Lexical conservatism and the notion *base of affixation*

There are at least three distinct interpretations of the statement in (1):

(1) "*indentify* is the base of *identifiable.* "

On the first interpretation, *identify* is the morphosyntactic base of the derived word *identifiable*: this statement is justified by the fact that a morphosyntactic property of *identify* - the fact that it is a verb - is a prerequisite for the affixation of *-able*. On a second interpretation, *identify* is the semantic base of *identifiable*: the semantics of the *-able* form are a function of those of the verb contained within it. Finally, a third interpretation of the statement in (1) has to do with the phonology of the *-able* form. The phonological shape of the complex word is a function of the shape of its inner constituent. In this case, and many others, the morphosyntactic and semantic base of the complex expression are the same as the phonological base. If one considers the extended lexical paradigm of *identify (identity, identification)* one observes that the phonology of the *-able* form can be accounted for only by reference to the verb *identify.* And this suggests that the morphosyntactic base is of necessity identical to the phonological base of the word. It is for this reason that the normal use of statements like (1) conflates the two or three senses that I have distinguished above.

The first point of this paper is to show that the distinction above has empirical consequences and therefore that the usual conflation of the concepts of phonological and morphosyntactic or semantic base cannot be maintained. This can be shown by pointing to the existence of a large class of cases in which complex expressions are formed by reference to several distinct reference terms. I call this *the split-base effect*. For instance, I will show that in order to form a complex adjective such as *remédiable*, the speaker must consult not only the verb *rémedy*, which serves as its morphosyntactic base, by satisfying *-able's* subcategorization needs, but also the adjective *remédial*, which serves its phonological base, in that it lends its stress pattern to the *-able* form. Why does the stress pattern of the adjective intrude upon that of the verb *rémedy* in what is a clearly deverbal formation, *remédiable*? The answer offered here will be that by adopting the stress pattern of *remédial*, the *-able* form is improved stress-wise: the lapsed string is diminished. I will argue that split-base effects arise always when the morphosyntactic base of affixation lacks a phonological property that is desirable in the derivative: in such cases, the derivative may adopt the phonology of a distinct listed allomorph.

More generally, I will suggest that the process by which speakers compute the phonological properties of complex expressions engages in principle *all* members of a given lexical paradigm, or even larger sections of the lexicon, rather than some one unique base form. And I will propose an explanation for the existence of the split-base effect which relates to and subsumes better known lexical phenomena such as Aronoff-style blocking (Aronoff 1976). The empirical discussion will bear mostly on French adjectival liaison and the morphology of English *-able* affixation, but effects of lexical conservatism can be documented more widely.

My second goal is to outline an analytical framework for the characterization of phonological correspondence that can extend to split base effects. Note that most current work on correspondence, as well its precursors in derivational frameworks, assume a strictly bilateral base-derivative relation: a unique base for each derivative. It is non-trivial to adjust the system in such a way as to permit an explicit characterization of formations with multiple bases[1].

2. Listedness

I would first like to clarify a question relating to the contents of the lexicon. If the derived form *remédial* contributes its stress pattern to the nonce form *remédiable*, then this implies that *remédial* is accessible to the speaker for purposes of on-line word formation. Exactly what is accessible in this way needs to be spelled out. The assumptions I make in this respect are adopted, with some modifications, from Halle (1973). Any non-nonce word, any non-hapax form is, I assume, accessible as a base of affixation for the creation of a novel form. In other terms, I assume that any non-nonce form is lexically recorded, as Halle had anticipated. One clear motivation for this assumption is the very fact that speakers have relatively reliable intuitions of listedness which are independent of the intuition of grammaticality: speakers can tell whether words like *matchable* or *identifiable* are pre-existing, established words or not and their linguistic behavior is affected by this distinction, as shown by Bradley 1980, Anshen and Aronoff 1978 and others. I'll use the term *listed* here: an accessible form is listed, not just in the sense of being familiar but in the sense that its linguistic use is sanctioned by the past use of other speakers. Other terms conveying the same thing are *received, institutionalized, item* (as against *type*) *familiar* (Bauer 1983). It is likely that the distinction between listed words and nonce formations is not categorical, so that simply placing some words in the mental lexicon and excluding others may not be a realistic representation of this distinction. For instance, use frequency differences may insure that the properties of certain forms are readily accessible under most circumstances to most speakers, while other forms might not be, even though the speaker may recognize them as familiar when presented with them: *necromancy* is familiar to me but hardly present in my consciousness and I might hesitate about stressing it, while *cartomancy* was unheard-of until encountered in a dictionary. In contrast with both of these, *pestilence* is both listed and reasonably familiar: I would not hesitate about its stress pattern. For simplicity, however, I will ignore the gradient aspect of the distinction between listed and unlisted forms and treat is as categorical.

A distinct source of possible bases for the purpose of on-line word formation are *potential* expressions, those whose existence is guaranteed by principles of word formation: for instance, all English *-id* adjectives have *-ídity* derivatives: *lánguid* implies potential *languídity*. Similarly, all English *-id* nouns have *-idic(al)* derivatives: *ephémerid* implies potential *ephemerídical*. Such forms as *langúidity*, or although unlisted, may occasionally function in the same way as the listed forms, in their ability to lend their phonological properties to novel words. I will argue, for instance, that the nonce form *languídify* ('to render lánguid') is likely to be stressed as indicated precisely because the speaker is aware of the potential form *langúidity*, which is in turn guaranteed by a fact of English word formation: the implicational relation between *-id* and *-idity*. By contrast, we know that not all English nouns possess adjectival *-al* derivatives: *párody* does not imply *paródial*, even though comparable pairs like *cústody - custódial* do exist. We will observe that occurring *custódial* endows its paradigm with certain properties that are

lacking in the paradigm of *párody*, precisely because a form like *paródial* is both unlisted and not implied by any principle of English word formation.

It is important to note that a listed word, in the sense adopted here, may have fully predictable morphological and phonological properties: thus the stress of *remédial* is predictable from the segmental shape of *rémedy* and the stress properties of the *-al* suffix, but nonetheless the fact that *remédial* is stressed as it is a lexically listed fact and thus may have an effect on word formation. This means then that even predictable properties of listed items may be lexically recorded, which implies a difference regarding the evidentiary value of listed words as against nonce formations. When we look for evidence of grammatical principles internalized by speakers we must rely primarily on the study of nonce or unlisted forms, since it is these forms whose properties must be generated de novo rather than being conjured in pre-compiled form from a mental dictionary. Consequently, I may ignore the sound patterns of listed forms to the extend that they diverge systematically from those of nonce fomations. In this respect, the evidence considered below may differ fundamentally from that used in previous studies, such as Aronoff (1976) and Burzio (1994).

3. Split base effects in French liaison

I begin by outlining the argument for split-base effects in French adjectival liaison. French masculine adjectives take on a special form when they occur in potential hiatus, as (2) shows. This form - which I call the liaison allomorph - is frequently identical to the citation feminine (cf. Tranel 1981, Perlmutter 1996 and Tranel 1996).

(2) French liaison: masculine determiner forms used to avoid hiatus

masc. citation form	fem. citation form	masc. liaison form
[nuvo]	[nuvEl]	[nuvEl a)]
nouveau	*nouvel le*	*nouvel an*
'new-masc' 'new-fem'		'new year'
[sœ]	[sEt]	[sEt a)drwa]
ce	*cette*	*cet endroit*
'this-masc'	'this-fem'	'this place'
[bo]	[bEl]	[bEl Om]
beau	*belle*	*bel homme*
'beautiful-masc'	'beautiful-fem'	'handsome man'
[p´ti];	[p´tit]	[p´tit a)fa)]
petit	*petite*	*petit enfant*
'little-masc'	'little-fem'	'little child'

However the liaison form is in fact distinct, for many French speakers, from the citation feminine, in the case of forms like those in (3). This is systematically the case with adjectives whose stressed vowel is predictably distinct in

feminine and masculine forms: in such cases, the masculine vowel appears in the masculine liaison form, followed by the feminine consonant, which is being used as hiatus buffer.

(3) Split base effects in Liaison: feminine C, masculine V in final VC

	masc. citation	fem. citation	masc. liaison form
(a)	[so]	[sOt]	[sot ami]
	sot	*sotte*	*sot ami*
	'silly-masc'	*'silly-fem'*	*'silly friend'*
(b)	[vE)]	[vEn]	[vE)n EspwaÂ]
	vain	*vaine*	*vain espoir*
	'vain-masc'	'vain-fem'	'vain hope'
(c)	[dEÂnje]	[dEÂnjEÂ]	[dEÂnjeÂ Om]
	dernier	*dernière*	*dernier homme*
	'last-masc'	'last-fem'	'last man'
(d)	[paÂfe]	[paÂfEt]	[paÂfet amuÂ] '
	parfait	*parfaite*	*parfait amour* 'perfect-masc' perfect-fem'
	'perfect love'		

The phonological basis for the vocalic alternations in (3) is outlined below, following Tranel 1981 and Dell 1974:

(4) Phonological basis for masculine-feminine alternations:

(a) A word final [O] disallowed: [so] *[sO]

(b) Tense o disallowed in closed syllables: [sOt], *[sot].

(c) Nasals disallowed in same syllable as nasal vowels: [vEn], [vE)] *[vE)n]

(d) Tense [e] disallowed in closed syllables: [dEÂnjEÂ], *[deÂnjeÂ]

Note that the liaison forms in (3) satisfy the conditions in (4), if we count the final consonant as resyllabified into the following syllable: thus *sot ami*, if syllabified *so.ta.mi*, satisfies (4.b). However reference to the conditions in (4) is not sufficient to explain why the masculine vowel appears in the liason forms: both [dEÂ.nje.Â Om] and [dEÂ.njE.Â Om] are consistent with (4.d); similarly both [vE.n Es.pwaÂ] and [vE.)n Es.pwaÂ] satisfy (4.c). The reason why the masculine vowel appears in these forms is that the adjective is, syntactically speaking, a masculine form: the vowel

quality encodes this fact. The reason why the feminine consonant appears in the liaison form is, as indicated, hiatus avoidance.

Before turning to the analysis, we should note the forms in (5), which provide just some of the reasons why we cannot rethink the masculine liaison allomorphs in terms of preservation of the underlying form. The arguments in this case are drawn from Tranel (1990, 1992, 1996) and Perlmutter (1996). The paradigms in (5) - from Tranel (1990) and Morin (1990) - show that adjectives ending in VCC - which alternate between VCC citation feminines and VC citation masculines - use their masculine form VC in liaison. In such cases, hiatus is not an issue since both the masculines and the feminines end in consonants. It is important to note that in such cases only *similarity to the gender-appropriate surface citation form matters*: the masculine liaison form is identical to the citation masculine.

(5) Lexical conservatism effect in liaison is not an input preservation effect

- *court espace de temps* 'brief period': [kuÂ Espas], *[kuÂt Espas]
  UR /kurt/; feminine citation [kuÂt]; masculine citation [kuÂ].
- *perverse attrait* 'perverse attraction': [pEÂvEÂ atÂE], *[pEÂvEÂs atÂE] UR /pEÂvEÂs/; feminine citation [pEÂvEÂs]; masculine citation [pEÂvEÂ]

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 general characterization of allomorphic choices in pre-V position involves hiatus avoidance: 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]. This is a case of suppletive allomorphy, because we cannot claim that either /ma/ or /mon/ represent the 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. The same point is made à propos of other suppletive pairs 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.

We have observed that the masculine liaison form is a blend of the surface properties of the citation feminine and masculine: the feminine consonant and the masculine vowel. This blend is used in cases where the feminine consonant can be used as a hiatus buffer. This then is the split base effect.

We can now ask: why do split base effects exist? Why must the French liaison form be a blend of masculine and feminine properties? It seems clear that the masculine vowel is used in these forms to encode, even partially, the grammatical gender of the adjective but what is not clear is why it should be the feminine C, as against some fixed inserted C, that is used to avoid hiatus.

The answer here is lexical conservatism. By this I refer to a set of grammatical preferences against the use of forms that are phonologically novel or lack lexical precedents. (6) outlines the role of lexical conservatism in the French case. The situation there is that hiatus must be avoided and that it cannot be avoided at the expense of the vowels (6.a-b). Therefore we must find a buffer C. Lexical conservatism dictates that *this C must be found in a pre-existing form belonging to the adjective's paradigm*. This C must have a precedent in some familiar form. Such a pre-existing form is the citation feminine. It's pre-existing or listed in the sense that speakers know citation forms, whereas they might not know the liaison forms, especially for unfamiliar, or less frequent adjectives or for adjectives that are not customarily preposed[2]. Thus the reason for blending feminine and masculine properties in the liaison form is that the feminine provides a phonologically desirable property - hiatus avoidance - while at the same time satisfying lexical conservatism, the need to avoid linguistic innovation.

(6) Summary and rationale for the split-base effect in liaison:

(a) **Avoid hiatus**: **dern[je Om]*.

(b) **Preserve stressed and initial** V: **dern[j Om], *derni[e m]*

(c) **Lexical conservatism**: every property of the novel form must find a precedent in a listed form.
Final C in liaison allomorph must find a lexical precedent in the adjective's paradigm, **dern[jet Om]*.

(d) **Phonological encoding of morphosyntax**: *der[njEÂ Om] fails to encode masculine gender.

(e) **Split-base form** *der[nje.Â Om]* satisfies conditions (a)-(b)-(c)-(d).

A further consequence of the lexical conservatism effect is that no hiatus breaking C is possible in forms where both the feminine and the masculine end in a vowel: thus *joli enfant* 'pretty child (fem. and masc.)' [Zoli A)fA)], is realized with hiatus, not with C insertion *[Zolit A)fA)] etc.

We can now sketch the formal analysis of the split base effect in French liaison. We begin by characterizing generally the class grammatical conditions that models lexical conservatism. This characterization appears in (7) [3]:

(7) The form of lexical conservatism conditions: Lex (P)

Let T (μ) be the allomorph of μ appearing in a form under evaluation.

Let L(μ) be a listed allomorph of μ.

Let P be a phonological property.

T(μ) is characterized by P only if *some* L(μ) is characterized by P.

The statement of lexical conservatism conditions in (7) adopts certain of the properties of McCarthy and Prince's 1995 correspondence conditions. Central among these is the notion of *correspondence relation* between some property P in one form and its not-necessarily-identical counterpart P' in a related form. The difference between Lex(P) constraints and McCarthy and Prince's correspondence conditions involve the range of phonological properties for which correspondence is sought and the mode of evaluation of lex P conditions: P in Lex P may be any phonological property,

including global properties like overall stress pattern, syllabic shape or global identity; second, Lex (P) evaluates candidates by searching the lexicon for some, non-specific, listed allomorph of μ that possesses P, whereas McCarthy and Prince's Ident F or MAX/DEP evaluation involves a bilateral comparison between target and UR or target and a specific surface representation.

In the French case several distinct lexical conservatism conditions play a role in the final analysis. The fundamental one is the condition in (8) which requires the final consonant of the liaison allomorph to possess a lexical precedent - a featurally identical counterpart - among the listed allomorphs of the adjective.

(8) Lex C]
The absolute final C in the target allomorph of morpheme μ has a correspondent C' in some listed allomorph of μ and is featurally identical to C'.

(a) listed allomorphs: *joli*

	Lex C]	*Hiatus
joli enfant		*
jolit enfant	*!	

(b) listed allomorphs: [vE)], [vEn]

	Lex C]	*Hiatus
i. [vE)n] espoir		
ii. [vE)] espoir		*!
iii. [vEn] espoir		

The assumption that I will make - here following Tranel's and Perlmutter's work - is that the citation forms of an adjective (both the masculine and the feminine) are listed allomorphs. To know a French adjective one must know the rough shape of both of these forms, even if details of their pronunciation may be predictable by general principles. The masculine liaison form may also be a listed allomorph - at least for some speakers - to the extent that the adjective is common in prenominal position or occurs in lexicalized phrases before a vowel initial noun (as, for example, *bon homme* 'good man', *bel homme* 'handsome man' . Otherwise however, the shape of the liaison allomorph must be computed on line as any sandhi variant. Going beyond earlier proposals, I will also assume that the listed allomorphs may contain as part of their lexical representations certain predictable properties, such as the laxness or tenseness of their vowels. This assumption is made necessary by the fact that the masculine vowel quality, even when entirely predictable, may be imposed upon the liaison allomorph, as in the case of *sot ami* [so.ta.mi].

Returning to (8.b) , we note now that this tableau contains in fact two winners: [vE)n] *espoir* and [vEn] *espoir*. Both pronunciations are attested but I will now discuss the speech of those where the one in (i) prevails. This is the pronunciation in which the split base effect is manifest, because the feminine C and the masculine V are combined in one form. Informally, (i) is preferred because it is reminiscent of a masculine form, and it thus partially encodes the required gender agreement. I will assume that the gender agreement condition is not at stake in liaison: whether or not the liaison form sounds like masculine, syntactically it is a masculine. What is at issue is only whether the masculine gender on the adjective is or is not phonologically encoded. Therefore what we need is a class of conditions that promote encoding morphosyntactic properties through phonological identity to listed forms that indubitably possess those properties. This class is schematized in (9), under the name of Lex P-M:

(9) Lex P-M conditions

Let T (μ) be the allomorph of μ appearing in a form under evaluation.

Let M be a morphosyntactic feature required in T(μ).

Let L(μ) be a listed allomorph of μ.

Let P be a phonological property.

T(μ) is characterized by P if some L(μ) is characterized by P and M.

listed allomorphs: [vE)], [vEn]

	Lex 'V gender
i. [vE)n] espoir	
iii. [vEn] espoir	*!

The two candidates in (10) are the two winners of the earlier tableau in (8.b). The Lex 'V-gender condition differentiates them: [vE)n] *espoir* passes - since the stressed [E)] of [vE)n] is identical to that of the citation masculine [vE)], while [vEn] *espoir* fails. Note also that there is a consideration that favors [vEn] *espoir* - namely the fact that the adjective in this phrase is identical *in its entirety* to a listed allomorph, the citation feminine. Indeed some French speakers accept only forms of this type: [vEn] *espoir,* [sOt] *ami,* [dEÂnjEÂ] *homme,* etc. This shows that global or maximally stringent conditions of lexical conservatism will also have a role to play in the analysis and that they conflict with the class of Lex PM conditions. For the moment we will say only that speakers who prefer [vEn] *espoir* rank a global condition of Lexical conservatism Lex σ' ((*The stressed syllable of the allomorph under evaluation must be identical in its entirety to the corresponding stressed syllable of a listed allomorph*) above Lex 'V-gender, whereas the speakers whose patterns we have focussed on so far rank Lex 'V-gender above Lex σ'.

4.Interim conclusions

Let me comment on what we have seen so far. In the customary sense of base, the base of a liaison masculine is either the citation masculine, a surface form, or the underlying representation of the adjective. We have seen that the underlying representation is not the relevant lexical reference term for computing the liaison allomorph. But we have also seen that the citation masculine is not alone in deciding the shape of novel liaison form - the feminine contributes as well. The feminine's contribution is to provide the C that helps avoid hiatus, while satisfying lexical conservatism. So the root of the split-base effect in this case is the conflict between phonological wellformedness, the encoding of morphosyntactic features, like gender, and lexical conservatism: split-base effects represent one way of adjudicating the conflict between these desiderata. We can generalize our observations as follows: suppose that two properties are desirable in a novel

form of some morpheme μ, a phonological property P and a morphosyntactic property M. Suppose also that the morpheme μ has a set of listed allomorphs, as described below, such that none possesses both P and M, although some possess P and some possess M:

(11) Listed allomorphs of μ:
μ_1: has M, does not have P
μ_2: has P, does not have M
μ_3: does not have P, does not have M

The split-base effect is the choice of realizing the novel form of μ by adopting P - the desired ph- property - from 2 and some phonological property identifying μ_1, through which the presence of M - which characterizes μ_1 - is encoded.

Note that the analysis proposed so far does allow us to characterize the customary use of the term base, namely *the morphosyntactically appropriate base*, as in (12), without falsely implying that the morphosyntactic base is always the unique lexical reference term in word or phrase formation.

(12) Standard notion *base form* (or *morphosyntactically appropriate base form*) defined:
Base = a listed allomorph L(μ) possessing morphosyntactic features required for T(μ) and not possessing

Base identity conditions invoked recently in the literature on surface correspondence correspond, in the present context, either to Lex P or, more frequently, to Lex PM conditions taken individually or in combination. For instance, consider briefly the Levantine interaction of stress and syncope (Brame 1974, Kager 1996, Kenstowicz 1995, Steriade 1996) whereby the existence of forms such as 3rd sg. *fíhim* 'he understood' block syncope in derivatives such as *fihímna* 'he understood us' (**fhímna*). In this case, we may assume that the basic subject-inflected forms, (e.g. *fíhim*) are lexically listed and that Lex PM conditions help predict the shape of the object inflected forms (e.g. *fihímna*) based on them. The Lex PM condition relevant here states that the allomorph of μ (here μ = *fhm*) under evaluation must be phonologically similar to a listed allomorph of μ with which it shares tense, aspect, and subject agreement features: it must be identical, specifically, in that the stressed vowel of the listed allomorph must find a featurally identical counterpart in the allomorph under evaluation. This condition will have the effect of blocking the syncope of base-stressed vowels in the unlisted derivative. One advantage of the present re-formulation of base-identity conditions is that it spells out the difference between the so called base and other listed allomorphs in a given lexical paradigm: previous work left unspecified just what identifies *fíhim* as the base of *fihímna*. However, the fundamental point we make here is that *all* listed allomorphs are in principle capable of playing a role as phonological bases, whether or not they can serve as morphosyntactic bases.

Let us consider now what gives rise to the appearance of single-base derivations, the cases in which the phonological base is the same as the morphosyntactic or semantic one. This situation has a variety of possible sources. One of them is the existence of an impoverished paradigm: thus if a French adjective has exactly one listed allomorph (e.g. *immonde* 'filthy' or *joli* 'pretty') rather than two, there will be no occasion of mixing phonological properties of the feminine and masculine in the liaison form. Alternatively, if no conflict obtains between between phonological well-formedness and a Lex PM condition, then again no opportunity will arise for a Lex P constraint distinct from Lex PM to assert itself. A third source of single-base situations involves cases in which stricter requirements of phonological identity of target to morphosyntactic base outrank conflicting phono-constraints: for instance some French speakers prefer [so ami] to [sot ami], [sOt ami] etc., thus displaying a willingness to violate hiatus if this allows them to unambiguously mark the masculine gender. Finally, global, as against piece-meal, Lex P conditions will also militate against the split-base effect: a Lex P condition such as Lex σ' (*The stressed syllable of the allomorph under evaluation must be identical in its entirety to that of a listed allomorph*) will prohibit mixed forms like [vE)n] *espoir* 'vain hope' and will select either [ve§)] *espoir* (using the masculine form, to satisfy Lex PM) or [vEn] *espoir* (using the feminine to avoid hiatus)

But, to reiterate, it is important to view the split-base phenomenon not as pathological (as implied by traditional terms like *contamination*) but as one of several analytical options predicted by the framework.

5.Lexical conservatism in English level 2 phonology

I now turn to instances of English lexical conservatism. This is a more complex situation, as it appears that different affixes are lexically conservative in their phonology, but to different degrees. My primary goal is just to show that split base effects occur in English and that their analysis follows the system of constraints developed so far: it involves the conflict between phonological well-formedness, Lex P conditions of lexical conservatism and Lex PM conditions, which encode morphosyntactic and semantic relations between lexical items through phonological similarity. The full reanalysis of English Level 2 phonology along the lines suggested here cannot be undertaken in this study, for reasons of space (cf. Steriade 1998).

Level 2 affixes are defined, since Siegel 1974, by the invariance of stresses in the stem, relative to the pronunciation of the stem in isolation: for instance *ínvalidism* is recognized as a Level 2 form because its stem does not differ accentually from that of *ínvalid*. This accentual resemblance between base and derivative is obtained at the cost of metrical deviance: *ínvalidism* contains a string of 4 stressless syllable, more than normally tolerated in English. But the same suffix *-ism* can generate accentually modified forms, as in *bureáucratism*, which differs stresswise relative to *búreaucrat*. Similarly *-able* generates Level 2 formations, mostly, as in *éducable* (on *éducate*; not **edúcable*) but occasionally it does generate apparent Level 1 formations, such as *demónstrable*, with shifted stress relative to the verb *démonstrate*. and *remédiable*, which differ similarly from *rémedy*. My point now will be to suggest that both the level 1/2 distinction and this level heterogeneity phenomenon (noted earlier by Aronoff 1976) are in fact just reflexes of lexical conservatism.

Let us compare first the *-able* forms based on two very similar verbs, *parody* and *remedy*. Both verbs are transitive and can yield corresponding *-able* forms, but their extended paradigms differ, as shown in (13): *rémedy* has both a dactylic listed allomorph and an amphibrachic one, which appears in *remédial*. *Parody*, has for all intents and purposes only the dactylic allomorph[3].

(13) *Parody* vs. *remedy* as bases for *-able*

Listed dactylic allomorph (σ!σσ)	Listed amphibrachic allomorph (σσ!σ)
remedy [rE!m´|i]	[r´mi!di-], as in *remedi-al*
parody [pœ!r´|i]	----

This difference in paradigmatic structure results in differences in the results of *-able* affixation. Item (14) records these differences. The speakers consulted were asked to read aloud (or, in some cases, record in writing) lists of *-able* adjectives based on each verb. The predominant responses in each category are in bold characters. The speakers consulted treated these as nonce forms: most were initially uncertain as to how the forms should be stressed.

(14) *-able* forms based on *remedy* and *parody*

stress pattern wrt base	based on *rémedy*	based on *párody*
same as base: σ!σσ	*rémedi-able (3/24)*	***párodi-able*** *(21/22)*
change to σσ!σ	***remédi-able*** *(W; 20/24)* [r´mi!di´bl`]	*paródi-able (0/22)* [p´ro!di´bl`]
change to σ~σσ!	*rèmedí-able (1/24)*	*pàrodí-able (1/22)*

Sources of data:

W = Webster's Third Int'l Unabridged 1966

S_1 = survey of 22 speakers (1996-1997) who read aloud lists of -able forms. Numbers refer to responses with a given stress pattern/total number of responses. Some speakers provided more than one pronunciation, so total no. of responses may exceed 22.

The first generalization we can infer from this data is that long sequences of stressless syllables (*Lapse violations: cf. Prince 1983 Kager 1989, Green and Kenstowicz 1995 ms.) are to be avoided. The rightward shift of stress is therefore desirable before the unstressable *-able*. It is for this reason that *remédiable* is generally preferred to *rémediable*. To make this precise we have to distinguish several varieties of Lapse: based on the **number** of stressless sylables (more is worse), on the **location** of lapsed string (final lapse better tolerated than medial) and on syllable **size** (onsetless offends less). Thus *remédiable* (3 stressless syllables one onsetless) is better stressed than *rémediable* (4 stressless, one onsetless) or than *erádicable* (3 stressless, all with onsets). We assume here distinct ***Lapse** conditions corresponding to these effects, though the optimal analysis may have to reflect the continuous nature of the variation. The *Lapse σσσ condition used below penalizes only sequences of three stressless syllables with onsets.

The second generalization involves lexical conservatism: a change in [±stress] status of sylables relative to the verbal base is acceptable only when there is lexical precedent for the shifted stress pattern in the same lexical paradigm: a listed allomorph with same [±stress] distribution. We formulate the relevant condition, preliminarily, as the Lex P condition in (15):

(15) **Lex [±stress]**: For any stressed syllable σ in the target form T(μ), there is a correspondent σ' in some listed allomorph, L(μ), such that σ' is stressed and the nuclear vowels of of σ and σ' are identical.

The ranking Lex [stress] >> *Lapse σσσ is illustrated below:

(16) a.listed allomorphs: include *rémedy, remédi- in remédial*

	Lex [±stress]	*Lapse σσσ
i. remédiable		
ii. rémediable		*!

b.listed allomorphs: include *párody*, not *paródial*

	Lex [±stress]	*Lapse σσσ
i. párodiable		*
ii. paródiable [p´rodi´bl`]	*!	

Some speakers who have encountered of *paródic* [p´rA!dIk] accept *paródiable* [p´rA!di´bl`], but this form violates CiV Tensing (No stressed non-high lax V before CiV) and thus is not generally acceptable. Nuclear identity between

, deviates in two ways from the listed allomorph: by removing stress from the initial syllable and by stressing the final.

The data in (14) indicates that there do exist speakers who prefer *rémediable*. For these, we must invoke a Lex PM condition whose effect will be to require that the verbal nature of the stem of *-able* be signalled through accentual similarity to a listed verbal stem. By listed verbal stem I mean one occurring in a recognizable verb: thus *civilize* [aIz] is a verbal stem because it occurs in the verb but *civiliz* [Iz] (as in *civilization*) is not. The vowel shortening is associated with a change of lexical category. The relevant Lex PM condition appears in (17) .

(17) Lex ([±stress], lexcat):

For any vowel V in T(μ), V has a correspondent V' in a listed allomorph of μ, L(μ), and is identical wrt [±stress] to V', if L(μ) and T(μ) are identical wrt lexical category.

listed allomorphs: *remédi-* (Adj.), *rémedy* (Vb)

	Lex [±stress], lexcat
remédiable	*
rémediable	

The ranking of Lapse σσσ relative to Lex [±stress, lexcat] must obviously differ across speakers, as indicated below:

(18) Lex [±stress] >> *Lapse σσσ >> Lex ([±stress], lexcat): *remédiable*
Lex [±stress] (>>) Lex ([±stress], lexcat) >> *Lapse σσσ : *rémediable*

A distinct effect of Lex ([±stress, lexcat) is to exclude non-verbal allomorphs from the *-able* stem in cases where *Lapse σσσ (or other phonotactics) are moot. An instance of this sort appears below: both *describe-* and *descript-* are listed allomorphs, both are final stressed, hence equally able to avoid Lapse violations, but only the unambiguously verbal *describe* is used with *-able*[4]:

(19) listed allomorphs: *descríbe* (Vb.), *descrípt-* (N or Adj.)

	*Lapse σσσ	*Lex ([±stress, lexcat])
descríbable		
descríptable		*!

Next, we need to establish that the lexical conservatism effect is fully general for *-able* formations. I have done this by using a large corpus of *-able* words based on the 1700 *-able/-ible* forms in Lehnert's 1971. Among these, I selected all *-able/-ible* forms possessing independently occurring polysyllabic bases with final stressless syllable in verbal stem (a total of 186 words). Thus *hospitable* was excluded because it lacked an independently occurring base; *avoidable, identifiable* were excluded because the last syllable in the verbal form (*avóid, idéntify~*) is stressed. On the other hand, *compensable* was not excluded, since [p´n] is stressless in the verbal allomorph contained in *cómpensàte.* The corpus of 186 *-able/-ible* forms thus created was divided into the three categories described below and illustrated in (21):

(20) Classification of *-able* corpus by paradigmatic properties

(a) **homogeneous paradigms**: *-able* words based on acentually invariant paradigms, with the same distribution of stressed and stressless syllables in all allomorphs of the stem

(b) **heterogeneous paradigms**: *-able* words based on accentually variable paradigms, in which different allomorphs of the stem have different [±stress] distributions.

(c) **pseudo-heterogeneous paradigms**: *-able* words whose stem has variably stressed allomorphs but whose meanings diverge significantly so that each semantically congruent subparadigm is in fact accentually uniform.

(21) Examples of the paradigmatic classification of *-able* forms

Homogeneous	**Heterogeneous**	**Pseudo-heterogeneous**
(160 forms)	(17 forms)	(9 forms)
•*execrable*	*illustrable*	*integrable*
(only *éxecrate*)	(both *íllustràte*	(*íntegràte*; unrelated
	and *illústratìve*)	*intégrity*)
•*administrable*	*compensable*	*concentrable*
(only *admínister*)	(*cómpensàte*	(*cóncentràte*, unrelated
	and *compénsatory*)	*concéntric*, *céntral*)

The analysis presented earlier predicts that stress in the homogeneous *able* forms will remain unchanged and forms with modified stress will be disfavored. For instance, we predict that speakers will accept *éxecrable* despite the lapse, and will reject**exécrable*, since there is no precedent for the *exécr-* stress pattern within this paradigm; similarly, speakers will accept only *admínistrable*, not accentually improved **adminístrable*. For heterogeneous paradigms, the analysis predicts that speakers who can access the non-verbal allomorphs with rightwards shifted stress will adopt the stress of that allomorph in the *-able* form. For instance in forming the *-able* form based on *óbfuscàte* speakers who can access *obfúscatory* may adopt its stress pattern and prefer *obfúscable*; speakers who can access *illústrative* (when producing the *-able* form on *íllustràte*) may opt for *illústrable*; and those who access *prégnant* may say *imprégnable*, on *ímpregnàte*: both the dictionary and survey data confirms these predictions[5], as we see below. Our prediction however is not that all speakers who access *illústrative* will opt for *illústrable*: we simply predict that for this class of forms either shifted stress (relative to the verbal base) or unshifted stress is possible. These two options correspond to the two rankings anticipated in (18): *Lapse σσσ >> Lex ([±stress], lexcat) (yielding) and Lex ([±stress], lexcat) >> *Lapse σσσ (yielding *rémediable*). Only speakers tolerating or preferring *remédiable* are predicted to tolerate or prefer *illústrable*. Finally we predict that stress in pseudo-heterogeneous *-able* forms is less likely to be based on the non-verbal allomorph than in a heterogeneous paradigm: for instance if the speaker focusses on the semantic difference between *íntegràte* and *intégrity*, then their answer is likely to be *íntegrable*; if however this difference is not attended to, then we may expect *intégrable*.

We now compare these predictions with relevant data available in dictionaries (Webster's, Daniel Jones's and Harrap's), earlier linguistic work (Aronoff 1976) and data obtained through the survey outlined in (15). Item (22) provides data illustrating the behavior of homogeneous paradigms: as predicted, the *able* forms maintain the [±stress] distribution found in the rest of the stem's paradigm.

(22) Homogeneous *-able*: unchanged [±stress] pattern

able forms based on truncated verbs ending in *-σ!σ** ate (sources: W or S_1) *ábdicable, prédicable, prognósticable, állocable, nóminable, éducable, anníhilable, póllinable, dédicable, própagable,*

able forms based on verbs ending in -σ!σ¤ (sources: Aronoff 1976 and S_1)

bállastable, challengeable, lícenceable, flúmmoxable, pátentable

able forms based on truncated verb ending in -σ!σ* σ* ate (source S_1)

péregrinable

• -*able* forms on verbs ending in -σ!σ* σ* (sources: Aronoff 1976 and S_1)

dísciplinable, próphesyable

The only exceptions to this pattern (encountered during the subject survey) involve the occasional stress shift in forms with heavy presuffixal syllable (*corúscable, persíflable, extírpable*): in these cases however the base verb was either unknown to the subjects or very infrequent. My conjecture is that infrequent - or never heard - *coruscate, persiflate, extirpate* exert little pressure on the accent of the -able adjective, which is therefore free to satisfy metrical well-formedness conditions.

The results concerning heterogeneous -*able* forms appear below: the numbers under S_1 refer to the ratio of accentually changed to unchanged responses. Thus the figure 22/4 next to *remédiable* indicates that 22 responses differed stresswise relative to the verb *rémedy*, while 4 did not: this means 22 *remédiable* to 4 *rémediable*.

(23) Heterogeneous -*able* : [±stress] pattern based on non-verbal allomorph

	démonstrate, demónstrative	(S_1: 24/1; W, DJ, H)
compénsable	cómpensate, compénsatory	(S_1: 22/0; W)
equílibrable	equílibrate, equilíbrium	(S_1: 17/5)
sequéstrable	séquestrate, sequéster	(S_1: 22/0, W)
remédiable	rémedy, remédial	(S_1: 20/4, W, DJ)
incúlpable	ínculpate, cúlpable	(W)
obfúscable	óbfuscate, obfúscatory	(S_1: 19/5, W)
confíscable	cónfiscate, confíscatory (físcal)	(S1: W, DJ)
imprégnable[6]	ímpregnate, prégnant	(W)
illústrable	íllustrate, illústrative	(W)
contémplable	cóntemplate, contémplative	(S_1: 22/1, W)

The forms in (24) were identified too late for inclusion in S_1: the predicted stresses (for speakers who accept forms like *remédiable* and who can access the non-verbal allomorphs listed) are as indicated below:

(24) **remónstrable** rémonstrate, remónstrance
influénceable (ínfluence, influéntial)
indícable (índicate, indícative)

One class of heterogeneous paradigms did however behave unexpectedly:

(25) Heterogeneous -able forms with accentual and final C alternations

attríbutable (attríbute, attribútion)	(W)
contríbutable (contríbute, contribútion)	(W, DJ)
recíprocable (recíprocate, reciprócity)	(S_1: 8/1)
domésticable (domésticate, domestícity)	(S_1: 8/0, W)
intúitable (intúit, intuítion)	(W)
óbligable (óbligate, oblíge)	(W)

These are analyzed below. Data illustrating the behavior of pseudo-heterogeneous forms appears in (26). It is difficult to discern a pattern in this case, in large part because the number of relevant forms is small.

(26) Pseudo-heterogeneous *-able* :

désignable (désignate, sígnatory)	(W)
íntegrable (íntegràte, intégrity)	(W)
cóncentrable (cóncentrate, concéntric, céntral)	(W)
assímilable (assímilate, simílitude)	(S1: 22/0, W)
intérrogable (intérrogate, ìnterrógative)	(W)
incórporable (incórporate, corpóreal)	(W)
infítrable (ínfiltràte, fílter, fíltrable)	(S1: 19/5)
expúrgable/ible (éxpurgate, púrge, púrgatory, púrgative)	(S1: 20/2)
custódiable ((take) cústody, custódial, custódian)	

While the bulk of the data supports the analysis presented earlier, the data in (25) requires an interesting revision. We have to ask why, given the attested pre-suffixal stress in forms like *intuítion*, the verb *intúit* cannot give rise to an *-able* form such as *intuítable*; similarly, given *domestícity*, with stress shifted rightwards, why is *domestícable* impossible (from *domésticate*)? Note that all heterogeneous forms raising this question have verbal allomorphs whose last stem consonant differs from that of the non-verbal allomorph with rightwards shifted stress: for instance, compare *intúi[t]* with non-verbal *intuí[S]ion*. Why then is a consonantal alternation of this sort an impediment to the adoption

of the non-verbal stress pattern? We suggest that two distinct preferences play a role here: first, speakers prefer to adopt the last stem consonant of the *-able* adjective from the corresponding verbal allomorph, hence *domesti[k´bl`]*, not **domesti[s´bl`]*, *attribu[t´bl`]*, rather than **attribu[S´bl`]*, and *intui[t´bl`]* not *intui*[S´bl`]*. The identity of the last stem C is one way to signal the fact that the stem allomorph to which *-able* is affixed is verbal, and thus that it satisfies the suffix's subcategorization requirements. The relevant Lex PM condition appears below:

(27) **Lex (C], lexcat):** The final consonant C in T(μ) has a correspondent C' in a listed allomorph L(μ) and is featurally identical to C' if T(μ) and L(μ) have the same lexical category.

Lex (C], lexcat) is sufficient to rule out **attri[bju!S´bl`]* but fails to exclude forms such as **attri[bju!t´bl`]*, which combines the nominal stress of *attribútion* with the verbal consonantism of *attríbute*. I attribute the rejection of such forms to the preference for adopting the stressed rime (or syllable) **in its entirety** from a single given allomorph. The condition in (28) requires that the stressed syllable of the novel form find, in its entirety, an identical counterpart in some listed allomorph. Thus (28) is a Lex P condition, the pure expression of lexical conservatism:

(28) **Lex s '**: For any stressed syllable σ in the target form T(μ), there is a correspondent σ' in some listed allomorph, L(μ), such that σ' is stressed, and σ and σ' are segmentally identical.

Both (27) and (28) outrank *Lapse σσσ, as shown below: relevant stressed syllables in each candidate appear in bold characters.

(29) Ranking Lex (C], lexcat) and Lex σ' above Lapse

	Lex (C], lexcat) Lex	*Lapse σσσ
i. ´**trI!b**jut´bl`		*
ii. œ~tr´**bju!S**´bl`	*!	
iii. œ~tr´**bju!t**´bl`	*!	

With these additions, the analysis of *-able* as a lexically conservative affix is confirmed.

We have substantiated here the initial claim that nonce-words forms like *remédiable* are formed on two distinct bases: the verb *rémedy* and the adjective *remédial*. The independent existence of *remédial* permits lapse avoidance in *remédiable*. Where such an end-stressed allophone is lacking, the desirable stress pattern cannot be adopted: so, in the absence of *paródial*, **paródiable* is impossible. We can also show that the verb *rémedy* also plays a role in *remédiable*: the independent existence of (verbal) *rémedy* is what allows *-able* affixation in the first place. Paradigms

lacking a verbal form do not possess *-able* derivatives: for instance**melódiable* 'which can be provided with, turned into melody' **paróchiable* 'which can be made parochial', **testimóniable* 'which can be turned into a testimony or testimonial' are all ill-formed. This judgment is supported by the observation that among the 1700 -able forms in Lehnert's only one (*kítchenable,* considered uninterpretable by speakers consulted) belongs unambiguously to a paradigm lacking a verb. In that sense, the presence of the verb *rémedy* is an indispensible factor in *-able* affixation, even if its phonological properties are not being carried over in their entirety to the *-able* form.

Also notable is the fact that the lexical conservatism analysis adopted for split-base effects in French extends to the *-able* data. Generalizing, we may state that split-base effects emerge under any ranking of the form in (30):

(30) {Lex P, Phono-constraint on P} >> Lex PM

In English, one instantiation of this ranking schema is Lex σ'>> *Lapse σσσ >> Lex ([±stress], lexcat). In French, the relevant schema is Lex C] >> *Hiatus >> Lex (σ', gender). Effects of the same sort can be enountered in reduplicative constructions and lend themselves to parallel analysis (Steriade 1997 ms.)

We have also observed that the accentual properties of *-able* forms reflect lexical conservatism, not boundary or level differences. Thus, a level 2 (#-boundary) analysis predicts only **démonstrable, rémediable;* while a level 1 (+-boundary) analysis predicts **challéngeable, edúcable* (cf. Aronoff 1976). Finally, a mixed-level analysis (Aronoff's own) fails to predict the correlation between paradigm structure and accentual optimization: it does not recognize the central fact that *only an independently existing stress pattern from the same paradigm can replace that of the verbal base*. It is this fundamental observation that supports the notion of lexical conservatism which lies at the core of the split base phenomenon.

Notes

1. On correspondence within Optimality Theory, see McCarthy and Prince 1995. On correspondence between surface forms: Burzio 1993, 1994 , 1997 LSRL, Benua 1995, Flemming 1995, Kenstowicz 1995, 1996, McCarthy 1995, Steriade 1995, 1996.

2. Hiatus avoidance is a strong effect in Specifier-Noun constituents and less so in Noun-Adjective structures: hence the difference between, for instance *savant anglais* [savA) A)gle] 'an English scholar' (a Noun-Adjective structure) vs. *savant Anglais* [savA)t A)gle] 'a scholarly Englishman' (an Adjective-Noun structure). The difference in liaison possibilities is due to the fact that Spec-Noun structures represent a single phonological phrase while Noun-Adjective constituents can be phrased as two separate phonological phrases. The phrasing difference can be independently tested by considering the possibility of a pause or an an intonational rise: this exists only between the noun and the postposed adjective.

3.The form *paródial* is listed in Webster's (1966) Unabridged but is not known to any speaker consulted; *paródic* is extremely infrequent (it is missing from most major dictionaries: Lehnert's *Rückläufiges*, Harrap's and D.Jones) and unknown to most speakers. See below.

4.Aronoff (1976: 120ff) notes forms such as *defénsible, percéptible, defénsible* which contain, on his analysis, the Level 1 (+ boundary) *able*. This use of the non-verbal allomorph is limited to lexically listed forms, as indicated by the fact that that the following are grossly ill-formed: **offensible* (cf. *offénsive; offendable*), **erosible* (cf. *errosion; erodable*), **inclusible* (cf. *inclusive; includeable*), **decisible* (cf. *decisive; decideable*). Forms like *defensible* have counterparts in French and/or Latin and there is no tendency to extend the pattern within English.

5. In testing these predictions of the analysis, we were able to take advantage of the possibility of truncating *-ate* before *-able*, noted by Aronoff (1976). Non-truncated forms such as *obfuscateable* are irrelevant here, because their stem-final syllable (-ate-) is stressed in the base word. Many speakers, however, are clearly reluctant to drop *-ate*; thus truncated forms such as *confíscable* (or *cónfiscable*) were uttered under protest by several subjects in the survey.