiatus (1.a). Similarly *ce* [s"] is the basic masculine form of 'this' (1.d), replaced by the liaison allomorph [sEt] in potential hiatus situations, such as (1.c).

- |        |                        |                   |                               |
|--------|------------------------|-------------------|-------------------------------|
| (1) a. | <i>le bel homme</i>    | [l" bEl Om]       | 'the handsome man'            |
| b.     | <i>l'homme beau</i>    | [lOm bo]          | 'the man (that is) handsome'  |
| c.     | <i>cet ancien pays</i> | [sEt a)sjE§] pei] | 'this old country'            |
| d.     | <i>ce pays ancien</i>  | [s" pei A)sjE§]   | 'this country (which is) old' |

We follow here Tranel's (1996) and Perlmutter's (1996) arguments to the effect that hiatus avoidance is the rationale for using the liaison forms in (1.a) and (1.c). The comparison of NP's containing preposed and postposed adjectives (2) indicates that hiatus is tolerated in the latter cases

(2.b), perhaps because the [NA] structures form two distinct accentual phrases, whereas the [AN] structure must be incorporated into a single accentual phrase. The urgency of hiatus avoidance depends then on the prosodic closeness of the two vowels<sup>2</sup>.

- (2) a. *savan[t] Anglais* 'a learned Englishman'                      one accentual phrase  
       b.            *sav[A]] Anglais* 'a scholar who is English'            two accentual phrases

All liaison allomorphs inspected so far are identical to the feminine forms of the adjective. Thus the nouns in (1) and (2) are masculine - as indicated by the choice of definite article (*le bel homme*, vs. feminine *la*), the choice of adjective (*cet ancien pays*, vs. feminine *ancienne*) and the form of the noun itself (masc. *Anglais* 'Englishman'; vs. feminine *Anglaise*) - but the pre-vocalic determiners found in (1.a) and (1.c) are the same as those required within feminine noun phrases, where hiatus is not an issue:

- (3) a. *la belle femme* [la bEl fam]                      'the handsome woman'  
       b. *la femme belle* [la fam bEl]                      'the woman (who is) handsome'  
       c. *cette théorie nouvelle* [sEt teOri nuvEl]            'this theory (which is) new'  
       d. *cette nouvelle théorie* [sEt nuvEl teOri]            'this new theory'

The point developed in this section will be that the liaison form of the masculine adjective is computed by reference to constraints promoting lexical conservatism. These constraints require that every element of the liaison allomorph - or more generally every element of any novel allomorph - possess a lexical precedent in some listed allomorph.

Among the liaison allomorphs of masculine adjectives, there exist listed forms, which will not concern us here (cf. Tranel 1981, 1990). Our focus will be on adjectives that do not normally occur in prenominal position, hence do not possess a generally known liaison allomorph, or else are infrequent across the board and thus unlikely to occur in potential hiatus positions, because of the relative lack of frequency of vowel-initial nouns<sup>3</sup>. For these adjectives, the speakers could not have memorized a

---

<sup>2</sup>Cf. Kammans (1950: 241) among many other codifiers of French liaison "La liaison est interdite d'un groupe rythmique à l'autre".

<sup>3</sup>There is no other context - aside from the prenominal position - where a French adjective could generate hiatus within the accentual phrase, hence no other context where hiatus avoidance would be as urgent.

The policy followed here of concentrating on the liaison form of adjectives lacking a listed liaison allomorph diverges from that of earlier writers (e.g. Tranel 1981) who focus primarily on enumerating the uses of listed liaison forms like

solution to the hiatus problem: if the problem arises, they will have to project such a solution, based on their general understanding of French phonology. We focus on these because the formation of their liaison allomorphs is more likely to illustrate general principles rather than idiosyncratic lexical properties. Examples of this sort are the three adjectives below, all of which are uncommon or unknown in prenominal position:

- (4)
- |                                     |  |
|-------------------------------------|--|
| (a) <i>sot, sotté</i> 'silly':      | listed allomorphs: [so] (masc.), [sOt] (fem.)        |
| (b) <i>vain, vaine</i> 'vain':      | listed allomorphs: [vE)§] (masc.), [vEn] (fem.)      |
| (c) <i>dernier, dernière</i> 'last' | listed allomorphs: [dEÂnje] (masc.) [dEÂnjEÂ] (fem.) |
| (d) <i>dodu, dodue</i> 'plump':     | listed allomorph: [dOdy] (masc. and fem.)            |

When asked to form noun phrases in which these adjectives are preposed to a V-initial noun, French speakers compute an answer based on the following considerations:

- (5) Factors in the formation of the liaison allomorph of the masculine:

- Phonological: Avoid hiatus<sup>4</sup>.
- Syntactic: Mark gender agreement with masculine head N.  
Avoid appearance of gender conflict betw. N and Adj.
- Lexical: Avoid proliferation of allomorphs. Use existing forms.

Item (6) lists some conceivable responses to this task, classified by their degree of linguistic conservatism<sup>5</sup>:

- (6) Three reactions to the task of forming liaison allomorphs to (4)

---

cet 'this' or the occurrence of liaison in lexicalized phrases such as *divin Enfant* 'divine child' It should also be noted that we owe to Tranel (1990) the important distinction between suppletive and regular liaison

<sup>4</sup>The existence of *liaison sans enchainement* (Encrevé 1988, Klein 1995) as well as the interesting paradigm of liaison in right dislocated environments presented by Tranel 1992 make it clear that the phonotactic condition triggering the appearance of the liaison consonant is hiatus avoidance rather than the need for onsets.

<sup>5</sup> The term *conservatism* employed in this description of contemporary French has nothing to do with the historical priority of one pattern over the other but rather strictly with the issue of lexical conservatism, the preference for using words possessing generally known precedents. The history of French nasal vowels in liaison has been sketched by Tranel (1981) and reveals that some of the forms labelled here as *moderately conservative* is in fact the earliest attested ones: thus *anci[E)n] ami* 'old friend' occurs earlier than the ultra-conservative *anci[En] ami*.

	ultra-conservative	moderately conservative	innovative	gloss
(a)	sOt ami	sot ami	sot ami	"silly friend"
(b)	vEn EspwaR	vE)§n EspwaÂ	vE)§n EspwaÂ	"vain hope"
(c)	dEÂnjEÂ Om	dEÂnjeÂ Om	dEÂnjeÂ Om	'last man'
(d)	dOdy elefA)	dOdy elefA)	dOdy t elefA)	"plump elephant"

The ultra-conservative pattern selects for the pre-V position an allomorph that is strictly identical to a listed allomorph. If a listed allomorph exists that ends in a C, then that one is selected, to avoid hiatus. This means, in the case of *sot*, *dernier* and *vain* that the pre-V form of the masculine is strictly identical to the feminine adjective, in apparent violation of gender agreement. If, as in the case of *dodu*, such an allomorph does not exist, hiatus is violated.

The moderately conservative pattern selects a form that agrees with the masculine in the quality of the final vowel and with the feminine in the presence of a final C, to avoid hiatus. This pattern is also lexically conservative, to the extent that the hiatus breaking C appears only when it has a lexical precedent in the feminine form. In this case however, the pre-V allomorph is being generated by using two distinct lexical reference points: the listed masculine form and the listed feminine, each of which contributes a property to the final product. The result then is a mixed allomorph that is in effect novel, since its last syllable as a whole is not found among the listed forms. The virtue of the mixed allomorph is that it avoids hiatus while also signalling its connection to the masculine: its stressed vowel is identical in quality to that of the listed masculine<sup>6</sup>.

The innovative pattern will insert a C to break hiatus regardless of whether this C has a lexical precedent among the listed allomorphs.

---

<sup>6</sup>One clear argument supporting our policy of concentrating on adjectives that lack listed liaison allomorphs is that very frequent prenominal adjectives, like *bon* [bO§] 'good', fem. *bonne* [bOn], possess liaison forms whose phonological behavior deviates from that of the infrequent majority: for instance, speakers who generally follow the moderate conservative pattern (hence [vE§)n EspwaÂ]) may nonetheless use an oral vowel, like the ultraconservatives, in phrases such as *bon ami* [bOn ami]. This is because the sheer frequency of prenominal *bon* allows more speakers to become aware of the prescriptive standard, which in this case is denasalization. The same may hold for lexicalized phrases like *prochain arrêt* 'next stop' or *divin Enfant* 'divine child'

The actual situation attested in French is as follows: the ultra-conservative pattern is the only one sanctioned by prescriptive grammar (Kammans 1950, Fouché 1959, Arrivé, Gadet, Galmiche 1986). Recent work on French allomorphy (e.g. Perlmutter 1996) recognizes only this set of idiolects. But the moderately conservative pattern is also attested and in fact predominates with the younger generation: Prunet 1987, Encrevé 1988:204ff, Morin 1991, Tranel 1981, 1987, 1992, F.Dell p.c. The impression one gathers from soliciting data of this sort from educated French speakers is that the nasal vowel in forms like [vE]n *espoir* would be much more wide-spread, were it not for the prescriptive pressure, which supports the oral form [vEn]. Innovative forms like *dodu-t-éléphant* are attested, dialectally (Tranel 1981) and perhaps among the very young, but have always been stigmatized. Such forms are common enough to have technical terms associated with them: insertion of non-etymological (or in our terms, lexically unprecedented) [z] is referred to as "velours". Insertion of a non-etymological [t] is labelled "cuirs" or "pataquès" (Arrivé, Gadet, Galmiche 1986, Tranel 1987). Thus *donne-moi [z] en* 'give me some' or *reviendra-z-à Pâques* 'will return at Easter' is a case of velours, while *donne-moi [t] en* is an instance of pataquès. This indicates that innovative speakers may be more common than the literature suggests, but in the absence of complete records about their speech I will simply note their existence. Finally, under circumstances to be discussed below, French speakers of either the moderate or ultra-conservative sort, are reduced to accept hiatus in accentual phrase medial position, even when the adjective in question does possess a listed allomorph containing a final consonant: *fin expérimentateur* 'subtle experimenter' can be produced as [fE] EkspeÂimA)tat''Â], with hiatus, even by speakers who reject the hiatus in *prochain arrêt* 'next stop' [pÂoSEn aÂe], not \*[pÂoSE] aÂe]. Thus it is likely that the three categories of speakers recognized above must be augmented by yet others, in order to characterize differences in the tolerance of hiatus.

A critical detail concerning the ultra-conservatives must be settled now. For this class of speakers we attribute the orality of the adjective's vowel in *vain espoir* and the lax quality in *sot ami* to the influence of the related feminine form, [vEn] and [sOt] respectively. But how do we know that it is the feminine that is responsible for the vowel quality of the liaison masculine? We compare the liaison possibilities of homophonous or similar forms which differ only in the composition of their paradigm of listed forms: for instance *bien* 'well', *rien* 'nothing' possess a unique listed allomorph, the phrase-final form [bjE]), [ÂjE]), whereas *sien* 'his/hers-masc.' [sjE]) relates to *sienne* 'his/hers-fem.' [sjEn]. The liaison properties of these forms differ correspondingly: ultraconservatives may produce phrases like *un sien ami* [') sjEn ami] 'a friend of his/hers', adopting the feminine oral vowel, but will maintain the nasalized vowel in *bien aimable* [bjE])n emabl] 'very kind' or *rien à faire* [ÂjE])n a fEÂ], 'nothing to do' (cf. Fouché 1959 on a clear statement to this effect). Compare also *bon* [bO]) 'good', fem. [bOn], liaison masc. *bon ami* [bOn ami], with possessive adjectives like *mon* 'my' [mO]). In the possessive, the suppletive form *ma* is used as a feminine, thus pre-empting regular \*[mOn]: the

absence of \*[mOn] has the consequence that in liaison contexts the masculine *mon* maintains its nasalized vowel for all speakers: *mon ami* [mO§n ami]. Therefore, the correspondence between oral vowels in the liaison form and nasalized vowels in the citation form is limited to paradigms in which the oral vowel has an independent lexical precedent, typically in the form of a feminine. The same may hold for the tense-lax alternation: thus ultra-conservative [sOt ami] owes its lax vowel to the feminine [sOt], but the invariable adverb *trop* 'too (much)' [tÂo] maintains a tense vowel in liaison, since no \*[tÂOp] form exists: [tÂop Eme] 'loved too much'<sup>7</sup>. Similarly, *dernier homme* 'last man' is realized as [dEÂnjEÂ Om] by ultras, under the influence of the feminine [dEÂnjEÂ], whereas invariant infinitive forms such as *aimer* 'to love' [Eme] may give rise to liaison variants (based on the orthographic r) but primarily with a tense vowel, e.g. *aimer ainsi* 'to love in such a way' [EmeÂ E§si]. This is because no citation form [EmEÂ] exists in this case<sup>8</sup>.

A further observation supports the role of the feminine forms in the vocalism of masculine liaison: plural nouns and adjectives employ liaison forms in which [z] - the plural marker - is used as hiatus buffer: *bons amis* [bO§z ami] 'good friends-masc.' and *bonnes amies* [bOnz ami]. The principles regulating the occurrence of this [z] are different from the ones discussed here: a high priority factor in this case is the overt expression of number, which favors the realization of [z] even when hiatus is not an issue, as in *bonnes amies*. Since [z] is thus independently available in the plural, the use of the feminine consonant is unnecessary to break hiatus. And since the feminine consonant is not being used, the feminine vowel is not used either, hence masculine plural phrases like [bO§z ami], [dEÂnjez Om] - with the masculine vowels [O§], [e] of [bO§], [dEÂnje] - rather than the feminine vowels [O], [E] of [bOn], [dEÂnjeÂ]. Once again we conclude that the use of feminine vowels in masculine liaison forms is a direct and exclusive consequence of the use of feminine consonants to break hiatus.

What have we learned so far ? The first conclusion we reach is that lexical conservatism plays a role in all adult varieties of standard French: no adult speaker of the standard language will consistently insert a consonant completely lacking in paradigmatic - or at least orthographic - support in order to break hiatus in phrases like *odou éléphant* or *joli enfant*. The hiatus breaking consonants always possess a lexical precedent in the shape of the corresponding feminine or in the orthographic representation for invariant forms such as *trop*, *rien*. The phenomenon is referred to by

---

<sup>7</sup>Tranel (1987:174) records the infrequent pronunciation [tÂOp Eme], but this may be due to factors independent from liaison since *trop plein* 'too full' is also recorded as [tÂOp lE§] (Harrap's Dictionary 1977).

<sup>8</sup>Tranel (1987) is alone among my sources to mention at all the option of lax vowels in the liaison form of infinitives. Neither Fouché's 1959 extensive listing of liaison forms nor my other sources, including the French speakers consulted, accept lax vowels at the end of infinitives.



Perlmutter (1996) as Lexical Sourcing and characterized as follows: "The input set [of allomorphs] is supplied by the lexicon." In the context of Perlmutter's analysis of pre-vocalic allomorphy, which recognizes only the ultra-conservative pattern, this statement must be interpreted as follows: the adjectival candidates considered in the realization of [A N] phrases must be strictly identical, in their entirety, to a listed allomorph of the adjective.

### 3.2. A new analysis of Lexical Conservatism effects in French liaison

We have noted however that the mode of implementation of what Perlmutter calls Lexical Sourcing differs across generations: the younger generation uses pieces of the listed masculine and feminine allomorph to cobble together a novel allomorph that avoids both hiatus and the more severe violation of gender concord inherent in phrases like [sOt ami]. Perlmutter's analysis - and that of Tranel 1996 - is based on the assumption of Lexical Sourcing and the interaction of two constraints (Onset >> Gender Concord) but this mechanism cannot characterize the behavior of the moderately conservatives (e.g. [sot ami] ) and the difference between their speech patterns and those of ultra conservative speakers (e.g. [sOt ami]). Varying the ranking of Onset relative to Gender Concord is insufficient to yield the attested variation: under Onset >> Gender Concord, plus the assumption of Lexical Sourcing, we can describe the ultra-conservative [sOt ami], [vEn EspwaÂ], whereas under Gender Concord >> Onset we can describe only [so ami], [vE§) EspwaÂ]. We will also observe below that Lexical Sourcing is the incorrect assumption even for the ultra-conservative speakers. We must therefore model the notion of lexical conservatism in a way that departs from earlier analyses.

The informal terms *lexical conservatism* and *lexical reference term* correspond to grammatical conditions that can be differently prioritized. My proposals on this score are introduced in the form of schemas by reference to which we can generate correspondence constraints. (On correspondence theory see McCarthy and Prince 1995; on the idea of correspondence between surface forms see Benua 1995, Burzio 1997 and refs. there, Downing 1997, Flemming 1995, Kager 1996,1997, Kenstowicz 1995, McCarthy 1995, Orgun 1995, Ito & Mester 1996, Steriade 1995, 1996). I assume first that there exists a family of grammatical conditions that require any form to be identical in various respects to **some** (non-specific) listed allomorph, whether or not that allomorph carries appropriate morphosyntactic features. A general schema for writing such conditions is (7).

(7) **Ident P:** element  $x$  of the target allomorph of morpheme  $\mu$  has a correspondent  $x'$  in some listed allomorph of  $\mu$  and is identical wrt P (a phonological property) to  $x'$ .

(7) refers to the target allomorph: this is the form that is being generated, various candidates for which are being evaluated. An **Ident P** condition will require identity between this form and some

listed allomorph of the same morpheme with respect to some phonological property. What is the range of such properties is an interesting question that cannot be addressed now: it will suffice to note that the most commonly documented effects of surface analogy involve segmental identity of the morpheme's edges and identity of the string carrying cues to stress, such as the vowel quality, quantity, syllabicity and pitch contour within the stressed syllable (Steriade 1996). Both of these effects are found in the French data: the stressed V quality and the quality of the final C are subject to strict paradigmatic identity conditions. It is these properties that are extended from listed allomorphs such as the citation masculine or the citation feminine to the masculine liaison form. The two instances of **Ident P** we refer to in the analysis of French appear below:

- (8) **Ident (C#)**: The last C in the target allomorph of morpheme  $\mu$  has a correspondent C' in some listed allomorph of  $\mu$  and is featurally identical to C'.
- (9) **Ident (V')**: The stressed V in the target allomorph of morpheme  $\mu$  has a correspondent V' in some listed allomorph of  $\mu$  and is featurally identical to V'.

The ranking **Ident (C#)** >> \*Hiatus is sufficient to characterize the difference between *odou éléphant* (with obligatory hiatus) and *vain espoir* (where hiatus is disfavored or impossible):

(10) listed allomorphs: [dody]

	Ident (C#)	>>	*Hiatus
[dOdy] elephant			*
[dOdy t] elephant	*!		

(11) listed allomorphs: [vEn], [vE)§]

	Ident (C#)	>>	*VV
[vEn] espoir			
[vE)§n] espoir			
[vE)§] espoir			*!

Constraint rankings discussed below will decide the winner in (11), depending on the dialect: what matters for the moment is the mechanism that excludes hiatus in \*[vE)§] *espoir*. **Ident (V')** is undominated in French, but fails to conflict with any of the conditions we will investigate here.

A second class of correspondence conditions (12) requires identity for some phonological property between the target allomorph and some listed allomorph, provided that the two share one or more specified morphosyntactic properties:

- (12) **Ident P under morphosyntactic identity:** element  $x$  of the target allomorph for morpheme  $\mu$  has a correspondent in  $x'$  in a listed allomorph of  $\mu$ ,  $a(\mu)$ , and is identical wrt  $P$  to  $x'$ , and the target allomorph share a morpho-syntactic feature  $Q$ .

For example: if  $a(\mu)$  and the target allomorph share the gender feature [masculine], then they must be identical with respect to the quality of final C; or they must be identical in the quality of their stressed vowels. The class of conditions in (12) - phonological identity under morphosyntactic identity - will be used here to model the phonological encoding of grammatical agreement. The conditions in (12) can also be used to refer to the phonological encoding of subcategorization for syntactic features. In general we shall see that it is possible and desirable to use such conditions to dispense entirely with the notion of *base of affixation*. Two French instantiations of (12) will be needed:

- (13) **Ident (C#) under morphosyntactic identity** (abbreviated **Ident (C# ms)**): Two words are C-distinct iff one ends in a consonant  $C$  and the other fails to end in the same consonant  $C$ . If the target allomorph of morpheme  $\mu$  is identical in morphosyntactic features to a listed allomorph  $a(\mu)$ , then the target allomorph and  $a(\mu)$  are not C-distinct.

This condition is violated by pronunciations such as [sot ami], [sOt ami], [vEn espwaÂ] and : it is violated in virtue of the ranking **\*Hiatus >> Ident (C# ms)**. For instance:

(14)

	*Hiatus	>>	Ident (C#, ms)
[vEn] espoir			*
[vE)§] espoir	*		

That **Ident (C#, ms)** is an active constraint in French will become apparent below. The second French instantiation of (12) is a similar condition involving the vocalic quality in stressed syllables:

- (15) **Ident (V') under morphosyntactic identity** (abbreviated **Ident (V', ms)**): The stressed  $V$  in the target allomorph of morpheme  $\mu$  possesses a correspondent  $V'$  in a listed allomorph  $a(\mu)$

and is featurally identical to V', if a( $\mu$ ) is identical in morphosyntactic features<sup>9</sup> to the target allomorph.

The description of the moderate conservatives, who prefer [sot] *ami*, [vE)§n] *espoir*, will have to invoke (15): the target allomorphs below are masculine, hence morphosyntactically identical to the masculine listed allomorphs [so] and [vE§]) respectively. Given this, the constraint **Ident (V', ms)** penalizes differences in vowel quality between these listed allomorphs and the target allomorph.

(16)

Ident (V', ms)		
[sot] ami		
[sOt] ami	*	
=====		
[vE)§n] espoir		
[vEn] espoir	*	

A further type of condition will be needed to characterize the preference against mixing properties borrowed from different listed allomorphs.

(17) **Ident P (to-allomorph<sub>i</sub>) if ident Q (-to-allomorph<sub>j</sub>)** (abbreviated **Ident P if Q**): if the target allomorph is identical wrt P to some listed allomorph a( $\mu$ ), it is also identical wrt the phonological property Q to a( $\mu$ ).

For instance, in the speech of ultra-conservatives the last C and stressed V quality must come from the same listed allomorph: in terms of (17), the situation can be characterized by requiring that, if some a( $\mu$ ) and the target allomorph have identical last consonants, they must have identical vowel quality. The French ultraconservatives need the final C of *sot* in *sot ami* to block hiatus, but (17) prohibits them from using the feminine C without also adopting the stressed vowel quality of the feminine.

(18) **Ident (V' if C#)**: The target allomorph of morpheme  $\mu$  and a listed allomorph a( $\mu$ ) have identical stressed V's if they are not C-distinct.

---

<sup>9</sup> Strictly speaking we only need to mention grammatical gender identity here: but the condition can be generalized to require identity for all morphosyntactic features without apparent harm.

The ranking **Ident (V' if C#)** >> **Ident (V', ms)** characterizes the choice made by ultra conservatives between [sOt ami] and disfavored [sot ami], [vEn EspwaÂ] and rejected [vE)n EspwaÂ].

(19) listed allomorphs: [vEn], [vE)§]

	Ident ( V' if C#)	>>	Ident (V', ms)
[vEn] espoir			*
espoir	*!		

The opposite ranking **Ident (V', ms)** >> **Ident (V', if C#)** will characterize the moderates. A second use of the **Ident P if Q** conditions will be to characterize the fact that identity between allomorphs for some phonological property engenders identity for yet others: the more similar two allomorphs are to begin with, the more similar they become. Conversely, lack of similarity on some dimension is sufficient to block the analogical extension of unrelated properties.

A possible objection to (17) is that this schema vastly increases the set of possible constraints, since any arbitrary pairing of P and Q can in principle give rise to an **Ident P if Q** condition. One solution to this is that, in the French case at hand, the particular pairing of P and Q we require happens to converge on a unified property, the global quality of the last, accented demisyllable of the French word, the rhyme of a French word. We may conclude from this that the critical condition takes the form in (20) - an instantiation of the **Ident P** class in (7) - in which P refers to the composition of the word's rhyme: it is the rhyme of the word, in its entirety, that must possess a lexical precedent in some listed allomorph, for the ultra-conservative speakers.

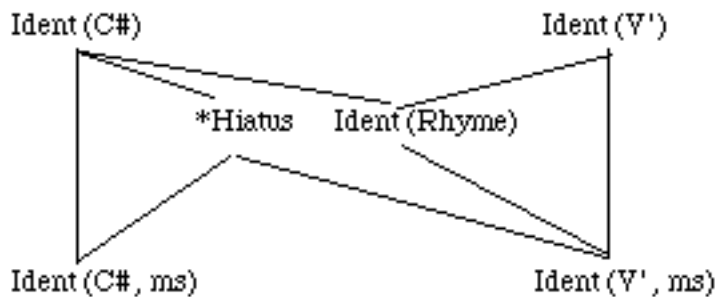
(20) **Ident Rhyme:**

The word's rhyme (in French: the string beginning with the last non-schwa vowel and ending with the last segment of the word<sup>10</sup>) in the target allomorph of  $\mu$  must be identical to the rhyme of some listed allomorph of  $\mu$ .

Under this interpretation, no new type of condition is necessary here. We retain however the possibility of **Ident P if Q**, the necessity for which may re-emerge in different circumstances. We note further that the ranking between **Ident V'**, **Ident C#**, on the one hand, and **Ident Rhyme**, on the other, has the following property: the strings for which identity is mandated in virtue of either **Ident V'** or **Ident C#** are proper substrings of the word rhyme. Therefore any candidate that violates **Ident C#** necessarily violates **Ident Rhyme**; and similarly for candidates violating **Ident V'**; while the converse is not true. Any constraint hierarchy in which the ranking relation between these constraints is distinct from **Ident C#, Ident V' >> Ident Rhyme** will fail to provide any evidence for either **Ident C#** or **Ident V'**. Since however we claim that **Ident C#** and **Ident Rhyme** are active in French, we are therefore committed to the ranking **Ident C#, Ident V' >> Ident Rhyme**.

The rankings characterizing the speech of ultra-conservative French speakers are given below:

(21) Constraint rankings for ultra-conservatives:



The summary of ranking arguments is as follows: **Ident (C#)** >> **\*Hiatus** is needed to block non-etymological C insertion (cf. (10)). **\*Hiatus** >> **Ident (C#, ms)** is necessary to enforce hiatus resolution through selection of a C-final listed allomorph in forms like *vain espoir, sot ami* (cf. (11)). By transitivity, the ranking **Ident (C#)** >> **Ident (C#, ms)** obtains as well. The necessity for ranking **Ident (Rhyme)** and **\*Hiatus** above **Ident (V', ms)** is shown below:

<sup>10</sup> We provide a non-syllabic definition of the word's rhyme in order to characterize both varieties of liaison, with and without enchainement (i.e. resyllabification): cf. Encrevé 1988.

(22)

	Ident (Rhyme), *Hiatus	>>	Ident (V', ms)
[vEn] espoir			*
[vE§] espoir	*!		
[vE§n] espoir	*!		

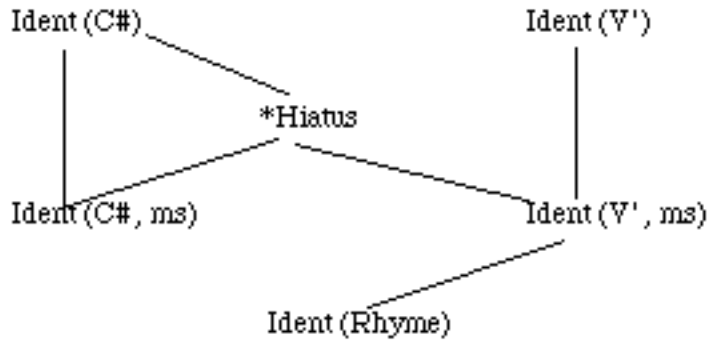
The need for the remaining rankings (**Ident (C#)** >> **Ident (Rhyme)** and **Ident (V')** >> **Ident (Rhyme)**) is based on the the discussion above. The tableau in (23) illustrates the fact that the same ranking accounts for the pronunciation of vowel laxing and nasalization in masculine liaison contexts for this class of dialects.

(23) listed allomorphs: [so], [sOt]

	Ident (Rhyme), *Hiatus	>>	Ident (C#, ms), Ident (V', ms)
[sOt] ami			* *
[so] ami	*!		
[sot] ami	*!		*

The dialect of the moderate conservatives is characterized by re-ranking **Ident (Rhyme)** below **Ident (V', ms)**. We must further assume either that **\*Hiatus** outranks **Ident (Rhyme)** or else that **\*Hiatus** outranks **Ident V', ms**, hence, by transitivity, **Ident Rhyme** as well, for the opposite ranking **Ident Rhyme** >> **\*Hiatus** will predict hiatus for all phrases like *vain espoir*, *dernier homme*, *sot ami.*, while lack of ranking between these constraints will predict variation between hiatus and its absence. The hierarchy we adopt in (24) minimizes the ranking differences between ultras and moderates: **\*Hiatus** continues to outrank **Ident (V', ms)** and only the latter's position relative to **Ident Rhyme** changes.

(24) Rankings for moderate conservative dialects:



Since the relative ranking of the top constraints is invariable in French, we illustrate the effects of ranking **Ident (V', ms)** and **\*Hiatus** relative to **Ident (Rhyme)**:

(25) Listed allomorphs: [sOt], [sOt]

	*Hiatus	Ident (V', ms)	>>	Ident (Rhyme),	Ident (C#, ms)
[sOt] ami		*!			*
[so] ami	*!				
[sot] ami				*	*

Listed allomorphs: [vEn], [vE]§

	*Hiatus	Ident (V', ms)	>>	Ident (Rhyme),	Ident (C#, ms)
[vEn] espoir		*!			*
[vE]§ espoir	*!				
[vE]§n espoir				*	*

The constraints discussed so far will be able to describe other patterns of concord and hiatus resolution in French. For the rankings **Ident (C#) >> Ident (C#, ms)**; **Ident (V') >> Ident (V', ms)** and **Ident (V'), Ident (C#) >> Ident (Rhyme)** ranking is intrinsic. Hence, these pairs of constraints will either be ranked as demonstrated in the dialects discussed above or else they will fail to provide evidence for the less specific of the two constraints involved in a pair. Ranking variation may however be expected for other pairs of constraints and is indeed found. Thus the innovative class of patterns outlined earlier (*pataquès* or *velours*) is obtained under **\*Hiatus >> Ident (C#)**. The possibility of hiatus in phrases such as *sot ami* (mentioned by Prunet 1987, who favors hiatus in phrases of this sort, containing an adjective that is normally postnominal) is characterized by the ranking **Ident (Rhyme), Ident (V', ms) >> \*Hiatus**.



(26) listed allomorphs: [so], [sOt]

	Ident (Rhyme), Ident (V', ms) >>	*Hiatus
so ami		*
sot ami	*!	
sOt ami		*!

One must then explain why speakers like Prunet, who prefer [so] *ami*, accept *bel homme*, *cet homme* etc. with the feminine consonant used as a hiatus buffer in phrases where the adjective is normally prenominal, as against impossible *\*beau homme*, *\*ce homme*, etc. The explanation must involve the difference between suppletive and normal liaison: frequent adjectives that are predominantly or exclusively prenominal may possess listed masculine liaison allomorphs, precisely because of their frequent use in situations potentially leading to hiatus. Thus *bel*, *cet* count as masculine forms to be used in hiatus contexts, not (or no longer) feminine forms borrowed for the purpose of hiatus avoidance. The orthography, which distinguishes the feminine *belle*, *cette* from the liaison masculine *bel*, *cet* appears to recognize this point. These suppletive masculine liaison allomorphs can, like any other form of unproductive morphology, be learned only through overt and repeated exposure: this accounts for the fact that the adjectives belonging to this class are both frequent and typically prenominal<sup>11</sup>. We conclude then that, the preference expressed by Prunet (1987) for [so] *ami*, with hiatus, can coexist, under our analysis, with structures like *bel homme*, under the ranking **Ident (Rhyme), Ident (V', ms) >> \*Hiatus**, and on the assumption that common prenominal adjectives like *beau /bel*, possess listed prevocalic allomorphs.

This exhausts all empirically observable effects of re-ranking the constraints proposed here.

### 3.3. An abstract isogloss: orality and laxing in adjectival liaison

Implicit in the discussion so far is the fact that speakers' preference to denasalize the final vowel in liaison contexts like *vain espoir* will correlate with the preference for a lax vowel in phrases like *sot ami*. Conversely, the preference for nasal vowels in *vain espoir* correlates with a preference for tense vowels in *sot ami*. This correlation has not, to my knowledge, been mentioned in the literature, although many writers on liaison present their data in a way that suggests that it holds. I note first that the core of our proposal - the existence of lexical conservatism conditions **Ident P** and of the **Ident (P, ms)** conditions, the substitutes for base-identity statements - is independent of this correlation. But if it does hold, as it seems to, this adds further support for **Ident P**, since approaches lacking this

<sup>11</sup> Postnominal adjectives like *sot* lack listed liaison allomorphs because any hiatus they might generate will typically occur across the accentual phrase boundary: in such cases hiatus is normally tolerated.

class of conditions lack a direct expression for the generalization noted: the segmental quality of the word's rhyme must possess a lexical precedent.

The first element supporting our correlation is Fouché's 1959 description of the cultivated standard pronunciation of French in his time. Fouché's dialect corresponds exactly to the ultra-conservative pattern described above. Fouché states first (p.435) that the standard requires orality in the vowel of liaison adjectives whose citation masculines end in nasal vowels: thus *certain auteur* [sEâtEn Ot'Â] 'certain author' with the feminine [En] ending, as against \*[E§)n]. In the same passage, Fouché indicates that adjectival liaison also has an effect on the tense/lax quality of vowels in cases like *léger* 'light', *premier* 'first'. These are produced with a final tense vowel in the citation masculine [leZe], [pÂ'mje], but with a lax vowel in the feminine and the liaison masculine: [leZEÂ] (fem.) [leZEÂ A)nÁi] (masc.) 'slight trouble', [pÂ'mjEÂ] (fem.), [pÂ'mjEÂ Om] (masc.) 'first man'. This is the front counterpart of the tense/lax alternation observed earlier with o in the case of *sot ami* [sOt ami]. V.Ivanov (p.c. 1997) informs me that the Moscow French dialect - the French spoken natively or near-natively by educated Russians - is also of the ultra-conservative sort, in that it denasalizes and laxes the vowels of liaison masculines.

Second, Prunet (1987), in his discussion of adjectival liaison, points out that his dialect gives preference to nasal vowels in phrases like *certain auteur* [sEâtE§)n Ot'Â] and tense vowels in phrases like *sot ami* [sot ami]. The pattern described by Prunet can be therefore identified as moderately conservative.

Third, the speakers of Parisian and Quebecois French that I have interviewed in the course of this study have spontaneously produced sets of forms identifiable as either ultra or moderately conservative, rather than mixed sets: they have not spontaneously produced either mixed sets like {[sEâtE§)n Ot'Â], [pÂ'emjEÂ Om]}, or {[sEâtEn Ot'Â], [pÂ'emjeÂ Om]}. Most speakers are aware that alternate ways of doing liaison exist and few are willing to declare ungrammatical any one of the forms cited but each one's own spontaneous productions seem to fall squarely into one or the other of the two patterns we have described<sup>12</sup>.

### 3.4. Height effects: further evidence for Ident (P, ms) conditions

---

<sup>12</sup>One possible exception is Tranel's own dialect, judging from his discussion of liaison (1987): Tranel seems to consider tense vowels as more common in cases like *premier homme*, and mentions the lax alternative only in passing, whereas he seems to favor the oral vowel in forms like *certain auteur*. However, no explicit statement is provided in his description that would disconfirm the correlation we propose.

A different detail of French liaison lends further support to the class of conditions in (12), **Ident P under morphosyntactic identity**. Nasalized vowels are subject in French to neutralization of their height distinction: there are no nasalized [i], [u]. The existing nasalized vowels contrast only as to front/back and round/unrounded. The result is that adjectives like *fin/fine* ([fE§], [fin]) possess listed allomorphs in which the masculine vowel differs in both height and nasality from the feminine vowel. Perceptually, the distance between surface allomorphs is greater here than in cases like *vain/vaine* ([vE§], [vEn]), whose vowels differ primarily in nasality.

This difference in perceptual distance between allomorphs has surprising and interesting effects: feminines like *fine*, unlike *vaine*, cannot be used as the pre-vocalic allomorphs of the corresponding masculine. This has been noted by Tranel (1981). The generalization is blurred by the fact that lexicalized phrase *divin enfant* 'divine child' is known to most French speakers in its archaic pronunciation [divin A)fA], used in the Catholic version of a popular Christmas song. (The Protestant versions of this hymn have [divE)§n A)fA] according to Encrevé 1988.) Some speakers extend the use of masculine [divin] to other phrases such as *divin Homère*, *divin amour*. but in such cases the alternate form [divE)§n] is either preferred or accepted. Aside from this case, the use of the feminine is impossible or disfavored in all cases where the feminine and masculine differ in the height of their last (stressed) vowels. Adjectives like *fin/fine* are compared in (27) to adjectives like *vain/vaine*, whose vowels do not differ (or not significantly so) in height<sup>13</sup>.

(27) **Height differences between masculine and feminines and effects on liaison**

(Key: M = contains masculine ending; F = word contains feminine ending;  
MF = liaison allomorph contains masculine V and feminine C)

Spelling	Pronounced	Gloss	Feminine form
<i>certain épisode</i>	<b>sEÂtE§)n</b> epizOd (MF) <b>sEÂtEn</b> epizOd (F) <b>sEÂtE§)</b> epizOd (M)	'certain incident'	<i>certaine</i> [sEÂtEn]
<i>prochain épisode</i>	<b>pÂoSE§)n</b> epizOd (MF) <b>pÂoSEn</b> epizOd (F) <b>pÂoSE§)</b> epizOd (M)	'next incident'	<i>prochaine</i> [pÂoSEn]

<sup>13</sup> Thanks to F.Dell for suggesting some of the phrases in (15) and for help with the generalization suggested in the text.

<i>commun accord</i>	<b>kom''n</b> akOÂ (MF) * <b>komyn</b> akOÂ (F) <b>kom''</b> akOÂ (M)	'mutual agreement'	<i>commune</i> [komyn]
<i>aucun espoir</i>	<b>Ok''</b> EspwaÂ (MF) <b>Ok''n</b> EspwaÂ (M) * <b>Okyn</b> EspwaÂ (F)	'no hope '	<i>aucune</i> [Okyn]
<i>divin archer</i>	<b>divE§n</b> aÂSe (MF) * <b>divin</b> aÂSe (F) <b>divE§</b> aÂSe (M)	'divine archer'	<i>divine</i> [divin]
<i>fin expérimentateur</i>	<b>fE§n</b> EkspeÂimA)tat''Â (MF) * <b>fin</b> EkspeÂimA)tat''Â (F) <b>fE§</b> EkspeÂimA)tat''Â (M)	'subtle experimenter'	<i>fine</i> [fin]

This restriction on the use of the feminine as a liaison allomorph of the masculine is also readily expressible in the language proposed here, by reference to the class of conditions in (12).

(28) **Ident ( $\pm$ high in V')** under morphosyntactic identity: The stressed vowel of the target allomorph of morpheme  $\mu$  is identical wrt  $[\pm$  high] to its correspondent in a listed allomorph  $a(\mu)$ , if  $a(\mu)$  and the target allomorph share morphosyntactic features.

The reference to stress is added in (28) in order to allow alternations for unaccented morphemes such as the prefix *in-* (e.g. *inattendu* with [in] vs. *impossible* with [E§]). It is conceivable that other means of accounting for the [E§] / [in] alternation exist.

The condition in (28) will prohibit use of the feminine vowel in masculine liaison allomorphs, when this vowel differs in height from the stressed vowel of the listed masculine. If undominated, the effect of (28) will be to yield two possible pronunciations for phrases like *divin archer*: [divE§] aÂSe] or [divE]§n aÂSe], depending on the relative ranking of \***Hiatus** and **Ident (Rhyme)**. Two ranking possibilities are illustrated below; the results of a third - no ranking between \***Hiatus** and **Ident [ $\pm$ high]** - are non-distinct from those of the ranking in (29.a).

(29) a. Deriving [divE)n] *archer*

	Ident ( $\pm$ high in V', ms) >>	*Hiatus >>	Ident (Rhyme)
divin archer	*!		
divE)n archer			*
divE) archer		*!	

b. Deriving [divE)] *archer*

	Ident ( $\pm$ high in V', ms), Ident (Rhyme) >>	*Hiatus
divin archer	*!	
divE)n archer	*!	
divE) archer		*

What guarantees the non-occurrence of [divin] *archer* is the undominated status of **Ident** [ $\pm$ high]. Note that the ranking variation in terms of which we describe the difference between moderates and ultra conservatives is irrelevant to this data.

It appears that height, the source of most basic vowel contrasts, is the property with respect to which the target allomorph must not differ from its listed, gender-appropriate counterpart. Yet another way of stating this is that the masculine feature of the adjective must be phonologically encoded through identity between target and a known listed masculine: and while identity between target and the listed masculine allomorph is obligatory for properties like the height of stressed vowels, identity for other properties, such as the quality of the final consonant, is negotiable. Clearly a better version of the analysis presented here will be one in which the ranking of correspondence for different features is made to follow from independent considerations, such as the relative perceptual salience of the contrasts being generated. A possible implementation of this idea - for much simpler forms of correspondence - is found in Jun (1995) and Steriade (1995).

## 4. Alternatives

### 4. 1. A derivational analysis

The impossibility of forms like [kOmyn] *accord* - with use of the feminine vowel height in the masculine liaison allomorph - provides a clear argument against a rule-based analysis of this data. Many liaison facts can be described by formulating a set of ordered rules mapping underlying representations onto surface forms. Such an approach - modelled on Dell's 1974 and Prunet's 1987 analyses - will consist of rules that nasalize vowels next to tautosyllabic nasal consonants, rules that lower nasalized vowels, rules that resyllabify prevocalic consonants and delete liaison consonants when they are not resyllabified. The ordering of such rules is expected to yield a characterization of the dialect

differences discussed. In this case, however, mere rule ordering differences cannot do this<sup>14</sup>. Note that in the speech of moderates, the vowel nasalization rule must apply on the word level, prior to resyllabification, to obtain [vE)n EspwaR] as the realization of *vain espoir*. Thus, for this set of speakers, we must assume the order: Nasalization before tautosyllabic n (word level) > Resyllabification > Deletion of (non-onset) Liaison C.

(30) Moderates:

		/ven espwaR/	/sot ami/
Word level:	Nasalization (nasalize V in same rime with nasal)	ve)n. es.pwaR	n/a
Phrase level:	Resyllabify (VC#V -> V#CV)	vE§). nes. pwaR	so.ta.mi
	Laxing (Lax mid V in closed syllable)	n/a	n/a
	Liaison C deletion	n/a	n/a
Anywhere:	Nasal Lowering (nasal V must be [-high,-tense])	vE§)n. es.pwaR	n/a

For all speakers, Nasalization precedes - because it feeds - Nasal Lowering, the rule that yields [E] from intermediate [i].

For ultra-conservatives, Nasalization must follow - and be bled by - Resyllabification, in order to derive the oral vowel in [vEn EspwaR]. Laxing must precede resyllabification, to obtain [sOt ami] with the same lax vowel as the feminine [sOt]. Since Nasalization is bled by Resyllabification, Lowering ( i.e. [i] -> [E]§) will be bled as well in cases of liaison.

<sup>14</sup> The first persuasive arguments against the derivational treatment of liaison along these lines have been presented by Tranel (1981). Tranel's arguments also hold, mutatis mutandis, against the alternative discussed in section 4.2.

## (31) Ultra-conservatives

		/ven espwar/	/sot ami/
Word level:	Laxing	vEn. Es.pwar	sOt. ami
Phrase level:	Resyllabification	vE. nEs.pwar	sO.ta.mi
	Nasalization	n/a	n/a
	Nasal Lowering	n/a	n/a

This necessary ordering makes it impossible to characterize the facts about [E]/ [in] and [yn]/ [œ] allomorphy. Recall that all speakers, including ultra-conservatives, reject or are reluctant to use feminines like [kOmyn] as the liaison allomorphs of the masculine, in forms like *commun accord*. The ordering Resyllabification > Nasalization > Lowering predicts that the ultraconservatives will in fact use such forms, or more precisely that any speaker that uses [vEn Espwa<sup>h</sup>] will also use [kOmyn akO<sup>h</sup>]:

## (32) Ultra-conservatives

		/kOmyn akO <sup>h</sup> /	
Word level:	Laxing	n/a	
Phrase level:	Resyllabification	kO.my.na.kO <sup>h</sup>	
	Nasalization	n/a	n/a
	Lowering	n/a	n/a
Output:		*kO.my.na.kO <sup>h</sup>	

A further relevant observations is this: if the oral vowel of ultra-conservative [vEn] *espoir* is attributed to the existence of underlying oral /ven/ - as in the derivation in (31) - then one wonders why the underlying oral vowels in -Vn words are strictly limited to adjectives whose feminine contains on the surface an oral vowel. The relevant descriptive point here has been mentioned earlier: the only morphemes of contemporary French to possess liaison allomorphs with oral vowels corresponding to citation forms with nasal vowels are those adjectives whose feminine ends in [En] or [On] (e.g. *vaine*, *bonne*) and whose citation masculine ends in [E<sup>h</sup>], [O<sup>h</sup>] (e.g. *vain*, *bon*). There are many other morphemes whose citation form ends in [E<sup>h</sup>], [O<sup>h</sup>] but none of these possess liaison allomorphs containing an oral vowel: recall *bien aimable* 'very kind' [bjE<sup>h</sup>nEmabl], *rien à voir* 'nothing to do

(with) [ʔjE§navwaʔ]. The impossibility of \*[bjEnEmabl], \*[ʔjEnavwaʔ] is predicted by our analysis: the oral vowel in *vain, bon* is imposed by **Ident (Rhyme)**, which requires that the entire rime of the word find a lexical precedent in some listed allomorph. This listed allomorph is the feminine *vaine, bonne*. There is no comparable listed variant in [En] in the case of uninflected forms like *rien* or *bien*, hence no possibility for the liaison allomorphs of these words to acquire an oral vowel<sup>15</sup>. Thus the explanation for the generalization noted hinges on the properties of the surface set of allomorphs of a given paradigm: this supports treatments like Perlmutter's (1996), Tranel's (1996) and ours, in which the analysis invokes reference to the set of surface allomorphs of a given morpheme. The facts remain unexplained on an analysis like (30)-(31), in which the properties of liaison forms are computed by reference to an underlying representation. There is no reason why *bien* cannot have the underlying representation /bien/ with an oral vowel in the same dialect in which *sien* is underlying /sien/.

We should also note that the derivational analysis provides only a clumsy account of the difference between moderates and ultras: two independent stipulations about rule ordering are necessary to characterize the difference between the [sot] *ami*, [vE§n] *espoir* dialects and the prescriptive [sOt] *ami*, [vEn] *espoir* forms. We have seen earlier that this dialectal difference follows from a single re-ranking of surface-oriented correspondence conditions: **Ident (V', ms) >> Ident (Rhyme)** vs. **Ident (Rhyme) >> Ident (V', ms)**. There are two distinct issues here, one of which is the relative descriptive simplicity of the two analyses. The more important point however is that the analysis we propose expresses directly an understandable source of difference between speakers: some give higher priority to an extreme form of lexical conservatism (**Ident (Rhyme)**) while others are more concerned to provide a fuller expression of gender concord (**Ident (V', ms)**).

#### 4.2 The Input Faithfulness alternative

The constraint-based alternative we consider next assumes that the liaison form of the masculine is identical to the underlying representation of the adjective. It is also necessary to show why such an analysis is wrong, because the thrust of the argument so far has been that *surface listed allomorphs* are the ones that determine the realization of novel forms, not the underlying representation. Let us flesh out this contender before showing how it fails. It is reasonable to hypothesize that the underlying representation of {[so], [sOt]} is /sot/ or /sOt/ or /sOt/ with a vowel lacking laxness values. We can formulate phonotactic conditions that will allow the underlying /t/ to surface only when prevocalic. Their exact formulation is irrelevant here, but I will posit **C#/\_V**: a word final coronal must be followed by a V within the same accentual phrase. This analysis can be made explicit by borrowing ideas from Dell's 1974 analysis of French, under which the feminine nouns and

<sup>15</sup> Cf. footnote 11 on the source of the hiatus breaking n in such forms.



adjectives possess a final schwa: it is this vowel that allows /t/ to surface in feminine forms like [sOt]. The masculines lack this vowel. On this theory, the /t/ surfaces in masculine phrases like [sot ami] or [sOt ami] not because it is needed to break hiatus, but because it *can* surface - given its prevocalic position. Clearly preserving more of the input is better than preserving less. The variation between [o] and [O] may represent speaker uncertainty as to the nature of the underlying vowel. These remarks lead to the analysis sketched below, where **C#/\_V** is seen to outrank an input faithfulness condition, **Preserve C**.

(32) Effects of C#/\_V

Input /sot/	C#/_V	>> Preserve C
l'ami [so]		*
l'ami [sot]	*!	

(33) Effects of Preserve C

Input /sot/	C#/_V	>> Preserve C
le [sot] ami		
le [so] ami		*!

There are many reasons not to pursue such an analysis: it is but a limited version of the derivational alternative explored earlier, stripped of some of its descriptive power. Therefore all arguments enumerated earlier against the derivational alternative hold of its translation into input-faithfulness language.

An additional reason identified by Tranel (1981) and Morin (1992), is that adjectives ending in two consonants (e.g. *court* 'short', masc. [kuÂ], fem. [kuÂt] or *vert* 'green' [vEÂ], fem. *verte* ) always display the masculine allomorph in liaison contexts, regardless of the speakers' degree of conservatism in *sot ami* or *vain espoir*- type phrases. The phrase *court espace de temps* 'brief period' is realized as [kuÂ Espas] not \*[kuÂt Espas]. Thus, even though the /t/ *can* be preserved in this phrase, it is not. The input faithfulness analysis predicts it will be:

(34) Input /kurt/, predicted \*[kuÂt Espas]

	C#/_V >> Preserve C	
[kuÂ Espas]		*
[kuÂt Espas]		

Our analysis accounts for this datum: [kuÂ Espas] avoids hiatus, as would \*[kuÂt Espas], and satisfies both **Ident (C#)** and **Ident (Rhyme)**, as does \*[kuÂt Espas]. But the former is superior to the latter, because it satisfies **Ident (C#, ms)** and violates no conditions that the latter fails to violate. This result is predicted to hold under all ranking variations considered, and this is in fact true: no French speakers say \*[kuÂt Espas].

(35) Effect of Ident (C#, ms)

listed allomorphs: kuÂ, kuÂt

	Ident (C#) >>	*Hiatus >>	Ident (C#, ms)
[kuÂ Espas]			
[kuÂt Espas]			*!

If identity to the underlying representation plays any role in the analysis, Morin's data shows that it is subordinated to the surface-oriented conditions encoding morphosyntactic agreement such as **Ident (C# ms)**. As (35) shows, it is this surface correspondence condition that decides the issue.

A second reason to reject an analysis of French liaison allomorphy based on input-output correspondence has been discovered by Perlmutter (1996), who notes that the most general characterization of allomorphic choices in prevocalic position involves hiatus avoidance<sup>16</sup>: Perlmutter notes that the feminine *ma* 'my' is replaced by *mon*, realized as [mO§n], in liaison contexts such as *mon amie* 'my friend-fem' [mO)§n ami]. Note that this is a case of suppletive allomorphy, where we cannot claim that either /ma/ or /mon/ represent the unified underlying representation for the 1st singular possessive adjective. *Ma* cannot be derived from *mon* and *mon* cannot be derived from *ma*. Thus the surfacing of *mon* in a gender-inappropriate context is not due to faithfulness to input but rather to hiatus avoidance<sup>17</sup>. The same point is made by Perlmutter à propos of other suppletive pairs

<sup>16</sup>Perlmutter refers to Onset - not Hiatus - as the relevant phonotactic condition, but this is unnecessary and insufficient, as shown by Tranel (1992), Encrevé (1988) and Klein (1995).

<sup>17</sup> Left unanalyzed in the text is the fact that the hiatus breaker n is not in fact present in either listed allomorph, *mon* or *ma*. My suggestion is that a general condition allows nasal vowels to extrude n in hiatus contexts, as in *il y en a* 'there is...' [il i A)n a], *on y va* 'we go there, let's go' [O)n i va], whereas oral vowels do not have this property. Thus the reason for choosing *mon* in *mon amie* is that this allomorph can extrude a hiatus breaker, according to general rules of French, whereas *ma*, the gender-appropriate allomorph cannot.

such as *beau, belle* ([bo], [bEl]) 'beautiful' and *vieux, vieille* ([vjP], [vjej]): the liaison allomorphs in masculine noun phrases are [bEl] - as in *bel enfant* 'beautiful child' - and [vjej] - as in *vieil éléphant* 'old elephant'. This is not the resurfacing of an underlying form but the use of a surface C-final allomorph, to avoid hiatus.

A final challenge to the analysis of French adjectival liaison based on input-correspondence involves the isogloss identified in section 3.3. The nasality of the *ai* nucleus in *vain espoir* appears to correlate with the laxness of *o* in *sot ami*. The analysis presented here characterizes this fact by reference to a single surface-oriented correspondence condition involving the featural identity of the stressed vowel with its correspondent in the gender appropriate allomorph (**Ident V', ms**). How will this same correlation be guaranteed in an analysis that identifies the liaison allomorph of the masculine with the underlying representation? No obvious answer presents itself.

## 5. Interim summary; base-derivative relations revised

We have observed so far that lexical conservatism must be modelled as the set of **Ident P** constraints requiring the target allomorph to be identical wrt P to some listed allomorph. For every identifiable P, the novel allomorph should, ideally, possess a lexical precedent. Ideally also these properties should cluster into a package of properties that identify the novel allomorph entirely with some one listed allomorph.

Empirical evidence for the phenomenon of lexical conservatism emerges only under the ranking **Ident P (>>) Phono-Constraint targetting P >> Ident (Q, ms)**, where  $Q = P$  or else P refers to a set of properties and  $Q \neq P$ . In the case of French two such rankings have been examined: **Ident (C #) >> \*Hiatus >> Ident (C#, ms)** characterizes all adult French dialects and accounts for the use of the feminine consonant in adjectival liaison. The ranking **Ident (Rhyme), \*Hiatus >> Ident (V', ms)** characterizes the ultra-conservative standard and accounts for the use of the feminine vowel in the masculine liaison forms in which the feminine consonant has been used as a hiatus buffer. In this case the word's rhyme (P) includes the identity of the accented vowel (Q), hence  $Q = P$ . Under rankings distinct from **Ident P (>>) Phono-Constraint targetting P >> Ident (Q, ms)**, it is not possible to observe the contribution of the lexical conservatism constraint **Ident P** as distinct from the more specific **Ident (Q, ms)** or **Ident (P, ms)**. We will observe below that lexical conservatism effects observed in English, Sanskrit, and Bantu, require the same ranking schema.

In the analysis presented here, the interaction of allomorphy with concord is modelled by imposing constraints which couple requirements of phonological identity with demands for morphosyntactic identity. Thus if there is identity of gender between two allomorphs there should be identity for certain other phonological properties, such as the height of the stressed vowel. I would like to comment more generally on this aspect of the analysis.

When we state that some linguistic expression is derived from another - as in '*happiness* is derived from *happy* ' - we are making two distinct statements, one about the morphosyntactic properties of the derivative and one about its phonological composition. The proposition "*Happiness* is derived from *happy* " means, on the one hand, that a linguistic expression possessing the morphosyntactic and lexico-semantic properties of *happy* is contained within *happiness*. Similarly, the statement, "*Vain espoir* is derived from *vain* " means that the morphosyntactic and lexico-semantic properties of *vain* - including its masculine gender feature - are contained within the phrase *vain espoir*. The other, distinct, side of the statement "x is derived from y" is that the phonological properties of the derivative x are computed by reference to those of base y: for instance, the phonological composition of *happiness* is a function of the phonology of *happy*.

The analysis presented here is based on the essential observation that these two sides of the statement "x is derived from y" must be kept distinct: it is possible for the phonological properties of x to be computed by reference to those of y, even if the morphosyntactic features of y are absent from x. Thus the phrase *vain espoir* contains one or more of the phonological properties of the feminine [vEn] but does not contain the feminine gender associated with [vEn] (pace Perlmutter 1996, Tranel 1996). The statement that *vain espoir* contains feminine gender features would render unintelligible both the idiolectal variation described here ([vEn] *espoir* vs. [vE§]n *espoir* ) and the relatively invariant properties of French liaison, such as the rejection of feminines in \*[divin] *archer*, \*[kOmy]n *accord*. What is the gender of the adjective in phrases like [sot] *ami* or [vE§]n *espoir* or [divE§]n *archer* ? Is it masculine, as its vowel suggests, or is it feminine, like its final consonant? It seems best to formulate an analysis under which such puzzles simply do not arise. The analysis we suggest subscribes to the following statements: the gender feature [+masculine] is required through agreement in phrases like *sot ami*. The presence of this and all other morphosyntactic features must be phonologically encoded. Its encoding is effected through conditions which state: "If it's a masculine, it must sound like a **known** masculine." These are the **Ident (P, ms)** conditions.

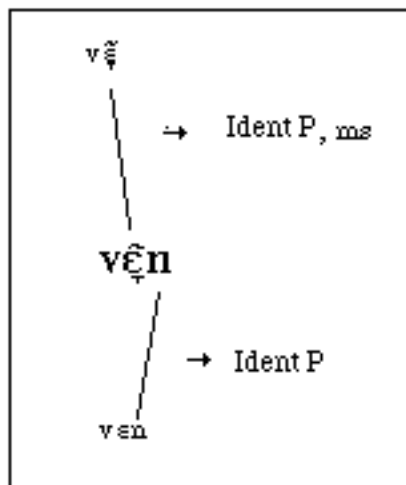
We draw from this discussion the conclusion that the phonological relations between bases and their morphosyntactic derivatives take the form of **Ident (P, ms)** conditions. The unanalyzed concept of Base of x reduces to two distinct classes of expressions: (a) the morphosyntactic components of x;

and (b) the linguistic expressions whose phonological properties are mentioned in the **Ident (P, ms)** conditions active in the computation of x's surface properties. As mixed expressions like [sot] *ami* or [vE§]n] *espoir* or [divE§]n] *archer* show, it is occasionally necessary to compute the phonological properties of a complex form by reference to several distinct phonological "bases", in this case by reference to the feminine as well as the masculine citation form of an adjective. The combination of **Ident P** and **Ident (P, ms)** conditions provides a general framework in which all types of base-derivative relations can be analyzed, whether they represent single-base relations (as in (36.a)) or multiple-base relations (36.b):

(36) Base derivative relations:

(a) Single-base relations: *happy* - *happiness*

(b) Multiple-base relations:



We now turn to two superficially different phenomena that illustrate other uses of the analysis of lexical conservatism sketched so far.

## 6. Lexical conservatism in English level 2 phonology

Earlier analyses of the level distinction in English have employed as the primary criterium for level assignment the effect of affixation on the stress pattern of the stem. Thus *-ness* is considered a Level 2 affix because it causes neither a change of the [ $\pm$ stressed] status of stem syllables nor a shift in the location of primary stress. In contrast *-ity* counts as Level 1 suffix because it changes stressed syllables into stressless, and vice-versa, and because it can shift the position of primary stress. (We distinguish stress *change* from stress *shift*: *change* refers to the [ $\pm$ stressed] status of a syllable with its consequences for segmental structure while *shift* refers to the location of primary stress only.) Since

Aronoff (1976), the fact has been noted that some suffixes generate heterogeneous formations level-wise: thus *-ism*, which generally fails to cause either shifts or changes in the stress of its stems (e.g. *invalid*, *invalidism*) does possess occasional formations like *buréaucratism*, whose stress pattern differs from that of its presumptive base, *búreaucrèt*. Later writers have either agreed with Aronoff that a level 2 suffix may occasionally generate words with level 1 properties (Kiparsky 1983, Anderson 1993) or have ignored the issue.

It is this phenomenon that we consider in this section: we propose here a new explanation for the phenomenon of level heterogeneity, for the fact that some level 2 suffixes do occasionally change the stress of their bases. We shall see that the level distinction is, in this case, nothing but an imperfect label for the accentual behavior. We should try to explain the accent pattern of Level 2 formations, both in the stress shifting forms (*bureáucratism*) and in the non-shifting cases (*invalidism*). Understanding the causes and limitations of the stress shift may render the level ordering labels unnecessary. The analysis we propose will draw on the notion of lexical conservatism illustrated earlier and will employ essentially the same analytical apparatus.

Our central claim is that level 2 suffixes should be defined not by their inability to change or shift the stress of their bases but rather by their lexical conservatism: for many speakers, the formations generated by these suffixes cannot lead to the creation of a stem variant that is distinct stresswise from some listed allomorph of that stem. But, if the addition of a Level 2 suffix renders some stress pattern metrically desirable, that stress pattern is adopted even if it differs from that of the morphosyntactic base, *provided that a lexical precedent for the desired stress pattern exists among the listed allomorphs of the stem*.

### **6.1. Custody, remedy and parody**

This point can be illustrated in abbreviated form through the comparison of three nonce forms of English: the *-able* adjectives derivable from the prosodically similar verbs *remedy*, *parody* and from the verbal expressions *take custody of*, *have custody of*. (Adjectives in *-able* require as their base a transitive verb but, for many speakers, *custody* can serve this purpose because of its association with idiomatic verb phrases like *take custody of*.)

We consider first the behavior of *custody* and *remedy* as possible bases for the *-able* formation. Each of these words has two listed allomorphs, which we refer to by their rhythmic properties, as the dactylic ('--') and the amphibrachic ('-') allomorph.

(37) Listed allomorphs of *custodi-* and *remedi-*

	listed dactylic allomorph	listed amphibrachic allomorph
<i>custody</i>	[kø!st´di]	[k´sto!di-], as in <i>custodi-al</i>
<i>remedy</i>	[rE!m´di]	[r´mi!di], as in <i>remedi-al, remedi-ate</i>

Before the unstressable suffix *-able*, it is preferable to have stress close to the right edge of the verbal stem, in order to avoid a long string of stressless syllables. English speakers, when asked to produce an *-able* adjective on *custody* and *remedy*, display three reactions, listed below in order of popularity:

(38) Possible *-able* forms based on (verbal) *custody* and on *remedy*

(a) based on the amphibrachic stem allomorph:

*custodi-able* [k´sto!di´bl`], *remedi-able* [r´mi!di´bl`]<sup>18</sup>

(b) based on the dactylic stem allomorph:

*cústodi-able* [[kø!st´di´bl`], *rédi-able* [rE!m´diabl`]

(c) forms using a novel stem allomorph:

*cùstodí-able* [kø~st´di!´bl`], *rèmedí-able* [rE~m´di!abl`]

Let us consider the advantages and disadvantages of the three options: the factors that play a role in this choice are enumerated below:

(39) Some factors relevant in computing the pre-*able* allomorph of the stem:

- Phonotactics: \*Lapse: a longer string of stressless syllables is dispreferred<sup>19</sup>.
- Morphosyntactic: *-able* is a deverbal adjective. Stem of *-able* form must be overtly marked as (i.e. phonologically identical with ) a verbal stem.

<sup>18</sup> *Remediable* was treated as a nonce form by the speakers I consulted but the 1970 edition of the Webster's Dictionary does record this form with the pronunciation [r´mi!di´bl`].

<sup>19</sup> On Lapse avoidance: Selkirk 1984, Prince 1983, Kenstowicz and Greene 1995. Mild Lapse violations at the end of words are made necessary (e.g. *América*) though the effect of constraints on avoidance of final and penult stress (Non-finality). These constraints are themselves outranked by others, so the effect of neither \*Lapse or Nonfinality is transparently present in English as a surface generalization.

• **Lexical: Avoid proliferation of allomorphs. Use existing forms.**

The advantages of the amphibrachic option (38.a) involve the phonotactic and lexical factors: the forms *custódiable* [k'ʰsto!di'bl̥] and *remédiable* [r'mi!di'bl̥] employ listed stem allomorphs and succeed in avoiding lapsed spans longer than two or three syllables (the syllabic scansion of the string [i'] appears to vary). The disadvantage of the amphibrachic option resides in the use of a stem allomorph that is either distinct in lexical category from the verbal stem (verbal *cústody*, *rémedy* vs. adjectival *custódial*; *remédial*) or else carries lexico-semantic connotations distinct from those of the intended verbal form (*rémedy* vs. *remédiate*<sup>20</sup>).

The dactylic option (38.c) has the advantage of using a listed stem allomorph and that of signalling the deverbal nature of the formation by selecting the same allomorph as the one appearing in the verb *rémedy*, (*take*) *cústody*. Its disadvantage lies in the greater length of the lapsed span<sup>21</sup>.

Hardly anyone prefers the third option (38.c) and this is clearly related to the fact that a novel or unlisted stem allomorph is being employed in *cùstodíable* [kø~st'ði!bl̥], *rèmedíable* [rE~m'di!abl̥]. Note that this disfavored pronunciation of the *-able* forms would have given optimal satisfaction to the accentual phonotactics, since the stresses are evenly spaced in this case and the initial syllable is stressed. Thus metrical well-formedness cannot be the reason for their lack of popularity. Note also that the stem allomorph employed in *cùstodíable* [kø~st'ði!bl̥], *rèmedíable* [rE~m'di!abl̥] is less different stresswise relative to the verbal allomorph than the stem allomorph employed in the preferred pronunciations *custódiable* [k'ʰsto!di'bl̥], *remédiable* [r'mi!di'bl̥]. We can quantify these differences by counting (a) the number of syllables whose stress status differs in the two sets of pairs and (b) the number of differences in relative prominence, i.e. the number of pairs of adjacent syllables the height of whose grid columns is different in the two sets of pairs.

(40) Measuring the metrical distance between verbal allomorph and *-able* allomorphs

(a) *rémedy* vs. *remédi-able*

- (i) two syllables (*re* and *me*) differ in their [ $\pm$ stress] status
- (ii) two pair of adjacent syllables differ in their relative prominence

<sup>20</sup> To remediate is defined as .... in ....; in contrast to remedy is ....

<sup>21</sup> That Lapse is a relevant consideration for this class of forms may be illustrated by the non-productive adjectives *formíable*, *despícable*, *hospítale* which lack verbal bases. There is no overt *\*fórmidate*, to prevent speakers from changing *fórmidable* into non-lapsed *formíable*.



(*reme* and *medy*)

(b) *rémedy* vs. *rèmedí-able*

- (i) one syllables (*di*) differs in its [ $\pm$ stress] status
- (ii) one pair of adjacent syllables (*medi*) differs in its relative prominence

The comparison carried out in (40) reveals that the disfavored forms *cùstodíable*, *rèmedíable* are in fact only negligibly different relative to *cústody*, *rémedy*, whereas relatively favored *custódiable*, *remédiable* are significantly different. Yet, despite this, the amphibrachic forms are preferred: the only reason for this preference must be the lexical conservatism factor.

A further indication of the relevance of lexical conservatism in the selection of stem allomorphy emerges from the comparison of *rémedy* and *cústody* with the (denominal) verb *párody*, whose paradigm differs from that of the other two in lacking a listed amphibrachic allomorph. There are no *paródial* or *paródiate* forms in English. The significant consequence of this fact is that the nonce *-able* form based on *párody* has fewer options: indeed, no speakers volunteer or accept *paródiable* and most are reduced to accept *párodiable*. (The same two speakers who accepted *cùstodíable* accepted *pàrodíable* as well.)

We conjecture then that the factors responsible for stress change in the *-able* adjectives are the dispreference for extended lapse and the existence of a listed stem allomorph with identically changed stress. The comparison between *remedy*, *parody* and *custody* indicates the need for conditions requiring novel allomorphs to be similar, in toto or in parte, to listed ones. Allomorphs like *custódi-*, *remédi-* are used in *-able* words not because they best satisfy the accentual phonotactics - *cùstodí-*, *rèmedí-* would have been better in this regard - but because they afford a measure of Lapse avoidance without violating Lexical Conservatism. The universally rejected *paródiable* is both novel and suboptimal accentually, hence never used.

## 6.2. The analysis

We now sketch the main lines of the analysis based on these generalizations. As in the case of French liaison we employ two types of correspondence conditions: **Ident P** expresses lexical conservatism in its pure form, the simple requirement that some lexical precedent exist for a specific property of the target allomorph. **Ident (P, ms)** expresses in the *-able* case the need to signal the verbal nature of the stem of *-able* through similarity to a listed verbal stem. It is the same class of conditions that was employed to signal gender agreement in French. In the present case, P will be the

(43) **Ident (stress under morphosyntactic identity)**. Abbrev.: **Ident (stress, ms)**):

The target allomorph of morpheme  $\mu$  has identical stress to a listed allomorph of  $\mu$ ,  $a(\mu)$ , if  $a(\mu)$  and the target allomorph have identical values for all morphosyntactic properties.

Like its French counterparts, the constraint in (43) penalizes differences in phonological composition between a candidate and listed allomorphs of the same morpheme, if the two forms share all morphosyntactic features. (For our analysis a constraint referring only to shared lexical category between allomorphs would have been sufficient but it is possible to generalize.) We assume uncontroversially that any form in *-able* contains a verbal stem. This appears to be a constraint on the syntactic category of the expression subcategorized for by *-able*, not on its semantic features: but this

---

<sup>22</sup> A simpler [ $\pm$ stress] correspondence constraint that considers only the stressed or stressless status of individual syllables is necessary in general but useless in the present instance: we want to insure that the stress pattern of a listed allomorph is adopted *in its entirety* in the target, not simply that the stressed or stressless status of some individual syllable possesses a lexical precedent.

aspect of the analysis is not essential and, if it turns out that lexico-semantic identity is what's at stake, (43) can be revised accordingly.

The lexical conservatism effects considered so far - can be analyzed as due to the ranking **Ident (stress) >> \*Lapse xxx >> Ident (stress, ms)**. **Lapse xxx** refers to a sequence of three stressless syllables: to simplify matters, we assume that post-tonic *-iable* is scanned disyllabically as [j'bl'] hence that **\*Lapse xxx** is violated in *rémediable* but not in *remédiable*. (Alternatively, we can recognize **\*Lapse xxxx** as distinct from **\*Lapse xxx** and invoke it in such cases<sup>23</sup>.) A tableau for the amphibrachic form *remédiable* - the form favored by most of my speakers and recorded in the Webster's is provided below.

(44) listed allomorphs: *rémedy*; *remédi-*

	Ident (stress)	>>	*Lapse xxx	>>	Ident (stress, ms)
rémedi-able			*!		
remédi-able					*
rèmedí-able	*!				*

The same ranking generates *párodiable*, as seen below:

(45) listed allomorph: *párody*

	Ident (stress)	>>	*Lapse xxx	>>	Ident (stress, ms).
párodi-able			*		
paródi-able	*!				*
pàrodí-able	*!				*

We obtain the pure dactylic pattern {*rémediable*, *cústodiable*, *párodiable*} under the ranking **Ident (stress, ms) >> \*Lapse xxx**, in which case the effects of **Ident (stress)** cannot be observed.

<sup>23</sup> Quantitatively different versions of **\*Lapse** can be obtained through local conjunction (Smolensky 1995) from the original **\*Lapse**. In all such cases the conjoined versions of a given constraint outrank the original, hence **\*Lapse xxxx >> \*Lapse xxx >> \*Lapse**.

(46) listed allomorphs: *rémedy*; *remédi-*

	Ident (stress, ms)	>>	*Lapse xxx
rémedi-able			*
remédi-able	*!		
rèmedí-able	*!		

(47) listed allomorphs: *párody*

	Ident (stress, ms)	>>	*Lapse xxx
párodi-able			*
paródi-able	*!		
pàrodí-able	*!		

### 6.3. The *-able* survey

We now proceed to consider the broader evidence for lexical conservatism in the formation of *-able* adjectives. The data involves the verbal bases in (48).

(48) Some verbal bases for *-able* adjectives.

*marriage, exterminate, educate, eradicate, govern, challenge, license, xerox, analyze, disciplin, caricature, demonstrate confiscate, sequestrate, contemplate, designate, infiltrate, equilibrate, obfuscate, persiflate, coruscate, promulgate, exacerbate, peregrinate, domesticate, reciprocate, communicate, assimilate, implicate, flummox, annihilate, procrastinate*

Some of these verbs end in *-ate*. Here, a relevant fact is a morphotactic preference of English not to have the affix *-ate* before another affix, such as *-ee* or *-able* (Aronoff 1976). Therefore the *-able* adjectives will, all else equal, not contain the affix sequence *-at-able*.

(49) Morphotactics: \*-ate<sub>aff</sub> - deriv. suffix

<i>nomin-ate</i>	<i>nomin-ee</i>	* <i>nomin-at-ee</i>
<i>evacu-ate</i>	<i>evacu-ee</i>	* <i>evacu-at-ee</i>

Thus *confisc-ate* may generate *confisc-able*, which satisfies the filter in (49), or *confisc-at-able* which violates (49) but whose component parts enjoy the advantage of greater recoverability.

There is also a list of purely phonotactic considerations, beyond \*Lapse. One of these is that English heavy syllables tend to be stressed. I refer to this, following Prince 1990, as the Weight-to-Stress Principle. In English this preference does not express itself in double clash situations, i.e. when both the preceding and the following syllable are stressed. The effect of the WSP will be to encourage pronunciations such as *pròmúlg-able* as against *prómúlg-able*. Finally, we must bear in mind the general preference is to place main stress as close as possible to the right edge of the word (cf. **Rightmost** in Prince and Smolensky 1993). Hence the stress shift from *ánaly~ze* to *àny!z-able*, \**ánaly~z-able*.

To investigate how speakers of English face the task of *-able* formation, I have done a survey of 24 speakers of American English at UCLA. Two groups, of 11 and 13 subjects each were asked to read aloud a list of *-able* adjectives based on verbs in (48) plus others. One group had, in parallel to the *-able* list, the list of corresponding verbs, with stresses marked. The other did not have the full list of verbs but had a partial list of less common verbs (e.g. *promulgate*), without stress marks. All *-able* forms on the written list conformed to the \**ate-able* filter: e.g. *promulgable*, not *promulgateable*. Subjects were invited to comment on the quality of all *-able* formations: thus some would say that *confiscable*, *promulgable* (with any stress pattern) are too ill-formed to utter. Others would indicate several possible pronunciations of the *-able* form, including ones with *-ate-able*, if they were uncertain of their preference. After reading the *-able* forms, subjects were asked to read aloud some of the less common verbs (e.g. *sequestrate*, *promulgate*) to verify where they locate stress in those forms.

The general interest of the experiment lies in the fact that most of the words considered were decidedly nonce formations (e.g. *equilibrable*) or else so infrequent (e.g. *educable*) that it was plausible to assume that the speakers would compute their stress pattern on-line, rather than look it up in their lexicon. Our findings then promise to tap directly speakers' knowledge of constraints and their interactions, as against knowledge of rote-learned lexical facts.

The survey allowed us to extend the analysis of the *remediable/parodiabile* pair to other instances of the lexical conservatism schema: **Ident P >> PhonoConstraint on P >> Ident P, ms**. In the case of *remediable*, the relevant phono-constraint was \***Lapse**. The ranking **Ident stress >> \*Lapse >> Ident stress, ms** was verified by considering the behavior of other *-able* forms. It is also possible that the WSP gives rise to a similar effect (i.e. **Ident stress >> WSP >> Ident stress, ms**) although the contribution of WSP independently of \***Lapse** is hard to establish. The results of the survey bearing on these rankings are shown below in (50). They can be summarized as follows: when faced with the conflict between accentual well-formedness, lexical conservatism, and the need to

encode subcategorization phonologically (**Ident, (P, ms)**) most subjects chose to optimize accentual well-formedness at the expense of the subcategorization, subject to lexical conservatism. They used in the *-able* adjectives any listed allomorph of target stem, if one existed that was better stressed than the verbal allomorph. Very few resorted to accentually improved but lexically unprecedented stem allomorphs.

The tables in (50) tabulate responses in which the stress of the *-able* adjective corresponding to a particular verb is different (or *shifted*) from the stress on the verb. When a subject offered more than one pronunciation for a given form, all responses were separately counted: this accounts for cases where the total number of responses exceeds 24. Table (50.a) gives the survey results for the *-able* words drawn from paradigms in which a verbal allomorph with initial or paen-initial stress coexists with a non-verbal allomorph with rightward-shifted stress. Table (50.b) provides the data for paradigms that lack the critical rightward-shifted allomorph.

(50) Some results of the *-able* survey

(a) *-able* forms based on paradigms containing an allomorph with rightward shifted stress

Shifted = stress on the syllable preceding <i>-able</i> . Non-shifted = stress on the same syllable as <i>-ate</i> verb.			
	shifted	non-shifted	root allomorphs with shifted stress
demonstrable	24	1	démonstrative
contemplable	22	1	contéplative
sequestrable	22	0	sequéster
compensable	23	1	compénsatory
expurgable	20	2	púrge, púrgatory
equilibrable	27	5	equilíbrium
remediable	8	1	remédial, remédiante
custody	5	3	custódial, custódian
	shifted	non-shifted	root allomorphs with shifted stress
obfuscalable	19	5	obfúscatory
infiltrable	19	5	fíler, fíltràte

(a) *-able* forms based on paradigms lacking an allomorph with rightward shifted stress

	shifted	non-shifted	listed allomorphs with shifted stress
challengeable	0	22	-
licenceable	0	22	-
governable	1	22	-
flummoxable	0	8	-
eradicable	0	24	-
educable	0	24	-
annihilate	0	24	-
domesticate	0	24	(domesticity <sup>24</sup> )
reciprocate	0	8	(reciprocité <sup>25</sup> )
assimilate	0	8	(similitude)
communicate	0	8	-

As can be seen, the vast majority of responses are consistent with the analysis sketched earlier for the pair *remédiable/párodiable*. We may therefore answer the question raised at the beginning of this section: is *-able* a Level 1, a Level 2 or a mixed-level morpheme? The answer anticipated earlier is that the best description of its accentual properties is provided not by reference to Levels but by lexical conservatism: *-able* adjectives are lexically conservative. Only this characterization accounts for the pattern of discrepancies between verbal stress and the stress of the *-able* form. A mixed-level analysis fails to explain why all speakers adopt "Level 2" *domesticable* (rejecting the shifted, "Level 1" *\*domesticable*) but many reverse their level preference systematically in cases like "Level 1" *compensable*, *promulgable*, *equilibrable*, *contéplable*. Since these are all nonce forms we must assume that speakers are guided by some general principle in their choice: rote memorization is out of question here. Lexical conservatism is the principle.

#### 6.4. Lexical conservatism elsewhere in English

A question that remains unresolved is to what extent the property of lexical conservatism characterizes more generally English word formation. We note first that only productive formations need be tested, since unproductive affixation is necessarily conservative, being limited by definition to a closed set of listed forms. Most Level 1 affixes are of this type. The lone exception here are the learned *-ian* adjectives based on proper names: stress shifts in *Spenglerian* (cf. *Spéngler*), *Mendelian* (cf. *Méndel*),

<sup>24</sup> *Domesticity* does not refer to the state of affairs brought about by the action of domesticating someone, hence we do not expect that this noun should have an effect on the stress pattern of the *-able* adjective. Recall that the task of the survey was to form and pronounce aloud *-able* adjectives based on the verbs listed.

<sup>25</sup> *Reciprocité* refers to a state of affairs that can be brought about through the action of the verb *reciprocate*. The allomorph *recipro[s]-* is therefore in principle useable in the context of the deverbal *-able* adjective. However none of the subjects volunteered *recipró[s]`ble* or *recipró[k]`ble*. After the survey, several subjects were asked their judgments about these forms and while *recipró[s]`ble* was considered acceptable by some, *recipró[k]`ble* was universally rejected. This is probably due to the fact that this form mixes the stress pattern of *reciprocité* with the segmental composition of *recipro[k]ate*. No analysis of this fact will be attempted but it seems clear that the constraint discussed earlier (17) **Ident P if Q** is relevant here. See also below section 6.4.2.

and many others, without apparent restriction. For the moment I see no better account of the difference between *-able* and *-ian* that a ranking difference: **Ident stress** >> **\*Lapse** >> **Ident (stress, ms)** for conservative-*able* vs. **\*Lapse** >> **Ident stress** (>>) **Ident (stress, ms)** for innovative-*ian*<sup>26</sup>.

A more urgent question is whether the failure to change stem stress typically associated with other Level 2 formations also reduces to lexical conservatism. The preliminary answer appears to be positive. Stress change and its absence in the relevant Level 2 formations can be characterized by extensions of the lexically conservative ranking **Ident (stress)** >> **Stress related constraints** >> **Ident (stress, ms)**. This means that stress shift is in fact observed at Level 2 in many cases where this serves the goal of accentual improvement, but without violating lexical conservatism.

#### 6.4.1. *-ism* nouns

In the case of the productive denominal *-ism* nouns, the relevant accentual constraint will be not **\*Lapse** but **\*Clash**, since *-ism* carries a secondary stress. This means that nouns with pre-final stress (like *catástrophe*) will give rise to accentually well-formed *-ism* formations (such as *catástrophèism*). There is therefore no reason to expect the choice of an alternate stem allomorph in such cases: for instance *\*càtastròphèism*, obtained by using the stem *càtastròph-* of *càtastròphic*, violates **\*Clash** and **Ident (stress, ms)**, and possesses no redeeming quality. However, when the noun ends in a stressed syllable, affixing *-ism* leads to clash: in such cases the data suggests that a search for better-stressed listed allomorphs does in fact take place. The data is grouped in two sections: (51.a) contains final-stressed nouns whose paradigm possesses an allomorph without final stress; (51.b) contains final-stressed nouns lacking an accentually distinct listed allomorph. The data indicates that only the nouns in (51.a) yield "Level 1"-type *-ism* forms with changed stress.

(51) a. Final stressed nouns whose corresponding *-ism* forms differ stress-wise:

Noun	-ism form	Listed allomorphs lacking final stress
<i>arístocràt</i>	<i>àristócratèism</i>	<i>àristócracy</i>
<i>démocràt</i>	<i>demócratèism</i>	<i>demócracy</i>
<i>búreaucràt</i>	<i>bureáucratèism</i>	<i>bureáucracý</i>

b. Final stressed nouns whose corresponding *-ism* forms do not differ stress-wise:

<sup>26</sup>. Other properties of *-ian* attributed to the level difference (for instance by Kiparsky 1983) have been shown by Fabb (1988) to involve factors that cannot be attributed to the order of affixation



Noun	-ism form	Listed allomorphs
<i>mónàd</i>	<i>mónàdism</i>	<i>mònádíc</i>
<i>réfugèe</i>	<i>réfugèeism</i>	--- <sup>27</sup>
<i>dòctrinàire</i>	<i>dòctrinàirism</i>	--- <sup>28</sup>
<i>démagogue</i>	<i>démagoguism</i>	<i>démagógic, démagògy</i>

For this class of cases we may therefore posit **Ident (stress) >> \*Clash >> Ident (stress, ms)**, a lexically conservative ranking. Note that for paradigms like that of *invalid*, we must assume further that **\*Clash >> \*Lapse**, a ranking that can be defended more generally in English: the *-ism* form is *invalidism* (with lapse), rather than *\*invalidism* (based on the allomorph of *invalidity*; with clash).

#### 6.4.2. -Ify verbs

Lapse avoidance appears to be at issue in the pattern of stress changes encountered with denominal or deadjectival-*ify* verbs. Lapsed strings of two syllables are common and perhaps even favored at the end of English words; only the longer final strings of three or more unstressed syllables are avoided through stress change or allomorph selection, as we have seen in the discussion of *-able* forms. Initial lapse does not appear to exist in English. In the formation of *-ify* verbs, the potential strings violating **\*Lapse** are medial, flanked on one side by the secondary stress on *-fy* and on the other by the rightmost stem stress. It appears that these medial lapses are strongly avoided, wherever avoidance is consistent with lexical conservatism. The data below shows stress changes in the *-ify* form relative to the verb, only for those paradigms offering a lexical precedent to the desired stress pattern<sup>29</sup>.

(52) a. Bases whose corresponding *--ify* forms differ stress-wise:

Base	<i>-ify</i> form	Listed allomorphs with final stress
<i>rígíd</i>	<i>rigíd-ify</i>	<i>rigíd-ity</i>

<sup>27</sup> *Réfuge* is accentually helpful but fails to convey the connection to the base *refugee*.

<sup>28</sup> *Dóctrine* poses the same problem as *réfuge*: it fails to convey the lexico-semantic features of *doctrinaire*.

<sup>29</sup> One form, *éthérify*, appears to be an exception to this generalization. The Webster's Dictionary records the pronunciation [iTE!r'fa~I] and at least two speakers have volunteered this form. The adjective *éthéreal* may represent a lexical precedent in this case, but its meaning is in fact unrelated to that of the verb.

<i>flúid</i>	<i>fluíd-ify</i>	<i>fluíd-ity</i> <sup>30</sup>
<i>sóllemn</i>	<i>soll[E!mn]-ify</i>	<i>soll[E!mn]-ity</i>
<i>cálorie</i>	<i>calór-ify</i>	<i>calór-ic</i>
<i>hístory</i>	<i>histór-ify</i>	<i>histór-ical</i>
<i>íamb</i>	<i>íá[mb]-ify</i>	<i>íámb-ic</i>
<i>vítriol</i>	<i>vitriól-ify</i>	<i>vitriól-ic</i>
<i>stable</i>	<i>stabíl-ify</i>	<i>stabíl-ity</i>

b. Bases whose corresponding *-ify* forms do not differ stress-wise:

Noun	<i>-ify</i> form	Listed allomorphs with final stress
<i>résin</i>	<i>résin-ify</i>	-----
<i>púmpkin</i>	<i>púmpkin-ify</i>	----

Several forms in (52.a) - *stabilify*, *iambify*, *sollemnify* - deserve comment. They show not only a stress change relative to the presumptive base but also associated segmental changes. Thus *stabílify* is clearly based on the optimally stressed *stabíl-* allomorph present in *stability*. This form cannot have resulted from simply shifting stress from the first to the second syllable of the citation form [ˈsteɪbəl]. A shift in stress would have resulted in the impossible [ˈÆsteɪˈbəl ˈÆfəl] - with stress on the syllabic l. Rather, the stem allomorph [stˈbɪl] of *stability* was adopted in toto, stress and segmentals. The same fact is observed in *íambify* and *sollémnify*: the stem allomorph occurring before *-ity* maintains a final cluster, which the adjective cannot realize. This final cluster is immaterial to the metrical well-formedness of the *-ify* form, but it is adopted nonetheless. This confirms that forms whose stem stress differs from that of the morphosyntactically appropriate base (*démonstrate*, *démónstr-able*; *stáble*, *stabíl-ify*) are not the result of recomputing de novo the location of stress: rather they result from a search of listed allomorphs, a search for pre-existing better-stressed versions of the same morpheme<sup>31</sup>.

<sup>30</sup> Other *-id* paradigms work similarly: *liquid*, *Druid*, *morbid*, *turgid* change stress in the corresponding *-ify* verbs, because they can adopt the stem allomorphs occurring before the *-ic* or *-ity* affixes.

<sup>31</sup> The existence of *sollemnify* (and the rejection of *\*sollé[m]ify*) as well as that of *recípro[k]able*, (*\*recípró[k]able*) suggests a condition of the form **Ident (P if Q)**, which insures that if the stress pattern of an allomorph is adopted, the segmental composition of its right edge is adopted as well. But it remains unclear why this condition is not enforced in the case of *-ism* forms like *buréaucratism*, where it should lead to the adoption of

The characterization of the stress changes in-*ify* words requires reference to **\*Lapse**, a constraint penalizing sequences of two stressless syllables, rather than the earlier **\*Lapse xxx or xxxx**. We assume that the preference for final disyllabic lapse is enforced by constraints irrelevant here and having to do with the end-of-word properties (cf. Prince and Smolensky's 1993 **Nonfinality**) . We assume further that the tolerance for initial dactyls in words like *Tàtamagóuchi* is due to lexical conservatism: the speaker has no choice of allomorphs in this case and is bound to use the one version that he has encountered. It is therefore in cases like *vitriólify* - a nonce word without received pronunciation, whose extended paradigm offers a choice of allomorphs - that the effect of **\*Lapse** can manifest itself. The constraint hierarchy proposed earlier can be augmented as: **Ident stress** >> **\*Lapse xxx** (>>) **\*Lapse** >> **Ident (stress, ms)**. The relevant parts of this ranking are illustrated below:

(53) listed allomorphs: *vítriol*, *vìtriól-*

	Ident stress	>> *Lapse	>> Ident (stress, ms).
vi~trio!lify~			*
vi!triolify~		*!	

listed allomorphs: *résin*

	Ident stress	>> *Lapse	>> Ident (stress, ms).
resínify~	*!		*
résinify~		*	

In closing, I should note that the comments from several speakers indicate that in paradigms like {*expurgate* *purge*, *purgatory* } several forms represent simultaneously the lexical reference points in the formation of the -able adjective. This point was made clear to me by the subjects who saw *expurgable* in my list but read it as *expúrgible*, clearly bearing in mind the verb *purge*, which provides in this case a better lexical reference point than *éxpurgate*, because it has final stress and thus can yield a better-stressed -able form. (*Expurgate*, will tend to yield *éxpurgable*, with a **\*Lapse xxx**

---

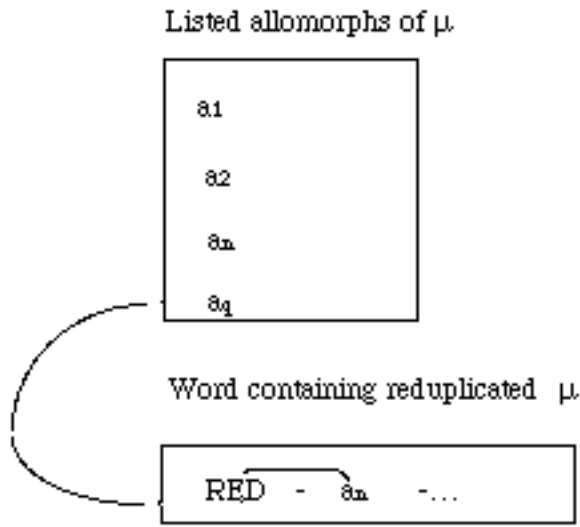
*buréacra[s]ism*. Perhaps the right approach here is the same as that taken in adopting for French the condition **Ident (Rhyme)**: we identify a substring - here the main stressed vowel and following interlude - which must be strictly identical, in its entirety, to its correspondent in some listed allomorph. Thus the main stressed syllable in *buréacratism* is identical to that found in *bureáucracy*, the differences between allomorphs bearing only on the composition of the unaccented string [kr't] vs. [kr's]: therefore this condition is satisfied. Our condition is also satisfied in the forms *sollénnify*, *iámbify* but violated in *\*sollé[m]ify*, *\*ia[m]ify*. Thus we may assume that such a condition will successfully differentiate acceptable blends like *buréacratism* from unacceptable ones like *\*sollé[m]ify*.

is deverbal because it doesn't mean 'being able to be or become remedial' but rather 'being able to be remedied'. The role of *remédial* in the formation of *remédiable* is limited strictly to providing a suitable stress contour for an expression that maintains a reference to verbal action. It is in the context of observations of this sort that one becomes aware of the ambiguity inherent in statements like 'x is the base of y', or 'y is derived from x', as noted earlier. The unambiguous language proposed here doesn't mention bases and derivatives, only morphosyntactic or phonological identity between various root allomorphs.

It is natural then that we should consider the other current use of the term *base*, the use associated with reduplication. The expression *base of reduplication* (initiated I believe by Broselow and McCarthy's 1983 study of infixing reduplication) is also ambiguous: it means first 'the allomorph that contributes phonological properties adopted by the reduplicant', the lexical reference term in the computation of the reduplicant. A distinct sense of this expression is 'the allomorph of morpheme  $\mu$  that is adjacent to a reduplicant based on  $\mu$ '. In most cases there is no empirical reason to distinguish these two senses, presumably because it is optimal for the reduplicant's identity to be computed relative to an adjacent allomorph of its root. Similarity is best computed at close range (Pierrehumbert 1993).

But it is also sometimes the case, as we observe in this section, that the adjacent allomorph is not the one responsible for the shape of the reduplicant. The data we consider here do not involve the distinction between the underlying vs. the surface form of the base: on this point, see McCarthy and Prince 1995. Rather, we consider cases in which two surface allomorphs of a root compete to impose their phonological properties upon the reduplicant, and the one that wins this competition is the non-adjacent allomorph. The situation is abstractly depicted below:

(54) Competition between adjacent and non-adjacent allomorphs for the reduplicant's shape



The interest of such cases is two-fold: first, these phenomena also require reference to **Ident P**, the basic expression of lexical conservatism. Non-adjacent allomorphs of a root - like a<sub>q</sub> in (54) - come into play only when they contribute a property that is desirable in reduplicant but absent from the adjacent allomorph, a<sub>n</sub>. They come into play because they can improve the shape of the reduplicant without violating lexical conservatism.

The second point of interest is that these reduplicative phenomena cannot be analyzed if we continue to conflate the two meanings of the expression "base of reduplication": both the adjacent and the non-adjacent allomorphs of some morpheme  $\mu$  may serve as bases for the reduplicant of  $\mu$ , but in the two different senses noted above, since one allomorph is responsible for the shape of the reduplicant while the other is responsible for its position within the word. We suggest that the language of reduplication studies be disambiguated as follows:

(55) Reduplicant adjacency :

The reduplicant of  $\mu$  must be (left/right) adjacent to an allomorph of  $\mu$ .

This statement makes explicit the localizing function of the base. Identity between the reduplicant and the adjacent allomorph is favored by condition (56):

(56) Identity under adjacency (abbreviated **Ident-adjacent**):

If the reduplicant of  $\mu$  and an allomorph of  $\mu$  are adjacent, they must be identical.

Under (56) we abbreviate a set of conditions stipulating **Ident P under adjacency** for various phonological properties. For simplicity, we operate here with the conflation of this set, bearing in mind that the evaluation of this constraint is non-categorical: one must record every source of difference between the target and the adjacent allomorph. In this respect, **Ident-adjacent** (56) must have an empirical coverage similar to the combination of McCarthy and Prince's (1995) **MAX Base-Reduplicant**, **DEP Base-Reduplicant** and **Ident F Base-Reduplicant**: we fail to distinguish these elements of correspondence only because these distinctions do not matter here. What matters is the following point: the lexical conservatism constraint class **Ident P** can only be evaluated categorically, since the evaluation consists of determining if some listed allomorph exists that is identical with respect to P to the target. The answer in such cases is always yes ( ) or no (\*), not a list of violations. In contrast, **Ident-adjacent** does have to be evaluated gradiently, by detailing the extent to which the reduplicant deviates from the adjacent allomorph. This has been shown by McCarthy and Prince (1995) and subsequent work and follows from the logic of the situation: when one compares only two elements, one can examine how closely they resemble. The difference then between the two approaches to reduplication lies in the fact that more than one possible "base" is recognized here and more than one type of base-identity condition: **Ident P** and **Ident-adjacent**. The contents of **Ident-adjacent** may well be exactly as McCarthy and Prince argue.

The effects of non-adjacent allomorphs upon the reduplicant can now be modelled through rankings such as this: **Ident P >> Phono-constraint favoring P in RED** (or a subset of the set P) >> **Ident-adjacent**. Thus, suppose that allomorph  $a_1$  in (54) possesses the property P, which is desirable in the reduplicant, and that P is absent from the adjacent allomorph  $a_2$ . Then this ranking will insure that the reduplicant adopts P. The same ranking will forbid adoption of P in the reduplicants issued from paradigms lacking any listed allomorphs endowed with P, in virtue of the condition on lexical conservatism **Ident P**. We now consider a concrete case of this sort. The following sections explore the factorial typology of the system proposed.

## 7.2. Lexical conservatism in the Sanskrit perfect reduplication

Basic generalizations on the interaction between Sanskrit root ablaut and perfect reduplication appear in Kiparsky 1986 and Steriade 1988. I summarize the relevant ones here. Verb roots are of a strictly monosyllabic shape, with typical segmental composition being  $C_0a(\cdot)(R)C_0$ , where R = sonorant and the  $C_0$  strings represent permissible word-beginnings and (at some level of abstraction) word-endings. Some examples of roots include<sup>32</sup>: *pat*, *svap*, *smai*, *ma:*, *gam*, *mard*, *vraj*, *stambÓ*,

---

<sup>32</sup> Standard notation for Sanskrit data: <v> = [w], <y> = [j], <j> = [dZ], <c> = [tS]. The sign identifies the underlying representation of the root. Surface <e>, <o> are always long [e:], [o:] and always issued from underlying diphthongs /ai/, /au/. High vowels and corresponding glides are in near complementary distribution and alternate.

*vakṣ*, *krand*, *bÓranç*. Roots ablaut. This means that in positions where they are unaccented or accented by default only (as against possessing a lexical accent) they lose their nuclear *a*, with subsequent syllabic reorganization: examples of the effect of ablaut on some roots appear below.

(57) Ablaut:	full grade (accented)	zero grade (unaccented, or accented by default)
	<i>smai</i>	-> <i>smi</i>
	<i>marj</i>	--> <i>mr`j</i>
	<i>svap</i>	-> <i>sup</i>
	<i>krand</i> ->	<i>krrn`d</i> [ <i>krad</i> ]
	<i>aiḅ</i>	-> <i>iḅ</i>
	<i>yaj</i>	-> <i>ij</i>

Note that pairs like *aiḅ/yaj* indicate that the underlying representation of the root must be similar in its segmental and syllabic composition to the full allomorph, not the zero-grade one. The zero grade can be considered a listed allomorph for the vast majority of roots, in the sense that essential and textually frequent elements of most roots' paradigm (for instance all athematic plural formations and the past participle) require the zero grade. It is thus plausible to think that to know a root meant, for Sanskrit speakers, to know both its full and zero grade forms.

Many roots fail to ablaut. Although all roots containing the nuclei *ai* or *au* ablaut as C<sub>0</sub>iC<sub>0</sub> or C<sub>0</sub>uC<sub>0</sub>, and no roots containing long *a*: ablaut in the usual manner, most other root types show idiosyncratic variation. One may rationalize this in various ways (cf. Steriade 1988, for one attempt, Kiparsky 1986 for another) but the fact is that it is unpredictable whether a given root will or will not lose its unaccented *a*. Some quasi-minimal pairs appear below:

(58) Ablaut variability:	full	zero
	<i>svaj</i>	-> <i>svaj</i>
	<i>svap</i>	-> <i>sup</i>
	<i>gras</i>	-> <i>gras</i>
	<i>grabÓ</i>	-> <i>gr`bÓ</i>
	<i>vas</i>	-> <i>uḅ</i>

---

Underlying or intermediate syllabic n becomes [a]. The effects of some sound alternations, such as Brugmann's Law, are omitted here.

<i>vaç</i>	->	<i>vaç</i>
<i>sar</i>	->	<i>sr`</i>
<i>svar</i>	->	<i>svar</i>

Ablaut variability accounts for the fact that root paradigms differ from each other in the number of listed allomorphs they contain: *svaj* possesses only a single allomorph, *svaj*, whereas *svap* possesses two, *sup* and *svap*<sup>33</sup>. This difference in the structure of paradigms has consequences for the analysis of reduplication, to which we now turn.

The perfect reduplicant is a light (C)V syllable whose nucleus is always a vocoid: *a*, *i*, or *u*. Onset clusters found within the root are simplified in reduplication: the reduplicant maintains only the lowest sonority segment. Post-nuclear segments are also necessarily eliminated. These effects indicate the general trend towards optimization in the syllabic shape of the reduplicant (Steriade 1988, McCarthy and Prince 1995).

(59) Perfect reduplication I<sup>34</sup>

root	perfect full grade	perfect zero grade	gloss
<i>pat</i>	<i>pa-pát-a</i>	<i>pa-pt-úr</i>	fly
<i>pratÓ</i>	<i>pa-pra!tÓ-a</i>	<i>pa-pratÓ-é</i>	spread
<i>kβad</i>	<i>ka-kβa!d-a</i>	<i>ka-kβad-é</i>	divide
<i>mna:</i>	<i>ma-mná:-u</i>	<i>ma-mn-úr</i>	note
<i>syand</i>	<i>sa-syánd-a</i>	<i>sa-syad-é</i>	move on
<i>sarj</i>	<i>sa-sárj-a</i>	<i>sa-sr`j-é</i>	send forth

<sup>33</sup> More precisely *svaj* possesses one class of allomorphs which are identical with respect to their vocalism, whereas *svap* possesses two distinct classes, differentiated by nuclear quality and by correlated accentual properties. There is considerably more allomorphy in Sanskrit than we discuss here.

<sup>34</sup> I ignore other modifications related to the perfect, e.g. Brugmann's Law or reduplicant dissimilations.



The vocalism of perfect reduplication is not determined by the adjacent root allomorph. In roots containing the sequence *ai* , *au* , reduplication copies always the high vowel:

(60) Perfect reduplication II

root	perfect: full grade	perfect: zero grade	gloss
<i>baudÓ</i>	<i>bu-báudÓ-a</i> [bu-bó:dÓ-a]	<i>bu-budÓ-é</i>	awake
<i>tvaiβ</i>	<i>ti-tváíβ-a</i> [titve:βa]	<i>ti-tviβ-é</i>	be stirred up
<i>auc</i>	<i>u-áuc-a</i> [uvo:ca]	<i>u-uc-úr</i> [u:cúr]	please

The data in (60) indicates that the reduplicant's nucleus may be identical to that of the zero grade allomorph, where one exists, *even when the adjacent root is not in its zero grade* as in *bu-bo!dÓ-a*. Where idiosyncratic ablaut differences are observed, we note corresponding differences in the vocalism of perfect reduplication: invariant *svaj*, for instance, reduplicates as *sa-svaj-a* whereas alternating *svap* reduplicates as *su-svap-a*.

(61) Correlations between ablaut and reduplication possibilities

	root	perfect full grade	perfect zero grade	gloss
a.	<i>svap</i>	<i>su-sváp-a</i> [suβva!pa]	<i>su-sup-úr</i> [suβupu!r]	sleep
	<i>vas</i>	<i>u-vás-a</i>	<i>u-us-úr</i> [u:βúr]	shine
	<i>yaj</i>	<i>i-yaj-a</i>	<i>i-ij-é</i> [i:jé]	'offer'
	<i>myakβ</i>	<i>mi-mya!kβ-a</i>	<i>mi-mikβ-u!r</i>	'be situated'
b.	<i>svaj</i>	<i>sa-sváj-a</i>	<i>sa-svaj-é</i>	embrace
	<i>vaç</i>	<i>va-váč-a</i>	<i>va-vaç-é</i>	be eager

<i>yat</i>	<i>ya-yat-a</i>	( <i>yet-úr</i> <sup>35</sup> )	'stretch'
<i>kÓya...</i>	<i>ka-kÓya!...-u</i>	<i>ka-kÓya...-u!r</i>	'see'
		[ <i>ca-kÓyu!r</i> ]	

The generalization then is that the reduplicant's vowel will be high, if and only if the vocalism of the zero grade allomorph contains a high vowel. In this sense then, the reduplicant is "derived from" the zero grade, as argued in Steriade 1988. The zero grade alternant of the root need not be adjacent to the reduplicant: it is not in cases like *su-βvap-a*, *bu-bodÓ-a*, *i-yáj-a*. But some zero grade form with syllabic *i* or *u* must exist in the set of listed allomorphs in order for the perfect reduplicant to contain *i* or *u*.

The next step in the analysis - following here a suggestion by Bruce Hayes - is to propose that the high vowel is desirable in the perfect reduplicant: *svap* reduplicates as *su-svap-a* because a phonological condition favors the high vowel over the low one. There are several hypotheses one can entertain on this score, all of which are equivalent for the point under discussion here. One possibility, discussed by Alderete et al. (1996) is that high vowels are inherently unmarked, regardless of context. The other is that they are especially good in light syllables, though not necessarily elsewhere. Sanskrit possesses four CV reduplicants (for the perfect, present, aorist, desiderative) and one CVX reduplicant (for the intensive, X being a syllabic or non-syllabic sonorant). The vocalism of the light CV reduplicants is invariably or preferably high, that of the heavy CVX reduplicant is invariably non-high. Other languages possessing invariant high vowels in reduplication associate these with the light template. This suggests that [+high] is preferably associated with light syllables, because of their typically shorter duration.

(62) **Light -high**: Light syllables contain high vowels.

In the Sanskrit perfect, **Light-high** is enforced only when consistent with lexical conservatism, via the ranking **Ident [high]** >> **Light-high** >> **Ident-adjacent**. The CV shape of the reduplicant is enforced through conditions irrelevant here. We remind the reader that the interpretation of **Ident P** is that some property - here the height of the nuclear vowel - must be present in some, non-specific, listed allomorph of the root. Therefore **Ident [high]** is satisfied by *su-* as a reduplicant of *svap*,

---

<sup>35</sup> *Yétur* belongs to a set of forms in which lack of accent on the root results in idiosyncratic alternations synchronically unrelated to zero grade formation. The root *yat* lacks a zero grade allomorph \*it: cf. the participle *yat-tá*.

sa - sváp-	*!	*(u a)
------------	----	--------

(64) listed allomorphs: *svaj*

	Ident [high]	>>	Light-high	>>	Ident-adjacent
sa-svaj-			*		
su-svaj	*!				*

Can we generalize **Ident [high]** or must we continue to narrowly refer to nuclear height? A possibility that must be excluded is that of generalizing **Ident [high]** to a broader condition requiring that the reduplicant be strictly identical to some listed allomorph, in all respects, not just the height of its nucleus. We call this generalized condition **Ident (root)**: there is a listed allomorph of the root such that it and the target allomorph are strictly identical. Clearly then the templatic conditions that insure the CV shape of the reduplicant must outrank **Ident (root)** while **Ident (root)** itself outranks **Light-high**. The derivation of *sa-svaj* and *su-svap* proceeds below: (66) shows that *sa-svaj-* can no longer be derived under this interpretation. Alternate rankings also fail. We conclude that **Ident (root)** - which will be useful for other systems discussed below - cannot be the critical constraint in Sanskrit.

(65) listed allomorphs: *svap*, *sup*

	RED=CV	>>	Ident (root)	>>	Light-high	>>	Ident-adjacent
su-svap			*				*
sa-svap			*		*!		*

(66) listed allomorphs: *svaj*

	RED=CV	>>	Ident (root)	>>	Light-high	>>	Ident-adjacent
sa-svaj			*		*!		
su-svaj-			*				*

It is however possible to achieve a more limited generalization of **Ident [high]**: we may require that the global quality of the root nucleus (though not its quantity) possess a lexical precedent. We dub this constraint **Ident (Nuc)** and define it as below. The verification follows.

(67) **Ident Nuc**:

There is a listed allomorph of morpheme  $\mu$  such that its nuclear vowel is identical for all feature values to that of the target allomorph.

(68) listed allomorphs: *svap, sup*

	Ident [Nuc]	>>	Light-high	>>	Ident-adjacent
su - sváp-					*(u a)
sa - sváp-			*!		

(69) listed allomorphs: *svaj*

	Ident [Nuc]	>>	Light-high	>>	Ident-adjacent
sa-svaj-			*		
su-svaj	*!				*(u a)

Note that in characterizing the high vowels of the other CV reduplicants of Sanskrit (e.g. desideratives such as *di-dÓvans-iβa-*, *di-dÓak-β-*, *di-dr`k-β-*, *du-dÓuk-β-*, *ji-jan-iβ-*) we must assume a reversal of the perfect ranking, namely **Light-high >> Ident [high]**, or **Ident (Nuc)**, since the high vowels appear in this reduplication type regardless of whether the root possess independently an allomorph with a high nucleus.

## 7.2. Global identity conditions in Austronesian reduplication

McCarthy and Prince's (1990) study of prosodic circumscription mentions a number of Austronesian reduplication patterns sharing the following properties: reduplication is limited to two syllables, while the root can exceed two, and the reduplicants based on disyllabic roots reduplicate differently from those based on longer roots. The critical detail is that reduplicants based on disyllabic roots are strictly identical to the root, while other reduplicants deviate in various ways from the segmental structure of the root. The analysis required for this data is a natural extension of that proposed for Sanskrit. We consider here the Austronesian data because it contributes to motivating a global and categorical identity condition similar to the **Ident (root)** mentioned earlier: we will then observe that this same type of global identity constraint has critical uses in the analysis of the lexical conservatism phenomena encountered in Bantu reduplication (Downing 1996, 1997 a,b).

The Cebuano data in (70.a) show that disyllabic roots possess strictly identical reduplicants. In contrast, in plus-quam-disyllabic roots (70.b), only the first root segment appears in the reduplicant, its VCV remainder being taken up by the invariant string *-ulu*.

(70) Cebuano. Reduplicating morpheme = 2 syll. (McCarthy and Prince 1990)

	Base	Reduplicated form	Gloss
a.	<i>sulti</i>	<i>sulti-sulti</i>	talk
	<i>balik</i>	<i>balik-balik</i>	come back
	<i>higda/</i>	<i>higda/-higda/</i>	lie
b.	<i>balibad</i>	<i>bulu-balibad</i>	refuse offering
	<i>paNutana</i>	<i>pulu-paNutana</i>	ask question
	<i>pahulay</i>	<i>pulu-pahulay</i>	rest
	<i>panaNhid</i>	<i>pulu-panaNhid</i>	ask to leave

**Ident (Root)** plays a role here<sup>36</sup>, through the ranking: **RED=2s** >> **Ident (Root)** >> **Red =ulu]**. Consider the derivations of *bulu-balibad*, which establishes that **RED= 2s** >> **Ident (Root)** and that of *balik-balik*, which provides the argument for ranking **Ident (Root)** >> **Red= ulu]**.. The latter constraint penalizes reduplicants that fail to end in *-ulu*.

(71) relevant listed allomorphs: *balibad*

	RED=2σ	>>	Ident (Root) >>	RED = ulu]
<i>bulu-balibad</i>			*	
<i>bali-balibad</i>			*	*!
<i>balibad-balibad</i>				*

(72) relevant listed allomorphs: *balik*

	Ident (Root)	>>	Red =ulu]
<i>balik-balik</i>			*
<i>bulu-balik</i>	*!		

<sup>36</sup> I owe this proposal to suggestions by Bruce Hayes and Paul Kiparsky.

Similar effects are observed in Tagalog and Macassarese reduplication, the latter analyzed first by Aronoff, Basri and Broselow (1987). Reduplicants here - also invariably disyllabic - end in / unless they happen to be based on disyllabic roots, in which case they are strictly identical to the root.

(73) Tagalog (McCarthy and Prince 1990)

	Base	Reduplicated form	Gloss
a.	<i>mag-li...nis</i>	<i>mag-li...nis-li:nis</i>	clean
	<i>mag-walis</i>	<i>mag-walis-walis</i>	sweep
	<i>pantay</i>	<i>pantay-pantay</i>	level
b.	<i>tahi...mik</i>	<i>tahi/-tahimik</i>	quiet
	<i>baluktot</i>	<i>balu/-baluktot</i>	bent
	<i>kalansiN</i>	<i>kala/-kalansiN</i>	jingle of coins

A similar analysis is required here: **RED=2s** >>**Ident (Root)** >> **RED = -/ ]**.

The argument for the global identity condition **Ident (Root)**, as against more piecemeal conditions of lexical conservatism such **Ident (nuc)** or **Ident (stress)**, is that the same condition must detect a change in syllable count and metrical structure as well as a change in the segmental structure of the root allomorph. No single piecemeal condition will have the decisive role **Ident (Root)** has in (72), while at the same time allowing *Red-ulu]* to assert itself in (71). The reader is invited to experiment with **Ident (stress)** or similar syllable-counting correspondence constraints to observe that this is so. Note also that **Ident (Root)** carries over painlessly from Macassarese to Tagalog, thus allowing us to characterize the essential similarity of these two systems.

This said, note that the Austronesian data does not directly argue for identity to some listed allomorph as against identity to the adjacent allomorph: what it does argue for is a mode of constraint evaluation in which the extent of the target's deviation from the target does not matter. This mode of evaluation, we have argued, is proper to lexical conservatism conditions: it is not proper to other correspondence conditions, such as **Ident-adjacent**. If the magnitude of deviation between the reduplicant *bulu* and the root *balibad* did matter, i.e. if a violation mark was assessed for every phonological property that distinguishes the two strings, then the analysis proposed could not derive *bulu-balibad*, as seen below. We assess one \* for every segment of *balibad* that is either missing from the reduplicant or distinct from its positional correspondent in the reduplicant:

(74) Mis-evaluation of *bulu-balibad* by **Ident-Adjacent**

	Ident-adjacent	>>	Red = <i>ulu</i> ]
bali-balibad	*(b),*(a),*(d)		
bulu-balibad	*(a), *(i), *(b)!		*
	*(a), *(d)		

We conclude then that it is the categorical mode of evaluation inherent in **Ident (root)** that is needed in the Austronesian analysis, not the gradient mode of **Ident-Adjacent**.

A final comment is that it should be possible to modify the syntactic scope of the global identity condition we have referred to so far as **Ident (root)**: in Sanskrit we must refer to roots and root allomorphs, but in other languages discussed here we may assume that the syntactic constituent targeted by reduplication is broader. Thus we may consider **Ident (word)**: such a condition identifies a larger paradigm, a larger set of words sharing a root, and states that the target allomorph is globally identical to some listed member of this larger set.

### 7.3. Lexical conservatism in Bantu reduplication

Having established the need for global conservatism conditions such as **Ident (Root)** or **Ident (word)**, we now consider a family of reduplication systems encountered in Bantu languages, which combine Sanskrit-like effects of lexical conservatism with the Austronesian preference for disyllabism in the reduplicant.

In a series of important studies, Downing (1997 a, b, c) has demonstrated that several Bantu languages possess reduplication patterns in which a non-adjacent allomorph dictates the shape of the reduplicant. This section is based on many of Downing's insights, in particular her idea that the reduplicant's shape is computed relative to the surface representation of a non-adjacent allomorph; as well as on earlier work by Mutaka and Hyman (1990) and by Odden (1996).

Consider first the simplified Swati paradigm in (75), based on Downing (1997a). (75.a) shows that one Swati option is to reduplicate a disyllabic substring of the adjacent stem (root + derivational affix). (75.b) illustrates an alternative disyllabic pattern, in which the reduplicant consists of the root, followed by the general purpose vowel *-a*, which is normally present at the end of Swati verbs.

(75) Swati reduplication (Downing 1997a)

(a) Reduplicating the adjacent stem

<i>u-ya-dlal-a</i>	<i>u-ya-dlala-dlal-a</i>	'you play'
<i>ni-ya-dlal-is-an-a</i>	<i>ni-ya-dlali-dlal-is-an-a</i>	'you pl. play with e.o.'
<i>u-ya-tfuts-el-a</i>	<i>u-ya-tfutse-tfuts-el-a</i>	'you move for s.o.'
<i>u-ya-lalel-a</i>	<i>u-ya-lale-lalel-a</i>	'you listen'
<i>u-ya-cebul-a</i>	<i>u-ya-cebu-cebul-a</i>	'you skin'

(b) Alternative, Red2: root+a, when root = 1syll.

<i>u-ya-bong-a</i>	<i>u-ya-bonga-bonga</i>	'you thank'
<i>u-ya-bong-is-a</i>	<i>u-ya-bongi-bong-is-a</i>	'you make s.o.thank'
	Red2 <i>u-ya-bonga-bong-is-a</i>	same
<i>u-ya-bik-a</i>	<i>u-ya-bika-bik-a</i>	'you report'
<i>u-ya-bik-el-a</i>	<i>u-ya-bike-bik-el-a</i>	'you report to'
	Red2 <i>u-ya-bika-bik-el-a</i>	same

Downing refers to the alternative reduplicating pattern as Red2 and notes that it is only available in monosyllabic roots: the polysyllabic roots in (76) accept only reduplicants that are identical to the first two syllables of the adjacent string.

(76) Red2 is unusable with polysyllabic roots

<i>u-ya-khulum-a</i>	<i>u-ya-khulu-khulum-a</i>	'you listen'
	not <i>*u-ya-khula-khulum-a</i>	
<i>u-ya-phuphut-a</i>	<i>u-ya-phuphu-phuphut-a</i>	'you are blowing out'
	not <i>*u-ya-phupha-phuphut-a</i>	
<i>u-ya-sebent-is-a</i>	<i>u-ya-sebe-sebent-is-a</i>	'you are using'
	not <i>*u-ya-seba-sebent-is-a</i>	

For Downing, the essential observation is that verbal forms consisting of the root plus the vowel *a* (*dlal-a*, *cebul-a*, *bong-a*, *phuphut-a*) are canonical stems. Indeed, they appear to represent the citation form for every verb, probably because they do not contain any identifiable inflectional or derivational affixes. Based on this fact, Downing proposes a correspondence condition that requires the



reduplicant to be identical to the canonical stem. It is the interaction between this condition of correspondence to the canonical stem and other conditions on the shape of the reduplicant that provide the analysis for this data.

The alternative we consider here relates to Downing's ideas but employs **Ident (word)**, a constraint that can generalize to the reduplication patterns discussed earlier. We consider forms like *dlala*, *cebula*, *bonga*, *phuphuta* to be surface allomorphs of the verb: the final vowel *a* signals here no specific morphemic function or feature. One may argue about the underlying presence of this vowel, if one's theory of underlying representations excludes predictable elements from lexical entries, but this is immaterial here, since we are talking about *surface* root allomorphs<sup>37</sup>. We assume, following in this Downing, that every Swati verb root possesses at least one listed allomorph and this listed allomorph is of the form Root-a, the citation or canonical form. The generalization that suggests itself as an alternative to the canonical stem analysis is that Red2 (e.g. *-bonga-bongisa*) represents the pattern in which preference is given to a reduplicant that is strictly identical to *some* listed allomorph, not necessarily a canonical one. On the other hand, the more common Red1 (*-bongi-bongisa*) is the pattern in which the preference for relative similarity to the adjacent allomorph overrides that for total identity to a listed allomorph. In other words we propose to allow the ranking between **Ident (Word)** and **Ident-adjacent** to vary freely; both are outranked by Red = 2σ, the disyllabism condition. We evaluate **Ident-adjacent** in the same way as above (74).

(77) *-bonga-bongisa*: listed allomorphs of the verb *bonga* include: *bongisa*, *bonga* but not *bongi-*

	Red = 2σ	>>	Ident (Word) >>	Ident-adjacent
bonga-bong-is-a				*(i), *(s), *(a)
bongi-bong-is-a			*!	*(s), *(a)
bongisa-bong-is-a	*!			

(78) *-bongi-bongisa*: listed allomorphs as above.

	Red = 2σ	>>	Ident-adjacent >>	Ident (word)
bongi-bongisa			*(s), *(a)	*
bonga-bongisa			*(i), *(s), *(a)	

<sup>37</sup> It is in any case not always possible to eliminate predictable information from lexical entries: all Sanskrit roots represent strings amounting to one syllable. The monosyllabism is therefore predictable but doesn't correspond to any one property that can be harmlessly taken out of lexical entries.

We observe next that either ranking between **Ident (word)** and **Ident (adjacent)** is sufficient to exclude Red2 forms like *\*-khula-khuluma-a* (cf. 76): the restricted behavior of polysyllabic roots falls out naturally from the analysis:

(79) *-khulu-khuluma-a*: listed allomorphs of *khuluma* include *khuluma* but not *\*khula*, *\*khulu*

	Red = 2σ	>>	Ident (word)	>>	Ident-adjacent
khulu-khuluma			*		*(m), *(a)
khula-khuluma		*			*(u), *(m), *(a)
khuluma-khuluma	*!				

	Red = 2σ	>>	Ident -adjacent	>>	Ident (word)
khulu-khuluma			*(m), *(a)		*
khula-khuluma			*(u), *(m)!, *(a)		*

Further data analyzed by Downing (1997a) indicate that inflectional suffixes are strictly forbidden within the reduplicant. Thus *ba-bik-ile* 'they have reported' contains the past tense *-ile* and can reduplicate only according to the Red2 pattern: *ba-bika-bik-ile*, *\*ba-biki-bik-ile*. Unlike the derivational suffixes like causative *-is-* (cf. *-bongi-bong-is-a*), inflectional affixes must thus be kept out of reduplication. The nature of the constraints having this result remains unclear but it seems in any case unlikely that the reference to canonical stem will help here, as Downing argues. The other two languages discussed by Downing (1997b, c: Kikerewe and Kinande) possess each conditions that exclude certain classes of morphemes from the scope of reduplication. These conditions differ from the one active in Siswati and hence cannot be deduced from any generalized hypothesis about the structure of the canonical stem.

Kinande, one of the other languages analyzed by Downing (1997b; cf. also the original study of Mutaka and Hyman 1990), differs from Swati primarily in ranking **Ident (word)** strictly above **Ident-adjacent**. This results in paradigms such as the following, where the benefactive *eri-hum-ir-a* has no choice but to reduplicate as *eri-huma-hum-ir-a*.

(80) Kinande verbal reduplication (Mutaka and Hyman 1990, Downing 1997b)

Verb form	Reduplicated	Gloss
<i>eri-hum-a</i>	<i>eri-huma-huma</i>	'to beat'
<i>eri-hum-ir-a</i>	<i>eri-huma-hum-ir-a</i> (* <i>eri-humi-hum-ir-a</i> )	'to beat for'
<i>eri-hum-an-a</i>	<i>eri-huma-hum-an-a</i>	'to beat each other'

*eri-hum-ir-an-a*      *eri-huma-hum-ir-an-a*      'to beat each other for'  
 (\**eri-humi-hum-ir-an-a*)

The data so far is consistent with **Red = 2s (>>) Ident (word) >> Ident-adjacent**

We note further a class of forms which is consistent only with this interpretation: these involve disyllabic subjunctives like *(tu-)hum-e* which always reduplicate in their entirety. As Downing points out, longer subjunctives like *(tu)-hum-ir-e* adopt a different pattern: note that only the reduplications indicated below are acceptable in Kinande.

(81) Reduplication of Kinande subjunctives (Mutaka and Hyman 1990, Downing 1997b)

	Verb form	Reduplicated	Gloss
(a)	<i>tu-hum-e</i>	<i>tu-hume-hum-e</i>	let's beat
	<i>tu-tum-e</i>	<i>tu-tume-tume</i>	let's send
	<i>tu-sw-er-e</i>	<i>tu-swere-sw-er-e</i>	let's grind for
(b)	<i>tu-hum-ir-e</i>	<i>tu-huma-hum-ir-e</i>	let's beat for
	<i>tu-tum-ir-e</i>	<i>tu-tuma-tum-ir-e</i>	let's send for

This data is accounted for by adding to the ranking established a condition forbidding affixal material in the reduplicant (**\*Aff-in-Red**): the overall hierarchy is therefore **Red = 2s (>>) Ident (word) >> Ident-adjacent >> \*Aff-in-Red**. Note that the final *-a* vowel present in most verbs is not an affix since it fails to convey morphosyntactic features of any sort. (It is a member of the **\*Aff-in-Red** family that probably accounts for the Swati restriction mentioned earlier: recall *ba-bika-bik-ile*, *\*ba-biki-bik-ile*.)

(82) verb forms of *huma*'s paradigm include *huma*, *hume*, *humira*, but not *humi-*

	Red = 2σ (>>)	Ident (word) >>	Ident-adjacent >>	*Aff-in-Red
-huma-humira			*(i), *(r), *(a)	
-humi-humira		*!	*(r), *(a)	*
-hume-humira			*(i), *(r), *(a)	*!
-humira-humira *!				*

	Red = 2σ (>>) Ident (word) >>	Ident-adjacent >>	*Aff-in-Red
-hume-hume			*
-huma-hume		*!	

	Red = 2σ (>>) Ident (word) >>	Ident-adjacent >>	*Aff-in-Red
-huma-humire		*(i), *(r), *(a)	
-hume-humire		*(i), *(r), *(a)	*!
-humi-humire	*!	*(r), *(e)	*

The analysis adopted here differs more significantly from that proposed by Downing: reliance on the concept of canonical stem identity leads Downing to split **MAX Base-RED** (corresponding to one of the components of **Ident-adjacent** here) into two distinct constraints differing only in their mode of evaluation, a distinction that seems hard to defend independently.

For derived stems the generalizations discussed by Mutaka and Hyman (1990) seem to indicate idiosyncratic variation in the ranking between **Red = 2s** and **Ident (adjacent)**, in the sense that each lexical item is characterized by one or the other of the rankings definable between these constraints: thus *eri-gambula* 'to talk' reduplicates as *eri-gamba-gambula*, as required by **Red = 2s, Ident (word) >> Ident (adjacent)** whereas *eri-bindula* 'to change' reduplicates as *eri-bindula-bindula*, a result obtainable via **Ident (word) >> Ident (adjacent) >> Red = 2s**. Note that the lexically conservative condition **Ident (word)** remains undominated throughout: every reduplicant is identical to some, independently occurring, verbal form.

Downing's discussion of reduplication in passive and causative verbs is also highly relevant. This data seems largely consistent with the analysis presented here but appears to involve a preference not so far detected for identity between the last syllable of the adjacent stem and the last syllable of the reduplicant:

(83) Reduplication of Kinande causatives (Mutaka and Hyman 1990, Downing 1997b)

	Verb form	Reduplicated	Gloss
(a)	<i>eri-bul-y-a</i>	<i>eri-bulya-bul-y-a</i> <i>*eri-bula-bul-y-a</i>	to ask
(b)	<i>eri-bul-ir-ya</i>	<i>eri-bula-bul-ir-ya</i> <i>eri-bulya-bul-ir-ya</i>	to ask for
(c)	<i>eri-bul-ir-an-y-a</i>	<i>eri-bula-bul-ir-an-y-a</i>	to ask for each other

The existing ranking predicts the unique reduplicant in (a) and the *eri-bula-* reduplicants of (b) and (c): the evaluation of *eri-bulya-bul-y-a* will be identical to that of *-hume-hume* in (82). We consider the other variants below. For the moment, let us note again an effect of lexical conservatism here: as Downing notes, "the causative *-y-* only occurs in RED if that variant of the causative is possible for the Base stem" (1997b: 7). Thus *bulya* is a possible reduplicant because *bulya*, in its entirety is an actual verb. In verbal paradigms like (84), which lack the Root-*y* form of the causative, i.e. for verbs where the causative *-y-* follows always some other suffix, reduplication differs:

(84) Reduplication in Kinande causatives lacking the Root-*y* form of the causative:

Causative forms	Reduplicated	Gloss
<i>eri-hum-is-y-a</i>	<i>eri-huma-hum-is-y-a</i>	to cause to beat
* <i>eri-hum-y-a</i>	* <i>eri-humya-hum-is-y-a</i>	

Rejecting *\*eri-humya-hum-is-y-a* is clearly the effect of **Ident (word)**: since *\*humya* does not exist independently as a word in *huma*'s paradigm, it cannot exist as a reduplicant either. The only remaining option is then the attested *eri-huma-humisya*. Our analysis still needs to explain why *eri-bulya-bulirya* and *eri-bulya-buliranya* are possible forms, in addition to *eri-bula-bulirya*, *eri-bula-buliranya*, the only ones we allow so far. (On this point, Downing's analysis remains unclear.) It appears that for this set of data calls for two modifications in the analysis. First, **Ident-Adjacent** must be broken down, as anticipated, into a set distinct corespondence conditions: so far the only real function of **Ident-Adjacent** in Kinande has been to insure that the last vowel of the stem is also the last vowel of the reduplicant. It is this effect that we need in order to guarantee, for instance, *tu-hume-hume*, as against *\*tu-huma-hume*. In the case of *eri-bulya-buliranya*, the acceptable alternative to *eri-bula-buliranya* we note that the rightmost vocoid string *-ya* is identical in the reduplicant and the stem. Suppose that the **Ident-Adjacent** condition that is active in Kinande is that of identity between the right (Glide)V edges of the reduplicant and the adjacent stem. The GV string can perhaps be identified with the nucleus. We refer to this interpretation as **Ident-Adjacent (GV)**. The second necessary modification appears to the ranking of **Ident-Adjacent (GV)** relative to **\*Aff-in-red**: the causatives appear to allow these two constraints to be freely ranked, unlike the subjunctives. The reasons for the ranking variation, which is more widely attested in Kinande, remain mysterious under all analyses proposed so far, including ours.

(85) existing forms for *bula* include: *bula*, *bulya*, *bulira*, *bulirana*, *bulirya*, *buliranya*, not *\*buli-*

	Ident (word)	>>	*Aff-in-red	Ident-Adjacent (GV)
	bula-buliranya			*
	bulya-buliranya		*	
	buli-buliranya	*!	*	*

(86) existing forms for *huma* include: *huma, humira, humirana, humisya*, but not *\*humya, humi*

	Ident (word)	>>	*Aff-in-red	Ident-Adjacent (GV)
	huma-humisya			*
	humya-humisya	*!	*	
	humi-humisya	*!	*	*

Note that the acceptability of forms such as *bulya-buliranya* indicates that the identity to the canonical stem plays, if any role, a minor one: *bulya* is not a citation form and cannot be considered canonical in any definable sense. The undominated condition in Kinande is however Ident (word), the requirement that every reduplicant be identical to some paradigmatically related word. This is a condition of lexical conservatism of the sort discussed throughout this study.

## 8. Conclusions

Let me summarize the proposals made here. The paradigm of any morpheme contains a pool of listed surface allomorphs. When we compute the phonological properties of a novel form based on this morpheme, the listed allomorphs compete to impose their own phonological properties upon the target allomorph. The outcome of the competition is decided only in part by the phonotactics and, in the cases considered here, not at all by the underlying representation<sup>38</sup>. A significant share of the final decision appears to rest with the degree of similarity between properties required of the target allomorph and properties present in any one of the listed allomorphs. The notion of phonological base of affixation reduces to the statement that morphosyntactically similar allomorphs prefer to become phonologically similar. There is no single reference point in computations of this sort, rather all listed allomorphs appear to be relevant in the process. The degree to which they are real contenders depends on their similarity to properties that are desirable in the target. Finally the linguistic creativity of speakers is limited by grammatical conditions that impose a measure of lexical conservatism on novel formations.

<sup>38</sup> See Flemming 1995 and Burzio 1997 on the conjecture that underlying representations should be dispensed with in favor of surface correspondence conditions. I note only that cases of the sort considered here and in the cyclicity-analogy literature are unlikely to shed light on this issue, since they show only that surface-to-surface relations are a necessary ingredient in the analysis of correspondence, not also a sufficient one.

## Incomplete References

S. Anderson 1992 *A-morphous Morphology*

M.Aronoff 1992

M.Aronoff 1995 *Morphology by itself*

X.(?) Arrivé, Gadet, Galmiche

L.Burzio 1997 "*Vincitori e vinti: participial morphology and correspondence theory*", LSRL, Irvine.

F.Dell 1974 *Les regles et les sons*, Paris, Hermann

P(?) .Encrevé 1988 *La liaison avec et sans enchainement*

E.Flemming UCLA Diss

W.Johnson 1987 "Lexical levels in French phonology," *Linguistics*, 25 (5).

J.Jun 1995 UCLA diss.

M.Kenstowicz 1995 "Base Identity and Uniform Exponence," MIT ms.

M.Kenstowicz and T.Greene 1995 "The Lapse constraint," MIT ms.

E.Klein 1994 in PASC 3

R. Lieber 1981 MIT diss.

J. McCarthy 1995 "Rotuman revisited" UMass ms.

J.McCarthy and A.Prince 1995 "Faithfulness and reduplicative identity," in J. Beckman et al. UMOP

Y.-C. Morin 1992 "Un cas méconnu de la déclinaison de l'adjectif français: des formes de liaison de l'adjectif antéposé," in André Clas (ed.) *Le mot, les mots, les bons mots*, Les Presses de l'Université de Montréal

O. Orgun 1996 UC Berkeley diss.

D.Perlmutter 1996 "Interfaces: Explanation of Allomorphy and the Architecture of Grammars" to appear in S.Lapointe et alii (eds.) *Morphology...*, CSLI.

A. Prince 1983 "Relating to the grid", LI.

- J.-F. Prunet 1987 "Liaison and nasalization in French," in C. Neidle and F. Nuñez-Cedeño (eds.) *Studies in Romance Languages*, Dordrecht, Foris.
- E. Selkirk 1984 *Phonology and Syntax*, MIT Press
- D. Steriade 1995 "Positional Neutralization" UCLA ms.
- D. Steriade 1996 "Paradigm uniformity and the phonetics phonology boundary", *Labphon* 6
- B. Tranel 1976 "A generative treatment of the prefix *in-* of modern French, *Language* 52 (2)
- B. Tranel 1981 *Concreteness in Generative Phonology*, UC Press Berkeley.
- B. Tranel 1990 "On Suppletion and French Liaison," in *Probus*, reprinted in P. Hirschbuhler and K. Koerner (eds.) *Romance Languages and Modern Linguistic Theory*, J. Benjamins